Technische Universität München
TUM School of Medicine and Health

# Opportunities and challenges for digital interventions in the prevention of depression

Janika Thielecke

Vollständiger Abdruck der von der TUM School of Medicine and Health der Technischen Universität München zur Erlangung einer

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Joachim Hermsdörfer

Prüfer*innen der Dissertation:

1. Jun.-Prof. Dr. Claudia Buntrock
2. Prof. Dr. Antonius Schneider

Die Dissertation wurde am 20.04.2023 bei der Technischen Universität München eingereicht und durch die Fakultät für Sport- und Gesundheitswissenschaften am 28.09.2023 angenommen.

Für Oma.

*"Which is more important," asked Big Panda*
*"the journey or the destination?"*
*"The company." said Tiny Dragon.*

— James Norbury[1]

## ACKNOWLEDGMENTS

I once read it would be a dangerous business to step out of my door because I could not know where I would be swept off too[2]. When I am thinking about all the people — named and unnamed — I met and who accompanied me on this PhD journey, I am more than happy that I made that step.

My most heartfelt thanks belongs to Jun.-Prof. Dr. Claudia Buntrock. Claudia, you have been guiding me on this way from the beginning and along all the ups and downs that come with a long journey. You inspire me with your ability to advise, listen and give the room for personal development. I am grateful for your support over the years and enjoy learning from you.

I would like to express my gratitude to my examination committee, Prof. Dr. Hermsdörfer and Prof. Dr. Schneider, for supporting and accompanying me on the last steps of this PhD project as well as to the TUM Department of Sport and Health Sciences' Graduate Center for their support.

I am deeply grateful to Prof. Dr. D. David Ebert. Thank you, David, for giving me the opportunity to pursue my PhD in your team and setting me up for meaningful work. Your vison on how to improve mental health care by technology is truly inspiring.

This journey started with the project on farmers' mental health. Huge thanks goes to the whole SVLFG-Team and especially to Johanna, Ingrid and Jenny for the inspiring teamwork on a day-to-day basis. Many thanks also goes out to our student assistants and collaborators at the University of Ulm and at the SVLFG without whom this project would not have been possible.

A big thanks goes out to the whole ProtectLab, live in the offices in Erlangen and Munich and in countless online meetings. It makes me proud to be part of team that is so empowering and supporting each other, no matter what. Special thanks goes to Anna-Carlotta – I appreciate that you are always open for discussion and giving new perspectives on topics. Further, Fanny and Christina, I am grateful for

1 Norbury, J. Big Panda and Tiny Dragon. UK, Middlesex: Michael Joseph. 2021. Instagram: @bigpandaandtinydragon

2 Tolkien, JRR. The fellowship of the ring. USA, New York, NY: HarperCollins. 1991

your warm welcome and your hospitality when my way led me to Munich. Another big thanks goes to you, Mathias and Lea for motivating me to embrace R and especially to Paula for the great collaborations in the analyses. Moreover, thank you, Sandra, for keeping up the Rhineland humor in the office and for the long nights of writing or crafting.

I would further like to thank the colleagues from the Open Science Brown Bag Meeting at the Psychology Department in Erlangen who sparked my interest in open science and motivated me to improve my own research to be more transparent and open.

In my experience, every journey profits from good preparation and the courage to face the unknown. I deeply want to thank my family, for preparing me for my journey, supporting me on the way and celebrating the milestones with me. You set me up to succeed on this path by teaching me to believe in myself and by always doing the same.

Lastly, thank you for joining my journey, Robin. I could not have wished for a more patient, encouraging and supportive travel companion – on this journey and many more to come.

# PUBLICATIONS

This dissertation is based on the following published articles and submitted or prepared manuscripts:

Article 1:
Thielecke J, Buntrock C, Titzler I, Braun L, Freund J, Berking M, Baumeister H, and Ebert DD. Clinical and Cost-Effectiveness of Personalized Telephone Coaching for Farmers, Foresters and Gardeners to Prevent Depression: Study Protocol of an 18-Month Follow-Up Pragmatic Randomized Controlled Trial (TEC-A). Frontiers in Psychiatry. 2020; 11:125. doi: 10.3389/fpsyt.2020.00125

Article 2:
Thielecke J, Buntrock C, Titzler I, Braun L, Freund J, Berking M, Baumeister H, and Ebert DD. Telephone Coaching for the Prevention of Depression in Farmers: Results from a Pragmatic Randomized Controlled Trial. Journal of Telemedicine and Telecare. 2022; online first. doi: 10.1177/1357633X221106027

Article 3:
Thielecke* J, Titzler* I (*shared first authorship), Braun L, Ebert DD, Freund J, Baumeister H, Berking M, and Buntrock C. Does telephone coaching improve farmers' mental health in the long term? Results of the 12- and 18-month follow-up of a pragmatic randomized controlled trial. In preparation

Article 4:
Thielecke J, Buntrock C, Freund J, Braun L, Ebert DD, Berking M, Baumeister H, and Titzler I. How to Promote Usage of Telehealth Interventions for Farmers' Mental Health? A Qualitative Study on Supporting and Hindering Aspects for Acceptance and Satisfaction with a Personalized Telephone Coaching. Internet Interventions. submitted

Article 5:
Thielecke J, Kuper P, Ebert DD, Cuijpers P, Smit F, Riper H, Lehr D, and Buntrock C. Does Outcome Expectancy Predict Outcomes in Online Depression Prevention? Secondary Analysis of Randomized Controlled Trials. Prevention Science. submitted

Article 6:
Thielecke J, Kuper P, Schuurmans L, Harrer M, Ebert D D, Lehr D, Behrend D, Brückner H, Horvath H, Riper H, Cuijpers P, and Buntrock C. Who benefits from indirect prevention and treatment of depression using online interventions for insomnia? Results from an individual-participant-data meta-analysis. In preparation

Reprint Permissions can be found in Appendix A.

## SUMMARY

Major depressive disorder (MDD) is a common psychological disorder associated with a high individual and societal burden, increased mortality, and immense healthcare costs. Although evidence-based treatments exist, structural barriers (e.g. limited service providers, long waiting times) and attitudinal barriers (e.g. perceived stigma, preference to solve one's problems alone) limit their use and effectiveness, and they cannot counter all depression burden.

Preventive measures are highly warranted, and interventions targeting individuals with elevated symptoms of depression who do not meet all criteria for MDD (i.e. subthreshold depression) have shown promising effects in reducing depressive symptoms and preventing or at least delaying the onset of MDD. With the rise of modern technology in every aspect of modern life, new possibilities for preventive interventions have emerged.

Digital mental health interventions (DMHIs) have the potential to improve mental healthcare accessibility by overcoming structural and attitudinal barriers to mental healthcare use. They can in most cases be used independently of time and location, eliminate the need for travel, and can be more anonymous and scalable. Additionally, DMHIs can be more easily tailored to individual needs and risk profiles, and be adapted to different language and cultural contexts.

Nevertheless, several challenges remain in general healthcare and arise with new technology. Implementing new interventions remains slow, and uptake and use in routine care are low. Skepticism about effectiveness, privacy, and security persists in clinicians and users. From a research perspective, not enough is known about which individuals will most likely benefit from DMHIs and what factors predict intervention outcomes, which limits the potential of DMHIs in reducing depression burden.

This dissertation aims to demonstrate the opportunities of DMHIs by focusing on personalized telephone coaching and web-based interventions and addressing some of the challenges still faced by DMHIs. To showcase the opportunities, in Articles 1 – 3, I report on the randomized controlled trial (TEC-A) that evaluates the effectiveness of personalized telephone coaching in farmer and related professions. The personalized telephone coaching could effectively reduce depressive symptom severity and improved secondary mental health outcomes such as stress, anxiety, and quality of life. However, the generalizability of the effectiveness of telephone coaching is limited by the high degree of personalization and the restriction to one occupational group.

To support implementation and address the challenge of low uptake, aspects that might influence acceptance of and satisfaction with personalized telephone coaching were investigated in Article 4. Recommendations for further implementation were derived from qualitative content analysis of 20 transcribed interviews with participants who completed their coaching. These included preserving the logistical simplicity, the high degree of personalization and flexibility, as well as expanding personalization by giving more choices of media and training coaches in occupation-specific topics.

The other challenges addressed in this dissertation, namely unknown predictors of treatment outcomes and differential treatment effects, were investigated in more standardized web-based interventions, which offer unique opportunities to study these aspects. In Article 5, the role of outcome expectancy as a potential predictor of treatment outcome is explored. Outcome expectancy is an established common factor in classic psychotherapy which supports treatment effectiveness, but less is known about how it transfers to prevention or online contexts. Results hint at a possible delayed effect on outcome in individuals with higher outcome expectancy using a web-based intervention for depression prevention. Since more methodologically sound studies are needed in the new context before building on it to enhance the effectiveness of DMHIs, a research agenda for exploring outcome expectancy in web-based prevention is presented.

Finally, in Article 6, I aimed to establish potential differential effects of the emerging paradigm of indirect prevention and treatment of depression. In an indirect approach, a common, less stigmatized problem like sleep is addressed, and by improving this can affect depression. In an individual participant data meta-analysis of four web-based sleep intervention trials, the indirect approach was shown to reliably reduce depressive symptoms. Multivariable moderation analysis revealed that from a wide range of sociodemographic, clinical and work-related variables, only depressive symptom severity at baseline showed to moderate the treatment effect.

Overall, all studies support the idea of using DMHIs for reducing depressive symptom severity. However, no effects on MDD onset were observed across studies, most likely due reduced sample sizes and statistical power. Methodological implications for prevention research as well as future directions for the field are discussed in the last part of this dissertation.

# CONTENTS

## LIST OF FIGURES



## LIST OF TABLES



## ABBREVIATIONS

| | |
|---|---|
| ANCOVA | Analysis of covariance |
| AQoL-8D | Assessment of Quality of Life |
| AUDIT-C | Alcohol Use Disorder Identification Test – Consumption Questions |
| BDI-II | Beck Depression Inventory II |
| BfArM | German Federal Agency for Drugs and Medical Devices |
| BPD | Bipolar disorder |
| CBT | Cognitive Behavioral Therapy |
| CBT-I | Specialized CBT for insomnia |
| CEQ | Credibility and Expectancy Questionnaire |
| CES-D | Center for Epidemiological Studies Depression Scale |
| CG | Control Group |

| | |
|---|---|
| CI | Confidence Interval |
| CIDI | Composite International Diagnosis Interview |
| CONSORT | Consolidated Standards of Reporting Trials |
| COREQ | Consolidated Criteria for Reporting Qualitative Research |
| COS | Core Outcome Sets |
| CSQ-8 | Client Satisfaction Questionnaire |
| CSQ-I | Adapted CSQ-8 for internet interventions |
| DiGA | Digitale Gesundheitsanwendung (Engl. Digital Health Application) |
| DMHI | Digital mental health intervention |
| DRKS | German Clinical Trial Register |
| DSM | Diagnostic and Statistical Manual of Mental Disorders |
| DSS | Depressive symptom severity |
| ERI-S | Effort-Reward Imbalance Scale – Short form |
| E | Expectation |
| FAIR | Research data principle: findability, accessibility, interoperability, and reusability |
| FDA | Food and Drug Administration |
| GAD | Generalized Anxiety Disorder |
| GAD-7 | Generalized Anxiety Disorder Screener |
| GLMM | Generalized Linear Mixed Models |
| HF | Hindering factor |
| HR | Hazard Ratio |
| IAPT | Improving Access to Psychological Therapy |
| iCBT | internet-delivered CBT |
| ICD | International Statistical Classification of Diseases and Related Health Problems |
| ICER | Incremental Cost-Effectiveness Ratio |
| ICHOM | Consortium for Health Outcomes Measurement |
| IG | Intervention Group |
| INEP | Inventory for the Assessment of Negative Effects of Psychotherapy |

| | |
|---|---|
| IPD | Individual participant data |
| IRR | Incidence Rate Ratios |
| ISI | Insomnia Severity Index |
| ITT | Intention-to-treat principle |
| JITAI | Just-in-Time Adaptive Mechanisms |
| MAP | Maximum a posteriori |
| MAR | Missing At Random |
| MBI-GS | Maslach Burnout Inventory - General Survey |
| MBI-CY | MBI Cynicism |
| MBI-EE | MBI Emotional Exhaustion |
| MBI-PE | MBI Professional Efficacy |
| MCAR | Missing Completely At Random |
| MCID | Minimal clinically important difference |
| MDD | Major depressive disorder |
| MDE | Major depressive episode |
| MI | Multiple imputation |
| MICE | Multiple imputation by chained equation |
| NICE | National Institute for Health and Care Excellence |
| NMAR | Not Missing At Random |
| NML-P | Nijmegen Motivation List for Prevention |
| NNH | Number-Needed-to-Harm |
| NNT | Number-Needed-to-Treat |
| OR | Odds Ratio |
| OSF | Open Science Framework |
| PAS | Panic and Agoraphobia Scale |
| PHQ-9 | Patient Health Questionnaire |
| PMM | Predictive Mean Matching |
| PSS-10 | Perceived Stress Scale |
| PWP | Psychological Wellbeing Practitioner |
| QIDS-SR16 | Quick Inventory of Depressive Symptomatology |
| RCI | Reliable Change Index |
| RCT | Randomized controlled trial |
| RoB | Risk of bias |
| RQ | Research question |
| RR | Risk Ratio |

| | |
|---|---|
| RWD | Real world data |
| RWE | Real world evidence |
| SCID | Structured clinical interview |
| SF | Supporting factor |
| SPE | Subjective Prognostic Employment Scale |
| SSS-8 | Somatic Symptom Scale-8 |
| SVLFG | German Social Insurance Company for Agriculture, Forestry and Horticulture |
| TAU | Treatment-as-usual |
| TAU+ | Enhanced treatment-as-usual |
| TiC-P | Trimbos Institute and Institute of Medical Technology Questionnaire for Costs Associated with Psychiatric Illness |
| UTAUT | Unified Theory of Acceptance and Use of Technology |
| UWES | Utrecht Work Engagement Scale |
| WAI-SR | Working Alliance Inventory - Short Revised |
| WAI-SRT | Working Alliance Inventory - Short Revised – Therapist |
| WLC | Waiting list control (group) |
| ZPP | Zentrale Prüfstelle Prävention (Engl. German Prevention Course Registry) |

# Part I

# BACKGROUND

# 1

# INTRODUCTION

*Melancholia*, nowadays better known as *depression*, is a broad term to describe a mental health condition of lowered mood that has been known to humankind since ancient times [1]. Over time our understanding of the condition has changed and developed from a theorized "unevenness of bodily fluids" to a more scientific bio-psycho-social model of mental health [2]. However, even in our digital age of rapid information exchange, smartphones and artificial intelligence, we still have not fully understood this condition comprehensively or found the best way how to treat it [3]. *Major depressive disorder* (MDD) is formally defined in the current Diagnostic and Statistical Manual of Mental Disorders (DSM-5) as showing either depressed mood or loss of interest for most of the days over two weeks and at least four additional symptoms such as changes in weight, sleep, activity level and cognitive processes. The symptoms have to cause distress or impairment in functioning and cannot better be explained by other conditions in order to present a *Major depressive episode* (MDE) [4].

As a psychological disorder, MDD is associated with high individual [5, 6] and societal burden [7, 8], increased mortality [9, 10] and immense health care cost [11, 12]. In terms of years lost to disability, depression is the largest single-disease contributor worldwide, accounting for approximately 7.5% of all years lost to disability [13]. Lifetime prevalence ranges from 6.1% (Japan) to 21.0% (France) in high-income countries with a prevalence of 9.9% estimated for Germany [14].

Evidence-based treatments for MDD are available in the form of psychotherapy and pharmacology, however, in primary care only half of MDD cases are recognized as such [15] and only an estimated 20% of individuals with MDD receive specialized mental health care [14]. This is not solely due to structural barriers like limited health care availability, because even in a hypothetical scenario of full coverage and compliance to evidence-based treatments, only one-third of disease burden attributed to MDD could be averted [16]. Attitudinal barriers like the preference to help onesself, perceived ineffectiveness of treatment options and stigma of mental health problems and help-seeking are especially important barriers for individuals with mild-moderate problem severity [17]. Even with the observed improve-

ments in treatment and availability in the last decades, prevalence in the general population has not decreased [18]. In consequence, preventive measures are warranted to face this global mental health challenge [19].

Preventive approaches can be distinguished by their main target group. Universal prevention targets the whole population, regardless of individual characteristics. Selective prevention, on the other hand, targets individuals with known risk factors (e.g. based on personal history, occupation or social group) while interventions for indicated selection target individuals with elevated symptom levels [20]. Individuals with elevated symptoms of depression who do not fulfill all criteria for MDD are considered cases of *subthreshold depression* [21]. Even though subthreshold depression is associated with fewer symptoms than MDD, the condition is also accompanied with poorer quality of life [21, 22], increased mortality [23, 9], high economic costs [24] and health care services use [25] as well as a high risk of developing a full MDE [26, 27]. Psychological interventions for the indicated prevention of depression have been shown to reduce depressive symptom levels in subthreshold depressed individuals (g = 0.35 [95%–CI: 0.23, 0.47]) [28]. Further, they can prevent (or at least delay) onset of depression with a reduced risk of MDD onset by 19% (Risk Ratio = 0.81 [95%–CI: 0.72, 0.91]) [29] and selected and indicated psychological interventions are likely to be cost-effective in comparison to treatment-as-usual (TAU) [30].

## 1.1 THE OPPORTUNITIES FOR DIGITAL MENTAL HEALTH INTERVENTIONS IN THE PREVENTION OF DEPRESSION

As our understanding and treatment of depressive symptoms changed over time, so did the technology available to us. With the rise of technology and the progress of digitalization, new options for treatment and prevention of depression emerged ranging from call-in radio shows with psychological counsellors in the 1980's [31] to just-in-time adaptive intervention on a smartphone based on personal risk factors emerging today [32]. Over time, the field of remotely delivered (psychological) interventions developed and the term *digital health* can be employed to describe the *"use of information and communications technologies to improve human health, healthcare services, and wellness for individuals and across populations"* [33, p.1].

Zooming in on the interventions that are expected to lead to this improvement on an individual level, all digital health interventions have in common that they connect potential users and health care services remotely – over a distance. In terms of mental health interventions, digital solutions might be a valuable addition to existing care, due to their opportunities in of overcoming at least some of the structural and attitudinal barriers to mental health care use.

In terms of structural barriers, digital interventions can most often be used independent of time and location and eliminate the need to travel which makes them more accessible and flexible than conventional in-person health care. Where no or only few resources from health care professionals are needed, digital interventions have the potential to be extremely scalable at low cost [34, 35]. Muñoz et al. [34] suggest a classification of four types of preventive interventions taking into account human resources needed for delivery. Type 1 and Type 2 interventions describe face-to-face interventions with Type 2 being enhanced by technological means. Type 3 (guided) and Type 4 (unguided) interventions on the other hand are considered pure self-help. The authors describe the latter Type 4 interventions as "non-consumable" interventions because availability and scalability of the intervention is not dependent on the amount of human resources available.

Presumably even more important, digital interventions might be able to overcome attitudinal barriers to health care use as well [17]. Digital interventions focusing on (guided) self-help might be more acceptable to individuals who want to help themselves instead of asking others for help. They often can be used more discretely and anonymously, avoiding potential stigmatization or discrimination [36].

Digital interventions also offer new ways for tailoring intervention to individual needs and risk profiles [37, 38], for example by adapting the intervention content or by using technology to derive physiological measurements to better understand and monitor the course of disease and treatment [39]. They can also be more easily adapted to serve different language and cultural contexts [40, 41, 42].

Notably, digital interventions can use a wide variety of different technologies as well as intervening approaches, resulting in a huge variety of possible interventions, employing those advantages to varying extent. This opens up possibilities to serve individuals with different needs (e.g. accounting for visual or hearing impairments) and prerequisites (e.g. internet access, living situation). At the same time, this heterogeneity also highlights the need for clear definitions and descriptions of those interventions in order to allow comparison and distinction.

De Witte et al. [43] recently developed two taxonomies to describe interventions based on systematic reviews in the field. The first taxonomy encompasses the types of intervention: preventive, measuring, supportive, behavioral, gamified, and psychotherapeutic interventions as well as skills training. The second taxonomy specifies the technology used which can either be a replacement for conventional delivery (synchronous or asynchronous), a standalone program (for the computer, smartphone or with mixed technology) or emerging technologies adding new aspects (including extended reality, social media, wearables and others).

In this work, I focus on two technologies to deliver psychological interventions to individuals with (subthreshold) depressive symptoms: the telephone for coaching and the computer for standalone web-based interventions. Therefore, I will use the term *"digital mental health intervention"* (DMHI) for any remotely delivered intervention and name the delivery medium and intervention characteristics where necessary for the context.

### 1.1.1 *Telephone coaching*

Traditionally established as an in-person intervention, telephone coaching, according to the taxonomies described above, can be considered a synchronous DMHI which is "conventional, but digital" and can be used in different intervention settings including additional support to other interventions or as a standalone preventive, behavioral or psychotherapeutic intervention. Psychological coaching — in distinction to sports or executive coaching — can be defined as

*"a conversational process of facilitating positive development and change towards optimal functioning, well-being and increased performance in the work and personal life domains, in the absence of clinically significant mental health issues, through the application of a wide range of psychological theories and principles."* [44, p.8].

Health coaching in particular has been defined as

*"a patient-centered approach wherein patients at least partially determine their goals, use self-discovery or active learning processes together with content education to work toward their goals, and self-monitor behaviors to increase accountability, all within the context of an interpersonal relationship with a coach. The coach is a healthcare professional trained in behavior change theory, motivational strategies, and communication techniques, which are used to assist patients to develop intrinsic motivation and obtain skills to create sustainable change for improved health and well-being"* [45, p.52]

Neither *coach* nor *coaching* are regulated terms or qualifications. The qualification of the coach (i.e. coaches have to be health care professional trained to support behavior change) is often considered a crucial distinction from life coaching, which can also include non-health related topics or general well-being. In terms of depression prevention, a coaching could encompass aspects of both life and health coaching, since individuals in indicated prevention are not yet considered ill, but their goals might mostly be health related. These coaches therefore should be especially trained to support individuals with mental health problems and uphold high standards for ethics and safety. However, given the undefined nature of the labels coach & coaching, no unified requirements for qualification exist and conduct can vary greatly from standardized procedures to more commonly free-flowing sessions [45, 46].

Anthony Grant, a pioneer in evidence-based coaching, together with Robert Green, an accredited psychotherapist, strictly distinguished coaching from counselling and therapy as aiming at non-clinical psychopathology [46], even though some of the aforementioned definitions use the word "patient". In addition, coaching is assumed to be more focused on attainment of individual, client-determined goals and personal development while clinical approaches focus on restoring function and relieving distress [46]. Therefore, the aim of the coaching as well as the process and methods to reach it vary greatly across coaching offers. Depending on the coaches' qualification and because of the often highly individualized nature of the interventions, it can sometimes be difficult to distinguish health coaching from therapeutic interventions and to synthesize the evidence for coaching [45, 47, 48].

While telephone coaching is well established in the support of (chronic) somatic conditions including diabetes, hypertension, cancer and cardiovascular problems [49, 50, 51, 52, 53], research on (standalone) telephone coaching for mental health remains scarce. Generalizability of its efficacy is not yet possible due to the high heterogeneity in interventions, differing specific populations and suboptimal study quality (concerning e.g. small samples, high dropout, per-protocol analysis). In a work context, a telephone coaching by professional counsellors, which included care coordination, cognitive behavioral therapy (CBT) via telephone in combination with a workbook and workplace coaching showed to be effective in reducing depressive symptoms after four months ($d_{between}$ = 0.60 [95%–CI: 0.39, 0.80]) when compared to treatment-as-usual (TAU) for individuals with MDD [54]. In privately insured individuals with MDD, health coaching delivered by the insurance company's health professionals significantly reduced depression scores compared to TAU ($d_{between}$ = 0.45 [95%–CI: 0.14, 0.75]. Coaching consisted of motivational interviewing and psychoeducation and aimed to strengthen self-help abilities in participants [55]. Van Mierlo et al. [56] report that for informal caregivers, telephone coaching in addition to a supportive respite care significantly reduced mental health problems (General Health Questionaire, $d_{between}$ = 0.26 [95%–CI: -0.54, -1.06]) when compared to a respite care alone. However, there was no difference in mental health between care givers receiving only telephone coaching and those having only respite care. One identified study investigated solely healthcare cost from an insurance perspective and reported significantly reduced overall and in-patient costs after 12-months of telephone-coaching aiming to increase self-help in MDD patients [57].

Literature on more manualized psychotherapy via telephone (almost exclusively CBT) is richer and indicates their efficacy in subthreshold depressed workers after 4 months (g = 0.69 [95%–CI: 0.32, 1.05]) [58]. In a systematic review, Coughtrey et al. [59] summarize the

efficacy for telephone therapy for depression with effect sizes between Cohen's d of 0.25 – 1.98 (median: 0.58). The most recent meta-analysis included four RCTs comparing telephone psychotherapy to TAU or no treatment for depression revealed pooled effects for of g = -0.85 [95%–CI: -1.56, -0.15] and showed comparable efficacy when pooling effects from six trials comparing telephone therapy to active controls (in-person therapy, medication and self-help; g = -0.18 [95%–CI: -0.45, 0.09]) [60]. Further studies also suggest that telephone interventions might have at least similar [60] if not higher adherence than face-to-face interventions [61, 62] and a comparable quality in therapeutic interactions [63].

Taken together, telephone coaching can be considered a promising approach in the prevention of depression, which will be further investigated in Article 1 – 3 of this dissertation. In the world of DMHIs, telephone interventions offers a location independent and more flexible intervention while leveraging the interpersonal relation and support from a personal coach. It might also be one of the most acceptable interventions to primary care patients [64]. Yet, they do not offer the same scalability or anonymity as for example unguided web-based interventions due to the high reliance on health care workers' capacity.

### 1.1.2 *Web-based interventions*

With the continuing spread of internet access into the everyday life of many people since the 1990s, completely new options for digital interventions emerged. With the therapeutics methods from (CBT) face-to-face interventions and experiences from bibliotherapy, "standalone programs" delivered over the internet were able to emerge [65]. Web-based intervention are often described using the definition by Barak et al., stating:

*A web-based intervention is a primarily self-guided intervention program that is executed by means of a prescriptive online program operated through a website and used by consumers seeking health- and mental-health related assistance. The intervention program itself attempts to create positive change and or improve/enhance knowledge, awareness, and understanding via the provision of sound health-related material and use of interactive web-based components* [66, p.5]

Barak et al. further distinguishes between educational and therapeutic interventions, online counselling and further online activities with the same goal. In the current work, I will focus on both self-help web-based therapeutic interventions and human supported web-based therapeutic interventions. When the intervention is human supported (i.e. guided), contact with the coach can be asynchronous (e-mail, messaging system) or synchronous (via chat, phone or video call).

Evidence for web-based interventions for depression is good, with them being recommended in the 2022 German national treatment

guidelines for mild cases of depression when embedded in a therapy context [67]. The NICE guidelines (by the National Institute for Health and Care Excellence in the UK) recommend web-based CBT programs as first-line treatment for subthreshold and mild depressive cases [68]. Meta-analyses consistently show significant symptom reduction in individuals with mild-moderate symptoms of depression or diagnosed MDD in comparison to different control condition (predominantly waiting list, but also TAU, attention control or active interventions) [69, 70, 71]. Evidence suggests that guided interventions yielded larger effects than unguided ones [70, 72]. Guided interventions are likely to be cost-effective, while cost-effectiveness for unguided and preventive intervention is still insufficiently evaluated [35].

In individuals with subthreshold depression, web-based interventions have likewise been shown effective in reducing depressive symptom [73]. A recent individual participant data (IPD) meta-analysis in this population summarized data from seven individual trials from Germany, the Netherlands and Japan [74]. The authors reported an average effect for symptom reduction of Hedges' g = 0.39 [95%–CI: 0.25, 0.53] at post-intervention assessment, g = 0.30 [95%–CI: 0.15, 0.45)] after 3 – 6 months and g = 0.27 [95%–CI: 0.07, 0.47] at 12 months. Three of the included studies also assessed depression onset and Reins et al. [74] identified a reduced risk for depression onset in the group receiving web-based intervention compared to TAU or online psychoeducation equivalent to a Hazard Ratio (HR) of 0.72 [95%–CI: 0.58, 0.89]) over a 12-month period.

One of the most pronounced advantages of web-based interventions as DMHIs for depression prevention is the autonomy and flexibility given to the users in terms of location, time and, in most cases, anonymity. At the same time, the reduced time needed by health care professionals carries great potential for scalability. So far, the human contact included in the intervention might be the biggest distinction of the two delivery formats addressed here — the telephone and the web-based training. Telephone coaching is still very close to face-to-face delivery and highly depends on "consumable" coach time and the improved flexibility and ease of use is more prevalent for the client than for the coach. Web-based interventions allow for similar or greater flexibility for the client but also the coaches need much less time per client. This is important to consider because human contact is one of the most crucial points when it comes to the acceptability of digital interventions in clinicians and users [64, 75, 76] and can influence the efficacy [77]. These key differences in characteristics of the intervention and the degree and the role of human support in each interventions will be considered throughout this dissertation.

## 1.2 THE CHALLENGES FOR DIGITAL MENTAL HEALTH INTERVENTIONS IN THE PREVENTION ON DEPRESSION

The opportunities outlined in the last section indicate the enormous potential of DMHIs to reduce depression burden by delivering easily accessible, evidence-based, effective health interventions to a broad and diverse audience at relatively low cost. However, despite their existence for over two decades now [65] and the general progress in mental health interventions, no detectable difference was made in depression burden [18] at last partly due to several new challenges which became apparent with the increase of digitalization in healthcare.

The access to communication media cannot be considered a huge limitation, even though this can be the case in some regions and populations [78]. In Germany, where all research for this dissertation was conducted, 99.9% of households have a telephone [79] and 91.0% have internet access [80] in 2022. In the population aged $16 - 74$, 94.5% are considered internet users, with use being lower in less educated and older groups [80]. However, implementation of health interventions that make use of these technologies in routine care varies highly between European countries and health care systems [36] and is generally considered to be slow [81]. Granja et al. [82] conducted a systematic review on 221 studies investigating various implemented digital health interventions across different health care fields and aimed to synthesize known factors for success and failure of implementation from perspectives of patients, health care providers and health systems (e.g. management staff, national health plans). They found that success was most commonly characteristic by patient empowerment and improved self-management (patients), improved health care quality (health care providers) and by policies governing costs (health care system representatives). Failures on the other hand were most often associated with privacy concerns (patients), disruptions in the established workflow (health care providers) and costs during implementation (health care system representatives).

Similar aspects were reported in relation to DMHI for depression. In a systematic review on 48 studies focusing on video-counselling as well as stand alone web-based programs, Vis et al. [83] identified several barriers and facilitators for implementation. Reviewed studies including again views of users, providers, organizations and health care systems and summarized 1) acceptance by patients and professionals, 2) the therapeutic interactions in the DMHIs and 3) the technological prerequisites and compatibilities with other aspects of health care as the three most often reported determinants for successful implementation.

However, even in implemented and routinely available in-person prevention offers, uptake might be as low as under 1% of targeted individuals with subthreshold depression [84]. Similarly, in the online

context, study-advertisement for a preventive online intervention employing ads on Google keyword searches yielded a similar response (i.e. study inclusion) [85]. This might be a crucial point to consider in DMHIs, since an observed low acceptability [76] of especially unguided DMHIs might exacerbate the problem of low uptake. However, attitudes towards DMHI intervention might be more positive when used for milder cases of depression or prevention [86, 75] and can improve with experience [75].

Closely relate to low uptake is the topic of intervention dropout, which is sometimes also considered a measure of acceptability. Even though no differences in efficacy between individual or group therapy with telephone-delivered therapy and guided self-help were observed in a recent network-meta-analysis of CBT as treatment for MDD [87], acceptability (defined as intervention dropout) was lower in self-help interventions comapred to individual or group CBT. In the same study, unguided self-help was more efficacious than a waiting list control group but was associated with higher dropout in the intervention than the waiting list.

Soley improving adherence to evidence-based interventions, would, however, not be enough to harvest the merits of DMHIs. Fact is that not every individual profits in the same way from an intervention that has been shown effective on a *population* level. Looking again at the study from Reins et al. [74], merely 40% of users of a web-based intervention showed a reliable change in their subthreshold depression symptoms at post and first follow-up. Two main approaches can be followed to address this conundrum: making interventions more effective and/or matching participants and interventions better.

When delivering psychological interventions via digital technologies it combines the questions of via which mechanisms psychological interventions lead to change in (depressive) symptomology [88, 89] and if mechanisms specific to the delivery medium might have an influence on the intervention's outcome. Common factors are hypothesized to have a positive effect regardless of the specific intervention (e.g. format, psychological tradition) in psychotherapy and include the patient's experience of support, their learning and change in action [90]. These might be to some extended transferrable to the digital sphere, however not all factors might be offered to the same extent (e.g. relationship to therapist, feedback). On the other hand, there might be characteristics specific to DMHIs (e.g. almost constant availability, tailoring) that contribute to their effectiveness and that can predict treatment outcome early on. Both factors need to be considered when trying to optimize the effectiveness of DMHIs [91, 92].

Besides the intervention's characteristics there are also user's characteristics that influence the course of a psychological disorder like depression as well as individual reaction to treatment. *Personalized medicine* aims to move away from a "one-size-fits-all" approach to-

wards individual treatment plans and interventions based on personal risk factors [93, 94]. Likewise, personalized prevention could help improve the effectiveness of preventive interventions by targeting the individuals with the highest risk of developing a disorder, based on personal risk indicators or the combination of such [95]. In in-person indicated depression prevention in youth, even the broad distinction between "cognitive" and "intrapersonal" risk factors and the match or no-match which a intervention focusing either on cognitive or interpersonal factors, revealed favorable results when adolescents received an intervention matching their risk profile [96, 97]. DMHIs can potentially support this by new ways of assessing risk profiles or by automatically adapting intervention content to individual needs. In order to realize this, more research is needed to identify who profits most from which kind of intervention (component) and especially more standardized, modular web-based interventions offer new possibilities to do so [92].

For the prevention of depression we can consequently summarize that in order to increase the impact of digital prevention interventions on overall depression burden, it is important to first establish robust evidence for the (long-term) effectiveness of preventive interventions using different digital solutions in order to profit from their differnt characteristics and possibilities. Second, these evidence-based interventions need to be sustainably implemented and actually used in routine care by the targeted group. Therefore more research is needed that support implementation efforts and intervention uptake. To enhance the impact of DMHIs on depression burden it is then – third – important to identify which individual is most likely to profit most from which intervention in order to achive the best outcome and not waste participants' and health care resources in ineffective interventions, and – fourth –, to gain more understanding of how treatment effects can be predicted and enhanced while minimizing potential negative effects.

## 1.3 OUTLINE OF THIS WORK

The aim of this work is therfore two-fold: First, I want to showcase the opportunities provided by DMHIs by reporting on a trial evaluating a personalized telephone coaching and add to the existing evidence for DMHIs in the prevention on the depression. Second, I want to address the three aforementioned problems currently faced by DMHIs, which have not yet been studied a lot in a depression prevention context, in the personalized telephone coaching as well as in more standardized web-based interventions.

The first study (Article 1 – 3) is a randomized controlled trial or RCT (TEC-A, **Te**le-based **C**oaching for **A**griculturists) evaluating the effectiveness of personalized telephone coaching in reducing depressive symptoms and onset of depression as part of the nation-wide

implementation project "With us in balance" [98] by the German Social Insurance Company for Agriculture, Forestry and Horticulture (SVLFG). Farmers and related professions are considered an at-risk occupational group for depression and suicide [99]. They combine a multitude of risk factors for mental health problems (including financial worries, weather uncertainty, long working hours, role conflicts) with often living in rural areas with limited health care options [99]. Additionally, they are supposed to have high stigma against psychological topics and low help seeking behavior for health complains [100, 101]. Therefore, the telephone seems to give a chance to deliver a preventive intervention to this at-risk group, since it eliminates travelling and can be used more flexibly and discretely as on-site services while allowing easy adaptation to the occupation's specific needs. To the best of my knowledge, there is currently no other research published on the long-term effectiveness of a personalized telephone coaching in the prevention of depression.

Moving on to the challenges, in the second study — again as part of the model project "With us in balance" — I aim to identify aspects that possible could support the implementation of the personalized telephone coaching and its uptake in routine care. The focus is on aspects of the intervention that influence the acceptance of the intervention as well as the participants' satisfaction with it (Article 4) and which could be use to inform further implemention activities.

The last two studies build on previous research on web-based interventions and aim to enhance the current knowledge on predictors of treatment outcomes and differential treatment effects, which are easier to study in more standardized interventions. In the third study (Article 5), I use data from two depression prevention trials evaluating a web-based intervention to examine the role of outcome expectancy as possible predictor of treatment outcome that could potentially be enhanced pre-treatment.

As highlighted above, average efficacy in RCTs does not allow deriving recommendations for individual treatment options. Therefore, in the last study (Article 6), the differential efficacy of an emerging prevention/treatment approach (i.e., indirect treatment, addressing related but less stigmatized problems) is evaluated using secondary depression data from four trials evaluating web-based interventions for insomnia. Potential differences in treatment efficacy on depression outcomes are evaluated across a wide range of sociodemographic, clinical and work-related characteristics with the aim who identify subgroups for which this approach is especially efficacious.

All pre-specified research questions are summarized in Table 1.1. In the next chapter, the methodology used to address the individual research questions is outlined and reasoned. The articles employing the described methods and answering the research questions are presented in chapters 3 – 8. In the closing discussion part, the main

contributions of the present research will be summarized and discussed in the light of what is needed to increase the impact of digital preventive intervention on depression burden in the future.

Table 1.1: Summary of the research questions examined in this dissertation

| **Showcasing the opportunities** | |
|---|---|
| ***Goal (Article)*** | ***Research question*** |
| Ealuating the effectiveness of personalized telephone coaching *(Article 1-3)* | RQ1: Is personalized TC effective in reducing depressive symptom severity compared to enhanced treatment as usual at post-treatment and follow-ups?<br>RQ2: Is personalized TC effective in preventing the onset of major depressive episodes as compared to TAU+ at follow-up?<br>RQ3: Is personalized TC effective in reducing the severity of various mental health outcomes (e.g., stress, anxiety, insomnia) compared to TAU+?<br>RQ4: What is the level of satisfaction with, adherence to and acceptance of personalized TC?<br>RQ5: Are there reported negative effects associated with TC? |
| **Addressing the challenges** | |
| ***Goal (Article)*** | ***Research question*** |
| Supporting implementation and uptake *(Article 4)* | RQ1: Which factors influence the acceptance of and satisfaction with a personalized telephone coaching for farmers? |
| Identifying outcome predictors in DMHIs *(Article 5)* | RQ1: Are outcome expectations and depression symptom severity at post-treatment and follow-up related?<br>RQ2: Are outcome expectations and time to depression onset within a 12-month follow-up period related?<br>RQ3: Are outcome expectations and "close-to-symptom-free"-status at post and follow-up related? |
| Exploring differential treatment effects *(Article 6)* | RQ1: What potential does an indirect treatment of depressive symptoms through targeting insomnia with internet interventions have in terms of symptom reduction, treatment response and MDD onset?<br>RQ2: Are treatment effects moderated by clinical, demographic or sleep intervention-related characteristics? |

## REFERENCES


1. Balmuth E. From Black Bile to the Brain: Tracing Melancholia and Depression. Max Planck PhDnet - Offspring Blog. 2017. Accessed on: 29-05-2023. Available from: https://www.phdnet.mpg.de/offspring-blog/2017/from-black-bile-to-the-brain
2. Nolen-Hoeksema S. Looking at Abnormality. *Abnormal Psychology*. Ed. by Nolen-Hoeksema S and Marroquín B. 8th ed. 2-21. USA, New York, NY: McGraw Hill Education, 2020
3. Nemeroff CB. The State of Our Understanding of the Pathophysiology and Optimal Treatment of Depression: Glass Half Full or Half Empty? American Journal of Psychiatry. 2020; 177:671–85. DOI: 10.1176/appi.ajp.2020.20060845
4. American Psychiatric Association APA. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. 5th ed. USA, Washington, D.C: American Psychiatric Association, 2013
5. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJ, Vos T, and Whiteford HA. Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. PLoS Medicine. 2013; 10:e1001547. DOI: 10.1371/journal.pmed.1001547
6. Lépine JP and Briley M. The Increasing Burden of Depression. Neuropsychiatric Disease and Treatment. 2011; 7:3–7. DOI: 10.2147/NDT.S19617
7. Greenberg PE, Fournier AA, Sisitsky T, Pike CT, and Kessler RC. The Economic Burden of Adults with Major Depressive Disorder in the United States (2005 and 2010). Journal of Clinical Psychiatry. 2015; 76:155–62. DOI: 10.4088/JCP.14m09298
8. Vos T, Haby MM, Barendregt JJ, Kruijshaar M, Corry J, and Andrews G. The Burden of Major Depression Avoidable by Longer-Term Treatment Strategies. Archives of General Psychiatry. 2004; 61:1097–103. DOI: 10.1001/archpsyc.61.11.1097
9. Ho CS, Jin A, Nyunt MSZ, Feng L, and Ng TP. Mortality Rates in Major and Subthreshold Depression: 10-Year Follow-up of a Singaporean Population Cohort of Older Adults. Postgraduate Medicine. 2016; 128:642–7. DOI: 10.1080/00325481.2016.1221319
10. Cuijpers P and Smit F. Excess Mortality in Depression: A Meta-Analysis of Community Studies. Journal of Affective Disorders. 2002; 72:227–36. DOI: 10.1016/S0165-0327(01)00413-X
11. König H, Rommel A, Thom J, Schmidt C, König HH, Brettschneider C, and Konnopka A. The Excess Costs of Depression and the Influence of Sociodemographic and Socioeconomic Factors: Results from the German Health Interview and Examination Survey for Adults (DEGS). PharmacoEconomics. 2021; 39:667–80. DOI: 10.1007/s40273-021-01000-1
12. König H, König HH, and Konnopka A. The Excess Costs of Depression: A Systematic Review and Meta-Analysis. Epidemiology and Psychiatric Sciences. 2020; 29:e30. DOI: 10.1017/S2045796019000180

13. World Health Organization. Depression and Other Common Mental Disorders: Global Health Estimates. Geneva: World Health Organization; 2017. Available from: https://apps.who.int/iris/handle/10665/254610

14. Bromet EJ, Andrade LH, Bruffaerts R, and Williams DR. Major Depressive Disorder. *Mental Disorders Around the World*. Ed. by Scott KM, de Jonge P, Stein DJ, and Kessler RC. 1st ed. 41-56. UK, Cambridge University Press, 2018. DOI: 10.1017/9781316336168.004

15. Mitchell AJ, Yadegarfar M, Gill J, and Stubbs B. Case Finding and Screening Clinical Utility of the Patient Health Questionnaire (PHQ-9 and PHQ-2) for Depression in Primary Care: A Diagnostic Meta-Analysis of 40 Studies. BJPsych Open. 2016; 2:127–38. DOI: 10.1192/bjpo.bp.115.001685

16. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P, and Saxena S. Scaling-up Treatment of Depression and Anxiety: A Global Return on Investment Analysis. The Lancet Psychiatry. 2016; 3:415–24. DOI: 10.1016/S2215-0366(16)30024-4

17. Andrade LH et al. Barriers to Mental Health Treatment: Results from the WHO World Mental Health Surveys. Psychological Medicine. 2014; 44:1303–17. DOI: 10.1017/S0033291713001943

18. Ormel J, Hollon SD, Kessler RC, Cuijpers P, and Monroe SM. More Treatment but No Less Depression: The Treatment-Prevalence Paradox. Clinical Psychology Review. 2022; 91:102111. DOI: 10.1016/j.cpr.2021.102111

19. Ebert DD and Cuijpers P. It Is Time to Invest in the Prevention of Depression. JAMA network open. 2018; 1:e180335. DOI: 10.1001/jamanetworkopen.2018.0335

20. Gordon RS. An Operational Classification of Disease Prevention. Public Health Reports. 1983; 98:107–9

21. Volz HP, Stirnweiß J, Kasper S, Möller HJ, and Seifritz E. Subthreshold Depression – Concept, Operationalisation and Epidemiological Data. A Scoping Review. International Journal of Psychiatry in Clinical Practice. 2022; online first:1–15. DOI: 10.1080/13651501.2022.2087530

22. Rucci P, Gherardi S, Tansella M, Piccinelli M, Berardi D, Bisoffi G, Corsino MA, and Pini S. Subthreshold Psychiatric Disorders in Primary Care: Prevalence and Associated Characteristics. Journal of Affective Disorders. 2003; 76:171–81. DOI: 10.1016/S0165-0327(02)00087-3

23. Cuijpers P, Vogelzangs N, Twisk J, Kleiboer A, Li J, and Penninx BW. Differential Mortality Rates in Major and Subthreshold Depression: Meta-analysis of Studies That Measured Both. British Journal of Psychiatry. 2013; 202:22–7. DOI: 10.1192/bjp.bp.112.112169

24. Cuijpers P, Smit F, Oostenbrink J, De Graaf R, Ten Have M, and Beekman A. Economic Costs of Minor Depression: A Population-Based Study. Acta Psychiatrica Scandinavica. 2007; 115:229–36. DOI: 10.1111/j.1600-0447.2006.00851.x

25. Wagner HR, Burns BJ, Broadhead WE, Yarnall KSH, Sigmon A, and Gaynes BN. Minor Depression in Family Practice: Functional Morbidity, Co-Morbidity, Service Utilization and Outcomes. Psychological Medicine. 2000; 30:1377–90. DOI: 10.1017/S0033291799002998

26. Cuijpers P, De Graaf R, and Van Dorsselaer S. Minor Depression: Risk Profiles, Functional Disability, Health Care Use and Risk of Developing Major Depression. Journal of Affective Disorders. 2004; 79:71–9. DOI: 10.1016/S0165-0327(02)00348-8

27. Lee YY, Stockings EA, Harris MG, Doi SA, Page IS, Davidson SK, and Barendregt JJ. The Risk of Developing Major Depression among Individuals with Subthreshold Depression: A Systematic Review and Meta-Analysis of Longitudinal Cohort Studies. Psychological Medicine. 2019; 49:92–102. DOI: 10.1017/s0033291718000557

28. Cuijpers P, Koole SL, Van Dijke A, Roca M, Li J, and Reynolds CF. Psychotherapy for Subclinical Depression: Meta-analysis. British Journal of Psychiatry. 2014; 205:268–74. DOI: 10.1192/bjp.bp.113.138784

29. Cuijpers P, Pineda BS, Quero S, Karyotaki E, Struijs SY, Figueroa CA, Llamas JA, Furukawa TA, and Muñoz RF. Psychological Interventions to Prevent the Onset of Depressive Disorders: A Meta-Analysis of Randomized Controlled Trials. Clinical Psychology Review. 2021; 83:101955. DOI: 10.1016/j.cpr.2020.101955

30. Conejo-Cerón S, Lokkerbol J, Moreno-Peral P, Wijnen B, Fernández A, Mendive JM, Smit F, and Bellón JÁ. Health-Economic Evaluation of Psychological Interventions for Depression Prevention: Systematic Review. Clinical Psychology Review. 2021; 88:102064. DOI: 10.1016/j.cpr.2021.102064

31. Paterson JG. Public Wellness–Preventive Mental Health. American Association for Counseling and Development Annual Convention. Vol. 1–8. Reno, Nevada, USA: Educationa Resources Infromation Center (ERIC), 1991

32. Teepe GW, Da Fonseca A, Kleim B, Jacobson NC, Salamanca Sanabria A, Tudor Car L, Fleisch E, and Kowatsch T. Just-in-Time Adaptive Mechanisms of Popular Mobile Apps for Individuals With Depression: Systematic App Search and Literature Review. Journal of Medical Internet Research. 2021; 23:e29412. DOI: 10.2196/29412

33. Kostkova P. Grand Challenges in Digital Health. Frontiers in Public Health. 2015; 3. DOI: 10.3389/fpubh.2015.00134

34. Muñoz RF, Pineda BS, Barrera AZ, Bunge E, and Leykin Y. Digital Tools for Prevention and Treatment of Depression: Lessons from the Institute for International Internet Interventions for Health. Clinica y Salud. 2021; 32:37–40. DOI: 10.5093/CLYSA2021A2

35. Kählke F, Buntrock C, Smit F, and Ebert DD. Systematic Review of Economic Evaluations for Internet- and Mobile-Based Interventions for Mental Health Problems. npj Digital Medicine. 2022; 5:175. DOI: 10.1038/s41746-022-00702-w

36. Ebert DD, Van Daele T, Nordgreen T, Karekla M, Compare A, Zarbo C, Brugnera A, Øverland S, Trebbi G, Jensen KL, Kaehlke F, and Baumeister H. Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health. European Psychologist. 2018; 23:167–87. DOI: 10.1027/1016-9040/a000318

37. Smit E, Linn A, and Weert J. Taking Online Computer-Tailoring Forward. The Potential of Tailoring the Message Frame and Delivery Mode of Online Health Behaviour Change Interventions. The European Health Psychologist. 2015; 17:25–31. Available from: https://www.ehps.net/ehp/index.php/contents/article/view/762

38. Lustria MLA, Cortese J, Noar SM, and Glueckauf RL. Computer-Tailored Health Interventions Delivered over the Web: Review and Analysis of Key Components. Patient Education and Counseling. 2009; 74:156–73. DOI: 10.1016/j.pec.2008.08.023

39. Huckvale K, Venkatesh S, and Christensen H. Toward Clinical Digital Phenotyping: A Timely Opportunity to Consider Purpose, Quality, and Safety. npj Digital Medicine. 2019; 2:88. DOI: 10.1038/s41746-019-0166-1

40. Schueller SM, Hunter JF, Figueroa C, and Aguilera A. Use of Digital Mental Health for Marginalized and Underserved Populations. Current Treatment Options in Psychiatry. 2019; 6:243–55. DOI: 10.1007/s40501-019-00181-z

41. Muñoz RF. Using Evidence-Based Internet Interventions to Reduce Health Disparities Worldwide. Journal of Medical Internet Research. 2010; 12:e1463. DOI: 10.2196/jmir.1463

42. Harper Shehadeh M, Heim E, Chowdhary N, Maercker A, and Albanese E. Cultural Adaptation of Minimally Guided Interventions for Common Mental Disorders: A Systematic Review and Meta-Analysis. JMIR Mental Health. 2016; 3:e44. DOI: 10.2196/mental.5776

43. De Witte NAJ, Joris S, Van Assche E, and Van Daele T. Technological and Digital Interventions for Mental Health and Wellbeing: An Overview of Systematic Reviews. Frontiers in Digital Health. 2021; 3:754337. DOI: 10.3389/fdgth.2021.754337

44. Interest Group of Coaching and Consulting Psychology IGCCP. Code of Practice for Registered Psychology Practitioners in Coaching. Accessed on: 27.12.2022. Available from: www.siopsa.org.za

45. Wolever RQ, Simmons LA, Sforzo GA, Dill D, Kaye M, Bechard EM, Southard ME, Kennedy M, Vosloo J, and Yang N. A Systematic Review of the Literature on Health and Wellness Coaching: Defining a Key Behavioral Intervention in Healthcare. Global Advances in Health and Medicine. 2013; 2:38–57. DOI: 10.7453/gahmj.2013.042

46. Grant AM and Green RM. Developing Clarity on the Coaching - Counselling Conundrum: Implications for Counsellors and Psychotherapists. Counselling and Psychotherapy Research. 2018; 18:347–55. DOI: 10.1002/capr.12188

47. Passmore J, Stopforth M, and Lai Y. Defining Coaching Psychology: Debating Coaching and Coaching Psychology Definitions. The Coaching Psychologist. 2018; 14:120–2. Available from: https://centaur.reading.ac.uk/81825/

48. Bachkirova T and Baker S. Revisiting the Issue of Boundaries between Coaching and Counselling. *Handbook of Coaching Psychology - a Guide for Practitioners*. Ed. by Palmer S and Whybrow A. 487-499. USA, New York, NY: Routledge, 2019

49. An S and Song R. Effects of Health Coaching on Behavioral Modification among Adults with Cardiovascular Risk Factors: Systematic Review and Meta-Analysis. Patient Education and Counseling. 2020; 103:2029–38. DOI: 10.1016/j.pec.2020.04.029

50. Yu-Mei Chen D, Wu XV, Chan EY, and Goh YS. Nurse-Led Tele-Coaching on Modifiable Cardiovascular Risk Factors in People with Type 2 Diabetes Mellitus: A Systematic Review and Meta-Analysis. Worldviews on Evidence-Based Nursing. 2019; 16:424–32. DOI: 10.1111/wvn.12409

51. Boehmer KR, Barakat S, Ahn S, Prokop LJ, Erwin PJ, and Murad MH. Health Coaching Interventions for Persons with Chronic Conditions: A Systematic Review and Meta-Analysis Protocol. Systematic Reviews. 2016; 5:146. DOI: 10.1186/s13643-016-0316-3

52. Meng F, Jiang Y, Yu P, Song Y, Zhou L, Xu Y, and Zhou Y. Effect of Health Coaching on Blood Pressure Control and Behavioral Modification among Patients with Hypertension: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. International Journal of Nursing Studies. 2023; 138:104406. DOI: 10.1016/j.ijnurstu.2022.104406

53. Barakat S, Boehmer K, Abdelrahim M, Ahn S, Al-Khateeb AA, Villalobos NÁ, Prokop L, Erwin PJ, Fleming K, Serrano V, Spencer-Bonilla G, and Murad MH. Does Health Coaching Grow Capacity in Cancer Survivors? A Systematic Review. Population Health Management. 2018; 21:63–81. DOI: 10.1089/pop.2017.0040

54. Lerner D, Adler DA, Rogers WH, Chang H, Greenhill A, Cymerman E, and Azocar F. A Randomized Clinical Trial of a Telephone Depression Intervention to Reduce Employee Presenteeism and Absenteeism. Psychiatric Services. 2015; 66:570–7. DOI: 10.1176/appi.ps.201400350

55. Schröder S, Fleckenstein J, and Wunderlich M. Effektivität Eines Telefonbasierten Coaching-Programms Für Patienten Mit Einer Depressiven Erkrankung. Monitor Versorgungsforschung. 2017; 10:65–8. DOI: 10.24945/MVF.04.17.1866-0533.2022

56. Van Mierlo LD, Meiland FJM, and Dröes RM. Dementelcoach: Effect of Telephone Coaching on Carers of Community-Dwelling People with Dementia. International Psychogeriatrics. 2012; 24:212–22. DOI: 10.1017/S1041610211001827

57. Gerlach-Reinholz W, Drop L, Basic E, Rauchhaus M, and Fritze J. Telefoncoaching bei Depression. Der Nervenarzt. 2017; 88:811–8. DOI: 10.1007/s00115-017-0316-0

58. Furukawa TA, Horikoshi M, Kawakami N, Kadota M, Sasaki M, Sekiya Y, Hosogoshi H, Kashimura M, Asano K, Terashima H, Iwasa K, Nagasaku M, Grothaus LC, and on behalf of the GENKI Project. Telephone Cognitive-Behavioral Therapy for Subthreshold Depression and Presenteeism in Workplace: A Randomized Controlled Trial. PLoS ONE. 2012; 7:e35330. DOI: 10.1371/journal.pone.0035330

59. Coughtrey AE and Pistrang N. The Effectiveness of Telephone-Delivered Psychological Therapies for Depression and Anxiety: A Systematic Review. 2018; 24:65–74. DOI: 10.1177/1357633X16686547

60. Castro A, Gili M, Ricci-Cabello I, Roca M, Roca M, Gilbody S, Perez-Ara MÁ, Maria Ángeles Perez-Ara, Pérez-Ara M, Seguí A, and McMillan D. Effectiveness and Adherence of Telephone-Administered Psychotherapy for Depression: A Systematic Review and Meta-Analysis. Journal of Affective Disorders. 2020; 260:514–26. DOI: 10.1016/j.jad.2019.09.023

61. Mohr DC, Ho J, Duffecy J, Reifler D, Sokol L, Burns MN, Jin L, and Siddique J. Effect of Telephone-Administered vs Face-to-Face Cognitive Behavioral Therapy on Adherence to Therapy and Depression Outcomes among Primary Care Patients: A Randomized Trial. JAMA - Journal of the American Medical Association. 2012; 307:2278–85. DOI: 10.1001/jama.2012.5588

62. Mohr DC, Vella L, Vella L, Hart SL, Heckman TG, and Simon GE. The Effect of Telephone-Administered Psychotherapy on Symptoms of Depression and Attrition: A Meta-Analysis. Clinical Psychology-science and Practice. 2008; 15:243–53. DOI: 10.1111/j.1468-2850.2008.00134.x

63. Irvine A, Drew P, Bower P, Brooks H, Gellatly J, Armitage CJ, Barkham M, McMillan D, and Bee P. Are There Interactional Differences between Telephone and Face-to-Face Psychological Therapy? A Systematic Review of Comparative Studies. Journal of Affective Disorders. 2020; 265:120–31. DOI: 10.1016/j.jad.2020.01.057

64. Mohr DC, Siddique J, Ho J, Duffecy J, Jin L, and Fokuo JK. Interest in Behavioral and Psychological Treatments Delivered Face-to-Face, by Telephone, and by Internet. Annals of Behavioral Medicine. 2010; 40:89–98. DOI: 10.1007/s12160-010-9203-7

65. Andersson G. Internet Interventions: Past, Present and Future. Internet Interventions. 2018; 12:181–8. DOI: 10.1016/j.invent.2018.03.008

66. Barak A, Klein B, and Proudfoot JG. Defining Internet-Supported Therapeutic Interventions. Annals of Behavioral Medicine. 2009; 38:4–17. DOI: 10.1007/s12160-009-9130-7

67. Bundesärztekammer BÄK, Kassenärztliche Bundesvereinigung KBV, and Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften AWMF. Nationale VersorgungsLeitlinie Unipolare Depression - Langfassung. Version 3.0. 2022. DOI: 10.6101/AZQ/000493

68. National Institute for Health and Care Excellence NICE. Depression in Adults: Treatment and Management. 2022. Available from: www.nice.org.uk/guidance/ng222

69. Sztein DM, Koransky CE, Fegan L, and Himelhoch S. Efficacy of Cognitive Behavioural Therapy Delivered over the Internet for Depressive Symptoms: A Systematic Review and Meta-Analysis. Journal of Telemedicine and Telecare. 2018; 24:527–39. DOI: 10.1177/1357633X17717402

70. Wright JH, Owen JJ, Richards D, Eells TD, Richardson T, Brown GK, Barrett M, Rasku MA, Polser G, and Thase ME. Computer-Assisted Cognitive-Behavior Therapy for Depression: A Systematic Review and Meta-Analysis. The Journal of Clinical Psychiatry. 2019; 80:18r12188. DOI: 10.4088/JCP.18r12188

71. Königbauer J, Letsch J, Doebler P, Ebert D, and Baumeister. Internet- and Mobile-Based Depression Interventions for People with Diagnosed Depression: A Systematic Review and Meta-Analysis. Journal of Affective Disorders. 2017; 223:28–40. DOI: 10.1016/j.jad.2017.07.021

72. Baumeister H, Reichler L, Munzinger M, and Lin J. The Impact of Guidance on Internet-based Mental Health Interventions — A Systematic Review. Internet Interventions. 2014; 1:205–15. DOI: 10.1016/j.invent.2014.08.003

73. Zhou T, Li X, Pei Y, Gao J, and Kong J. Internet-Based Cognitive Behavioural Therapy for Subthreshold Depression: A Systematic Review and Meta-Analysis. BMC Psychiatry. 2016; 16:356. DOI: 10.1186/s12888-016-1061-9

74. Reins JA, Buntrock C, Zimmermann J, Grund S, Harrer M, Lehr D, Baumeister H, Weisel K, Domhardt M, Imamura K, Kawakami N, Spek V, Nobis S, Snoek F, Cuijpers P, Klein JP, Moritz S, and Ebert DD. Efficacy and Moderators of Internet-Based Interventions in Adults with Subthreshold Depression: An Individual Participant Data Meta-Analysis of Randomized Controlled Trials. Psychotherapy and Psychosomatics. 2021; 90:94–106. DOI: 10.1159/000507819

75. Gun SY, Titov N, and Andrews G. Acceptability of Internet Treatment of Anxiety and Depression. Australasian Psychiatry. 2011; 19:259–64. DOI: 10.3109/10398562.2011.562295

76. Apolinário-Hagen J, Harrer M, Kählke F, Fritsche L, Salewski C, and Ebert DD. Public Attitudes Toward Guided Internet-Based Therapies: Web-Based Survey Study. JMIR Mental Health. 2018; 5:e10735. DOI: 10.2196/10735

77. Werntz A, Amado S, Jasman M, Ervin A, and Rhodes JE. Providing Human Support for the Use of Digital Mental Health Interventions: Systematic Meta-review. Journal of Medical Internet Research. 2023; 25:e42864. DOI: 10.2196/42864

78. Richardson S, Lawrence K, Schoenthaler AM, and Mann D. A Framework for Digital Health Equity. npj Digital Medicine. 2022; 5:119. DOI: 10.1038/s41746-022-00663-0

79. Statistisches Bundesamt DESTATIS. Laufende Wirtschaftsrechnungen - Ausstattung privater Haushalte mit ausgewählten Gebrauchsgütern. Fachserie 15 Reihe 2. 2022. Available from: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Ausstattung-Gebrauchsgueter/_inhalt.html#_t9yfhn9bi

80. Statistisches Bundesamt DESTATIS. Private Haushalte in der Informationsgesellschaft – Nutzung von Informations- und Kommunikationstechnologien (Mikrozensus-Unterstichprobe zur Internetnutzung). Fachserie 15 Reihe 4. 2022. Available from: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/IT-Nutzung/_inhalt.html#_1emt25loc

81. Proctor EK, Landsverk J, Aarons G, Chambers D, Glisson C, and Mittman B. Implementation Research in Mental Health Services: An Emerging Science with Conceptual, Methodological, and Training Challenges. Administration and Policy in Mental Health and Mental Health Services Research. 2009; 36:24–34. DOI: 10.1007/s10488-008-0197-4

82. Granja C, Janssen W, and Johansen MA. Factors Determining the Success and Failure of eHealth Interventions: Systematic Review of the Literature. Journal of Medical Internet Research. 2018; 20:e10235. DOI: 10.2196/10235

83. Vis C, Mol M, Kleiboer A, Bührmann L, Finch T, Smit J, and Riper H. Improving Implementation of eMental Health for Mood Disorders in Routine Practice: Systematic Review of Barriers and Facilitating Factors. JMIR Mental Health. 2018; 5:e20. DOI: 10.2196/mental.9769

84. Cuijpers P, van Straten A, Warmerdam L, and van Rooy MJ. Recruiting Participants for Interventions to Prevent the Onset of Depressive Disorders: Possibile Ways to Increase Participation Rates. BMC Health Services Research. 2010; 10:181. DOI: 10.1186/1472-6963-10-181

85. Morgan AJ, Jorm AF, and Mackinnon AJ. Internet-Based Recruitment to a Depression Prevention Intervention: Lessons From the Mood Memos Study. Journal of Medical Internet Research. 2013; 15:e31. DOI: 10.2196/jmir.2262

86. Topooco N, Riper H, Araya R, Berking M, Brunn M, Chevreul K, Cieslak R, Ebert DD, Etchmendy E, Herrero R, Kleiboer A, Krieger T, García-Palacios A, Cerga-Pashoja A, Smoktunowicz E, Urech A, Vis C, and Andersson G. Attitudes towards Digital Treatment for Depression: A European Stakeholder Survey. Internet Interventions. 2017; 8:1–9. DOI: /10.1016/j.invent.2017.01.001

87. Cuijpers P, Noma H, Karyotaki E, Cipriani A, and Furukawa TA. Effectiveness and Acceptability of Cognitive Behavior Therapy Delivery Formats in Adults With Depression: A Network Meta-analysis. JAMA Psychiatry. 2019; 76:700–7. DOI: 10.1001/jamapsychiatry.2019.0268

88. Kazdin AE. Treatment Outcomes, Common Factors, and Continued Neglect of Mechanisms of Change. Clinical Psychology: Science and Practice. 2005; 12:184–8. DOI: 10.1093/clipsy.bpi023

89. Kazdin AE. Mediators and Mechanisms of Change in Psychotherapy Research. Annual Review of Clinical Psychology. 2007; 3:1–27. DOI: 10.1146/annurev.clinpsy.3.022806.091432

90. Cuijpers P, Reijnders M, and Huibers MJ. The Role of Common Factors in Psychotherapy Outcomes. Annual Review of Clinical Psychology. 2019; 15:207–31. DOI: 10.1146/annurev-clinpsy-050718-095424

91. Domhardt M, Steubl L, Boettcher J, Buntrock C, Karyotaki E, Ebert DD, Cuijpers P, and Baumeister H. Mediators and Mechanisms of Change in Internet- and Mobile-Based Interventions for Depression: A Systematic Review. Clinical Psychology Review. 2021; 83:101953. DOI: 10.1016/j.cpr.2020.101953

92. Domhardt M, Cuijpers P, Ebert DD, and Baumeister H. More Light? Opportunities and Pitfalls in Digitalized Psychotherapy Process Research. Frontiers in Psychology. 2021; 12:544129. DOI: 10.3389/fpsyg.2021.544129

93. Cuijpers P, Ebert DD, Acarturk C, Andersson G, and Cristea IA. Personalized Psychotherapy for Adult Depression: A Meta-Analytic Review. Behavior Therapy. 2016; 47:966–80. DOI: 10.1016/j.beth.2016.04.007

94. Simon GE and Perlis RH. Personalized Medicine for Depression: Can We Match Patients With Treatments? American Journal of Psychiatry. 2010; 167:1445–55. DOI: 10.1176/appi.ajp.2010.09111680

95. Cuijpers P. Prevention: An Achievable Goal in Personalized Medicine. Dialogues in Clinical Neuroscience. 2009; 11:447–54. DOI: 10.31887/DCNS.2009.11.4/pcuijpers

96. Young JF, Jones JD, Gallop R, Benas JS, Schueler CM, Garber J, and Hankin BL. Personalized Depression Prevention: A Randomized Controlled Trial to Optimize Effects Through Risk-Informed Personalization. Journal of the American Academy of Child & Adolescent Psychiatry. 2021; 60:1116–1126.e1. DOI: 10.1016/j.jaac.2020.11.004

97. Jones JD, Long EE, Hankin BL, Gallop R, Davis M, and Young JF. Personalized Depression Prevention Reduces Dependent Stressors Among Adolescents: Results from a Randomized Controlled Trial. Journal of Clinical Child & Adolescent Psychology. 2023; online first:1–14. DOI: 10.1080/15374416.2023.2188562

98. Freund J, Titzler I, Thielecke J, Braun L, Baumeister H, Berking M, and Ebert DD. Implementing Internet- and Tele-Based Interventions to Prevent Mental Health Disorders in Farmers, Foresters and Gardeners (ImplementIT): Study Protocol for the Multi-Level Evaluation of a Nationwide Project. BMC Psychiatry. 2020; 20:424. DOI: 10.1186/s12888-020-02800-z

99. Yazd D, Wheeler SA, and Zuo A. Key Risk Factors Affecting Farmers' Mental Health: A Systematic Review. International Journal of Environmental Research and Public Health. 2019; 16:4849. DOI: 10.3390/ijerph16234849

100. Vayro C, Brownlow C, Ireland M, and March S. A Thematic Analysis of the Personal Factors Influencing Mental Health Help-seeking in Farmers. The Journal of Rural Health. 2022; 1–9. DOI: 10.1111/jrh.12705

101. Hagen BNM, Winder CB, Wootten J, McMullen CK, and Jones-Bitton A. A Systematic Review and Meta-Analysis of Depression among Farming Populations Worldwide. International Journal of Environmental Research and Public Health. 2020; 17:9376. DOI: 10.3390/ijerph17249376

# 2

# METHODS

The research questions raised in the introductory section are in need of different research methods. Article 1 – 3 operationalize a quantitative RCT design to evaluate the effectiveness of an intervention, while Article 4 uses a qualitative approach in the same target population in order to identify new aspects of interest for implementation of said intervention. The research questions in Article 5 & 6 on the other hand rely on secondary analysis of pooled quantitative data, in order to answer questions that are beyond a single RCT. In the following section, the rationale for the main methods used in each study as well as important reported characteristics and differences of the methods used are outlined. For detailed information on preregistrations, analysis models and individual instruments see the method sections in the related articles in part II.

## 2.1 EVALUATING THE EFFECTIVENESS – ARTICLES 1 - 3

To answer the question whether a personalized telephone coaching can effectively reduce mental health complaints, the first three articles describe the pragmatic randomized controlled trial TEC-A. The trial aimed to evaluate the effectiveness of a personalized telephone coaching for the prevention of depression in farmers and related professions 6, 12 and 18 months after randomization. The intervention group (n = 160) received personalized telephone coaching and was compared to a control group (n = 154) receiving psychoeducational information material on stress, depression and routine care options once via e-mail. Main outcome in this trial was reduction of depressive symptom severity at post-assessment 6 months after randomization with a wide range of secondary outcomes including self-reported depression onset, stress and quality of life. The study design and all addressed outcomes are described in detail in Article 1 – the study protocol.

### 2.1.1 *Personalized telephone coaching*

The personalized telephone coaching evaluated in the TEC-A trial was conducted by the third-party-company *IVPNetworks* and evaluated

by the study team of the University Erlangen-Nürnberg as part of the nationwide prevention project "With us in balance" initiated by the SVLFG. The whole project aims to implement evidence-based telephone- and internet-based intervention for depression prevention into routine care for farmers and related professions.

The telephone coaching was personalized in terms of overall length of the coaching as well as duration and frequency of the coaching sessions. Overall, 850 coaching minutes over 6 months were available for each participant with the option of prolonging the coaching by up to 150 minutes over 3 months. The topics and problems addressed in the coaching ware adapted to the individual participant's needs.

Another means of personalization included the possibility to expand the coaching by supporting participants to find (specialized) on-site support services like socioeconomic or family counseling for farmers or to switch to an onsite-coaching, if indicated.

No standardized treatment manuals were applied and coaching methods could vary according to the coach's background. Coaches were psychologists with a diploma or master's degree and with further training in different psychological methods (e.g., cognitive behavioral, systemic, hypnotherapeutic) with licensed psychotherapists available for supervision.

A more comprehensive picture of this personalized intervention can be drawn with numbers of actual intervention use, as described in detail in Article 2. Participants in the TEC-A trial received on average 13 sessions (M = 13.43, SD = 6.02, range 1 – 32) over a time of 6 months (M = 6.14, SD = 1.99, range: 0.36 – 11.11 months). Sessions lasted on average 48 minutes (M = 48.57, SD = 14.69, range: 1 – 120) and took place every two weeks (M = 14.09, SD = 5.87, range: 0 – 42 days) leading to an average of 652 coaching minutes (M = 652.52, SD = 325.76, range: 50 – 1598 minutes). The wide range and standard deviations indicate the timewise personalization. In total, 29/160 coaching trajectories (18.2%) were prolonged and early termination (e.g. because of lack of compliance, consent withdrawal) was reported in 17 cases (10.6%). In 47 cases (29.4%), further support was recommended or installed. No switch to on-site coaching was documented by the coaches [Article 2].

Coaches (N = 34 coaches, 31/34, 91.2% female) were characterized by IVPNetworks as having a background in systemic counselling (n = 18, 53.0%), cognitive behavioral methods (n = 11, 32.0%), hypnotherapy (n = 8, 24.0%), gestalt- or analytic therapy (n = 7, 20.5%) and/or other coaching and counseling methods (n = 19, 55.9%). Semi-structured interviews with 8 of the coaches (8/34, 23.5%) revealed conjoint goalsetting and psychoeducation as main components for the coaching. The coaching itself was generally described as a three-phase intervention with an "introduction and alliance building-phase", an active "working phase" and a concluding "stabilizing-phase" [Article 2].

### 2.1.2 *Randomized Controlled Trial*

The study design followed a pragmatic RCT approach. Participants (N = 314) were randomly assigned to either the personalized telephone coaching or the enhanced treatment as usual (TAU+). The RCT design is still widely considered the gold standard for evaluating the efficacy of therapeutic interventions [1, 2, 3] (for critiques see [4, 5, 6]). The random assignment of participants to either of the groups is supposed to eliminate all systematic influence of known and unknown confounding variables on the observed effect. Therefore observed differences between groups in assessed outcomes can be attributed to the administered intervention(s). Unsystematic effects, introduced by chance, can be quantified and expressed by reporting confidence intervals around observed estimates [2].

Pragmatic trials aim to evaluate the effectiveness of an intervention under condition as close to routine care as possible while following the strict methods of RTCs [7]. In the TEC-A trial this was operationalized by evaluating the intervention in question as it was conducted by the service provider without interference or control by the evaluating study team. Furthermore, we did not exclude participants based on depression symptoms at baseline, comorbidity and medication use as this will not be the case in routine care. However, the study enrolment process, randomization and assessments deviate from routine care use and might affect characteristics of included participants due to selection bias.

The Consolidated Standards of Reporting Trials (CONSORT) [8, 9] and their extension for pragmatic trials [7] aim to improve the conduct and reporting of RCTs. In accordance with these, the TEC-A trial was approved by the ethics committee of the Universität Erlangen-Nürnberg (No. 345_18 B), preregistered at the German Clinical Trial Registry (DRKS00015655) and procedures, sample size, primary and secondary outcomes were pre-specified in the study protocol Article 1. While treatment allocation was concealed at baseline assessment, participants could not be blinded in regards to the intervention they received, which is often the case in psychological intervention trials [10].

### 2.1.3 *Statistical analysis*

Differences in means between intervention and control group in primary and continous secondary outcomes were analyzed using separate analyses of covariance (ANCOVA) which controlled for unsystematic differences in baseline values. Analyses followed the intention-to-treat principle (ITT) in multiple imputed data (see section 2.4), meaning all participants included in the trial at baseline are considered in the analysis, regardless of adherence and treatment completion.

Additionally, sensitivity analyses with study completer was conducted to assess the rebustness of the results. A $p < .05$ was set as level of statistical significance. Since secondary outcomes were not confirmatory by nature, no adjustment of alpha level was done.

#### 2.1.3.1 *Interpretation of effects*

A statistically significant group difference alone is not sufficient to determine if a result can also be considered relevant for clinical purpose. Guidelines such as CONSORT [9] call for effect size estimates as well as their confidence intervals as indicator of their precision to be reported alongside significance levels. Estimates and indicators of effectiveness as reported in the study are explained in the following sections.

**EFFECT SIZE COHEN'S D** For group comparisons, Cohen's d is often used as the effect size [11]. Cohen's d describes the difference between group means, divided by their pooled standard deviations. The classic interpretation Cohen [12, 13] suggests an interpretation of d = 0.2 as small, d = 0.5 as medium and d = 0.80 as large, respectively. However, these values are not meant as a fixed ruling but should be interpreted in the context of the study or intervention setting and the relevant literature in the field. For a reduction in depressive symptoms, Cuijpers et al. [14] argue for a value of d = 0.24 to mark the threshold for a clinically significant change in MDD patients, based on the participants' rating of symptom improvement. In indicated prevention, meta-analyses have found effect sizes of d = -0.35 [95%–CI: -0.23, -0.47] for face-to-face interventions [15] and d = -0.39 [95%–CI: -0.25, -0.53) for online interventions [16]. These findings gave guidance for power calculation and effect interpretation in the trial.

**RELIABLE CHANGE INDEX** While Cohen's d describes the mean effect in the studied sample, the reliable change index (RCI) is based on changes on individual level. The RCI indicates whether different observed values of an outcome between baseline and later measurement represent a true change in values instead of noise in the data. Therefore, the RCI was proposed by Jacobson & Truax [17] as the observed difference in an individual divided by standard error of meassurement (usually based on the retest-reliability of the instrument). An RCI of $\geq 1.96$ would then indicate a reliable improvement in an individual. An RCI $< 1.96$ on the other hand marks a reliable deterioration in an individuals as a possible negative effect of an intervention. Differences between groups in numbers of participants with reliable improvement or deterioration were evaluated using $\chi^2$-tests.

**DEPRESSION ONSET** In the preventive context of this trial, depression onset and close-to-symptom-free status at post- and follow-up

were of additional interest. True depression onset – i.e. fulfillment of diagnostic criteria for MDD according to standardized classification like the DSM-5 or ICD-11 (International Statistical Classification of Diseases and Related Health Problems) - can only reliably be assessed by means of a clinical interview conducted by trained professionals. In the pragmatic TEC-A trial, this procedure was not reasonable because diagnostic interviews are not routinely applied in routine care prevention. Instead, a cut-off of $\geq$ 13 on the Quick Inventory of Depressive Symptomatology (QIDS-SR16) was set to estimate a clinically relevant level of depressive symptom severity indicating a MDD. The cut-off was chosen based on the recommendations of Lamoureux et al. [18] to yield the best balance of sensitivity and specificity. Additionally, a not-yet validated version of the self-reported items of Composite International Diagnosis Interview (CIDI) [19] was used to assessassess presence of MDE in the past year.

For individuals without indication for clinically relevant depression (QIDS-SR16 cut-off or CIDI questions) at baseline, group differences in depression onset are analyzed using Poisson Regression, using the time between baseline and date of follow-up assessment to account for the time in which onset could have occurred. Incidence rate ratios (IRR) are reported to describe differences in the rate of new MDD/MDE cases between groups. An IRR < 1 indicates that the incidence rate of MDD/MDE was lower in the intervention group compared to the control, while an IRR > 1 on the other hand would indicate that the incidence rate was higher in the intervention than in the control group. If the groups do not differ in incidence rate, the IRR would be 1.

For the subsample of individuals with potential MDD at baseline, group differences in numbers of individuals with symptoms reduced under this cut-off at post-treatment and follow-up and remission from potential MDE at follow-up were analyzed using logistic regression. Odds ratios (OR) are reported to compare the odds (i.e. the probability of remission divided by the probability of non-remission) when the treatment variable changes from control condition to treatment. An OR > 1 indicate an increased occurrence of remissions in the intervention compared to the control group. Likewise, an OR < 1 indicate a decreased occurrence of remissions in the intervention compared to the control group. Absence of group differences are indicated by an OR of 1.

CLOSE-TO SYMPTOM FREE STATUS Especially in the context of indicated prevention and increased symptom levels, a further measure of intervention success can be the number of participants reaching a close-to-symptom-free status after the intervention. In case of the QIDS-SR16 this equals a score of < 6 [18]. Group differences were analyzed using a $\chi^2$-test.

NUMBER-NEEDED-TO-TREAT When dealing with effect descriptors based on binary data (e.g. IRR, OR, counts), Number-Needed-to-Treat (NNT) are reported for easier interpretation in practical applications [20]. As the name suggests, NNT describes the number of individuals that needs to receive the treatment under investigation instead of the treatment in the control condition for one additional individual to profit from it (e.g. one more individual that shows reliable improvement). NTTs are presented as the inverted risk ratios of the counts in intervention and treatment group with corresponding confidence intervals [21]. In case the value of NNT is negative, it describes the number of individuals that need to receive the treatment in order for on additional person to experience negative effects; also called Number-Needed-to-Harm (NNH).

## 2.2 CREATING NEW IDEAS TO SUPPORT IMPLEMENTATION – ARTICLE 4

A qualitative approach was used in Article 4 to answer the question which factors influence acceptance of and satisfaction with the personalized telephone coaching formerly evaluated in the RCT. While the RCT already measured overall satisfaction with the intervention and engagement (i.e. session count and early termination) as a proxy for acceptance, only a qualitative approach could uncover reasons for the assessed acceptance and satisfaction. The study was therefore exploratory in nature and meant to inform and enhance ongoing implementation efforts [22].

Qualitative research has a long tradition and is an umbrella term for diverse methodological approaches and ideologies [23, 24]. In this part the reasoning for the chosen qualitative approach as well as the general procedure is described. The individual steps of the specific analyses are described in detail in Article 4. In addition, quality criteria for qualitative studies are explained in relation to their realization in the article.

### 2.2.1 *Qualitative content analysis*

Given that this study aimed to gain knowledge from the experience made by telephone coaching participants, the study design followed the phenomenological tradition by employing a content analysis focusing on acceptance and satisfaction. Phenomenology is one of the "big three" approaches in qualitative study designs with a distinct focus on exploring lived experiences from individuals. This is in contrast to Ethnography, which focuses on describing and understanding the culture in an observed group and Grounded Theory, which aims to create theories (about social phenomena) deriving from real world observations [25].

However, the sole term "content analysis" does not allow for a precise definition of a study methodology either [26, 27]. Overarching, qualitative content analysis refers to the process of *"systematically describing the meaning of qualitative material [...] by classifying material as instances of the categories of a coding frame"* [26, p. 3]. It is therefore an interpretive or hermeneutic approach instead of a pure description [28].

The content analysis approach used in Article 4 was originally proposed by Philipp Mayring [29, 30], who suggested using the term *"qualitatively orientated category-based content analysis"* instead of *"qualitative content analysis"* in order to avoid a strict dichotomization, especially when involving frequency analysis of identified categories. His approach to content analysis is characterized by rigorous guidance by a research question and the development of distinct categories in relation to that research question. Finally, the process is highly systematic and rule based and can be described by step-by-step-analyses models [30].

Depending on the research question and the material, different methods can be used to develop the category system which is the centerpiece of the content analysis. One important distinction is whether these categories are defined following a deductive approach (based in theoretic foundation), an inductive approach (derived from the material) or a combination of both [27].

### 2.2.2 *Theoretic foundation and interview guide*

In accordance with the research question, the interview guide was based on established models for acceptance and satisfaction. Each component of the model was represented with at least one question in the interview guide. The Unified Theory of Acceptance and Use of Technology (UTAUT) was chosen for the acceptance of the digital intervention [31]. It contains the elements performance expectancy, effort expectancy, social influence, facilitating conditions and behavioral intention.

The Evaluation Model, chosen to reflect different areas of participants experience, encompasses technical care quality, psychosocial care quality, accessibility, spatial and technical equipment, treatment outcome, continuity of care, financing and availability [32]. These aspects can also be classified into the dimensions of structure, process, and outcome in Donadebian's classic Quality of Care Model [33].

Additionally, the Discrepancy Theory [34] was applied by asking participants to rate expectation fulfillment with regard to the coaching, coach, and organization on a scale from 0% – 100% in order to reflect the degree of satisfaction. All model components, their definitions and example questions are presented in Article 4.

### 2.2.3 *Analysis*

The analysis was conducted on 20 verbatim transcribed interviews derived from telephone interviews with participants from the TEC-A trial (Article 1 – 3, n = 17) and the implementation study (ImplementIT [22], n = 3).

A deductive-inductive approach for structuring the content as suggested by Schreier [27] (and criticized by Mayring [30]) in relation to the constructs in questions was applied. As a starting point, the category system was deductively retrieved from the theoretical components of acceptance and satisfaction named above. Because of the overlap in the two constructs the theoretical components were – where reasonable – combined and reworked in their definitions in order to build the first broad dimensions. In the proceeding steps, aspects that potentially influenced the participants' evaluation to these dimensions were identified, paraphrased and generalized to form categories inductively under these dimensions. For each category a code name, definition and anchor example was defined. The category system was developed by a study team consisting of two psychology students, a PhD candidate in psychology and a postdoctoral researcher in the field of digital health in regular consensus meetings. For the development up to 50% of the material (10 transcribed interviews) was consecutively used. The category system was then applied to 100% of the material (20 transcribed interviews) in the final round of coding by two independent coders and the interrater agreement was calculated. Frequencies of the identified category, i.e. number of interviews containing a code and total count of the code, are reported. Identified themes were presented back to interviewees for validation purposes, indicating agreement with the theme as "yes" or "no" ratings.

### 2.2.4 *Quality criteria for qualitative studies*

In the present study the aim was to consider both the classic quality criteria (objectivity, reliability and validity) to the extent possible in qualitative research as well as more specific criteria for qualitative research [29], in study conduct, reporting and interpretation of the research material.

#### 2.2.4.1 *General quality criteria*

The major difference in considering the "classic" quality criteria for qualitative research is that they cannot simply be expressed in numbers like for example reliability measurements and associated values to determine to which degree those criteria are met. Instead, the process and decisions made during study conduct, analysis and interpretation have to be described in detail and discussed in the article. To support a comprehensive report of study conduct, the Consolidated Criteria

for Reporting Qualitative Research (COREQ) [35] guided the reporting of the study.

OBJECTIVITY Objectivity, being the independence of results from the person conducting the research, is less emphasized in qualitative research due to the nature of the research approach [26]. Most often in qualitative research — and also in the present study — researchers are involved in the context of data collection (e.g. conducting interviews) and the interpretative analyses of the material and it is therefore not reasonable to assume complete objectivity. In order to counter this dilemma, detailed descriptions of the circumstances of data collection and analyses (e.g. gender and qualification of involved researcher, relationship between researcher and participant) is necessary and ensured by the specific quality criteria and guidelines like the COREQ checklist.

INTERRATER RELIABILITY A main criterion for reliability, even though it could also be considered objectivity, is inter-rater agreement. In the present case, two raters independently coded all material according to the category system. Cohen's $\kappa$ was calculated according to Brennan & Prediger [36] by dividing the relative observed rater-agreement by the probability that agreement was by chance.

Coefficients, including Cohen's $\kappa$ and Krippendorff's $\alpha$ are thought to be ideally $\geq$ 0.80 but coefficients of minimum $\geq$ 0.67 are deemed acceptable for inter-rater agreement [37, 38].

CONSISTENCY (SPLIT-HALF RELIABILITY) The category system was developed aiming to reach theoretical saturation (i.e. no new codes are derived from adding more material, see "saturation") with only parts (at most half) of the material and applying it to the total volume of the material in the final coding. With this approach the category system developed with half of the material could be "tested" on the second half of the material.

VALIDITY First, well-established theories guided the data collection, analysis and interpretation of the study, assuring that the constructs in question were central to the process. The deductive-inductive approach also assured that no parts of the constructs were left out without notice. In addition, and more specific to qualitative research, results were presented back to the participants for validation (see "communicative validity").

#### 2.2.4.2 *Quality criteria for qualitative studies*

For the quality criteria taking into the account the qualitative study design, we followed the recommendations from Mayring [39], who suggests six key aspects that are coherent to the general approach

we applied in this study. In addition to that, data saturation, as a measure borrowed from Grounded Theory, is discussed [30]. Again, these cannot be expressed in numbers but described and interpreted within the context of the study.

RULE BASED PROCEDURE The first aspect of the quality criteria is a rule-based procedure that structures the process. Starting with the preparation of the material, rules for the transcription (based on recommendations by Dresing & Pehl [40], e.g. dealing with incomprehensible speech or pauses) ensured a good quality of material to be analyzed. Coding rules (e.g., about length of coding; coding of parts fitting multiple codes [29]) were developed alongside the category system and ensured a more transparent and objective coding procedure and allowed for better interrater reliability.

DOCUMENTED PROCEDURE Predefined and standardized steps, that nevertheless need to be adapted to the individual research endeavor, is the most important feature that distinguishes a content analysis from a mere summary. The steps for analyses were predefined by following the general approach by Mayring [29] and were documented and elaborated in all steps. Deviations from the planned procedure and the reasons why those were necessary, were highlighted and discussed in the manuscript.

PROXIMITY TO RESEARCH OBJECT The proximity to the research object is a main advantage of qualitative research because it does not rely (or even aims to avoid) a controlled research environment but instead aims to assess the situation in the individual's daily life. In this study, this was easily achieved by conducting telephone interviews which could very closely reassemble the telephone coaching situation and the medium which was of special interest.

COMMUNICATIVE VALIDITY The validation questionnaire mentioned before was sent out to all individuals participating in the interviews. The results are discussed in the article by comparing the numbers of interviews including a code with the number of participants agreeing with it in the validation questionnaire in order to assess how well the identified themes reflected participants's views.

TRIANGULATION Triangulation generally refers to a combination of different methods and data sources to try to understand the researched objective. This can be fulfilled by involving multiple different methods, investigators, theories and data sources in the research project [41].

In the present study, method triangulation was achieved by discussing the qualitative result in the light of the adjunctive RCT, which also assessed intervention use (as a proxy for acceptance) and

satisfaction as a global rating using a validated standard instrument (adapted German version of the Client Satisfaction Questionnaire for web-based interventions, CSQ-I [42]). Those measurements allowed to compare overall acceptance and satisfaction between RCT and interview results, however, only the qualitative research is able to give insight on how these ratings might have emerged and which factors contributed to it.

Investigator triangulation (i.e. including multiple individuals in the observation and interpretation of the study) was a core characteristic of the study from the beginning with always at least three individuals involved in data collection, analysis and interpretation. However, all individuals involved were female mental health researchers or psychology students, which were mostly naïve to the everyday life of the investigated farming community.

Theory triangulation was only applied for the construct of satisfaction, which was guided by three different theories for patient satisfaction (i.e. discrepancy theory [34], evaluation model [32] and quality of care model [33]), which had to be adapted to the digital context. Acceptance was based on solely one theory, but that one had already proven itself in the context of DMHIs [43, 44].

Data source triangulation was not deemed applicable, because the research question focused specifically on participants' experience, which is not reasonably assessed with other groups like the coaches or insurance counsellors. These groups and their experiences with the intervention are however considered in the implementation study [22, 45].

INTERPRETATION SAFEGUARDS The quality criterion "interpretation safeguard" describes the need to support interpretations and conclusions that are derived from the qualitative analysis with concrete arguments based on the material. While the interpretation of an effect size in quantitative settings, as described above, can be derived from cut-off scores and conventions, in qualitative research these have to be supported by the description of which part(s) fueled this interpretation and the consideration of opposing findings, which is done in the discussion part of the article.

SATURATION The concept of "data" or "theoretical" saturation is derived from theoretical sampling in Grounded Theory research in which it describes a criterion to determine if sufficient material has been collected and analyzed in order to define a fully coherent category [46, 47]. In other qualitative approaches a broader concept of "saturation" sometimes is used in the discussion about generalizability of the results and sample size justification [48, 49]. Content analyses rarely claim that the derived results are generalizable to a broader population. Nevertheless, concepts like "code" and "meaning" satura-

tion can be used to describe the characteristics of the category system [50, 51] and are reported and discussed in the article in order to evaluate the comprehensiveness of the category system in relation to our sample. Code saturation describes the point during the development of the category system where no more new codes were derived from additional material. Saturation of meaning on the other hand was employed to describe how well identified categories are founded in the research material (i.e. how many interviews/paraphrases informed the definition of a category) and can indicate if some identified aspects might not be fully represented by the material.

## 2.3 LEARNING FROM SECONDARY DATA ANALYSIS – ARTICLES 5 AND 6

Most research endeavors are planned and powered to answer one specific research question. In RCTs, this usually means they are powered to detect differences between an intervention and control group in relation to a primary outcome. The determined sample size however is usually not sufficient to answer additional questions such as differentiating effects for different subgroups.

Individual participant data (IPD) meta-analyses can help overcome this shortcoming by combining data from multiple trials and are therefore gaining popularity in the field of health sciences and personalized medicine in the recent decades [52, 53]. In comparison to traditional meta-analyses, which synthesize research evidence based on published, aggregated data, IPD meta-analyses have more flexibility and complexity in their analyses because they do not need to rely on the original methods and reporting but can use all the individual data of the trials while harmonizing the data and adjusting methodology to best fit the new research question [54]. Most commonly, IPD meta-analyses are used to examine differential treatment effect (moderation analysis) or prognostic factors of a disease outcome (prediction modeling).

In Article 5, an IPD-approach was used to pool datasets of two very similar studies (PREV-DEP I & II [55, 56, 57, 58]) and investigated the prognostic quality of outcome expectancy. We included 304 individuals with subthreshold depression who received the web-based intervention (*GET.ON Mood Enhancer*) as part of the intervention group in the RCTs. The predictive values of outcome expectancy was examined in relation to depressive symptoms, severity, close-to-symptom-free status and in a subset (n = 102) also on time to MDD onset. Exploratory moderation analyses were conducted to examine whether the predictive effect varies across sociodemographic subgroups. This should not be considered an IPD *meta-analysis* since no systematic search for studies on the topic was conducted. However, we wanted to use the statistical advantages the approach offers to consider the po-

tential varaince between the studies besides their similarity in design and conduct.

Article 6 embarked on a IPD meta-analysis of trials all investigating the same web-based intervention — *GET.ON Recovery* — focusing on insomnia. Its main aim was to investigate the potential of the intervention for indirect prevention and treatment of depression. Studies were identified by the scientific advisor at *GET.ON* and by searching the German Clinical Trail Register (DRKS) and IPD was obtained from four of seven eligable trials. A total of 563 working adults were included and group differences between the *GET.ON Recovery* web-based intervention and waiting list control groups were examined in depressive symptom severity as the primary outcome, next to potential depressive onset, close-to-symptom-free-status, RCI and a patient and clinician anchored metric for clinically relevant change (see section 2.1.3.1). Additionally, multivariable moderation analysis was carried out to investigate if effects differ across subgroups characterized by sociodemographic, clinical, and work-related variables. The article follows the Preferred Reporting Items for Systematic Review and Meta-Analyses of Individual Participant Data (PRISMA-IPD) [59] statement where applicable.

Since both studies had a very similar design, the core methodological aspects of the IPD analysis are presented combined in the following section with differences in the study designs highlighted where necessary.

### 2.3.1 *Individual participant data meta-analysis*

For IPD meta-analyses, two main different statistical approaches are available, namely a one-stage and a two-stage approach. The two-stage approach represents the workflow of a traditional meta-analytic approach more closely, since analyses are first conducted in each included trial separately and then pooled together in a second step. In the one-step approach, trial data are pooled and analyzed simultaneously, taking the nested structure of the data into account by using multi-level regression models, also called *Generalized Linear Mixed Model* (GLMM) [60, 61]. Recent recommendations by Riley at al. [62] recommend one-stage models for IPD analyses that contain studies with small number of participants or events of interest (e.g. depression onset) and was therefore used in Articles 5 & 6.

Besides the key advantage that combining data from multiple studies increases the sample size and allows for better observation of interaction effects and prognostic factors, further statistical advantages are highlighted by Riley et al. (2020) [63]. Working with IPD allows for more flexibility, as each component in the analysis including intercept, treatment effects and interactions, can be individually specified to account for underlying assumptions. These assumptions formulated

within the model parameters are crucial to understand in order to derive correct conclusions from the analysis and are therefore explained in detail in the following paragraphs.

Model parameters can be assumed to be *common*, meaning they are the same in each included trial, *stratified* allowing that they are different in each trial without making assumptions about the underlying distributions or *random* implying that they are different in each trial while assuming the same underlying distribution. This affects the interpretation for example of a treatment effect of interest and its usefulness for clinical practice. Even though common and random effects both yield one estimate in the end, the common effect describes the one single true effect for all trials, while the random effect describes the average effect across included studies with variance between studies. When assuming the treatment effect to be stratified, the treatment effect would be estimated individually for each study and would fail to give one estimate for the effect of interest and would limit the interpretability. However, stratification can be interesting for intercepts or when including covariates because no assumptions about the underling distribution(s) have to be made. This is important if one is interested in the adjusted effect (i.e. one that would add additional value to the prediction) of a newly proposed predictor such as outcome expectancy in addition to already known predictor, which in the example of predicing treatment effect were baseline severity, sex and age (Article 5).

Since these assumptions made for the analyses in Article 5 and Article 6 are crucial for the interpretation of the results, they were pre-specified in the respective preregistrations on the Open Sience Framework (OSF) [64, 65] and separate models were specified for each hypothesis. The overall model fit was described with $R^2$ and is presented alongside all results. $\tau^2$ as a measure of variability in the estimated true effect that is based on the heterogeneity between studies was reported where between study differences were of interest (Article 6) and not merely controlled for (Article 5).

#### 2.3.1.1 *Moderation analysis*

Both articles embarked into moderation analyses, however with slightly different aims. In Article 5 univariable moderation analyses were carried out to examine whether the predictive power of expectancy differs across sociodemographic groups while controlling for other potential covariates in order to estimate the additional or adjusted effect of outcome expectancy as a predictor. In Article 6 the aim was to investigate differential treatment effects across a wide range of sociodemographic, clinical and work-related variables using multivariable moderation and univariable interactions were only examined to select potential candidate moderators.

UNIVARIABLE MODERATION For one-stage IPD analyses, Riley et al. [66] highlight the need to be cautious about the amalgamation of within- and across-trial information when estimating (treatment) effect modifiers. The aim is to base estimates solely based on the within trial information so that subgroup analyses or interaction effects are not influenced by the distribution of participant characteristics of the individual studies. A drastic example would be if a study containing only female participants would still be used in estimating interactions between sex and treatment effect even though the interaction could not be observed within the individual study. To avoid this amalgamation, Riley et al. [66] recommend to either a) stratify all parameters outside of the interaction term or b) centering the covariates by their study-specific mean and including the covariate mean in the model term. The interpretation is then similar to single regression models, where the interaction describes how a one-step increase in the moderator effect the relation from the effect of interest (i.e. either the treatment effect or the effect of the predictor of interest) to the outcome.

MULTIVARIABLE MODERATION Multivariable moderation aims to incorporate more than one independent variable in the prediction of an outcome and consider the relationship between the different independent variables as well [67]. With the machine learning methods of *recursive partitioning* or also named *decision-tree methods* as presented by Fokkema et al. [68, 69] the idea of these rather complex analyses can be presented in an easy to interpret, dichotomous tree, that can help guiding decisions in clinical practice.

The general idea of algorithm-build decision trees is that in every step first a splitting variable (out of all variables introduced as potential partitioning variables) is identified based on the strongest statistical test statistic (Bonferroni corrected p-value) describing the association with the dependent variable. Then a cut-point for the splitting variable is selected based on an optimization process. This step is repeated until no more splitting variables can be identified based on a pre-specified $\alpha$-level [69]. The decision-tree ends in so called terminal nodes that are characterized by specific, interpretable values of the candidate predictors (e.g. females, over 25 years with at least moderate stress levels). The effect in the individual subgroups can then be tested by fitting the original model in the identified groups to estimate precise effects.

Since not much is known about moderators in indirect prevention and treatment of depression, a wide range of variables were considered. However, adding a large amount of candidate predictors diminishes the possibility to identify moderating effects, because splitting variables are chosen based on Bonferroni corrected p-values. Adding an abundance of candidate predictors would inflate the necessary p-level to an unrealistic degree, making an informed pre-selection

("pre-pruning") of candidate predictors necessary. Relying on statistical tests for variable selection, however, allows for a stopping criterion for the splitting algorithm and thus allows to receive meaningful groups. For the decision tree in Article 6 candidate predictors therefore were first selected based on potential relevance for depression outcome (e.g. focus on sociodemographic variables instead of specific insomnia subtypes) and then were chosen data-driven based on 1) statistically significant univariable moderation analysis or 2) a variable importance of > 0 in a random forest tree. In the later method described by Garge et al. [70], 300 bootstrap samples were created and a random forest tree was fit in each to determine how often candidate variables were used as splitting variables in the tree. Variable importance was then classified using the permutation accuracy method which ranks variables according to their relevance in predictive accuracy. Lower scores indicate that, averaged across trees, the variable is less predictive compared to the other candidates.

Despite their easy interpretation, results from decision-trees should be carefully interpreted and viewed as a primarily explorative tool until external validation. Recursive partitioning has a tendency to overfit the resulting model given that the same data is used for training the algorithm and predicting the outcome [69]. This was partly accounted for in the tree-model in Article 6 by reporting an optimism-corrected $R^2$ via a bootstrapping method [71, 72] and testing the initial effects in the identified subgroups to consider the imputation insecurity again which was not considered when fitting the tree in aggregated data. However, external validation of the decision tree in a new sample would be a better way to confirm results, but was beyond the scope of this dissertation.

#### 2.3.1.2 *Pseudo-Bayesian approach*

Both IPD analyses in this dissertation included a small number of studies ($k < 5$) which were conducted in similar contexts and with similar interventions. In such cases, it is difficult to correctly estimate the assumed variance between studies because it might be improperly estimated to be zero in the imputation and multi-level regression models. To avoid these singularity problems, a *pseudo-Bayesian* approach was used to employ an approximate maximum a posteriori (MAP) estimate instead of a frequentist maximum likelihood approach [73]. The priors assured that at least some heterogeneity was assumed between trials, as highly plausible based on the study designs, while still remaining largely uninformative themselves. In dependence of the model specifications in the studies either a weakly informative *Wishart* prior (for a random intercept) or *gamma* prior (for stratified intercepts) was applied to the covariance matrix. While this procedure was necessary for the specification and convergence of the imputation

and analyses models, it does not need to be specially considered when interpreting the model parameters and results.

### 2.3.2 *Outcomes & reported estimates*

The focus of Article 5 was the predictive value of outcome expectancy in explaining depressive outcomes after using a DMHI for insomnia. The outcomes selected for these prediction models nevertheless were similar again to the criteria used for treatment evaluation (compare section 2.1.3.1), namely depressive symptom severity (this time measured with the Center for Epidemiological Studies Depression Scale [CES-D] [74]), close-to-symptom-free status and MDD onset. Deviating from the method of self-reported MDD onset described before for Articles 1 – 3, MDD onset in this sample was measured with more objective diagnostic interviews and the time-to-onset could be established, allowing to conduct a survival analysis. The level of statistical significance was set to $p < .05$ and p-values were adjusted for multiple testing using the Bonferroni-Holm method [75].

In Article 6 on the other hand the (differential) treatment effect of insomnia interventions on depressive symptoms were in focus. Therefore, outcomes were again very similar to the ones reported for the treatment effect in the single RCT and included depressive symptom severity (CES-D) as the primary outcome, next to onset of clinically relevant depressive symptoms, close-to-symptom-free-status, RCI and additionally a more patient centered (anchored) metric for clinically relevant change. The predictive effect of outcome expectancy on the difference outcomes were adjusted for baseline symptom severity, sex and age. Analysis of group differences were controlled for differences in baseline severity and tested using a $p < .05$ as level of statistical significance. Significance values were adjusted using Bonferroni correction [76] to match the correction used in the decision tree.

OBJECTIVE DEPRESSION ONSET From the studies included in Article 5, one study did not only verify MDD absence at baseline via diagnostic interview but also assessed onset of MDD within the 12 months study period. Telephonic diagnostic interviews have shown to be comparable to their face-to-face counterparts [77, 78]. Time to onset of MDD was assessed using the Life Chart method by Lyketsos et al. [79] to avoid potential recall bias. The analysis of the data deviates a bit from the procedures described before, because data was derived from a single study and did not need to consider the nested structure and was analyzed using Cox regression and thus did not rely on imputed data (see section 2.4.)

A Cox regression analysis was conducted with right-censored data, meaning that participants were *censored* (i.e. no longer at risk) at

the time point when participants either developed a MDD or at the last assessment before they dropped out of the study. Hazard ratios with their 95% confidence intervals (CI) were reported. While HRs are typically used in group comparison, HRs can also be calculated from continuous predictors such as outcome expectancy in this specific analysis. In the interpretation the risk of MDD onset, changes according to the estimated HR with each unit increase in the predictor (i.e. one more point on the outcome expectancy scale). A HR of 1 would thus mean that an increase in expectancy has no effect on the risk of MDD onset, while a value < 1 would indicate that an increase in expectancy would lower the risk of developing MDD during the observation time. A HR > 1 would consequently indicate that an increase in the predictor heightened the risk to develop MDD in the observation timeframe.

To approximate overall goodness-of-fit measure, Nagelkerke's pseudo $R^2$ was reported alongside concordance, which is supposed to be a more adequate estimation for survival analysis [80]. Concordance describes how well a model can correctly rank survival times given individual risk factors or predictors. Concordance values between 0.50 – 0.55 are considered no better than chance with higher values indicating a better discriminatory power.

ANCHOR-BASED CLINICALLY RELEVANT CHANGE In section 2.1.3.1 the problem was discussed to determine whether a difference in pre- and post-treatment scores are clinically meaningful. In addition to the statistical approaches in that section, recently more focus is given on judgments from participants whether observed differences are meaningful to them [81]. For the 15-item short form of the CES-D, a change of 9 – 11 points on the instrument was determined as minimal clinically important difference (MCID) based on both patient- and clinician-rated change [82] and was recommended in the newest German depression treatment guideline [83]. This is equivalent to roughly 33% change in symptoms, which was used as a criterion in Article 6, which employed the long version of the CES-D. Numbers of participants exceeding their individual 33% symptom improvement based on their baseline scores were compared between groups. Group differences were expressed as OR with according CIs and NNT.

### 2.3.3 *Risk of bias assessment*

When summarizing evidence or including different data sources, it is important to assess the quality of the original evidence since all forthcoming results depend on it. Therefore, a risk of bias (RoB) assessment was carried out in Article 6. This was done using the revised version of the Cochrane risk of bias assessment tool for randomized trials "RoB2" [84] and the related Excel tool [85]. RoB ratings are completed with regard to the outcome of interested – in this case

depressive symptom severity at post-treatment and/or follow-up assessment. With the RoB2 tool five domains for potential bias are rated: "randomization process", "deviations from interventions", "missing data", "outcome measurement" and "selective reporting". For each domain, 3 – 5 signaling question guide a judgement for each domain which can be either "low risk", "some concern" or "high risk" for bias. The judgments of each domain are then combined to one overall RoB rating per study. The decision process is guided by an algorithm with defined rules on how to include all answers to the signaling question. An overall rating of "low risk" is only given when all domains are rated as such, "some concerns" is given when one of the domains is rated as such and "high risk" is given as an overall rating if at least one domain is considered to represent a high risk of bias or multiple domains are rated as "some concerns".

RoB assessment for the studies included in Articles 6 was conduced by the first two authors (who were not involved in the original studies) independently and in duplicate for all published articles and the trial registrations. Disagreement on ratings in the individual domains and the overall rating was dissolved by discussion. All studies were rated as having a high risk of bias, due to a rating of as "high risk" in the domain of outcome measurements. In all studies, depressive symptom severity was assessed using self-report measures from non-blinded participants. This is a known problem in the field of psychological intervention where blinding is difficult [10] and especially in studies on DMHIs where self-report can be readily integrated in the intervention or at least assessed in the same digital medium. Therefore, special attention was paid when comparing the study results to the existing literature in terms of assessments and control conditions used in prior studies. The choice of control conditions and their potential for blinding are discussed in section 9.2.3.3.

## 2.4 DEALING WITH MISSING DATA

Missing data is a common phenomenon in psychological intervention research, because it is highly likely that not all participants in a study will report back full data on all measurements and on all assessments for a wide range of reasons (e.g. burdensome assessment, technical problems, lack of motivation) [86]. This poses a challenge for analyzing mostly quantitative data[1], because assumptions have to made on why the missing data occurs. Rubin (1976) [87] put forward the idea to view missing data as a result of a probability process: data points could all have the same chance of being missing, which is

1 Missing data as such usually does not occur in qualitative (interview) data. However, selection of interview partners and material might bias available information and thus lead to "missing" information. The sampling technique and limitation deriving from it are discussed in Article 4.

considered to be Missing Completely at Random (MCAR) and with no relation to other (observed) data. Data points are considered Missing At Random (MAR) if they are partly dependent on observed data or Not Missing At Random (NMAR) if they are dependent at least partly on unobserved data [88].

For all missing data in the studies presented here, MAR was assumed as it is recommended as a starting point in clinical trials [88]. Multiple imputation by chained equation (MICE) was used to account for the missing data following the general approach presented by van Buuren [88]. In R this was implemented using the *mice* [89] and *miceadds* [90] packages.

Multiple imputation (MI) describes a process in which, based on the observed data with missings, multiple datasets are created which are completed by realistic values drawn from a distribution specified for each variable. MI is becoming the standard in clinical research because it allows to account for the insecurity about with which values the missing data should be replaced [88].

Therefore, first a predictor matrix is defined, specifying which variables should be used as predictors in the imputation model to inform potential values that replace the missings under the MAR assumption. Second, the underlying methods to use for each variable with missings has to be specified. In the studies reported in this work, this usually was *predictive mean matching* (PMM). PMM is a procedure where a small set so-called "donors" are identified in the complete cases which have a combination of predicting variables that is similar to the case with missing values and from which a possible value is randomly selected to replace the missing value. Because PMM takes a value from a complete observation, the method has the advantage that no unrealistic values (e.g. depression scores beyond the questionnaire's values) are created. For the studies using an IPD approach, the models defined to impute the missing data had to consider the hierarchical structure of the data, thus two-level PMM or two-level PAN (an imputation method specifically developed for nested data [91]) were used to account for this. Similar to the approach described in the section 2.3.1.2 the same weakly informative priors were used in the imputation models to avoid singularity during the prediction of missing values and ensure a better fit (congeniality) of imputation and analysis models.

The result of the imputation process are $m$ individual data sets in which missing values were replaced. Different recommendations are available for the number of imputed data sets that should be created, e.g. percentage of missing data [92] or something in the range of $m = 20 - 200$ sets [88]. The recommendation is to then fit the analysis models separately in all individual data sets and the pool the estimates according to the Rubin's rules [93, 94], which considers not only the variance in the underlying sample but also the variance introduced

by the imputation. For methods such as the random forest and the model-based recursive partitioning, no method has yet been developed to implement these in MI, therefore aggregated data had to be used.

## REFERENCES


1. Rothman KJ and Greenland S. Types of Epidemiologic Studies. *Modern Epidemiology*. Ed. by Rothman KJ, Greenland S, and Lash TL. 2nd ed. 74–75. USA, Philadelphia, PA: Lippincott Raven, 1998. Available from: https://scholar.google.de/scholar?q=Rothman+%26+Greenland,+1998&hl=de&as_sdt=0&as_vis=1&oi=scholart
2. Altman DG. Why We Need Confidence Intervals. World Journal of Surgery. 2005; 29:554–6. DOI: 10.1007/s00268-005-7911-0
3. Lilienfeld SO, McKay D, and Hollon SD. Why Randomised Controlled Trials of Psychological Treatments Are Still Essential. The Lancet Psychiatry. 2018; 5:536–8. DOI: 10.1016/S2215-0366(18)30045-2
4. Cartwright N. Are RCTs the Gold Standard? BioSocieties. 2007; 2:11–20. DOI: 10.1017/S1745855207005029
5. Cartwright N. Predicting What Will Happen When We Act. What Counts for Warrant? Preventive Medicine. 2011; 53:221–4. DOI: 10.1016/j.ypmed.2011.08.011
6. Grossman J and Mackenzie FJ. The Randomized Controlled Trial: Gold Standard, or Merely Standard? Perspectives in Biology and Medicine. 2005; 48:516–34. DOI: 10.1353/pbm.2005.0092
7. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D, and for the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups. Improving the Reporting of Pragmatic Trials: An Extension of the CONSORT Statement. BMJ. 2008; 337:a2390–a2390. DOI: 10.1136/bmj.a2390
8. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, and Altman DG. CONSORT 2010 Explanation and Elaboration: Updated Guidelines for Reporting Parallel Group Randomised Trials. Journal of Clinical Epidemiology. 2010; 340. DOI: 10.1016/j.jclinepi.2010.03.004
9. Schulz KF, Altman DG, and Moher D. CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials. BMC Medicine. 2010; 8:18. DOI: 10.1136/bmj.c332
10. Juul S, Gluud C, Simonsen S, Frandsen FW, Kirsch I, and Jakobsen JC. Blinding in Randomised Clinical Trials of Psychological Interventions: A Retrospective Study of Published Trial Reports. BMJ Evidence-Based Medicine. 2021; 26:109–9. DOI: 10.1136/bmjebm-2020-111407
11. Ferguson CJ. An Effect Size Primer: A Guide for Clinicians and Researchers. Professional Psychology: Research and Practice. 2009; 40:532–8. DOI: 10.1037/a0015808
12. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. USA, Hillsdale, NJ: L. Erlbaum Associates, 1988
13. Lakens D. Calculating and Reporting Effect Sizes to Facilitate Cumulative Science: A Practical Primer for t-Tests and ANOVAs. Frontiers in Psychology. 2013; 4:863. DOI: 10.3389/fpsyg.2013.00863
14. Cuijpers P, Turner EH, Koole SL, Van Dijke A, and Smit F. What Is the Threshold for a Clinically Relevant Effect? The Case of Major Depressive Disorders. Depression and Anxiety. 2014; 31:374–8. DOI: 10.1002/da.22249

15. Cuijpers P, Koole SL, Van Dijke A, Roca M, Li J, and Reynolds CF. Psychotherapy for Subclinical Depression: Meta-analysis. British Journal of Psychiatry. 2014; 205:268–74. DOI: 10.1192/bjp.bp.113.138784

16. Reins JA, Buntrock C, Zimmermann J, Grund S, Harrer M, Lehr D, Baumeister H, Weisel K, Domhardt M, Imamura K, Kawakami N, Spek V, Nobis S, Snoek F, Cuijpers P, Klein JP, Moritz S, and Ebert DD. Efficacy and Moderators of Internet-Based Interventions in Adults with Subthreshold Depression: An Individual Participant Data Meta-Analysis of Randomized Controlled Trials. Psychotherapy and Psychosomatics. 2021; 90:94–106. DOI: 10.1159/000507819

17. Jacobson NS and Truax P. Clinical Significance: A Statistical Approach to Defining Meaningful Change in Psychotherapy Research. Journal of Consulting and Clinical Psychology. 1991; 59:12–9. DOI: 10.1037/0022-006X.59.1.12

18. Lamoureux BE, Linardatos E, Fresco DM, Bartko D, Logue E, and Milo L. Using the QIDS-SR16 to Identify Major Depressive Disorder in Primary Care Medical Patients. Behavior Therapy. 2010; 41:423–31. DOI: 10.1016/j.beth.2009.12.002

19. Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, Demyttenaere K, Ebert DD, Green JG, Hasking P, Murray E, Nock MK, Pinder-Amaker S, Sampson NA, Stein DJ, Vilagut G, Zaslavsky AM, and Kessler RC. WHO World Mental Health Surveys International College Student Project: Prevalence and Distribution of Mental Disorders. Journal of Abnormal Psychology. 2018; 127:623–38. DOI: 10.1037/abn0000362

20. Kraemer HC and Kupfer DJ. Size of Treatment Effects and Their Importance to Clinical Research and Practice. Biological Psychiatry. 2006; 59:990–6. DOI: 10.1016/j.biopsych.2005.09.014

21. Bender R. Calculating Confidence Intervals for the Number Needed to Treat. Controlled Clinical Trials. 2001; 22:102–10. DOI: 10.1016/S0197-2456(00)00134-3

22. Freund J, Titzler I, Thielecke J, Braun L, Baumeister H, Berking M, and Ebert DD. Implementing Internet- and Tele-Based Interventions to Prevent Mental Health Disorders in Farmers, Foresters and Gardeners (ImplementIT): Study Protocol for the Multi-Level Evaluation of a Nationwide Project. BMC Psychiatry. 2020; 20:424. DOI: 10.1186/s12888-020-02800-z

23. Mey G and Mruck K. Handbuch Qualitative Forschung in Der Psychologie. 1st ed. Germany, Wiesbaden: Verlag für Sozialwissenschaften, 2010

24. LaMarre A and Chamberlain K. Innovating Qualitative Research Methods: Proposals and Possibilities. Methods in Psychology. 2022; 6:100083. DOI: 10.1016/j.metip.2021.100083

25. Korstjens I and Moser A. Series: Practical Guidance to Qualitative Research. Part 2: Context, Research Questions and Designs. European Journal of General Practice. 2017; 23:274–9. DOI: 10.1080/13814788.2017.1375090

26. Schreier M. Qualitative Content Analysis in Practice. USA, Los Angeles, CA: SAGE, 2012

27. Schreier M. Varianten Qualitativer Inhaltsanalyse: Ein Wegweiser Im Dickicht Der Begrifflichkeiten. Forum: Qualitative Social Research. 2014; 15:18. DOI: 10.17169/fqs-15.1.2043

28. Polit DF and Beck CT. Nursing Research: Generating and Assessing Evidence for Nursing Practice. 11th ed. USA, Philadelphia, PA: Wolters Kluwer, 2021

29. Mayring P. Qualitative Inhaltsanalyse. 12th ed. Germany, Weinheim: Beltz, 2015

30. Mayring P. Qualitative Content Analysis: Demarcation, Varieties, Developments. Forum: Qualitative Social Research. 2019; 20:16. DOI: http://dx.doi.org/10.17169/fqs-20.3.3343.

31. Venkatesh, Morris, Davis, and Davis. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly. 2003; 27:425–78. DOI: 10.2307/30036540

32. Ware JE, Snyder MK, Wright WR, and Davies AR. Defining and Measuring Patient Satisfaction with Medical Care. Evaluation and Program Planning. 1983; 6:247–63. DOI: 10.1016/0149-7189(83)90005-8

33. Donadebian A. *Evaluating the Quality of Medical Care*. 166-206. 1966

34. Fox JG and Storms DM. A Different Approach to Sociodemographic Predictors of Satisfaction with Health Care. Social Science and Medicine Part A Medical Psychology and Medical. 1981; 15:557–64. DOI: 10.1016/0271-7123(81)90079-1

35. Tong A, Sainsbury P, and Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): A 32-Item Checklist for Interviews and Focus Groups. International Journal for Quality in Health Care. 2007; 19:349–57. DOI: 10.1093/intqhc/mzm042

36. Brennan RL and Prediger DJ. Coefficient Kappa: Some Uses, Misuses, and Alternatives. Educational and Psychological Measurement. 1981; 41:687–99. DOI: 10.1177/001316448104100307

37. Krippendorff K. Reliability in Content Analysis: Some Common Misconceptions and Recommendations. Human Communication Research. 2004; 30:411–33. DOI: 10.1111/j.1468-2958.2004.tb00738.x

38. Mayring P. Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution. SSOAR Open Access Repository. Austria, Klagenfurt: Beltz, 2014. Available from: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173

39. Mayring P. Einführung in Die Qualitative Sozialforschung. Eine Anleitung Zu Qualitativem Denken. 6th ed. Germany, Weinheim: Beltz, 2016

40. Dresing T and Pehl Thorsten. Praxisbuch Interview, Transkription et Analyse: Anleitungen Und Regelsysteme Für Qualitativ Forschende. 6th ed. Germany, Marburg: self-publishing, 2015

41. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, and Neville AJ. The Use of Triangulation in Qualitative Research. Oncology Nursing Forum. 2014; 41:545–7. DOI: 10.1188/14.ONF.545-547

42. Boß L, Lehr D, Reis D, Vis C, Riper H, Berking M, and Ebert DD. Reliability and Validity of Assessing User Satisfaction with Web-Based Health Interventions. Journal of Medical Internet Research. 2016; 18:e234. DOI: 10.2196/jmir.5952

43. Philippi P, Baumeister H, Apolinário-Hagen J, Ebert DD, Hennemann S, Kott L, Lin J, Messner EM, and Terhorst Y. Acceptance towards Digital Health Interventions – Model Validation and Further Development of the Unified Theory of Acceptance and Use of Technology. Internet Interventions. 2021; 26:100459. DOI: 10.1016/j.invent.2021.100459

44. Harst L, Lantzsch H, and Scheibe M. Theories Predicting End-User Acceptance of Telemedicine Use: Systematic Review. Journal of Medical Internet Research. 2019; 21:e13117. DOI: 10.2196/13117

45. Freund J, Ebert DD, Thielecke J, Braun L, Baumeister H, Berking M, and Titzler I. Using the Consolidated Framework for Implementatiom Research to Evaluate a Nationwide Depression Prevention Project (ImplementIT) from the Perspective of Health Care Workers and Implementers: Results on the Implementation of Digital Interventions for Farmers. Frontiers in Digital Health. 2023; 4:1083143. DOI: 10.3389/fdgth.2022.1083143

46. Cho J and Lee EH. Reducing Confusion about Grounded Theory and Qualitative Content Analysis: Similarities and Differences. The Qualitative Report. 2014; 19:1–20. DOI: 10.46743/2160-3715/2014.1028

47. Glaser BG and Strauss AL. The Discovery of Grounded Theory: Strategies for Qualitative Research. 5th ed. USA, New Brunswick, NJ: Aldine Transaction, 1967

48. Mayring P. On Generalization in Qualitatively Oriented Research. Forum: Qualitative Social Research. 2007; 8:26. DOI: http://nbnresolving.de/urn:nbn:de:0114-fqs0703262.

49. O'Reilly M and Parker N. 'Unsatisfactory Saturation': A Critical Exploration of the Notion of Saturated Sample Sizes in Qualitative Research. Qualitative Research. 2013; 13:190–7. DOI: 10.1177/1468794112446106

50. Hennink MM, Kaiser BN, and Marconi VC. Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough? Qualitative Health Research. 2017; 27:591–608. DOI: 10.1177/1049732316665344

51. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, Burroughs H, and Jinks C. Saturation in Qualitative Research: Exploring Its Conceptualization and Operationalization. Quality and Quantity. 2018; 52:1893–907. DOI: 10.1007/s11135-017-0574-8

52. Stewart LA and Tierney JF. To IPD or Not to IPD?: Advantages and Disadvantages of Systematic Reviews Using Individual Patient Data. Evaluation & the Health Professions. 2002; 25:76–97. DOI: 10.1177/016327870202500100 6

53. Cuijpers P, Ciharova M, Quero S, Miguel C, Driessen E, Harrer M, Purgato M, Ebert D, and Karyotaki E. The Contribution of "Individual Participant Data" Meta-Analyses of Psychotherapies for Depression to the Development of Personalized Treatments: A Systematic Review. Journal of Personalized Medicine. 2022; 12:93. DOI: 10.3390/jpm1201 0093

54. Tierney JF, Vale C, Riley R, Smith CT, Stewart L, Clarke M, and Rovers M. Individual Participant Data (IPD) Meta-analyses of Randomised Controlled Trials: Guidance on Their Use. PLOS Medicine. 2015; 12:e1001855. DOI: 10.1371/journal.pmed.1001855

55. Buntrock C, Ebert DD, Lehr D, Cuijpers P, Riper H, Smit F, and Berking M. Evaluating the Efficacy and Cost-Effectiveness of Web-Based Indicated Prevention of Major Depression: Design of a Randomised Controlled Trial. BMC Psychiatry. 2014; 14:25. DOI: 10.1186/1471-244X-14-25

56. Buntrock C, Ebert D, Lehr D, Riper H, Smit F, Cuijpers P, and Berking M. Effectiveness of a Web-Based Cognitive Behavioural Intervention for Subthreshold Depression: Pragmatic Randomised Controlled Trial. Psychotherapy and Psychosomatics. 2015; 84:348–58. DOI: 10.1159/000438673

57. Buntrock C, Ebert DD, Lehr D, Smit F, Riper H, Berking M, and Cuijpers P. Effect of a Web-Based Guided Self-Help Intervention for Prevention of Major Depression in Adults with Subthreshold Depression a Randomized Clinical Trial. JAMA - Journal of the American Medical Association. 2016; 315:1854–63. DOI: 10.1001/jama.2016.4326

58. Ebert DD, Buntrock C, Lehr D, Smit F, Riper H, Baumeister H, Cuijpers P, and Berking M. Effectiveness of Web- and Mobile-Based Treatment of Subthreshold Depression With Adherence-Focused Guidance: A Single-Blind Randomized Controlled Trial. Behavior Therapy. 2018; 49:71–83. DOI: 10.1016/j.beth.2017.05.004

59. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, and Tierney JF. Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. JAMA. 2015; 313:1657. DOI: 10.1001/jama.2015.3656

60. Burke DL, Ensor J, and Riley RD. Meta-Analysis Using Individual Participant Data: One-Stage and Two-Stage Approaches, and Why They May Differ. Statistics in Medicine. 2017; 36:855–75. DOI: 10.1002/sim.7141

61. Kontopantelis E. A Comparison of One-stage vs Two-stage Individual Patient Data Meta-analysis Methods: A Simulation Study. Research Synthesis Methods. 2018; 9:417–30. DOI: 10.1002/jrsm.1303

62. Riley RD, Burke DL, and Morris T. One-stage versus Two-stage Approach to IPD Meta-Analysis: Differences and Recommendations. *Individual Participant Data Meta-Analysis*. Ed. by Riley RD, Tierney JF, and Stewart LA. 1st ed. 199-217. UK, Chichester: Wiley, 2021. DOI: 10.1002/9781119333784.ch8

63. Riley RD, Stewart LA, and Tierney JF. Individual Participant Data Meta-Analysis: A Handbook for Healthcare Research. UK, Chichester: Wiley, 2021

64. Thielecke J, Buntrock C, Harrer M, Schuurmans L, Ebert D, Lehr D, Behrendt D, Sander L, and Spanhel K. Online Sleep Trainings for the Prevention and Treatment of Depression – An Individual Patient Data Meta-Analysis. Open Science Framework. 2022. DOI: 10.17605/OSF.IO/XCUS5

65. Thielecke J, Kuper P, and Buntrock C. Is Outcome Expectancy a Predictor for Depression Symptoms in iCBT for Depression Prevention – a Secondary Data Analyses. Open Science Framework. 2022. DOI: 10.17605/OSF.IO/9XJ6A

66. Riley RD, Debray TP, Fisher D, Hattle M, Marlin N, Hoogland J, Gueyffier F, Staessen JA, Wang J, Moons KG, Reitsma JB, and Ensor J. Individual Participant Data Meta-analysis to Examine Interactions between Treatment Effect and Participant-level Covariates: Statistical Recommendations for Conduct and Planning. Statistics in Medicine. 2020; 39:2115–37. DOI: 10.1002/sim.8516

67. Hidalgo B and Goodman M. Multivariate or Multivariable Regression? American Journal of Public Health. 2013; 103:39–40. DOI: 10.2105/AJPH.2012.300897

68. Fokkema M, Smits N, Zeileis A, Hothorn T, and Kelderman H. Detecting Treatment-Subgroup Interactions in Clustered Data with Generalized Linear Mixed-Effects Model Trees. Behavior Research Methods. 2018; 50:2016–34. DOI: 10.3758/s13428-017-0971-x

69. Fokkema M, Edbrooke-Childs J, and Wolpert M. Generalized Linear Mixed-Model (GLMM) Trees: A Flexible Decision-Tree Method for Multilevel and Longitudinal Data. Psychotherapy Research. 2021; 31:329–41. DOI: 10.1080/10503307.2020.1785037

70. Garge NR, Bobashev G, and Eggleston B. Random Forest Methodology for Model-Based Recursive Partitioning: The mobForest Package for R. BMC Bioinformatics. 2013; 14:125. DOI: 10.1186/1471-2105-14-125

71. Smith GCS, Seaman SR, Wood AM, Royston P, and White IR. Correcting for Optimistic Prediction in Small Data Sets. American Journal of Epidemiology. 2014; 180:318–24. DOI: 10.1093/aje/kwu140

72. Harrell FE, Lee KL, and Mark DB. Multivariable Prognostic Models: Issues In Developing Models, Evaluating Assumptions And Adequacy, And Measuring And Reducing Errors. Statistics in Medicine. 1996; 15:361–87. DOI: 10.1002/(SICI)1097-0258(19960229)15:4<361

73. Chung Y, Gelman A, Rabe-Hesketh S, Liu J, and Dorie V. Weakly Informative Prior for Point Estimation of Covariance Matrices in Hierarchical Models. Journal of Educational and Behavioral Statistics. 2015; 40:136–57. DOI: 10.3102/1076998615570945

74. Hautzinger M, Bailer M, Hofmeister D, and Keller F. Allgemeine Depressionsskala. 2nd ed. Germany, Göttingen: Hogrefe, 2012

75. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics. 1979; 6:65–70. Available from: http://www.jstor.org/stable/4615733

76. Emerson RW. Bonferroni Correction and Type I Error. Journal of Visual Impairment & Blindness. 2020; 114:77–8. DOI: 10.1177/0145482X20901378

77. Lobbestael J, Leurgans M, and Arntz A. Inter-Rater Reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). Clinical Psychology & Psychotherapy. 2011; 18:75–9. DOI: 10.1002/cpp.693

78. Crippa JAS, de Lima Osório F, Del-Ben CM, Filho AS, da Silva Freitas MC, and Loureiro SR. Comparability Between Telephone and Face-to-Face Structured Clinical Interview for DSM-IV in Assessing Social Anxiety Disorder. Perspectives in Psychiatric Care. 2008; 44:241–7. DOI: 10.1111/j.1744-6163.2008.00183.x

79. Lyketsos CG, Nestadt G, Cwi J, and Heithoff K. The Life Chart Interview: A Standardized Method to Describe the Course of Psychopathology. International Journal of Methods in Psychiatric Research. 1994; 4:143–55

80. Therneau TM and Atkinson E. Concordance. 2020. Accessed on: 17.11.2022. Available from: https://rdrr.io/cran/survival/man/concordance.html

81. Wyrwich KW and Norman GR. The Challenges Inherent with Anchor-Based Approaches to the Interpretation of Important Change in Clinical Outcome Assessments. Quality of Life Research. 2022; preprint. DOI: 10.1007/s11136-022-03297-7

82. Haase I, Winkeler M, and Imgart H. Ascertaining Minimal Clinically Meaningful Changes in Symptoms of Depression Rated by the 15-item Centre for Epidemiologic Studies Depression Scale. Journal of Evaluation in Clinical Practice. 2022; 28:500–6. DOI: 10.1111/jep.13629

83. Bundesärztekammer BÄK, Kassenärztliche Bundesvereinigung KBV, and Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften AWMF. Nationale VersorgungsLeitlinie Unipolare Depression - Langfassung. Version 3.0. 2022. DOI: 10.6101/AZQ/000493

84. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, and Higgins JPT. RoB 2: A Revised Tool for Assessing Risk of Bias in Randomised Trials. BMJ. 2019; 366:l4898. DOI: 10.1136/bmj.l4898

85. Higgins JPT, Savović J, Page MJ, and Sterne JA. Revised Cochrane Risk-of-Bias Tool for Randomized Trials (RoB 2). 2019 [Accessed on: 09-04-2023]. Available from: riskofbias.info

86. Palmer MJ, Mercieca-Bebber R, King M, Calvert M, Richardson H, and Brundage M. A Systematic Review and Development of a Classification Framework for Factors Associated with Missing Patient-Reported Outcome Data. Clinical Trials: Journal of the Society for Clinical Trials. 2018; 15:95–106. DOI: 10.1177/1740774517741113

87. Rubin DB. Inference and Missing Data. Biometrika. 1976; 63:581–92. DOI: 10.1093/biomet/63.3.581

88. Van Buuren S. Flexible Imputation of Missing Data. 2nd ed. London, UK: Chapman & Hall/ CRC Press, 2018

89. Van Buuren S and Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 2011; 45:1–67. DOI: 10.18637/jss.v045.i03

90. Robitzsch A and Grund S. Miceadds: Some Additional Multiple Imputation Functions, Especially for 'Mice'. manual. 2022. Accessed on: 24.11.2022. Available from: https://CRAN.R-project.org/package=miceadds

91. Schafer JL and Yucel RM. Computational Strategies for Multivariate Linear Mixed-Effects Models With Missing Values. Journal of Computational and Graphical Statistics. 2002; 11:437–57. DOI: 10.1198/106186002760180608

92. Graham JW, Olchowski AE, and Gilreath TD. How Many Imputations Are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. Prevention Science. 2007; 8:206–13. DOI: 10.1007/s11121-007-0070-9

93. Rubin DB. Multiple Imputation for Nonresponse in Surveys. USA, Hoboken, NJ: John Wiley & Sons, Inc., 1987. DOI: 10.1002/9780470316696

94. Rubin DB. Multiple Imputation after 18+ Years. Journal of the American Statistical Association. 1996; 91:473–89. DOI: 10.1080/01621459.1996.10476908

# Part II

# THE STUDIES

# 3

## ARTICLE 1 — STUDY PROTOCOL FOR THE RANDOMIZED CONTROLL TRIAL TEC-A

Title:
Clinical and Cost-Effectiveness of Personalized Telephone Coaching for Farmers, Foresters and Gardeners to Prevent Depression: Study Protocol of an 18-Month Follow-Up Pragmatic Randomized Controlled Trial (TEC-A)

Summary:
Farmers and related professions are at risk for depression due to factors such as poor physical health, high workload, financial pressures and uncertain weather conditions. Personalized telephone coaching has the potential to be a cost-effective and easily accessible intervention to prevent depression in this rural population. This study is part of a nationwide model project ("With us in balance") by the German Social Insurance Company for Agriculture, Forestry and Horticulture (SVLFG). This study protocol describes the method to evaluate the (cost-)effectiveness of a personalized telephone coaching in the reduction of depressive symptom severity and prevention of depression onset.

A two-armed pragmatic randomized controlled trial will be conducted with 312 farmers with elevated levels of depressive symptoms (Patient Health Questionnaire; PHQ-9 $\geq$ 5). Participants will be randomly assigned to receive either a personalized telephone coaching (n = 156) or enhanced treatment as usual (n = 156) which consists of one time e-mailed psychoeducational material and information on preventive services in routine care. The telephone coaching will be conducted by trained psychologists. It can be adapted in length, frequency and topics based on the individual's needs. The coaching can consist of up to 34 sessions of 25 minutes and 50 minutes over a 6-months period. If indicated, it can be enhanced by on-site social services (e.g. socio-economic counselling for farmers) or switched to onsite-coaching if needed.

The assessments will take place at baseline, post-treatment (6 months) and at follow-ups at 12 and 18 months after randomization. The primary outcome is the change in depressive symptom severity

at post-treatment, assessed with the Quick Inventory of Depressive Symptomatology (QIDS-SR16). Secondary outcomes include reliable symptom change and potential onset of depression (QIDS-SR16 $\geq$ 13), as well as anxiety, stress, sleep, quality of life and other mental health outcomes. Group differences will be evaluated using analysis of covariance (ANCOVA) in order to control for baseline values. Evaluation of the cost-effectiveness will be conducted from a societal perspective.

This study protocol outlines the first evaluation concept to estimate the (cost)-effectiveness of a personalized telephone coaching as a preventive intervention for farmers and related profession. If the study shows the effectiveness of the intervention, the results would lay the basis to implement the intervention into routine care and potentially reduce the depression burden in this at-risk population.

Publication:
Thielecke J, Buntrock C, Titzler I, Braun L, Freund J, Berking M, Baumeister H, and Ebert DD. Clinical and Cost-Effectiveness of Personalized Telephone Coaching for Farmers, Foresters and Gardeners to Prevent Depression: Study Protocol of an 18-Month Follow-Up Pragmatic Randomized Controlled Trial (TEC-A). Frontiers in Psychiatry. 2020; 11:125. doi: 10.3389/fpsyt.2020.00125

Contribution:
Janika Thielecke contributed to the study conceptualization, design and methods. She prepared and submitted the research proposal for the local ethics committee. She contributed to the questionnaire programming, data curation plan, and statistical analysis plan. She was responsible for the recruitment and study administration throughout the trial. She drafted the original manuscript and revised it after feedback from her co-authors. She managed the submission process and reviewer comments while receiving feedback from her co-authors.

frontiers in Psychiatry

STUDY PROTOCOL
published: 04 March 2020
doi: 10.3389/fpsyt.2020.00125

# Clinical and Cost-Effectiveness of Personalized Tele-Based Coaching for Farmers, Foresters and Gardeners to Prevent Depression: Study Protocol of an 18-Month Follow-Up Pragmatic Randomized Controlled Trial (TEC-A)

*Janika Thielecke*[1]*, *Claudia Buntrock*[1], *Ingrid Titzler*[1], *Lina Braun*[2], *Johanna Freund*[1], *Matthias Berking*[1], *Harald Baumeister*[2] *and David D. Ebert*[1,3,4]

[1] Department of Clinical Psychology and Psychotherapy, Institute of Psychology, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany, [2] Department of Clinical Psychology and Psychotherapy, Institute of Psychology and Education, Ulm University, Ulm, Germany, [3] Department of Clinical, Neuro- & Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, Netherlands, [4] GET.ON Institute, Hamburg, Germany

OPEN ACCESS

**Edited by:**
Karsten Heekeren,
University of Zurich, Switzerland

**Reviewed by:**
Kazuki Ide,
Kyoto University, Japan
Daniel Jimenez,
Leonard M. Miller School of Medicine,
University of Miami, United States

***Correspondence:**
Janika Thielecke
janika.thielecke@fau.de

**Specialty section:**
This article was submitted to Public Mental Health, a section of the journal Frontiers in Psychiatry

**Received:** 08 August 2019
**Accepted:** 13 February 2020
**Published:** 04 March 2020

**Citation:**
Thielecke J, Buntrock C, Titzler I, Braun L, Freund J, Berking M, Baumeister H and Ebert DD (2020) Clinical and Cost-Effectiveness of Personalized Tele-Based Coaching for Farmers, Foresters and Gardeners to Prevent Depression: Study Protocol of an 18-Month Follow-Up Pragmatic Randomized Controlled Trial (TEC-A). Front. Psychiatry 11:125. doi: 10.3389/fpsyt.2020.00125

**Background:** Farmers show high levels of depressive symptoms and mental health problems in various studies. This study is part of a nationwide prevention project carried out by a German social insurance company for farmers, foresters, and gardeners (SVLFG) to implement internet- and tele-based services among others to improve mental health in this population. The aim of the present study is to evaluate the (cost-)effectiveness of personalized tele-based coaching for reducing depressive symptom severity and preventing the onset of clinical depression, compared to enhanced treatment as usual.

**Methods:** In a two-armed, pragmatic randomized controlled trial ($N = 312$) with follow-ups at post-treatment (6 months), 12 and 18 months, insured farmers, foresters, and gardeners, collaborating family members and pensioners with elevated depressive symptoms (PHQ-9 $\geq 5$) will be randomly allocated to personalized tele-based coaching or enhanced treatment as usual. The coaching is provided by psychologists and consists of up to 34 tele-based sessions for 25–50 min delivered over 6 months. Primary outcome is depressive symptom severity at post-treatment. Secondary outcomes include depression onset, anxiety, stress, and quality of life. A health-economic evaluation will be conducted from a societal perspective.

**Discussion:** This study is the first pragmatic randomized controlled trial evaluating the (cost-)effectiveness of a nationwide tele-based preventive service for farmers. If proven effective, the implementation of personalized tele-based coaching has the potential to reduce disease burden and health care costs both at an individual and societal level.

**Clinical Trial Registration:** German Clinical Trial Registration: DRKS00015655.

**Keywords: prevention, depression, tele-based coaching, RCT, farmers**

## INTRODUCTION

Major depressive disorder (MDD) is a highly prevalent condition with lifetime prevalence rates estimated between 10.6 and 19.8% (1). Moreover, MDD is associated with substantial individual and societal burden due to functional impairment (2), increased mortality (3) as well as high socioeconomic costs (4, 5).

Despite the availability of effective treatments like psychotherapy and psychopharmacological interventions (6, 7), depression remains under-recognized and under-treated in primary care (8). Moreover, even in a scenario of full coverage of and compliance to evidence-based treatments, estimates suggest that only approximately one-third of the total MDD-related disease burden could be averted (9, 10). Therefore, prevention of MDD evokes global interest (11). Recent meta-analytic evidence suggests that psychological interventions could prevent or at least delay the onset of MDD by targeting individuals at elevated risk (e.g., selective prevention) or individuals with subclinical depression (e.g., indicated prevention) (12).

A large population-based study on anxiety and depression showed that male farmers and workers in related occupations had the highest level of depression of all 10 occupational groups of the Classification of Occupations (e.g., armed forces, legislators/managers, professionals including teachers, technicians, craft, and trade workers) (13). The Norwegian HUNT study found a doubled risk for farm-workers compared to their non-farming siblings to report depressive symptoms (14). Risk factors for depression associated with working in agriculture include: work-related stress (15, 16), additional jobs off the farm (15) and financial instability (17) paired with attitudinal barriers for help-seeking especially in male farmers (18, 19). Considering these risk factors, effective, and accessible preventive measures for these individuals are highly warranted and have the potential to greatly improve the well-being of this vulnerable group.

Face-to-face as well as internet-based psychological interventions have been shown to be effective at treating subthreshold depression and in preventing MDD onset (12, 20–22). However, both delivery modes possess certain limitations [e.g., low participation rates (23) or less use of online services in general and more adverse attitudes toward internet use (24)]. Even though access to the internet has greatly increased in Germany over recent years, rural areas still struggle with regards to the use of internet services (24, 25). Therefore, tele-based coaching might be a feasible adjunct to existing preventive interventions for depression, especially in rural communities.

Additionally, some evidence suggests that tele-based coaching is, compared to treatment-as-usual, effective in reducing depressive symptom severity in mildly to moderately depressed individuals with between-group effect sizes ranging from 0.60 at a 4-month follow-up in an intention-to-treat sample (26) to 0.44 at 12-month follow-up based on study completer-only data (27). However, the clinical and cost-effectiveness of tele-based coaching in the prevention of MDD remains understudied.

Thus, the present trial will evaluate whether personalized tele-based coaching is (cost-)effective in reducing depressive symptom severity and preventing the onset of clinical depression. This study is embedded in an evaluation project of a nationwide rollout of preventive services called "With us in balance," carried out by the German social insurance for farmers, foresters, and gardeners (www.svlfg.de). The project aims to implement personalized tele-based coaching and internet-based interventions in conjunction with established, on-site prevention group workshops aimed to improve mental health amongst farmers, foresters, and gardeners.

The following research questions will be investigated in this pragmatic randomized controlled trial:

- Is personalized tele-based coaching effective in reducing depressive symptom severity compared to enhanced treatment as usual (TAU+) at post-treatment (primary outcome) and follow-ups (secondary outcome)?
- Is personalized tele-based coaching effective in preventing the onset of major depressive episodes as compared to TAU+ at follow-up?
- Is personalized tele-based coaching effective in reducing the severity of various mental health outcomes (e.g., stress, anxiety, insomnia) compared to TAU+?
- Is personalized tele-based coaching preferable to TAU+ in terms of costs and utilities in reducing depressive symptom severity?
- Which variables moderate and mediate the effects of personalized tele-based coaching on mental health outcomes?
- What is the level of satisfaction with, adherence to and acceptance of personalized tele-based coaching?
- Are there reported negative effects associated with personalized tele-based coaching?

## METHODS AND ANALYSIS

### Study Design

The study is designed as a two-armed pragmatic randomized controlled trial comparing the clinical and cost-effectiveness of a personalized tele-based coaching provided by IVPNetworks (intervention group, IG) to a control group (CG) receiving enhanced treatment-as-usual (enhanced = e-mailed psychoeducation on stress, depression and information about access to regular care, TAU+). Assessments will take place at baseline (T0), post-treatment (6-month, T1) and at follow-ups at 12-month (T2) and 18-month (T3) after enrolment.

This clinical trial has been approved by the Medical Ethics Committee of the Friedrich-Alexander University Erlangen-Nürnberg (No. 345_18 B) and is registered in the German clinical trial register under DRKS00014000. The results will be reported in accordance with the Consolidated Standards of Reporting

**Abbreviations:** AQoL-8D, Assessment of Quality of Life; AUDIT, Alcohol Use Disorder Identification Test—Consumption Questions; BDI-II, Beck Depression Inventory; CG, control group; CIDI, Composite International Diagnosis Interview; CSQ, Client Satisfaction Questionnaire; GAD, Generalized Anxiety Disorder; ICER, Incremental cost-effectiveness ratio; IG, intervention group; ISI, Insomnia Severity Index; MDD, Major depressive disorder; PHQ, Patient Health Questionnaire; QIDS-SR16, Quick Inventory Depressive Symptomology; SVLFG, Social Insurance for Agriculture, Forestry and Horticulture; TAU, Treatment as usual; TiC-P, Trimbos Institute and Institute of Medical Technology Questionnaire for Costs Associated with Psychiatric Illness; WAI, Working Alliance Inventory.

Trials (CONSORT) 2010 Statement and extension for reporting pragmatic trials (28–30).

## Participants and Procedure

### Inclusion and Exclusion Criteria

***Study participants***

We will include farmers, foresters, and gardeners in Germany who (a) are insured by the SVLFG, (b) are an entrepreneur, collaborating spouse, assisting family member or pensioner, (c) are 18 years or older, (d) show elevated symptoms of depression (PHQ-9 $\geq$ 5), (e) have internet (for online assessments) and telephone access (for coaching), and (f) are willing to provide informed consent. Persons will be excluded if they (a) currently receive psychotherapy or (b) are not willing to sign a non-suicide contract in case of suicidal ideation as indicated by a score of one or greater on the Beck Depression Inventory (BDI-II) item 9 ("I feel I would be better off dead") or a PHQ-9 total score of 20 or greater, or (c) primarily live in the German Federal States Bavaria or Schleswig-Holstein, since national roll-out into routine care started in these two states. As this is a pragmatic trial, we did not exclude comorbidities and current medication usage.

***IVPNetwork coaches***

Written informed consent to participate in online assessments will be obtained from coaches assigned to study participants. There are no specific eligibility criteria for participating coaches. Coaches are either self-employed or employed by IVPNetworks.

### Recruitment

Invitation letters including information about the study procedures are sent to 120,000 randomly selected policyholders via postal mailings by the insurance company. The study is also listed on associated websites and in the member journal of the SVLFG (reaching 1.3 million policyholders). Interested individuals can respond to the study invitation via mail, e-mail, fax, and/or telephone, or can directly access the screening assessment through a URL or QR-Code. The study flow is visualized in **Figure 1**.

## Assessment of Eligibility and Randomization

Study eligibility is assessed with an online screening questionnaire. Persons who fulfill the inclusion criteria will receive detailed study information by e-mail along in addition to information on data security and their right to withdraw from the study at any time. After providing informed consent, participants will be invited to complete the baseline assessment (T0). If any of the inclusion criteria are not fulfilled but the person has a valid insurance status, the person will be redirected to the SVLFG call center for preventive services.

Randomization will take place at an individual level. Participants will be randomly allocated to the intervention or control condition consecutively after completing baseline assessment. A randomization list was created before the start of the study with the web-based program Sealed Envelope (www.sealedenvelope.com) by using permuted block randomization with randomly arranged block sizes (4, 6, 8, 12) and an allocation ratio of 1:1. A person not otherwise involved in the study will independently carry out the randomization procedure. Study participants cannot be blinded in this trial design but data collectors will be blinded to group allocation. The coaches providing the tele-based coaching are aware of treatment allocation; however, the coaches are not otherwise involved in the study. Participants will be informed about group allocation via e-mail.

## Intervention

All participants in the study will have unrestricted access to treatment-as-usual (e.g., general practitioner). The German S3-Guideline/National Disease Management Guideline Unipolar Depression (31) recommends a stepped care model in which more intensive treatments (i.e., cognitive behavioral therapy, antidepressant medication) are only provided if depressive symptoms intensify (i.e., diagnosed major depressive disorder). TAU was not protocolized in this pragmatic trial but will be monitored with the TiC-P (see Assessments) in order to provide an accurate description of the TAU utilized.

### Intervention Condition

The personalized tele-based coaching is offered by IVPNetworks, which is a provider of integrated care in Germany. As evaluator, the study team has no influence on the design and content of the intervention. Even though the tele-based intervention does not focus on depression specifically but more general on mental health problems, it will be evaluated as part of a depression prevention program.

Participants assigned to the intervention group are registered with their contact details on the IVPNetworks management and documentation platform (IVPnet 2.0., www.ivpnet.de). A case manager assigns participants to their personal coach before the coach schedules the first session. The coaching consists of three major phases. (1) A beginning phase in which the first coaching session serves as an assessment of the participants' individual situation and provides time for mutual goal setting for the coaching. The individual goals guide the coaching process and are used to monitor participants' progress. Furthermore, building of a working alliance is emphasized in this first phase. (2) In the working phase, the coaching is intended to support participants in recognizing and understanding conflict patterns so that they can effectively cope with them by activating their own resources. Thus, the coaching focuses on the individual participant's personal situation and stressors (e.g., financial burden, family problems, work-related stress). (3) In the final phase of the coaching, the focus is on transferring learnt skills to everyday life and if needed further supporting offers (e.g., support groups or on-site group offers) are discussed and initiated to maintain coaching effects.

Personalized coaching implies that there are no fixed procedures or standardized manuals for the individual coaching process. Methods used during the coaching vary depending on the coach's therapeutic background. Coaches are qualified psychologists with at least a master's degree in psychology and completed a training in cognitive behavioral therapy, systemic therapy, psychodynamic therapy, hypnotherapy or

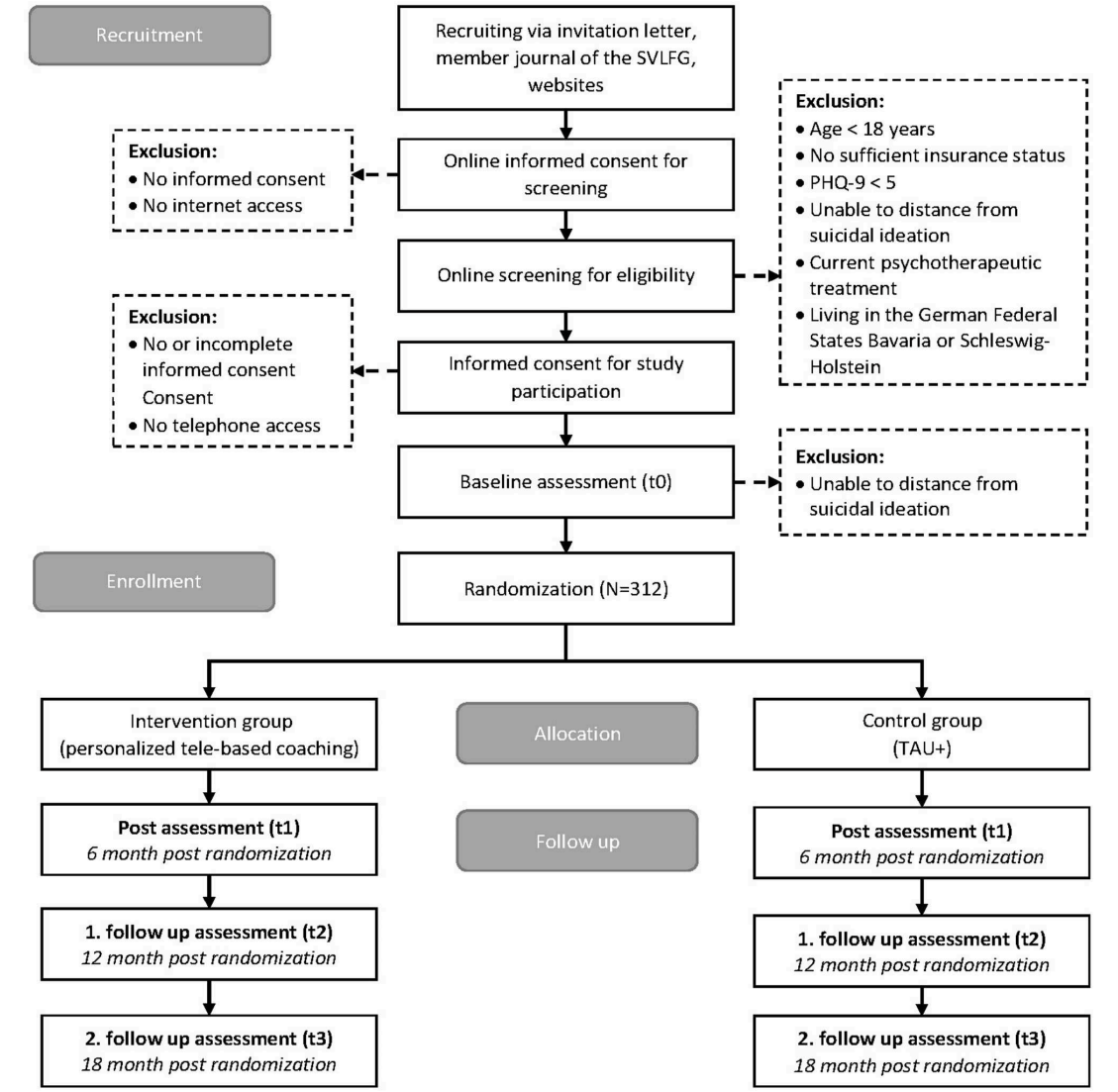



**FIGURE 1 |** Overview of study procedure.

other therapeutic trainings. All coaches, presumably 20–30, are managed by IVPNetworks and are either self-employed or employed by IVPNetworks. Licensed psychotherapists are available for supervision at any time. Furthermore, prior to interacting with any study participant, the coaches receive an introduction into common problems faced by farmers, foresters, and gardeners. The tele-based coaching intervention comprises a maximum of 34 sessions lasting for 25 or 50 min each over a 6 months period (maximum 850 min in total). The coaching duration and frequency are adapted to the participants' needs based on the individual coach's assessment. Theoretically, the coaching can be prolonged up to 9 months (resulting in up to an additional 150 min) after a new needs assessment has been done. If indicated, participants are supported in finding on-site social and health care services to complement the tele-based coaching (e.g., socioeconomic consultants, agricultural family counseling).

Alternatively, on-site coaching can be arranged if a participant no longer prefers tele-based coaching. In the case that a

participant's depressive symptoms worsen at any time during the coaching, the participant will be stabilized and further supporting steps are arranged.

### Control Condition

Participants in the control group (enhanced treatment as usual, TAU+) will receive psychoeducation materials via e-mail. Moreover, they will receive a link to an audiobook on how to deal with stress as well as a link to a platform for preventive services (https://17358.zentrale-pruefstelle-praevention.de/kurse/) in routine care in order to facilitate access to service use.

## Assessments

Assessments at screening, baseline (T0), post-treatment (T1), and follow-up (T2–T3) will be conducted via a secured online platform (www.unipark.de). For an overview of instruments, see **Table 1**. Participants in both study conditions will receive €15 for each completed post-treatment or follow-up assessment as an incentive to continue study participation, resulting in a total of €45 for study completion.

### Screening

The preliminary screening assesses sociodemographic variables (e.g., age, gender, residence, employment relationship, insurance status), current psychotherapy status and depressive symptomology to validate eligibility for participation.

Depressive symptomology is assessed using the German Version of the Patient Health Questionnaire (PHQ-9) (32). The screening inventory consists of nine items on a 4-point-scale with a rating scale ranging from 0 to 3 (0 = "not at all," 1 = "several days," 2 = "more than half the days," 3 = "nearly every day") with each item assessing one symptom criterion domain of MDD. Total score ranges from 0 to 27 with higher scores indicating more severe symptoms. Scores from 0 to 4, 5 to 9, 10 to 14, 15 to 19, and 20 to 27 indicate minimal, mild, moderate, moderately severe, and severe depression severity, respectively (33). An additional item assesses severity of daily life limitations associated with depressive symptoms. The computerized version of the PHQ-9 ($\alpha = 0.88$) and the paper-pencil version ($\alpha = 0.89$) show equally high internal consistency (34).

Additionally, the suicide item of BDI-II (35) will be applied to screen for suicidal ideation if the suicidal item of PHQ-9 is greater than or equal to one. Participant's contact information is recorded in order to monitor suicidal ideation independent of eligibility for study participation. Lastly, to better understand recruitment tactics, participants are asked which recruitment activity made them aware of the study.

### Outcome Measurements

***Primary outcom***

*Depressive symptom severity at post-treatment (T1).* Depressive symptom severity (primary outcome) at post-treatment (T1), is assessed with the German Version of the Quick Inventory Depressive Symptomology (QIDS-SR16) (36). With 16 items (4-point-scale ranging between 0 and 3), the self-report inventory covers all nine DSM-5 symptom criterion domains of MDD. When analyzing scores, only the highest rated item for sleep, weight, and psychomotor activity is included leading to total scores ranging between 0 and 27, with a higher score indicating more depressive symptom severity. Scores between 0 and 5, 6 and 10, 11 and 15, 16 and 20, and >20, indicate normal health status, or mild, moderate, severe, or very severe depressive symptom

**TABLE 1** | Overview of the assessments.

| Instruments | Description | Time of measurement | | | | |
|---|---|---|---|---|---|---|
| | | Screening | T0 | T1 | T2 | T3 |
| **Screening instruments** | | | | | | |
| PHQ-9 | Patient Health Questionnaire | x | | | | |
| **Primary outcome** | | | | | | |
| QIDS-SR16 | Quick Inventory for Depressive Symptomatology | | x | x[a] | x | x |
| **Secondary outcomes** | | | | | | |
| Adapted items from CIDI 3.0, CIDI-SC, and Epi-Q Screening Survey | Prevalence of major depression | | x | x | x | x |
| Adapted items from CIDI 3.0, CIDI-SC, and Epi-Q Screening Survey | Prevalence of bipolar disorder and anxiety disorder | | x | | | |
| PSS-10 | Perceived Stress Scale | | x | x | x | x |
| ISI | Insomnia Severity Index | | x | x | x | x |
| SSS-8 | Somatic Symptom Scale | | x | x | x | x |
| GAD-7 | Generalized Anxiety Disorder | | x | x | x | x |
| PAS | Panic and Agoraphobia Scale | | x | x | x | x |
| AUDIT-C | Alcohol use disorder identification test—Consumption questions | | x | x | x | x |
| MBI-GS | Maslach-Burnout-Inventory | | x | x | x | x |
| AQoL-8D | Quality of life | | x | x | x | x |
| SPE | Subjective Prognosis of Gainful Employment Scale | | x | x | x | x |
| **Cost measurement** | | | | | | |
| TiC-P | Utilization of health services, work-related productivity | | x | x | x | x |
| **Intervention-related outcomes** | | | | | | |
| WAI-SR[b], WAI-SRT[c] | Therapeutic relationship | | | x | x | |
| adapted CSQ-I | Patient satisfaction | | | x | x | |
| INEP[b] | Inventory of negative effects in psychotherapy | | | x | x | |
| Negative effects (open questions)[b] | Negative effects in psychotherapy | | | x | x | |
| **Other assessments** | | | | | | |
| Socio-demographics | Sex, age, farm type, family situation | | x | | | |
| Predictors | Predictors for development of major depression and anxiety | | x | | | |
| BDI-II | Suicidality-Item | x | x | x | x | x |
| Reporting data IVPNetworks | Adherence, drop-outs, Coaching length, amount of sessions, topics (keywords) | | x | x | | |

[a] *Primary outcome is the standardized mean difference between intervention and control group at T1. QIDS-SR16 will also be assessed at T0 and T2-T3.*
[b] *Recorded in intervention group only.*
[c] *Recorded in coaches only.*
*T0 Baseline, T1 6 months, T2 12 months, T3 18 months after randomization.*

severity, respectively (36). Compared to current and lifetime diagnosis based on Structured Clinical Interview for DSM-IV-TR Axis I Disorders (SCID), the QIDS-SR16 proved to be a reliable screening instrument for a diagnosis of clinical depression (37) with good in ($\alpha = 0.86$) (36).

***Secondary outcomes***

*Depressive symptom severity at follow-up (T2 - T3).* Additionally, depressive symptom severity at follow-up assessments at 12-months (T2) and 18-months (T3) will be measured with the QIDS-SR16.

*Onset of ICD-10 major depressive episode and bipolar disorder.* The onset of major depressive episodes and bipolar disorder, are assessed with adapted items from the web version of the Composite International Diagnosis Interview version 3.0 (CIDI 3.0) and Screening Scales (CIDI-SC) (38) and the Epi-Q Screening Survey (39) as used in the WHO World Mental Health Surveys International College Student Project (40). Bipolar disorder will be assessed at baseline (T0), while depressive episodes will be assessed at all time points (T0–T3).

In addition, the QIDS-SR16 is used to identify possible clinically relevant cases of depression. A total score of 13 and greater is defined as cut-off for possible cases of clinical depression for all time points (T0–T3). This value was selected as it yield best results for sensitivity (76.5%) and specificity (81.8%) and resulted in a correct classification of over 80% of participants (37).

*Perceived stress.* Perceived stress is assessed with the perceived stress scale (PSS-10) in which ten items are rated on a 5-point Likert scale (41, 42). The German Version of the perceived stress scale assess perceived stress in the past week and is reported with good internal consistency ($\alpha = 0.84$) (42, 43).

*Insomnia severity.* The Insomnia Severity Index (ISI) (44) is used to measure insomnia severity in participants. The questionnaire has been validated to identify cases of clinical insomnia. The ISI has adequate discriminative validity of the individual items and a high internal consistency ($\alpha = 0.90$–$0.92$). The self-report questionnaire consists of seven items on a 5-point Likert scale (45, 46). Examination of the German version of the ISI in three cross-sectional studies in different target groups revealed acceptable to good internal consistency ($\alpha = 0.76$–$0.81$) (47)

*Somatic symptom burden.* Somatic symptom burden will be assessed with the Somatic Symptom Scale, the short version of the PHQ-15 questionnaire (48). In 8 items (5-point Likert scale), eight symptoms common in primary care and relevant for mental health are assessed (gastrointestinal symptoms, pain, fatigue, cardiopulmonary symptoms). Additionally, a study in a German sample of participants reported good reliability ($\alpha = 0.81$) and correlation with depression and other health outcomes (49).

*Generalized anxiety disorder.* The GAD-7 (50), a 7-item self-report questionnaire with a 4-point Likert scale, is used for screening and severity measuring of Generalized Anxiety Disorder (GAD). Previous studies have found that the internal consistency of this instrument is only slightly lower in a German sample ($\alpha = 0.89$) (51) compared to a sample in the United States ($\alpha = 0.92$) (50). The GAD-7 is considered a reliable screening instrument for GAD (50, 51)

Additionally, the presence of GAD at baseline is assessed with adapted items from the Composite International Diagnosis Interview version 3.0 (CIDI 3.0), Screening Scales (CIDI-SC) (38) and the Epi-Q Screening Survey (39).

*Panic and agoraphobia scale.* The panic and agoraphobia scale (52, 53) will be assessed to measure severity of panic and agoraphobic symptoms.

Twelve items with a 5-point Likert scale assess five individual subscales: regarding panic attacks, agoraphobic avoidance, anticipatory anxiety, daily life limitations, and health concerns. An additional item (6-point Likert scale) differentiates between unexpectedness vs. expectedness of panic attacks and is not included in the calculation of the total score (52). Additionally, situations in which individuals believe panic attack are likely to occur are recorded. Good internal consistency is reported for the panic and agoraphobia scale ($\alpha = 0.88$) (52).

*Alcohol consumption.* Alcohol consumption will be assessed with the Alcohol Use Disorder Identification Test-Consumption Questions (AUDIT-C) (54). As previous research has shown, the first three items of the AUDIT can be used as a reliable stand-alone screening of risky alcohol consumption (AUDIT-C) (55). Good internal consistency is reported for the German version of the AUDIT-C ($\alpha = 0.80$) (56). In the present study, the German version of the AUDIT-C, as distributed by the German Medical Association (https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/AlkAUDITCFragebogen.pdf) is administered.

*Burnout.* The Maslach Burnout Inventory (57) is used to record burnout symptomology. The Inventory assess the three dimensions of burnout including "Emotional Exhaustion," "Cynicism," and "Professional Efficacy" with 22 items and a 7-point Likert scale. The reliability ($\alpha = 0.71$–$0.88$ for subscales) has been proven acceptable in cross-national comparison (58)

*Quality of life.* The self-report questionnaire Assessment of Quality of Life (AQoL-8D) will be used to assess health-related quality of life. With 35 items (5-point Likert scale) three physical dimensions ("independent living," "pain," and "senses") as well as the five psycho-social dimensions ("mental health," "happiness," "coping," "relationships," and "self-worth") are assessed, yielding two distinct sum scores (59). The questionnaire is reported with excellent reliability ($\alpha = 0.96$) and good psychometric properties (59).

*Subjective capacity to work.* The participants subjective capacity to work will be measured using the 3-item Subjective Prognosis of Gainful Employment Scale (60). The short self-report scale showed high internal consistency (Guttman scaling: rep = 0.99) and validity for assessment of subjective endangerment and prognosis of capacity to work (60).

*Cost measures.* For cost evaluation, the German version of the Dutch cost questionnaire "Trimbos Institute and Institute of Medical Technology Questionnaire for Costs Associated with Psychiatric Illness" (TiC-P) (61) was adapted to specifically assess farmers, foresters, and gardeners. The self-report questionnaire assesses direct medical costs (e.g., use of health care services, visits to the general practitioner, use of medications, sessions with psychotherapists or psychiatrists, inpatient hospital care), direct non-medical costs (e.g., patient and family costs), and indirect costs (e.g., productivity losses due to absenteeism and presenteeism). The German version has been widely used in health economic outcome evaluations alongside clinical trials (62–65).

### *Intervention-related outcomes*

As part of the evaluation, the process and content of the tele-based coaching will be documented and reported in detail during the intervention phase and at the end of the coaching (e.g., amount of completed sessions and their duration, common topics in coaching sessions, compliance of participants, referrals to onsite services). The qualification of coaches will be assessed with a sociodemographic questionnaire. In addition, to gain deeper insights the coaching procedure and methods, interviews in a subsample of coaches will be conducted.

*Intervention satisfaction.* At post-treatment (T1) and the first follow-up (T2) participants' satisfaction with the personalized tele-based coaching is measured using a German Version of the Client Satisfaction Questionnaire (CSQ-8, adapted version for internet interventions: CSQ-I) (66–68). In the IG, the CSQ-I will be adapted for telephone coaching (the original CSQ-8 is validated only in face-to-face contexts). The CSQ-8 as well the CSQ-I consist of eight items (4-point Likert scale) and report a high internal consistency ($\alpha = 0.87$ and 0.93) (67, 68). An adapted version of the CSQ-I will be used for the assessment of participant's satisfaction with the information material received in the CG.

*Working alliances.* The therapeutic alliance between participants and coaches will be addressed with the German short version of the Working Alliance Inventory (WAI-SR) (69). The three subscales "agreement on tasks," "agreement on goals," and "development of an affective bond" are assessed with 12 items on a 5-point Likert scale. The internal consistencies for the German version are between $\alpha = 0.81$ and $\alpha = 0.91$ for the subscales and between $\alpha = 0.90$ and $\alpha = 0.93$ for the total score (69, 70). The WAI-SR will be applied in the IG at T1 and T2, respectively. Furthermore, the coaches will be asked to complete the 10-item therapist version (WAI-SRT, developed by Adam O. Horvath, wai.profhorvath.com) at the same time. This approach provides the opportunity to evaluate the therapeutic relationship from two different points of view (i.e., the participant and the coach) and to report a differentiated and comprehensive picture of the experienced working alliance. The WAI-SR and the WAI-SRT were adapted for the current trial to assess coaching in a prevention setting. The items were adjusted to refer to "coaches" instead of therapists and to "telephone coaching" instead of therapy.

*Side effects of the intervention.* Side effects of the intervention will be assessed with a version of the Inventory for the Assessment of Negative Effects of Psychotherapy (71) adapted to the tele-based setting of the intervention. This 22-item questionnaire assesses whether any negative changes were experienced during or after the tele-based coaching that are causally attributed to the intervention. Furthermore, open-ended questions will be included for qualitative assessments of possible negative side effects of the personalized tele-based coaching. Symptom deterioration will be assessed using the Reliable Change Index based on QIDS at post-treatment and follow-ups.

### *Other assessments*

*Covariates.* Variables that could potentially moderate the expected effects, including sociodemographic as well as information about the agricultural business (e.g., farm size, area cultivated, number of workers) and the overall family and work situation (e.g., financial situation, number of relatives living and working together, general work load), will be assessed at baseline (T0).

Furthermore, we will explore variables associated with depression (72–74) or that have shown to predict treatment outcome for depression and anxiety (75). Clinical characteristics that shall be investigated include depressive symptom severity (76), lifetime history of MDD or any mental disorder (77–80), past or present suicidal thoughts or plans (35, 79–81), experience with psychotherapy (82), treatment motivation (83), treatment preference (84–86), family history of mental illness (77, 87–89) (chronic) illness, self-perceived health and energy (61, 78, 90–95), traumatic or adverse childhood experience (abuse, parental death, or divorce) (80, 87, 88, 96–101), body satisfaction and eating disorder (94, 102–107), sleep quality (44, 90, 94, 108, 109), and accidents and injuries (15, 110).

Personality characteristics include smoking (82, 94, 111), alcohol consumption and drug use (77, 91, 94), relationship quality (91, 112, 113) (partner) violence experience (96, 112, 114–116), threatening life events in the past 12 months (79, 80, 91, 117–120), physical activity and sedentary behavior (94, 106, 111, 121–123), stress (15, 43, 79, 124, 125), need for affect, emotional avoidance (126–128), anxiety sensitivity (129–131), sense of mastery, internal locus of control (79, 132–135), worry (136–138), self-worth (100, 104, 139–141), and loneliness (142–145).

Sociodemographic variables that shall be examined include age (76, 77, 111, 146, 147), sex (82, 92, 104, 110, 111, 117, 118, 146–149), origin country (148), migration (150–153), ethnicity, minority (111, 147, 154), discrimination (103, 155, 156), education level (92, 96, 109–111, 148), employment status (88, 111, 125), relationship status, living situation (88, 92, 111, 146), parenthood (82), caregiving living situation (94, 157, 158), economic status (90, 91, 118, 159), and social status (94, 160, 161).

*Safety monitoring.* Suicidal ideation will be screened in all assessments using PHQ-9 (screening) or QIDS-SR16 (T0–T3). If

the suicide item of PHQ-9 or QIDS-SR16 has a score of one or greater, the BDI-II (35) suicide item will additionally be assessed.

A score of one or greater on the BDI-II suicide item or a PHQ-9 score of 20 greater / QIDS-SR16 score of 16 or greater [indicating severe depressive symptoms (36)] will result in a standardized suicide protocol. Participants will receive e-mails with detailed information on 24/7 health services and are advised to seek professional help if symptoms persist or increase. The wording of this information is adapted to the severity of the indicated suicidality as given in the BDI-II Item or PHQ-9/QIDS-16 scores. In cases of a BDI-II scores of 2 or 3, and/or severe depressive symptoms, the persons are called by a psychologist or a psychotherapist (licensed or in training) within 3 days for evaluation of suicidal ideation severity and dissociation. The suicide protocol is developed and closely monitored by a licensed psychotherapist.

As part of the suicide protocol, at screening and baseline, all participants with suicidal ideation and/or severe depressive symptoms are asked to return a signed non-suicide contract if they are able to distance themselves from suicidal ideation (inclusion criteria). Participants are reminded that the coaching is not designed for persisting suicidal ideation and are provided necessary resources (as discussed previously).

## Statistical Analyses

### Sample Size/Power Calculation

The estimated effect size of d = 0.35 used for this study is based on the findings of the meta-analysis by Cuijpers et al. (20) regarding the effect of psychological interventions on subclinical depression ($g = 0.35$, 95%-CI: 0.23–0.47) and effect sizes commonly observed in studies on health outcomes (162).

Power calculation conducted with G*Power (Version 3.1.9.2) for a two-sided $T$-Test ($\alpha = 0.05$, $\beta = 0.80$, d = 0.35) resulted in $N = 260$. In order to account for dropouts during the 18-month follow-up, we calculated an extra 20% of participants, taking into account reported dropout rates between 20 and 23% from studies with telephone coaching in different settings (27, 163). Accounting for drop-outs resulted in the final $N = 312$.

### Clinical Analyses

Clinical analysis will be performed using R statistic software (164). All statistical analyses will be performed based on intention-to-treat principle. Missing data will be imputed using a Markov Chain Monte Carlo multivariate imputation algorithm. Group difference between pre- (T0) and post-assessment (T1) in the primary outcome will be analyzed using generalized linear modeling adjusting for baseline depressive symptom severity. Within- and between-group Cohen's d effect sizes [and their 95% CIs according to Hedges and Olkin (165)] controlling for baseline data (i.e., calculating change scores divided by the pooled standard deviation of change scores) will be reported. Improvements in the primary outcome at individual level will be examined by assessing the number of participants who displayed a treatment response defined by the Reliable Change Index as proposed by Jacobson and Truax (166) and a close-to-symptom-free status (e.g., QIDS $\leq$ 5). Close-to-symptom-free status will only be evaluated in the subgroup of participants who reported at least mild depressive symptom severity at baseline. Group differences in depression onset will be assessed using Poisson regression models in the subset of participants who did not meet the diagnostic depression status at baseline as assessed with the self-report web version of the adapted CIDI. Group differences in diagnostic status in the subset of participants who did not meet depression diagnosis criteria at baseline will be analyzed using logistic regression models. In sensitivity analyses, we will assess the influence of objective measures, such as therapeutic background of coaches and length of coaching, on intervention outcomes. Other secondary outcomes like anxiety, perceived stress, sleep quality, and quality of life will be analyzed as the primary outcome. Additionally, moderation and mediation analyses will be performed.

### Economic Evaluation

We will conduct and report the health economic evaluation in agreement with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement (167) and the guidelines from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) (168). The health-economic evaluation will involve a combination of a cost-effectiveness analysis and a cost-utility analysis (169). The economic evaluation will be performed from a societal perspective and from the perspective of the social insurance company with a time horizon of 18 months. In the cost-effectiveness analysis, the incremental cost-effectiveness ratio (ICER) will be calculated by dividing incremental costs (e.g., cumulative per-participant costs) by the unit of effect gained (e.g., reliably improved case or quality-adjusted life years based on the AQoL-8D). The corresponding equation is ICER = (CostsIG – CostsCG) / (EffectsIG – EffectsCG), with IG representing the intervention group and CG the control group (169). To handle sampling uncertainty, we will bootstrap the seemingly unrelated regression equations model (SURE model) to generate 2,500 simulations of cost and effect pairs while allowing for correlated residuals of the cost and effect equations and adjusting for potential confounding variables (170). Based on the bootstrapped SURE model, bias-corrected and accelerated 95% confidence intervals will be obtained for incremental costs and effects while 95% confidence intervals around ICERs will be obtained by the bootstrap acceptability method as proposed by Glick et al. (171). The bootstrapped cost and effect pairs will be plotted in a cost-effectiveness plane. The cost-effectiveness plane is used to visually represent the differences in costs and health outcomes between the tele-based coaching and TAU+ in two dimensions. Health outcomes (effects) are plotted on the x-axis and costs on the y-axis. TAU+ is plotted at the origin, and so the x and y values represent incremental health outcomes and incremental costs vs. routine care. The cost-effectiveness plane is divided into four quadrants: the north-east (NE) quadrant, in which new interventions generate more health gains but at higher costs; the south-east (SE) quadrant, in which new interventions costs less and create higher health effects (e.g., the new intervention dominates the alternative); the north-west (NW) quadrant, in which new interventions produce less health effects at higher costs (e.g., the new intervention is dominated

by the alternative) and the south-west (SW) quadrant, in which new interventions costs less but also generate less health effects (169). To disclose the probability that the intervention is cost-effective for a range of willingness-to-pay ceilings (e.g., the maximum a decision maker is willing to pay for a unit of effect), the bootstrapped cost and effect pairs will also be shown in a cost-effective acceptability curve (172). To test the robustness of the base-case findings, multi-way sensitivity analyses will be completed. The analysis will be performed using StataCorp (173).

## DISCUSSION

This study will be the first pragmatic randomized controlled trial to examine the clinical and cost-effectiveness of personalized tele-based coaching for farmers, foresters, and gardeners in Germany. The intervention is meant to prevent clinical depression and reduce depressive symptom severity in farmers, foresters, and gardeners.

The study is embedded in a large scale nationwide prevention project which does not only aim at the evaluation of the (cost-)effectiveness but also target nation-wide implementation at the same time. Stepwise national rollout of the intervention runs parallel to this randomized controlled trial. Confounding is controlled for via recruitment activities and eligibility criteria. Therefore, if proven effective, the personalized tele-based coaching will be sustainably introduced into regular care in Germany, which can be seen as a strength of this trial since most studies in the field of depression prevention lack a sustainable integration in the health care system (174, 175). Study findings can therefore directly contribute to improving the preventive service of the SVLFG.

Another major strength of tele-based coaching compared to face-to-face psychological interventions or mental health workshops is that delivery over the telephone greatly reduces the time and effort associated with people affected in the rural area (e.g., no travel time to visit a therapist).

Some limitations of the study should also be mentioned. First, the assumed advantage of tele-based coaching compared to internet-based interventions that telephone coaching does not rely on internet access and might therefore be more accessible in rural areas, cannot be verified with this study. Participants are only enrolled in the study if they are able to complete the online-based assessments and therefore require internet access for study participation. Nevertheless, participants in IG and CG will be asked about their preferences for treatment of mental health problems (e.g., face-to face, online, telephone, combined delivery modes) in order to get first insights in setting preferences.

Second, the operationalization of the personalized tele-based coaching intervention itself is difficult. As Ammentorp et al. (176) claims, detailed descriptions of coaching interventions and methods used are crucial in order to improve the field of life coaching research and its impact on (mental) health outcomes. Although this intervention is highly personalized with regard to contact duration, frequency, and covered coaching topics, we will be able to give a detailed and comprehensive description of the intervention by routine data documentation in addition to data collected from coaches and participants. Thereby, we will add to this growing field of research by providing strongly needed evidence from a randomized controlled trial regarding preventive coaching interventions. The information gained could also be used in the development of standardized coaching manuals.

Third, as mentioned earlier, previous research does not provide profound evidence about estimates regarding effect sizes and drop-out rates. We accounted for possible drop out by increasing the target sample by 20%. Nevertheless, we are at a risk to over- or underestimate dropout. Reporting effect sizes and drop-out rates will help researchers calculate target sample sizes in future studies.

To sum up, this pragmatic trial leads to a robust estimation of the (cost-)effectiveness of personalized tele-based coaching in the target population. Therefore, results from the study can be generalized to farmers, foresters, and gardeners in Germany and comparable coaching situations.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Friedrich-Alexander-University Erlangen-Nürnberg (no. 345_18 B). The patients/participants provided their written informed consent to participate in this study.

## TRIAL STATUS

The recruitment of participants started in December 2018 and has been finished in April 2019 while the first draft of this paper was composed. Data collection is still ongoing and no analyses have been performed, yet.

## AUTHOR CONTRIBUTIONS

DE, HB, and MB obtained funding for this study. DE, IT, CB, JT, and HB contributed to the study design. CB contributed to the design of the economic evaluation study. JT drafted the manuscript, supervised by CB, and is responsible for recruitment and coordination of the trial. JF and LB contributed to the acquisition of data and managing of recruitment. IT and CB are supervising the recruitment and trial management. All authors provided critical revision of the article and approved the final manuscript.

## FUNDING

The German insurance company SVLFG provided a financial contribution to the Friedrich-Alexander University Erlangen-Nuremberg and Ulm University as expense allowance. SVLFG had no role in study design, decision to publish, or preparation of this manuscript. SVLFG will not be involved in data collection, analyses, decision to publish, or preparation of future papers regarding the study.

## ACKNOWLEDGMENTS

Authors would like to thank Friederike Dietz for the creation of the Unipark surveys and the support in data security questions. Further, authors thank Sarah Banellis, Merle Bloom, Hanna Böckeler, Albina Chafisuf, and Tomris Ohloff for their engagement in enrolling and supporting the participants throughout the study and Lukas Fuhrmann and Marvin Franke for carrying out the randomization. Authors thank the IVPNetworks staff who provided details on their intervention and Allison Grace for proofreading the manuscript as well.

## REFERENCES

1. Kessler RC, Chiu WT, Demler O, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. *Arch Gen Psychiatry*. (2005) 62:617. doi: 10.1001/archpsyc.62.6.617
2. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJL, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med*. (2013) 10:e1001547. doi: 10.1371/journal.pmed.1001547
3. Gilman SE, Sucha E, Kingsbury M, Horton NJ, Murphy JM, Colman I. Depression and mortality in a longitudinal study: 1952-2011. *CMAJ*. (2017) 189:1304–10. doi: 10.1503/cmaj.170125
4. Greenberg PE, Fournier A-A, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry*. (2015) 76:155–62. doi: 10.4088/JCP.14m09298
5. Vos T, Haby MM, Barendregt JJ, Kruijshaar M, Corry J, Andrews G. The burden of major depression avoidable by longer-term treatment strategies. *Arch Gen Psychiatry*. (2004) 61:1097–103. doi: 10.1001/archpsyc.61.11.1097
6. DGPPN, BÄK, KBV, AWMF (Hrsg.) für die Leitliniengruppe Unipolare Depression. S3-AWMF-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression – Langfassung, 2. Auflage. Version 5. (2015). doi: 10.6101/AZQ/000364
7. National Institute for Health and Care Excellence. *Depression in Adults: Recognition and Management. Clin Guidel 90*. (2009) Available online at: https://www.nice.org.uk/guidance/cg90 (accessed July 17, 2019).
8. Lotfi L, Flyckt L, Krakau I, Mårtensson B, Nilsson GH. Undetected depression in primary healthcare: occurrence, severity and co-morbidity in a two-stage procedure of opportunistic screening. *Nord J Psychiatry*. (2010) 64:421–7. doi: 10.3109/08039481003786378
9. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders Record Status Study population. *Br J Psychiatry*. (2004) 184:526–33. doi: 10.1192/bjp.184.6.526
10. Chisholm D, Sanderson K, Ayuso-Mateos JL, Saxena S. Reducing the global burden of depression: population-level analysis of intervention cost-effectiveness in 14 world regions. *Br J Psychiatry*. (2004) 184:393–403. doi: 10.1192/bjp.184.5.393
11. Muñoz RF, Cuijpers P, Smit F, Barrera AZ, Leykin Y. Prevention of major depression. *Annu Rev Clin Psychol*. (2010) 6:181–212. doi: 10.1146/annurev-clinpsy-033109-132040
12. van Zoonen K, Buntrock C, Ebert DD, Smit F, Reynolds CF, Beekman AT, et al. Preventing the onset of major depressive disorder: a meta-analytic review of psychological interventions. *Int J Epidemiol*. (2014) 43:318–29. doi: 10.1093/ije/dyt175
13. Sanne B, Mykletun A, Dahl AA, Moen BE, Tell GS. Occupational differences in levels of anxiety and depression: The Hordaland Health Study. *J Occup Environ Med*. (2003) 45:628–38. doi: 10.1097/01.jom.0000069239.06498.2f
14. Torske MO, Bjørngaard JH, Hilt B, Glasscock D, Krokstad S. Farmers' mental health: a longitudinal sibling comparison - The HUNT study, Norway. *Scand J Work Environ Heal*. (2016) 42:547–56. doi: 10.5271/sjweh.3595
15. Onwuameze OE, Paradiso S, Peek-Asa C, Donham KJ, Rautiainen RH. Modifiable risk factors for depressed mood among farmers. *Ann Clin Psychiatry*. (2013) 25:83–90.
16. Logstein B. Predictors of mental complaints among norwegian male farmers. *Occup Med*. (2016) 66:332–7. doi: 10.1093/occmed/kqw019
17. Simkin S, Hawton K, Fagg J, Malmberg A. Stress in farmers: a survey of farmers in england and wales. *Occup Environ Med*. (1998) 55:729–34. doi: 10.1136/oem.55.11.729
18. Judd F, Jackson H, Fraser C, Murray G, Robins G, Komiti A. Understanding suicide in Australian farmers. *Soc Psychiatry Psychiatr Epidemiol*. (2006) 41:1–10. doi: 10.1007/s00127-005-0007-1
19. Hull MJ, Fennell KM, Vallury K, Jones M, Dollman J. A comparison of barriers to mental health support-seeking among farming and non-farming adults in rural South Australia. *Aust J Rural Health*. (2017) 25:347–53. doi: 10.1111/ajr.12352
20. Cuijpers P, Koole SL, van Dijke A, Roca M, Li J, Reynolds CF, III. Psychotherapy for subclinical depression: meta-analysis. *Br J Psychiatry*. (2014) 205:268–74. doi: 10.1192/bjp.bp.113.138784
21. Ebert DD, Cuijpers P, Muñoz RF, Baumeister H. Prevention of mental health disorders using internet- and mobile-based interventions: a narrative review and recommendations for future research. *Front Psychiatry*. (2017) 8:116. doi: 10.3389/fpsyt.2017.00116
22. Sander L, Rausch L, Baumeister H. Effectiveness of internet-based interventions for the prevention of mental disorders: a systematic review and meta-analysis. *JMIR Ment Heal*. (2016) 3:e38. doi: 10.2196/mental.6061
23. Cuijpers P, Van Straten A, Warmerdam L, Van Rooy MJ. Recruiting participants for interventions to prevent the onset of depressive disorders: possibile ways to increase participation rates. (2010) 10:181. doi: 10.1186/1472-6963-10-181
24. Initiative D21. *D21-Digital-Index 2018/2019*. (2019). Available online at: https://initiatived21.de/publikationen/d21-digital-index-2018-2019/ (accessed May 7, 2019).
25. Bundesministerium für Verkehr und digitale Infrastruktur. *Aktuelle Breitbandverfügbarkeit in Deutschland (Stand Mitte 2018)*. (2018). Available online at: https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/breitband-verfuegbarkeit-mitte-2018.html?nn=12830 (accessed May 7, 2019).
26. Lerner D, Adler DA, Rogers WH, Chang H, Greenhill A, Cymerman E, et al. A randomized clinical trial of a telephone depression intervention to reduce employee presenteeism and absenteeism. *Psychiatr Serv*. (2015) 66:570–7. doi: 10.1176/appi.ps.201400350
27. Schröder S, Fleckenstein J, Wunderlich M. Effektivität eines telefonbasierten coaching-programms für patienten mit einer depressiven erkrankung. *Monit Versorgungsforsch*. (2017) 10:65–8. doi: 10.24945/MVF.04.17.1866-0533.2022
28. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. (2010) 340:c332. doi: 10.1136/bmj.c332
29. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*. (2008) 337:a2390. doi: 10.1136/bmj.a2390
30. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. (2010) 340:c869. doi: 10.1136/bmj.c869
31. Bundesärztekammer B, Fachgesellschaften. S3-*Leitlinie/Nationale Versorgungsleitlinie Unipolare Depression*. (2015). Available online at: www.depression.versorgungsleitlinien.de (accessed May 31, 2019).
32. Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J Affect Disord*. (2004) 81:61–6. doi: 10.1016/S0165-0327(03)00198-8
33. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. (2001) 16:606–13. doi: 10.1046/j.1525-1497.2001.016009606.x

34. Erbe D, Eichert HC, Rietz C, Ebert DD. Interformat reliability of the patient health questionnaire: validation of the computerized version of the PHQ-9. *Internet Interv*. (2016) 5:1–4. doi: 10.1016/j.invent.2016.06.006
35. Kühner C, Bürger C, Keller F, Hautzinger M. Reliabilität und Validität des revidierten Beck-Depressionsinventars (BDI-II). Befunde aus deutschsprachigen Stichproben. *Nervenarzt*. (2007) 78:651–6. doi: 10.1007/s00115-006-2098-7
36. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry*. (2003) 54:573–83. doi: 10.1016/S0006-3223(02)01866-8
37. Lamoureux BE, Linardatos E, Fresco DM, Bartko D, Logue E, Milo L. Using the QIDS-SR16 to identify major depressive disorder in primary care medical patients. *Behav Ther*. (2010) 41:423–31. doi: 10.1016/j.beth.2009.12.002
38. Kessler RC, Üstün TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res*. (2004) 13:93–121. doi: 10.1002/mpr.168
39. Kessler R, Farley P, Gruber M, Harshaw Q, Jewell M, Sampson N, et al. Concordance of computerized self-report measures of DSM-IV-TR mood and anxiety disorders compared to gold standard clinical assessments in primay care. *Value Heal*. (2010) 13:118. doi: 10.1016/S1098-3015(10)72567-5
40. Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Ebert DD, et al. WHO world mental health surveys international college student project: prevalence and distribution of mental disorders WHO WMH-ICS collaborators. *Am Psychol Assoc*. (2018) 127:623–38. doi: 10.1037/abn0000362
41. Naik AD, White CD, Robertson SM, Armento MEA, Lawrence B, Stelljes LA, et al. Behavioral health coaching for rural-living older adults with diabetes and depression: an open pilot of the HOPE Study. *BMC Geriatr*. (2012) 12:37. doi: 10.1186/1471-2318-12-37
42. Reis D, Lehr D, Heber E, Ebert DD. The German Version of the Perceived Stress Scale (PSS-10): evaluation of dimensionality, validity, and measurement invariance with exploratory and confirmatory bifactor modeling. *Assessment*. (2017) 26:1246–59. doi: 10.1177/1073191117715731
43. Klein EM, Brähler E, Dreier M, Reinecke L, Müller KW, Schmutzer G, et al. The German version of the Perceived Stress Scale – psychometric characteristics in a representative German community sample. *BMC Psychiatry*. (2016) 16:159. doi: 10.1186/s12888-016-0875-9
44. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med*. (2001) 2:297–307. doi: 10.1016/S1389-9457(00)00065-4
45. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep*. (2011) 34:601–8. doi: 10.1093/sleep/34.5.601
46. Gagnon C, Belanger L, Ivers H, Morin CM. Validation of the Insomnia Severity Index in primary care. *J Am Board Fam Med*. (2013) 26:701–10. doi: 10.3122/jabfm.2013.06.130064
47. Gerber M, Lang C, Lemola S, Colledge F, Kalak N, Holsboer-Trachsler E, et al. Validation of the German version of the insomnia severity index in adolescents, young adults and adult workers: results from three cross-sectional studies. *BMC Psychiatry*. (2016) 16:174. doi: 10.1186/s12888-016-0876-8
48. Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med*. (2002) 64:258–66. doi: 10.1097/00006842-200203000-00008
49. Gierk B, Kohlmann S, Kroenke K, Spangenberg L, Zenger M, Brähler E, et al. The Somatic Symptom Scale−8 (SSS-8). *JAMA Intern Med*. (2014) 174:399–407. doi: 10.1001/jamainternmed.2013.12179
50. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. (2006) 166:1092–7. doi: 10.1001/archinte.166.10.1092
51. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. *Med Care*. (2008) 46:266–74. doi: 10.1097/MLR.0b013e318160d093
52. Bandelow B. Assessing the efficacy of treatments for panic disorder and agoraphobia. II. The Panic and Agoraphobia Scale. *Int Clin Psychopharmacol*. (1995) 10:73–81. doi: 10.1097/00004850-199506000-00003
53. Bandelow B, Hajak G, Holzrichter S, Kunert HJ, Rüther E. Assessing the efficacy of treatments for panic disorder and agoraphobia. I. Methodological problems. *Int Clin Psychopharmacol*. (1995) 10:83–93. doi: 10.1097/00004850-199506000-00004
54. Dybek I, Bischof G, Grothues J, Reinhardt S, Meyer C, Hapke U, et al. The reliability and validity of the Alcohol Use Disorders Identification Test (AUDIT) in a German general practice population sample. *J Stud Alcohol*. (2006) 67:473–81. doi: 10.15288/jsa.2006.67.473
55. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med*. (1998) 158:1789–95. doi: 10.1001/archinte.158.16.1789
56. Rumpf HJ, Wohlert T, Freyer-Adam J, Grothues J, Bischof G. Screening questionnaires for problem drinking in adolescents: performance of AUDIT, AUDIT-C, CRAFFT and POSIT. *Eur Addict Res*. (2013) 19:121–7. doi: 10.1159/000342331
57. Maslach C, Jackson SE, Leiter MP. Maslach burnout inventory. In: Zalaquett CP, Wood RJ, editors. *Evaluating Stress: A Book of Resources.* Lanham, MD: The Scarecrow Press (1997). p. 191–218.
58. Aguayo R, Vargas C, de la Fuente EI, Lozano LM. A meta-analytic reliability generalization study of the Maslach Burnout Inventory. *Int J Clin Heal Psychol*. (2011) 11:343–61.
59. Richardson J, Iezzi A, Khan MA, Maxwell A. Validity and reliability of the Assessment of Quality of Life (AQoL)-8D multi-attribute utility instrument. *Patient Patient Centered Outcomes Res*. (2014) 7:85–96. doi: 10.1007/s40271-013-0036-x
60. Mittag O, Raspe H. Eine kurze Skala zur Messung der subjektiven Prognose der Erwerbstätigkeit: Ergebnisse einer Untersuchung an 4279 Mitgliedern der gesetzlichen Arbeiterrentenversicherung zu Reliabilität (Guttman-Skalierung) und Validität der Skala. *Rehabilitation*. (2003) 42:169–74. doi: 10.1055/s-2003-40095
61. Bouwmans C, De Jong K, Timman R, Zijlstra-Vlasveld M, Van der Feltz-Cornelis C, Tan SS, et al. Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). *BMC Health Serv Res*. (2013) 13:217. doi: 10.1186/1472-6963-13-217
62. Ebert DD, Kählke F, Buntrock C, Berking M, Smit F, Heber E, et al. A health economic outcome evaluation of an internet-based mobile-supported stress management intervention for employees. *Scand J Work Environ Heal*. (2018) 44:171–82. doi: 10.5271/sjweh.3691
63. Buntrock C, Berking M, Smit F, Lehr D, Nobis S, Riper H, et al. Preventing depression in adults with subthreshold depression: health-economic evaluation alongside a pragmatic randomized controlled trial of a web-based intervention. *J Med Internet Res*. (2017) 19:1–16. doi: 10.2196/jmir.6587
64. Nobis S, Ebert DD, Lehr D, Smit F, Buntrock C, Berking M, et al. Web-based intervention for depressive symptoms in adults with types 1 and 2 diabetes mellitus: a health economic evaluation. *Br J Psychiatry*. (2018) 212:199–206. doi: 10.1192/bjp.2018.10
65. Ebert DD, Lehr D, Boß L, Riper H, Cuijpers P, Andersson G, et al. Efficacy of an internet-based problem-solving training for teachers: results of a randomized controlled trial. *Scand J Work Environ Heal*. (2014) 40:582–96. doi: 10.5271/sjweh.3449
66. Attkisson CC, Zwick R. The client satisfaction questionnaire. *Eval Program Plann*. (1982) 5:233–7. doi: 10.1016/0149-7189(82)90074-X
67. Schmidt J, Wittmann WW. Fragebogen zur messung der patientenzufriedenheit. In: Brähler E, Schumacher J, Strauß B. *Diagnostische Verfahren in der Psychotherapie.* Göttingen: Hogrefe Publishing GmbH (2002). p. 392–6.
68. Boß L, Lehr D, Schaub MP, Castro RP, Riper H, Berking M, et al. Efficacy of a web-based intervention with and without guidance for employees with risky drinking: results of a three-arm randomized controlled trial. *Addiction*. (2018) 113:635–46. doi: 10.1111/add.14085
69. Wilmers F, Munder T, Leonhart R, Herzog T, Plassmann R, Barth J, et al. Die deutschsprachige Version des Working Alliance Inventory-short

revised (WAI-SR)-Ein schulenübergreifendes, ökonomisches und empirisch validiertes Instrument zur Erfassung der therapeutischen Allianz. *Klin Diagnostik u Eval*. (2008) 1:343–58.
70. Munder T, Wilmers F, Leonhart R, Linster HW, Barth J. Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clin Psychol Psychother*. (2010) 17:231–9. doi: 10.1002/cpp.658
71. Ladwig I, Rief W, Nestoriuc Y. Welche risiken und nebenwirkungen hat psychotherapie? - Entwicklung des Inventars zur Erfassung Negativer Effekte von Psychotherapie (INEP). *Verhaltenstherapie*. (2014) 24:252–63. doi: 10.1159/000367928
72. Nigatu YT, Liu Y, Wang JL. External validation of the international risk prediction algorithm for major depressive episode in the US general population: The PredictD-US study. *BMC Psychiatry*. (2016) 16:1–8. doi: 10.1186/s12888-016-0971-x
73. King M, Bottomley C, Bellón-Saameño JA, Torres-Gonzalez F, Švab I, Rifel J, et al. An international risk prediction algorithm for the onset of generalized anxiety and panic syndromes in general practice attendees: predict A. *Psychol Med*. (2011) 41:1625–39. doi: 10.1017/S0033291710002400
74. Liu Y, Sareen J, Bolton J, Wang J. Development and validation of a risk-prediction algorithm for the recurrence of panic disorder. *Depress Anxiety*. (2015) 32:341–8. doi: 10.1002/da.22359
75. Kessler RC, Van Loo HM, Wardenaar KJ, Bossarte RM, Brenner LA, Cai T, et al. Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. *Mol Psychiatry*. (2016) 2:1366–71. doi: 10.1038/mp.2015.198
76. Støen Grotmol K, Gude T, Moum T, Vaglum P, Tyssen R. Risk factors at medical school for later severe depression: a 15-year longitudinal, nationwide study (NORDOC). *J Affect Disord*. (2013) 146:106–11. doi: 10.1016/j.jad.2012.08.047
77. Hölzel L, Härter M, Reese C, Kriston L. Risk factors for chronic depression - a systematic review. *J Affect Disord*. (2011) 129:1–13. doi: 10.1016/j.jad.2010.03.025
78. Bromberger JT, Schott L, Kravitz HM, Joffe H. Risk factors for major depression during midlife among a community sample of women with and without prior major depression: are they the same or different? *Psychol Med*. (2015) 45:1653–64. doi: 10.1017/S0033291714002773
79. Lewinsohn PM, Gotlib IH, Seeley JR. Adolescent psychopathology: IV. Specificity of psychosocial risk factors for depression and substance abuse in older adolescents. *J Am Acad Child Adolesc Psychiatry*. (1995) 34:1221–9. doi: 10.1097/00004583-199509000-00021
80. Ebert DD, Buntrock C, Mortier P, Auerbach R, Weisel KK, Kessler RC, et al. Prediction of major depressive disorder onset in college students. *Depress Anxiety*. (2019) 36:294–304. doi: 10.1002/da.22867
81. Barkow K, Maier W, Üstün TB, Gänsicke M, Wittchen H-U, Heun R. Risk factors for new depressive episodes in primary health care: an international prospective 12-month follow-up study. *Psychol Med*. (2002) 32:595–607. doi: 10.1017/S0033291702005263
82. Yanzón de la Torre A, Oliva N, Echevarrieta PL, Pérez BG, Caporusso GB, Titaro AJ, et al. Major depression in hospitalized Argentine general medical patients: prevalence and risk factors. *J Affect Disord*. (2016) 197:36–42. doi: 10.1016/j.jad.2016.02.066
83. Zuroff D, Koestner R, Moskowitz DS, McBride C, Marshall M, Bagby M. Autonomous motivation for therapy: a new common factor in brief treatments for depression. *Psychother Res*. (2007) 17:137–48. doi: 10.1080/10503300600919380
84. Johansson R, Nyblom A, Carlbring P, Cuijpers P, Andersson G. Choosing between Internet-based psychodynamic versus cognitive behavioral therapy for depression: a pilot preference study. *BMC Psychiatry*. (2013) 13:268. doi: 10.1186/1471-244X-13-268
85. Swift JK, Callahan JL, Vollmer BM. Preferences. *J Clin Psychol*. (2011) 67:155–65. doi: 10.1002/jclp.20759
86. Renn BN, Hoeft TJ, Lee HS, Bauer AM, Areán PA. Preference for in-person psychotherapy versus digital psychotherapy options for depression: survey of adults in the U.S. *NPJ Digit Med*. (2019) 2:1–7. doi: 10.1038/s41746-019-0077-1
87. Vousoura E, Verdeli H, Warner V, Wickramaratne P, Baily CDR. Parental divorce, familial risk for depression, and psychopathology in offspring: a three-generation study. *J Child Fam Stud*. (2012) 21:718–25. doi: 10.1007/s10826-011-9523-7
88. Heslin M, Desai R, Lappin JM, Donoghue K, Lomas B, Reininghaus U, et al. Biological and psychosocial risk factors for psychotic major depression. *Soc Psychiatry Psychiatr Epidemiol*. (2016) 51:233–45. doi: 10.1007/s00127-015-1131-1
89. Reinherz HZ, Giaconia RM, Hauf AMC, Wasserman MS, Paradis AD. General and specific childhood risk factors for depression and drug disorders by early adulthood. *J Am Acad Child Adolesc Psychiatry*. (2000) 39:223–31. doi: 10.1097/00004583-200002000-00023
90. Chan MF, Zeng W. Exploring risk factors for depression among older men residing in Macau. *J Clin Nurs*. (2011) 20:2645–54. doi: 10.1111/j.1365-2702.2010.03689.x
91. Salokangas RKR, Poutanen O. Risk factors for depression in primary care findings of the TADEP project. *J Affect Disord*. (1998) 48:171–80. doi: 10.1016/S0165-0327(97)00171-7
92. Anstey KJ, Von Sanden C, Sargent-Cox K, Luszcz MA. Prevalence and risk factors for depression in a longitudinal, population-based study including individuals in the community and residential care. *Am J Geriatr Psychiatry*. (2007) 15:497–505. doi: 10.1097/JGP.0b013e31802e21d8
93. Joo Y, Roh S. Risk factors associated with depression and suicidal ideation in a rural population. *Environ Health Toxicol*. (2016) 31:e2016018. doi: 10.5620/eht.e2016018
94. Chang SC, Pan A, Kawachi I, Okereke OI. Risk factors for late-life depression: a prospective cohort study among older women. *Prev Med*. (2016) 91:144–51. doi: 10.1016/j.ypmed.2016.08.014
95. Wang JL, Sareen J, Patten S, Bolton J, Schmitz N, Birney A. A prediction algorithm for first onset of major depression in the general population: development and validation. *J Epidemiol Community Health*. (2014) 68:418–24. doi: 10.1136/jech-2013-202845
96. Blanco C, Rubio J, Wall M, Wang S. Risk factors for anxiety disorders: common and specific effects in a national sample. *Depress Anxiety*. (2014) 31:756–64. doi: 10.1002/da.22247
97. Wilson S, Vaidyanathan U, Miller MB, McGue M, Iacono WG. Premorbid risk factors for major depressive disorder: are they associated with early onset and recurrent course? *Dev Psychopathol*. (2014) 26:1477–93. doi: 10.1017/S0954579414001151
98. Berg L, Rostila M, Hjern A. Parental death during childhood and depression in young adults – a national cohort study. *J Child Psychol Psychiatry Allied Discip*. (2016) 57:1092–8. doi: 10.1111/jcpp.12560
99. Appel CW, Frederiksen K, Hjalgrim H, Dyregrov A, Dalton SO, Dencker A, et al. Depressive symptoms and mental health-related quality of life in adolescence and young adulthood after early parental death. *Scand J Public Health*. (2018) 47:782–92. doi: 10.1177/1403494818806371
100. Pelkonen M, Marttunen M, Kaprio J, Huurre T, Aro H. Adolescent risk factors for episodic and persistent depression in adulthood. A 16-year prospective follow-up study of adolescents. *J Affect Disord*. (2008) 106:123–31. doi: 10.1016/j.jad.2007.06.001
101. Wingenfeld K, Spitzer C, Mensebach C, Grabe H, Hill A, Gast U, et al. Die deutsche Version des Childhood Trauma Questionnaire (CTQ): erste befunde zu den psychometrischen kennwerten. *Psychother Psychosom Medizinische Psychol*. (2010) 60:442–50. doi: 10.1055/s-0030-1247564
102. Brausch AM, Gutierrez PM. The role of body image and disordered eating as risk factors for depression and suicidal ideation in adolescents. *Suicide Life Threatening Behav*. (2009) 39:58–71. doi: 10.1521/suli.2009.39.1.58
103. Robinson E, Sutin A, Daly M. Perceived weight discrimination mediates the prospective relation between obesity and depressive symptoms in U.S. and U.K. adults. *Heal Psychol*. (2017) 36:112–21. doi: 10.1037/hea0000426
104. Czeglédi E, Urbán R. Risk factors and alteration of depression among participants of an inpatient weight loss program. *Psychiatr Hung*. (2012) 27:361–78.
105. Riquin E, Lamas C, Nicolas I, Dugre Lebigre C, Curt F, Cohen H, et al. A key for perinatal depression early diagnosis: the body dissatisfaction. *J Affect Disord*. (2019) 245:340–7. doi: 10.1016/j.jad.2018.11.032
106. Hoare E, Skouteris H, Fuller-Tyszkiewicz M, Millar L, Allender S. Associations between obesogenic risk factors and depression among adolescents: a systematic review. *Obes Rev*. (2014) 15:40–51. doi: 10.1111/obr.12069

107. Hilbert A, de Zwaan M, Braehler E. How frequent are eating disturbances in the population? Norms of the eating disorder examination-questionnaire. *PLoS ONE*. (2012) 7:e29125. doi: 10.1371/journal.pone.0029125
108. Ogawa S, Kitagawa Y, Fukushima M, Yonehara H, Nishida A, Togo F, et al. Interactive effect of sleep duration and physical activity on anxiety/depression in adolescents. *Psychiatry Res*. (2019) 273:456–60. doi: 10.1016/j.psychres.2018.12.085
109. Zhou X, Bi B, Zheng L, Li Z, Yang H, Song H, et al. The prevalence and risk factors for depression symptoms in a rural Chinese sample population. *PLoS ONE*. (2014) 9:e99692. doi: 10.1371/journal.pone.0099692
110. Tang B, Liu X, Liu Y, Xue C, Zhang L. A meta-analysis of risk factors for depression in adults and children after natural disasters. *BMC Public Health*. (2014) 14:1–12. doi: 10.1186/1471-2458-14-623
111. Strine TW, Mokdad AH, Balluz LS, Gonzalez O, Crider R, Berry JT, et al. Depression and anxiety in the United States: findings from the 2006 behavioral risk factor surveillance system. *Psychiatr Serv*. (2015) 59:1383–90. doi: 10.1176/ps.2008.59.12.1383
112. Bernard O, Gibson RC, McCaw-Binns A, Reece J, Coore-Desai C, Shakespeare-Pellington S, et al. Antenatal depressive symptoms in Jamaica associated with limited perceived partner and other social support: a cross-sectional study. *PLoS ONE*. (2018) 13:1–19. doi: 10.1371/journal.pone.0194338
113. Kliem S, Job A-K, Kröger C, Bodenmann G, Stöbel-Richter Y, Hahlweg K, et al. Entwicklung und Normierung einer Kurzform des Partnerschaftsfragebogens (PFB-K) an einer repräsentativen deutschen Stichprobe. *Z Klin Psychol Psychother*. (2012) 41:81–9. doi: 10.1026/1616-3443/a000135
114. Han K-M, Jee H-J, An H, Shin C, Yoon H-K, Ko Y-H, et al. Intimate partner violence and incidence of depression in married women: a longitudinal study of a nationally representative sample. *J Affect Disord*. (2019) 245:305–11. doi: 10.1016/j.jad.2018.11.041
115. Trevillion K, Oram S, Feder G, Howard LM. Experiences of domestic violence and mental disorders: a systematic review and meta-analysis. *PLoS ONE*. (2012) 7:e51740. doi: 10.1371/journal.pone.0051740
116. Infurna MR, Reichl C, Parzer P, Schimmenti A, Bifulco A, Kaess M. Associations between depression and specific childhood experiences of abuse and neglect: a meta-analysis. *J Affect Disord*. (2016) 190:47–55. doi: 10.1016/j.jad.2015.09.006
117. Nakulan A, Sumesh TP, Kumar S, Rejani PP, Shaji KS. Prevalence and risk factors for depression among community resident older people in Kerala. *Indian J Psychiatry*. (2017) 57:262–6. doi: 10.4103/0019-5545.166640
118. Kounali D, Zammit S, Wiles N, Sullivan S, Cannon M, Stochl J, et al. Common versus psychopathology-specific risk factors for psychotic experiences and depression during adolescence. *Psychol Med*. (2014) 44:2557–66. doi: 10.1017/S0033291714000026
119. Rajapakshe OBW, Sivayogan S, Kulatunga PM. Prevalence and correlates of depression among older urban community-dwelling adults in Sri Lanka. *Psychogeriatrics*. (2019) 19:202–11. doi: 10.1111/psyg.12389
120. Brugha TS, Cragg D. The list of threatening experiences: the reliability and validity of a brief life events questionnaire. *Acta Psychiatr Scand*. (1990) 82:77–81. doi: 10.1111/j.1600-0447.1990.tb01360.x
121. Stubbs B, Koyanagi A, Hallgren M, Firth J, Richards J, Schuch F, et al. Physical activity and anxiety: a perspective from the World Health Survey. *J Affect Disord*. (2017) 208:545–52. doi: 10.1016/j.jad.2016.10.028
122. Blough J, Loprinzi PD. Experimentally investigating the joint effects of physical activity and sedentary behavior on depression and anxiety: a randomized controlled trial. *J Affect Disord*. (2018) 239:258–68. doi: 10.1016/j.jad.2018.07.019
123. Schilling R, Schärli E, Fischer X, Donath L, Faude O, Brand S, et al. The utility of two interview-based physical activity questionnaires in healthy young adults: comparison with accelerometer data. *PLoS ONE*. (2018) 13:e0203525. doi: 10.1371/journal.pone.0203525
124. Owens SA, Helms SW, Rudolph KD, Hastings PD, Nock MK, Prinstein MJ. Interpersonal stress severity longitudinally predicts adolescent girls' depressive symptoms: the moderating role of subjective and HPA axis stress responses. *J Abnorm Child Psychol*. (2019) 47:895–905. doi: 10.1007/s10802-018-0483-x
125. Yeoh SH, Tam CL, Wong CP, Bonn G. Examining depressive symptoms and their predictors in malaysia: stress, locus of control, and occupation. *Front Psychol*. (2017) 8:1411. doi: 10.3389/fpsyg.2017.01411
126. Sydenham M, Beardwood J, Rimes KA. Beliefs about emotions, depression, anxiety and fatigue: a mediational analysis. *Behav Cogn Psychother*. (2017) 45:73–8. doi: 10.1017/S1352465816000199
127. Brockmeyer T, Kulessa D, Hautzinger M, Bents H, Backenstrass M. Differentiating early-onset chronic depression from episodic depression in terms of cognitive-behavioral and emotional avoidance. *J Affect Disord*. (2015) 175:418–23. doi: 10.1016/j.jad.2015.01.045
128. Appel M, Gnambs T, Maio GR. A short measure of the need for affect. *J Pers Assess*. (2012) 94:418–26. doi: 10.1080/00223891.2012.666921
129. Kashdan TB, Zvolensky MJ, McLeish AC. Anxiety sensitivity and affect regulatory strategies: individual and interactive risk factors for anxiety-related symptoms. *J Anxiety Disord*. (2008) 22:429–40. doi: 10.1016/j.janxdis.2007.03.011
130. Hovenkamp-Hermelink JHM, van der Veen DC, Oude Voshaar RC, Batelaan NM, Penninx BWJH, Jeronimus BF, et al. Anxiety sensitivity, its stability and longitudinal association with severity of anxiety symptoms. *Sci Rep*. (2019) 9:1–7. doi: 10.1038/s41598-019-39931-7
131. Kemper CJ, Ziegler M, Taylor S. Überprüfung der psychometrischen qualität der deutschen version des angstsensitivitätsindex-3. *Diagnostica*. (2009) 55:223–33. doi: 10.1026/0012-1924.55.4.223
132. Crowe L, Butterworth P. The role of financial hardship, mastery and social support in the association between employment status and depression: results from an Australian longitudinal cohort study. *BMJ Open*. (2016) 6:1–10. doi: 10.1136/bmjopen-2015-009834
133. Slotman A, Snijder MB, Ikram UZ, Schene AH, Stevens GWJM. The role of mastery in the relationship between perceived ethnic discrimination and depression: the HELIUS study. *Cult Divers Ethn Minor Psychol*. (2017) 23:200–8. doi: 10.1037/cdp0000113
134. Pahlevan Sharif S. Locus of control, quality of life, anxiety, and depression among Malaysian breast cancer patients: the mediating role of uncertainty. *Eur J Oncol Nurs*. (2017) 27:28–35. doi: 10.1016/j.ejon.2017.01.005
135. Pearlin LI, Schooler C. The structure of coping. *J Health Soc Behav*. (1978) 19:2–21. doi: 10.2307/2136319
136. Young CC, Dietrich MS. Stressful life events, worry, and rumination predict depressive and anxiety symptoms in young adolescents. *J Child Adolesc Psychiatr Nurs*. (2015) 28:35–42. doi: 10.1111/jcap.12102
137. Gorday JY, Rogers ML, Joiner TE. Examining characteristics of worry in relation to depression, anxiety, and suicidal ideation and attempts. *J Psychiatr Res*. (2018) 107:97–103. doi: 10.1016/j.jpsychires.2018.10.004
138. Topper M, Emmelkamp PMG, Watkins E, Ehring T. Development and assessment of brief versions of the Penn State Worry Questionnaire and the Ruminative Response Scale. *Br J Clin Psychol*. (2014) 53:402–21. doi: 10.1111/bjc.12052
139. MacPhee AR, Andrews JJW. Risk factors for depression in early adolescence. *Adolescence*. (2006) 41:435–66.
140. Muris P, Meesters C, Pierik A, de Kock B. Good for the self: self-compassion and other self-related constructs in relation to symptoms of anxiety and depression in non-clinical youths. *J Child Fam Stud*. (2016) 25:607–17. doi: 10.1007/s10826-015-0235-2
141. von Collani G, Herzberg PY. Eine revidierte fassung der deutschsprachigen skala zum selbstwertgefühl von rosenberg [A revised version of the german adaptation of rosenberg's self-esteem scale]. *Zeitschrift Differ Diagnostische Psychol*. (2003) 24:3–7. doi: 10.1024//0170-1789.24.1.3
142. Li J, Theng YL, Foo S. Depression and psychosocial risk factors among community-dwelling older adults in Singapore. *J Cross Cult Gerontol*. (2015) 30:409–22. doi: 10.1007/s10823-015-9272-y
143. Matthews T, Danese A, Wertz J, Odgers CL, Ambler A, Moffitt TE, et al. Social isolation, loneliness and depression in young adulthood: a behavioural genetic analysis. *Soc Psychiatry Psychiatr Epidemiol*. (2016) 51:339–48. doi: 10.1007/s00127-016-1178-7
144. Cacioppo JT, Hughes ME, Waite LJ, Hawkley LC, Thisted RA. Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. *Psychol Aging*. (2006) 21:140–51. doi: 10.1037/0882-7974.21.1.140
145. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based

studies. *Res Aging*. (2004) 26:655–72. doi: 10.1177/0164027504268574
146. Miletic V, Lukovic JA, Ratkovic N, Aleksic D, Grgurevic A. Demographic risk factors for suicide and depression among Serbian medical school students. *Soc Psychiatry Psychiatr Epidemiol*. (2015) 50:633–8. doi: 10.1007/s00127-014-0950-9
147. Lu W, Bian Q, Song YY, Ren JY, Xu XY, Zhao M. Prevalence and related risk factors of anxiety and depression among Chinese college freshmen. *J Huazhong Univ Sci Technol Med Sci*. (2015) 35:815–22. doi: 10.1007/s11596-015-1512-4
148. Rai D, Zitko P, Jones K, Lynch J, Araya R. Country- and individual-level socioeconomic determinants of depression: multilevel cross-national comparison. *Br J Psychiatry*. (2013) 202:195–203. doi: 10.1192/bjp.bp.112.112482
149. Sajjadi H, Mohaqeqi Kamal SH, Rafiey H, Vameghi M, Forouzan AS, Rezaei M. A systematic review of the prevalence and risk factors of depression among Iranian adolescents. *Glob J Health Sci*. (2013) 5:16–27. doi: 10.5539/gjhs.v5n3p16
150. Ruiz-Castell M, Kandala NB, Perquin M, Bocquet V, Kuemmerle A, Vögele C, et al. Depression burden in luxembourg: individual risk factors, geographic variations and the role of migration, 2013–2015 European Health Examination Survey. *J Affect Disord*. (2017) 222:41–8. doi: 10.1016/j.jad.2017.06.056
151. Neale EP. Migration and depression. *J R Stat Soc*. (1993) 96:663–5. doi: 10.2307/2341904
152. Lindert J, Ehrenstein OS von, Priebe S, Mielck A, Brähler E. Depression and anxiety in labor migrants and refugees – a systematic review and meta-analysis. *Soc Sci Med*. (2009) 69:246–57. doi: 10.1016/j.socscimed.2009.04.032
153. Dennis CLE, Janssen PA, Singer J. Identifying women at-risk for postpartum depression in the immediate postpartum period. *Acta Psychiatr Scand*. (2004) 110:338–46. doi: 10.1111/j.1600-0447.2004.00337.x
154. Assari S. Social determinants of depression: the intersections of race, gender, and socioeconomic status. *Brain Sci*. (2017) 7:E156. doi: 10.3390/brainsci7120156
155. Fussner LM, Luebbe AM, Mancini KJ, Becker SP. Emotion dysregulation mediates the longitudinal relation between peer rejection and depression: differential effects of gender and grade. *Int J Behav Dev*. (2018) 42:155–66. doi: 10.1177/0165025416669062
156. Davis AN, Carlo G, Schwartz SJ, Unger JB, Zamboanga BL, Lorenzo-Blanco EI, et al. The longitudinal associations between discrimination, depressive symptoms, and prosocial behaviors in U.S. Latino/a recent immigrant adolescents. *J Youth Adolesc*. (2016) 45:457–70. doi: 10.1007/s10964-015-0394-x
157. Torres Á, Blanco V, Vázquez FL, Díaz O, Otero P, Hermida E. Prevalence of major depressive episodes in non-professional caregivers. *Psychiatry Res*. (2015) 226:333–9. doi: 10.1016/j.psychres.2014.12.066
158. Pinquart M, Sörensen S. Differences between caregivers and noncaregivers in psychological health and physical health: a meta-analysis. *Psychol Aging*. (2003) 18:250–67. doi: 10.1037/0882-7974.18.2.250
159. Bøe T, Dearing E, Stormark KM, Zachrisson HD. Subjective economic status in adolescence: determinants and associations with mental health in the Norwegian Youth@Hordaland Study. *J Fam Econ Issues*. (2018) 39:323–36. doi: 10.1007/s10834-017-9553-4
160. Sargent JT, Crocker J, Luhtanen RK. Contingencies of self–worth and depressive symptoms in college students. *J Soc Clin Psychol*. (2006) 25:628–46. doi: 10.1521/jscp.2006.25.6.628
161. Adler NE, Epel ES, Castellazzo G, Ickovics JR. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. *Heal Psychol*. (2000) 19:586–92. doi: 10.1037/0278-6133.19.6.586
162. Rothwell JC, Julious SA, Cooper CL. A study of target effect sizes in randomised controlled trials published in the Health Technology Assessment journal. *Trials*. (2018) 19:544. doi: 10.1186/s13063-018-2886-y
163. Hawkes AL, Chambers SK, Pakenham KI, Patrao TA, Baade PD, Lynch BM, et al. Effects of a telephone-delivered multiple health behavior change intervention (CanChange) on health and behavioral outcomes in survivors of colorectal cancer: a randomized controlled trial. *J Clin Oncol*. (2013) 31:2313–21. doi: 10.1200/JCO.2012.45.5873
164. R Core Team. *R: A Language and Environment for Statistical Computing*. R Found Stat Comput (2019). Available online at: http://www.r-project.org/
165. Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis*. San Diego, CA: Academic Press (1985).
166. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol*. (1991) 59:12–9. doi: 10.1037/0022-006X.59.1.12
167. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated health economic evaluation reporting standards (CHEERS)-explanation and elaboration: a report of the ISPOR health economic evaluation publication guidelines good reporting practices task force. *Value Heal*. (2013) 16:231–50. doi: 10.1186/1741-7015-11-80
168. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-effectiveness analysis alongside clinical trials II - An ISPOR good research practices task force report. *Value Heal*. (2015) 18:161–72. doi: 10.1016/j.jval.2015.02.001
169. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart G. *Methods for the Economic Evaluation of Healthcare Programmes*. 3rd ed. Oxford: Oxford University Press (2005).
170. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ*. (2004) 13:461–75. doi: 10.1002/hec.843
171. Glick HA, Doshi JA, Sonnad SS, Polsky D. In: Polsky D, editor. *Economic Evaluation in Clinical Trials. Handbooks*. Oxford: Oxford University Press (2014).
172. Fenwick E, Marshall DA, Levy AR, Nichol G. Using and interpreting cost-effectiveness acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. *BMC Health Serv Res*. (2006) 6:52. doi: 10.1186/1472-6963-6-52
173. StataCorp. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LLC. (2015).
174. Dejonghe LAL, Becker J, Froboese I, Schaller A. Long-term effectiveness of health coaching in rehabilitation and prevention: a systematic review. *Patient Educ Couns*. (2017) 100:1643–53. doi: 10.1016/j.pec.2017.04.012
175. Wandersman A, Duffy J, Flaspohler P, Noonan R, Lubell K, Stillman L, et al. Bridging the gap between prevention research and practice: the interactive systems framework for dissemination and implementation. *Am J Community Psychol*. (2008) 41:171–81. doi: 10.1007/s10464-008-9174-z
176. Ammentorp J, Uhrenfeldt L, Angel F, Ehrensvärd M, Carlsen EB, Kofoed PE. Can life coaching improve health outcomes? - A systematic review of intervention studies. *BMC Health Serv Res*. (2013) 13:428. doi: 10.1186/1472-6963-13-42

**Conflict of Interest:** DE has served as a consultant to/on the scientific advisory boards of Sanofi, Novartis, Minddistrict, Lantern, Schoen Kliniken, Ideamed, and German health insurance companies (BARMER, Techniker Krankenkasse) and a number of federal chambers for psychotherapy. DE and MB are stakeholders of the Institute for health training online (GET.ON), which aims to implement scientific findings related to digital health interventions into routine care. HB reports to have received consultancy fees and fees for lectures/workshops from chambers of psychotherapists and training institutes for psychotherapists in the e-mental-health context. IT reports to have received fees for lectures/workshops in the e-mental-health context from training institutes for psychotherapists. She is implementation lead and project lead for the EU-project ImpleMent. All at the Institute for health training online (GET.ON).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# 4

## ARTICLE 2 — EFFECTIVENESS TELEPHONE COACHING AT 6 MONTHS (TEC-A)

Title:
Telephone Coaching for the Prevention of Depression in Farmers: Results from a Pragmatic Randomized Controlled Trial

Summary:
Farmers and related professions are at risk of suffering from depression due to several occupational risk factors, limited health services in rural areas and low help-seeking behavior. To reduce the depression burden in this population, low-threshold, remotely deliverable digital preventative programs are necessary. The present study aimed to evaluate the effectiveness of telephone coaching in reducing depressive symptoms and preventing onset of depression in farmers.

The study was designed as a two-armed pragmatic randomized controlled trial. A total of 314 farmers with elevated levels of depressive symptoms (Patient Health Questionnaire; PHQ-9 $\geq$ 5) from Germany were recruited and randomly assigned to receive either personalized telephone coaching (N = 160) or an psychoeducational e-mail in addition to treatment as usual (TAU+, N = 154). On average, the coaching consisted of 13 (SD = 7) sessions of 48 minutes each (SD = 15) over 6 months (SD = 2) and was conducted by psychologists with varying training in psychological methods. Frequency, duration and topics of the coaching were adapted to individual needs. Depressive symptom severity as the primary outcome was measured with the Quick Inventory of Depressive Symptomatology (QIDS-SR16) and secondary outcomes included depression onset and further mental health problems. Group differences were tested using analysis of covariance (ANCOVA) while controlling for baseline symptom severity.

The results showed that the telephone coaching was superior in reducing the severity of depressive symptoms compared to TAU+ with a small to moderate effect (d = -0.39, [95%–CI: -0.64, -0.15]). No differences in depression onset based on a QIDS-SR16 cut-off of $\geq$ 13 was observed (IRR = 0.36 [95%–CI: 0.10, 1.36]). Group differences in favor of the intervention were also observed in stress (d = -0.34 [95%–CI: -0.08, -0.59]), generalized anxiety (d = -0.30 [95%–CI: -0.07,

-0.53]), somatic symptoms (d = -0.39 [95%–CI: -0.16, -0.63]), burnout risk (d = -0.24 [95%–CI: 0.0, -0.47] to d = -0.39 [95%–CI: -0.63, -0.16)]) and quality of life (d = 0.28 [95%–CI: 0.52, 0.06]). No group differences were observed in sleep, panic and agoraphobic symptoms, alcohol consumption and subjective working capacity. The intervention was well-accepted by the participants, with only 10.6% (n = 17/160) of participants terminating intervention early and a mean satisfaction rating of M = 28.17 (SD = 5.58) on the Client Satisfaction Questionnaire (CSQ-I, range: 8 – 32).

A limiting factor is the inclusion of participants with higher baseline severity because no upper cut-off was applied. This also reduced the power to evaluate depression onset. Generalization is limited by the high degree of personalization in the intervention and the specific occupational target group.

The study results suggest that telephone coaching can be a promising way to reduce depressive symptoms and other mental health problems in farmers and related professions. It can serve as an easily accessible intervention for prevention and early treatment and could potentially facilitate further help-seeking if needed. Information of long term- and cost-effectiveness are warranted to estimate the full potential of personalized telephone coaching for farmers' mental health.

Publication:
Thielecke J, Buntrock C, Titzler I, Braun L, Freund J, Berking M, Baumeister H, and Ebert DD. Telephone Coaching for the Prevention of Depression in Farmers: Results from a Pragmatic Randomized Controlled Trial. Journal of Telemedicine and Telecare. 2022; online first. doi: 10.1177/1357633X221106027

Contribution:
Janika Thielecke contributed to the study conceptualization, design and methods. She prepared and submitted the research proposal for the local ethics committee. She was responsible for the recruitment, study administration, data curation and analysis. She drafted the original manuscript and revised it after feedback from her co-authors. She managed the submission process and reviewer comments while receiving feedback from her co-authors.

*RESEARCH/Original Article*

*Journal of Telemedicine and Telecare*
1–13

DOI: 10.1177/1357633X221106027
journals.sagepub.com/home/jtt
SAGE

# Telephone coaching for the prevention of depression in depression in farmers: Results from a pragmatic from a pragmatic randomized controlled trial

**Janika Thielecke[1,2] Claudia Buntrock[1] Ingrid Titzler[1] Lina Braun[3] Johanna Freund[1,2] Matthias Berking[1] Harald Baumeister[3] and David D. Ebert[2]**

## Abstract

**Introduction:** Farmers have a high risk for depression (MDD). Preventive measures targeting this often remotely living population might reduce depression burden. The study aimed to evaluate the effectiveness of personalized telephone coaching in reducing depressive symptom severity and preventing MDD in farmers compared to enhanced treatment as usual (TAU+).

**Methods:** In a two-armed, pragmatic randomized controlled trial ($N = 314$) with post-treatment at 6 months, farming entrepreneurs, collaborating family members and pensioners with elevated depressive symptoms (PHQ-9 $\geq 5$) were randomized to personalized telephone coaching or TAU+. The coaching was provided by psychologists and consists on average of 13 (±7) sessions a 48 min (±15) over 6 months. The primary outcome was depressive symptom severity (QIDS-SR16).

**Results:** Coaching participants showed a significantly greater reduction in depressive symptom severity compared to TAU+ ($d = 0.39$). Whereas reliable symptom deterioration was significantly lower in the intervention group compared to TAU+, no significant group differences were found for reliable improvement and in depression onset. Further significant effects in favor of the intervention group were found for stress ($d = 0.34$), anxiety ($d = 0.30$), somatic symptoms ($d = 0.39$), burnout risk ($d = 0.24$–$0.40$) and quality of life ($d = 0.28$).

**Discussion:** Limiting, we did not apply an upper cutoff score for depressive symptom severity or controlled for previous MDD episodes, leaving open whether the coaching was recurrence/relapse prevention or early treatment. Nevertheless, personalized telephone coaching can effectively improve mental health in farmers. It could play an important role in intervening at an early stage of mental health problems and reducing disease burden related to MDD.

**Trial registration number and trial register:** German Clinical Trial Registration: DRKS00015655

## Keywords

Depression, prevention, RCT, farmers, telephone coaching, telehealth, mental health

Date received: 25 February 2022; Date accepted: 20 May 2022

## Introduction

Lifetime prevalence rates for major depressive disorder (MDD) are estimated between 2% and 21.0% worldwide.[1] MDD is associated with high individual[2,3] and societal burden.[4,5] Yet, primary care providers fail to recognize MDD in nearly half of patients[6] and only 20% receive specialized mental health care.[7] But even in a hypothetical scenario of full coverage and compliance to evidence-based treatments, only one-third of MDD-related disease burden could be averted.[8,9] Preventive approaches are promising in reducing disease burden. According to recent meta-analyses,

[1]Department of Clinical Psychology and Psychotherapy, Institute of Psychology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[2]Professorship of Psychology and Digital Mental Health Care, Technical University of Munich, Munich, Germany
[3]Department of Clinical Psychology and Psychotherapy, Institute of Psychology and Education, Ulm University, Ulm, Germany

**Corresponding author:**
Janika Thielecke, Friedrich-Alexander-University Erlangen-Nürnberg, Department of Clinical Psychology and Psychotherapy, Nägelsbachstr. 25a, 91052 Erlangen.
Email: janika.thielecke@fau.de

psychological interventions could effectively prevent MDD onset when targeting people in risk groups or with subthreshold symptoms.[10,11]

A recent systematic review including mostly studies from English-speaking countries showed an elevated risk in farmers for mental disorders compared to people in nonfarming occupations.[12] Farming-related risk factors include poor physical health, high workload, exposure to pesticides, financial hardship, and uncertain weather conditions. Farmers also face barriers to accessing mental health services including few providers in rural areas,[12,13] preferring help from oneself, family or friends,[14] and anticipated stigma.[15,16]

Therefore, low-threshold preventive interventions for this vulnerable group are warranted. Telephone coaching could aid by reducing travel time and being easily accessible. One-on-one telephone interventions have been shown to be effective in supporting people with physical long-term conditions[17] and in delivering psychological interventions for depression (e.g. cognitive behavioral therapy, interpersonal psychotherapy).[18] In a recent meta-analysis, cognitive behavioral therapy delivered over phone has shown comparable outcome improvements compared to face-to-face therapy for MDD.[19,20] However, for coaching conducted over the telephone, only two studies could be identified that evaluated the effectiveness in reducing depressive symptom severity and revealed moderate effects in mostly white-collar workers.[21,22] Yet the overall effectiveness of telephone coaching in farmers as well as its effect on preventing MDD remains unclear.

The present study evaluated the effectiveness of telephone coaching personalized to the individual farmer's needs compared to enhanced treatment as usual in reducing depressive symptom severity and preventing MDD onset at a six-month follow-up. This study is part of a nationwide prevention project by the German social insurance for farmers, foresters, and gardeners (SVLFG), which aims to implement internet- and tele-based interventions into routine care.

## Methods

The study was a two-armed pragmatic randomized controlled trial comparing the effectiveness of personalized telephone coaching delivered by a third-party independent healthcare provider "IVPNetworks" to a control group receiving information material in addition to usual care (TAU + ) at post-treatment (six months after randomization).

The trial was approved by the Medical Ethics Committee of the Friedrich-Alexander University Erlangen-Nürnberg (No. 345_18 B) and registered in the German clinical trial register (DRKS00015655). The study protocol can be found elsewhere.[23]

### Study population

Individuals were included if they were (a) insured by the SVLFG, (b) entrepreneur, assisting family member or pensioned farmer, (c) 18 years or older, (d) showed elevated depressive symptoms (PHQ-9≥5), (e) had internet/telephone access for online assessments/coaching, and (f) provided written informed consent.

Applicants were excluded if they were (a) currently receiving psychotherapy, (b) unwilling to sign a nonsuicide contract in case of suicidal ideation or (c) living in the German Federal States Bavaria or Schleswig-Holstein, as roll-out of the intervention into routine care was ongoing there.[24] Due to the parallel recruitment for two RCTs evaluating online trainings in the same target group,[25,26] we (d) excluded participants from these studies. To mimic routine care, no diagnostic interviews were conducted.

### Procedure

Participants were recruited and enrolled in the study from December 2018 to April 2019. The main recruitment channel was personal invitation letters to randomly selected insured persons. The study was also advertised in the SVLFG members' journal and newsletter and associated websites.

Inclusion criteria were assessed online and eligible individuals who provided written informed consent and completed the baseline assessment were randomly allocated to the intervention (IG) or control group (CG) (see Figure 1).

Randomization was centrally done at an individual level by an independent researcher not involved in the study. Permuted block randomization with randomly arranged block sizes(4,6,8,12) and an allocation ratio of 1:1 was used based on a web-based program (Sealed Envelope). Study participants and coaches could not be blinded but data collectors were blind to group allocation.

Post-assessment was assessed six months after randomization, regardless of the coaching duration, between June 2019 and December 2019. Completion was rewarded with €15 in both groups. The primary outcome was assessed via telephone if participants did not complete the assessment within two months.

### Study arms

All participants had unrestricted access to routine care. Actual health service use was monitored with items of the TiC-P (see secondary outcomes).[27,28]

*Intervention group.* The study team registered participants on the management platform (IVPnet). Case managers at IVPNetworks assigned coaches to participants. Coaches were psychologists with a diploma or master's degree in psychology and trained in different psychological methods (e.g. cognitive behavioral, systemic, hypnotherapeutic), while licensed psychotherapists were available for supervision. Coaching is an individual-centered process, which is problem-based and solution-oriented. It aims to reactivate

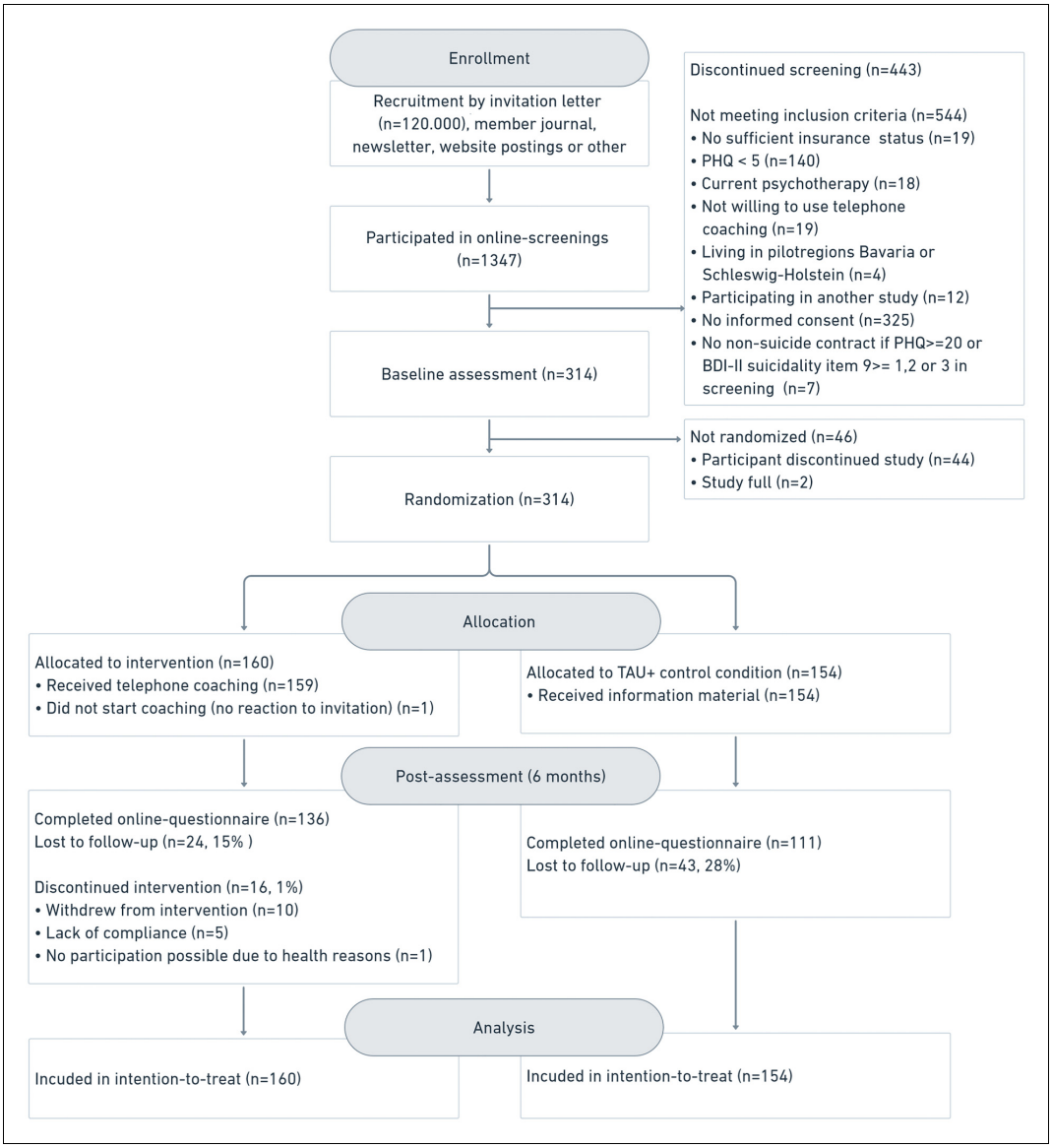



**Figure 1.** CONSORT study flow

and build individuals' resources by using psychoeducational and psychological methods to enable participants to cope with stress, acute problems, or general worries.

No fixed procedures or standardized manuals were applied in the coaching. Coaching methods varied depending on the coach's background, while timing and content were permanently adapted to the participants' needs. As a guideline, a coaching volume of 850 min or six months was set with the possibility of prolonging for an additional 150 min or three months, respectively. If coaches identified clinically relevant symptomatology, they could support participants in finding adequate help (e.g. contact general practitioner or psychiatric clinic). Additionally, participants were supported to find on-site support services (e.g. socio-economic or agricultural family counseling) or switch to an onsite coaching if indicated.

*Control group.* Participants in the control group received brief psychoeducation material on stress and depression by email, combined with information about access to regular care.

## Outcomes

Primary and secondary outcomes were assessed via online questionnaires (Unipark) at baseline and post-treatment (T1).

*Primary outcome.* Depressive symptom severity was assessed with the German version of the Quick Inventory Depressive Symptomology (QIDS-SR16) with scores ranging from 0 to 27. Scores between 0 and 5, indicate normal health status, while those between 6–10, 11–15, 16–20, and greater than 20, indicate mild, moderate, severe, or very severe depressive symptom severity, respectively.[29] Reliability in this study was acceptable[30] with $\alpha_{T0} = .72$ and $\alpha_{T1} = .78$.

*Secondary outcomes.* Lifetime history and major depressive episodes (MDE) onset within the past year, bipolar disorder (BPD), and general anxiety disorder (GAD) were self-reported with adapted items from the web version of the Composite International Diagnosis Interview version 3.0 (CIDI).[31] For depression onset at post-treatment, both CIDI rating and a score of 13 or greater on the QIDS-SR16 as a cutoff for possible acute cases of clinical depression[32] were applied.

Secondary outcomes further included perceived stress (Perceived Stress Scale, PSS, range 0–40, $\alpha_{T0} = .85$ – $\alpha_{T1} = .90$),[33] insomnia severity (Insomnia Severity Index, ISI, range 0–28, $\alpha_{T0} = .81$ – $\alpha_{T1} = .88$),[34] somatic symptom burden (Somatic Symptom Scale, SSS-8, range 0–32, $\alpha_{T0} = .71$ – $\alpha_{T1} = .76$),[35] severity of GAD (GAD-7, range 0–21, $\alpha_{T0} = .78$ – $\alpha_{T1}$ 0.83),[36,37] severity of panic and agoraphobic symptoms (Panic and Agoraphobia Scale, PAS, range 0–48, $\alpha_{T0} = .89$ – $\alpha_{T0} = .92$),[38] alcohol consumption (consumption questions from Alcohol Use Disorder Identification Test, AUDIT-C, range 0–12, $\alpha_{T0/1} = .59$)[39] and quality of life (Assessment of Quality of Life-8D, AQoL-8D, range 0–100, $\alpha_{T0} = .92$; $\alpha_{T1} = .94$).[40]

Work-related, the subjective capacity to work (SPE, Guttman scale categories 0–3, $rep_{T0/1} = .95$)[41] as well as burnout symptomology (Maslach Burnout Inventory, MBI-GS, subscales: "Emotional Exhaustion," EE, range 0–30; "Cynicism," CY, range 0–30; "Professional Efficacy," PE, range 0–36; $\alpha_{T0} = .70$ – .89, $\alpha_{T1} = .77$ – .91)[42,43] was assessed.

A context-adapted version of the cost questionnaire "Trimbos Institute and Institute of Medical Technology Questionnaire for Costs Associated with Psychiatric Illness" (TiC-P)[27,28] was used to monitor TAU.

*Intervention-related outcomes.* Intervention-related outcomes included satisfaction with the coaching or information material (Client Satisfaction Questionnaire for internet interventions, CSQ-I, range 8–32, $\alpha_{CG} = .95$ – $\alpha_{IG} = .96$),[44–46] working alliance (Working Alliance Inventory, participants: WAI-SR, range 12–60, $\alpha = .93$; coaches: WAI-SRT, range 10–50, $\alpha = .90$),[47] and negative effects (Inventory for the Assessment of Negative Effects of Psychotherapy, INEP).[48] Data on the coaching process (e.g. duration, sessions, topics) were retrieved from IVPnet.

## Data analyses

Based on a power of 80%, an alpha of 0.05 (two-sided), and an attrition rate of 20%,[21,22] 312 participants were needed to be able to detect an effect of $d = 0.35$[49] using an independent t-test (G*Power Version 3.1.9.2).

Results are reported according to the Consolidated Standards of Reporting Trials (CONSORT) 2010 Statement and its extension for reporting pragmatic trials.[50] Analyses were performed in R[51] based on intention-to-treat principles with a two-sided significance level of 0.05. Missing data were imputed using the R-package MICE[52] to apply Multiple Imputation by Chained Equations assuming data to be missing at random. Representing the percentage of incomplete cases at post-treatment, 21 imputed datasets were generated.[53] Intervention-related outcomes and health service use were not imputed. We used analysis of covariance (ANCOVA) to compare primary and secondary outcomes between study groups, adjusting for baseline scores. All analyses were run in each imputed data set and estimates were pooled using Rubin's Rule.[54,55] Results were reported as mean within- and between-group differences and as Cohen's *d* effect sizes with corresponding 95% CIs according to Hedges and Olkin.[56]

*Treatment response, reliable deterioration, and close-to-symptom-free status.* To assess improvements in the primary outcome at an individual level, we examined the number of participants who showed a treatment response as defined by a reliable change from baseline to post-treatment according to the reliable change index (RCI) of Jacobson and Truax.[57] Participants were defined as reliably improved if their QIDS-SR16 score declined from baseline to post-treatment with more than 1.96 standard units corresponding to a point change of at least 6, while taking into account the reliability of the measurement instrument. The same method was used to assess an increase in depressive symptoms to indicate reliable deterioration.

Close-to-symptom-free status was a priori defined as a score of $\leq 5$ on the QIDS-SR16.[29] Individuals below this score at baseline were excluded from this analysis. Differences in treatment response and close-to-symptom-free

**Table 1.** Baseline characteristics of the study population ($N = 314$).

| | IG ($n = 160$) | | CG ($n = 154$) | | Total | |
|---|---|---|---|---|---|---|
| Name | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| **Age** | 53.12 | 9.25 | 50.10 | 10.49 | 51.64 | 9.98 |
| **QIDS-SR16 (sum score)** | 9.81 | 4.13 | 10.26 | 4.40 | 10.03 | 4.26 |
| **QIDS-SR16 (categorical)** | **n** | **%** | **n** | **%** | **n** | **%** |
| 0–5 none | 20 | 12.5 | 21 | 13.6 | 41 | 13.1 |
| 6–10 mild | 76 | 47.5 | 61 | 39.6 | 137 | 43.6 |
| 11–15 moderate | 45 | 28.1 | 53 | 34.4 | 98 | 31.2 |
| 16–20 severe | 17 | 10.6 | 17 | 11.0 | 34 | 10.8 |
| ≥21 very severe | 2 | 1.3 | 2 | 1.3 | 4 | 1.3 |
| **Sex** | | | | | | |
| Male | 84 | 52.5 | 81 | 52.6 | 165 | 52.6 |
| Female | 76 | 47.5 | 73 | 47.4 | 149 | 47.5 |
| **Birthplace** | | | | | | |
| Germany | 158 | 98.8 | 154 | 100.0 | 312 | 99.4 |
| Other | 2 | 1.3 | 0 | 0.0 | 2 | 0.6 |
| **Ethnicity** | | | | | | |
| Caucasian | 158 | 98.8 | 153 | 99.4 | 311 | 99.0 |
| Other | 2 | 1.3 | 1 | 0.7 | 3 | 1.0 |
| **Relationship status** | | | | | | |
| With partnership | 148 | 92.5 | 133 | 86.4 | 281 | 89.5 |
| No partnership | 12 | 7.5 | 21 | 13.6 | 33 | 10.5 |
| **Education level**[a] | | | | | | |
| Low | 14 | 8.8 | 17 | 11.0 | 31 | 9.9 |
| Middle | 103 | 64.4 | 94 | 61.0 | 197 | 62.7 |
| High | 43 | 26.9 | 43 | 27.9 | 86 | 27.4 |
| **Employment status** | | | | | | |
| Entrepreneur | 94 | 58.8 | 103 | 66.9 | 197 | 62.7 |
| Entrepreneurs spouse | 35 | 21.9 | 25 | 16.2 | 60 | 19.1 |
| Pensioner farmer | 16 | 10.0 | 13 | 8.4 | 29 | 9.2 |
| Family member of entrepreneur | 14 | 8.8 | 12 | 7.8 | 26 | 8.3 |
| Incapacitated for work | 1 | 0.6 | 1 | 0.7 | 2 | 0.6 |
| **Second job off the farm**[b] | | | | | | |
| No | 42 | 85.7 | 33 | 89.2 | 75 | 87.2 |
| Yes | 7 | 14.3 | 4 | 10.8 | 11 | 12.8 |
| **Brutto Income**[c] | | | | | | |
| Low (<1000€) | 13 | 13.3 | 11 | 10.4 | 24 | 11.8 |
| Middle (1000–5000€) | 45 | 45.9 | 61 | 57.6 | 106 | 52.0 |
| High (>5000€) | 17 | 17.4 | 10 | 9.4 | 27 | 13.2 |
| Not disclosed | 23 | 23.5 | 24 | 22.6 | 47 | 23.0 |
| **Main farm branch** | | | | | | |
| Dairy farming | 34 | 23.1 | 35 | 24.1 | 69 | 23.6 |
| Arable farming | 27 | 18.4 | 30 | 20.7 | 57 | 19.5 |
| Animal farming | 33 | 22.5 | 18 | 12.4 | 51 | 17.5 |
| Wine growing | 19 | 12.9 | 25 | 17.2 | 44 | 15.1 |
| Vegetable growing | 9 | 6.1 | 11 | 7.6 | 20 | 6.9 |
| Horticulture | 11 | 7.5 | 7 | 4.8 | 18 | 6.2 |
| Fruit growing | 1 | 0.7 | 5 | 3.5 | 6 | 2.1 |
| Direct marketing | 3 | 2.0 | 3 | 2.1 | 6 | 2.1 |
| Other | 10 | 6.8 | 11 | 7.6 | 21 | 7.2 |
| **Physical illness** | | | | | | |
| Yes | 94 | 58.8 | 85 | 55.2 | 179 | 57.0 |
| No | 66 | 41.3 | 69 | 44.8 | 135 | 43.0 |
| **Current antidepressant medication** | | | | | | |
| No | 86 | 83.5 | 78 | 94.0 | 164 | 88.2 |

(continued)

**Table 1.** Continued

| | IG ($n = 160$) | | CG ($n = 154$) | | Total | |
|---|---|---|---|---|---|---|
| Name | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Yes | 17 | 16.5 | 5 | 6.0 | 22 | 11.8 |
| **Experience with psychotherapy** | | | | | | |
| No | 129 | 80.6 | 132 | 85.7 | 261 | 83.1 |
| Yes | 31 | 19.4 | 22 | 14.3 | 53 | 16.9 |
| **CIDI – MDE** | | | | | | |
| MDE in lifetime | 32 | 20.0 | 44 | 28.6 | 76 | 24.2 |
| MDE in past 12 month | 25 | 15.6 | 29 | 18.8 | 54 | 17.2 |
| **CIDI – BPD** | | | | | | |
| Broad mania in lifetime | 10 | 6.3 | 24 | 15.6 | 34 | 10.8 |
| Broad mania in past 12 month | 8 | 5.0 | 15 | 9.7 | 23 | 7.3 |
| **CIDI – GAD** | | | | | | |
| Anxiety in lifetime | 33 | 20.6 | 33 | 21.4 | 66 | 21.0 |
| Anxiety in past 12 month | 28 | 17.5 | 22 | 14.3 | 50 | 15.9 |
| Search for therapy | | | | | | |
| Not searching or waiting | 152 | 95.0 | 148 | 96.1 | 300 | 95.5 |
| **Searching for therapy** | 7 | 4.4 | 6 | 3.9 | 13 | 4.1 |
| On a waiting list for therapy | 1 | 0.6 | 0 | 0.0 | 1 | 0.3 |
| **Intervention preference** | | | | | | |
| Preference for telephone coaching | 72 | 45.0 | 74 | 48.1 | 146 | 46.5 |
| No preference | 76 | 47.5 | 69 | 44.8 | 145 | 46.2 |
| Preference for information material | 12 | 7.5 | 11 | 7.1 | 23 | 7.3 |

IG: Intervention group; CG: Control group; QIDS-SR16: Quick Inventory of Depressive Symptomatology; CIDI: Composite International Diagnostic Interview; MDE: Major Depressive Episode; BPD: Bipolar Disorder; GAD: Generalized Anxiety Disorder.
[a]Low: no formal education or lower secondary education, middle: finished upper secondary education, high: finished study or master's certificate.
[b]Only applied to spouses and family members of entrepreneur ($n = 86$).
[c]Only applied to entrepreneurs and spouses/family members with working contract in the company ($n = 204$).

status between study groups were assessed using $\chi^2$ tests. Numbers-needed-to-treat (NNT; with 95% CI) to achieve one treatment response and close-to-symptom-free status, respectively, were calculated as the inverse of the risk difference.[58,59]

*Onset and remission of potential MDD.* Group differences in depression onset were assessed with Poisson regression in the subsamples of individuals without self-reported MDE/MDD at baseline based on CIDI and QIDS-SR16 < 13, respectively, estimating incidence rate ratios (IRR) with 95% CIs. For individuals with potential MDD at baseline (QIDS-SR16≥13) group differences in remission were assessed using logistic regression. Odds Ratios (OR) and 95% CIs were reported.

*Sensitivity analyses.* To test the robustness of the findings, subgroup analyses were conducted with study completers. The influence of intervention (e.g. coaching duration, coaches' background) and individual characteristics (e.g. treatment preferences) on the primary outcome was analyzed using linear regression.

## Results

### Participants

A total of 1347 individuals completed the screening. Two individuals who completed the baseline assessment after study *N* was reached were randomized by coin-flip, resulting in a total sample of 314 participants, with 160 participants in IG and 154 in CG. In IG, 24 participants (15.0%) and in CG 43 participants (27.9%) were lost to follow-up. Study completers and study dropouts did not differ in any sociodemographic or clinical baseline characteristics.

The majority of participants were male ($n = 165$, 52.5%), lived in partnership ($n = 281$, 89.5%), had middle education ($n = 197$, 62.7%), and were entrepreneurs ($n = 197$, 62.7%) with an average age of 52 ($SD = 9.98$, range: 25–87). Average depressive symptom severity was mild to moderate ($M = 10.03$, $SD = 4.26$), with 88 participants (28.0%) scoring ≥13 on the QIDS-SR16 (Table 1).

### Intervention-related outcomes

*Coaches and intervention use.* During the study, 34 coaches ($n = 31$ female, 91.2%) worked with 1–20 participants ($M =$

**Table 2.** Intervention use in the study period based on IVP Networks reporting ($N = 160$).

| Name Session characteristics[a] | Coaching participants ($N = 160$) | | | |
|---|---|---|---|---|
| | *M* | *SD* | Min | Max |
| **Sessions** | 13.43 | 6.02 | 1 | 32 |
| **Frequency of sessions (days)** | 14.09 | 5.87 | 0 | 42.25 |
| **Session duration (minutes)** | 48.57 | 14.69 | 1 | 120 |
| **Duration of coaching (months inscribed)** | 6.14 | 1.99 | 0.36 | 11.11 |
| **Volume of coaching (minutes)** | 652.52 | 325.76 | 50 | 1598 |
| **Coaching addition[a]** | **n** | **%** | | |
| Prolonging of coaching | 29 | 18.2 | | |
| Transfer to on-site-coaching | 0 | 0.0 | | |
| **Documented coaching topics (keywords)[a]** | n | % | | |
| Workload | 110 | 68.7 | | |
| Partnership conflict | 75 | 46.9 | | |
| Distinction work – private life | 56 | 35.0 | | |
| Conflict with children | 44 | 27.5 | | |
| Health problem | 43 | 26.9 | | |
| Generation conflict | 39 | 24.34 | | |
| Farm succession | 28 | 17.5 | | |
| Role conflict | 26 | 16.3 | | |
| Conflict with employees | 22 | 13.8 | | |
| Care of family member | 18 | 11.3 | | |
| Other | 43 | 26.9 | | |
| **Documented discharge reason** | n | % | | |
| Improved outcome (coach rating) | 96 | 59.9 | | |
| Further support recommended or installed | 47 | 29.4 | | |
| Withdrew consent | 7 | 4.4 | | |
| Lacked compliance | 5 | 3.2 | | |
| Agreed on different needs for support | 3 | 1.9 | | |
| Physically too impaired to participate | 1 | 0.6 | | |
| Coaching not started | 1 | 0.6 | | |

[a]Data referring to $n = 159$ participants who had at least one coaching session.

4.71, $SD = 2.26$). According to IVPNetworks, coaches had a background in systemic counseling ($n = 18$, 53.0%), cognitive behavioral methods ($n = 11$, 32.0%), hypnotherapy ($n = 8$, 24.0%), gestalt- or analytic therapy ($n = 7$, 20.5%), or other coaching and counseling methods ($n = 19$, 55.9%). Additional sociodemographic and work-related information were assessed via an online questionnaire from coaches (17/34, 50%; Supplement 1). Analyses of semistructured interviews with eight coaches revealed psychoeducation, conjoint goalsetting, and a three-phase model ("introduction and alliance building"-, "working"-, and "stabilizing"-phase) as common elements for the individualized coaching. Intervention use is displayed in Table 2.

*Satisfaction and working alliance.* Satisfaction in IG was higher ($n = 135$; $M = 28.17$, $SD = 5.58$) compared to CG ($n = 109$; $M = 16.11$, $SD = 6.46$, $t(214.64) = -15.39$, $p < 0.001$). Working alliance between coach and participant was rated as good by participants ($n = 135$, 84.9%, $M = 4.05$, $SD = 0.66$) and by the 17 coaches who rated a total of 86/159 (54.1%) coachings ($M = 3.94$, $SD = 0.68$).

## Primary outcome

Both groups showed statistically significant reductions in depressive symptom severity indicated by changes in baseline to post-treatment scores on the QIDS-SR16 (IG: 3.37 points, $SD = 4.21$, $t(3597) = -7.17$, $p < 0.001$, $d = -0.92$ [95% CI: −0.67 to −1.17]; CG: 2.10 points, $SD = 5.01$, $t(1793) = -3.86$, $p < 0.001$, $d = -0.50$ [95% CI: −0.25 to −0.77]). Depressive symptom severity was statistically significantly reduced in IG compared to CG while adjusting for baseline symptoms ($F_{1,306.54} = 12.78$, $p < 0.001$) corresponding to a small to medium between-group effect size of Cohen's $d = -0.39$ [95% CI: −0.15 to −0.64].

*Treatment response, symptom deterioration, and close-to-symptom-free status.* Treatment response was observed in more participants in IG (58/160 = 36.3%) than in CG (44/154 = 28.6%) but without statistical significance ($\chi^2 = 2.11$, $p = 0.16$; NNT = 13.02, 95% CI: −38.71 to 5.57). In IG, statistically significantly more participants achieved a close-to-symptom-free status (61/140 = 43.6%) compared to CG (37/133 = 27.8%; $\chi^2 = 7.35$, $p = 0.007$; NTT = 6.35, 95% CI: 3.72–21.70). A significant difference was observed in reliable deterioration with one case (0.6%) in IG and nine cases (5.8%) in CG ($\chi^2 = 6.4$, $p = 0.02$).

## Secondary outcomes

*Onset and remission of depression.* Based on CIDI ($n = 260$), 15 cases (6.6%) of possible MDE onset were observed in CG compared to six cases (2.7%) in IG (IRR = 0.39, 95% CI: 0.14–1.09, $p = 0.07$). Likewise, no group difference was observed in potential MDD cases based on QIDS-SR16 ($n = 226$) with eight (3.5%) and four (1.8%) new cases of potential MDD in CG and IG, respectively (IRR = 0.36, 95% CI: 0.10–1.36, $p = 0.13$).

No statistically significant difference between groups in remission (IG: $n_{baseline} = 40$, $n_{post} = 11$, 27.5% remission, CG: $n_{baseline} = 48$, $n_{post} = 22$, 45.8% remission; OR = 1.84, 95% CI: 0.66–5.42, $p = 0.27$) was observed.

*Other mental health outcomes.* Analyses yielded significant differences in favor of IG for perceived stress ($d = -0.34$, 95% CI: −0.08 to −0.59), generalized anxiety ($d = -0.30$,

**Table 3.** Primary and secondary outcomes at 6-month post-treatment ($N = 314$).

| Outcome | Group | Baseline assessment (T0) M | Baseline assessment (T0) SD | 6-month post-treatment (T1) M | 6-month post-treatment (T1) SD | *F* | *p* | Between-group effect size Cohen's d (95% CI) |
|---|---|---|---|---|---|---|---|---|
| **QIDS-SR16** | IG | 9.81 | 4.13 | 6.44 | 3.95 | 12.78 | <0.001 | −0.39 (−0.64 to −0.15) |
| | CG | 10.26 | 4.40 | 8.16 | 4.62 | | | |
| **PSS-10** | IG | 21.76 | 5.78 | 17.06 | 7.31 | 8.06 | 0.01 | −0.34 (−0.59 to −0.08) |
| | CG | 21.95 | 6.36 | 19.73 | 8.31 | | | |
| **ISI** | IG | 10.95 | 4.98 | 7.91 | 5.21 | 3.47 | 0.06 | −0.20 (−0.44 to 0.05) |
| | CG | 10.72 | 5.05 | 8.97 | 5.55 | | | |
| **GAD-7** | IG | 8.28 | 3.67 | 5.87 | 3.78 | 9.72 | <0.001 | −0.30 (−0.53 to −0.07) |
| | CG | 8.70 | 4.05 | 7.07 | 4.11 | | | |
| **PAS** | IG | 6.08 | 6.67 | 4.26 | 6.47 | 0.39 | 0.53 | −0.07 (−0.30 to 0.17) |
| | CG | 6.65 | 7.10 | 4.67 | 6.18 | | | |
| **AUDIT-C** | IG | 3.04 | 1.77 | 2.89 | 1.77 | 1.28 | 0.26 | −0.07 (−0.30 to 0.15) |
| | CG | 3.19 | 2.05 | 3.02 | 1.96 | | | |
| **SSS-8** | IG | 11.03 | 4.66 | 8.36 | 4.60 | 17.98 | <0.001 | −0.39 (−0.63 to −0.16) |
| | CG | 11.74 | 5.50 | 10.46 | 5.84 | | | |
| **MBI (EE)** | IG | 15.33 | 7.17 | 12.92 | 7.71 | 10.76 | <0.001 | −0.28 (−0.51 to −0.05) |
| | CG | 16.44 | 7.58 | 15.14 | 8.07 | | | |
| **MBY (CY)** | IG | 8.50 | 5.84 | 8.11 | 6.69 | 5.73 | 0.02 | −0.24 (−0.47 to 0.00) |
| | CG | 9.41 | 5.86 | 9.72 | 6.68 | | | |
| **MBI (PE)** | IG | 26.69 | 7.19 | 28.33 | 6.56 | 13.96 | <0.001 | 0.40 (0.16 to 0.64) |
| | CG | 25.12 | 7.58 | 25.34 | 8.00 | | | |
| **AQoL-8D (total)** | IG | 66.89 | 9.38 | 72.17 | 10.39 | 13.71 | <0.001 | 0.29 (0.06 to 0.52) |
| | CG | 66.39 | 10.90 | 68.73 | 12.77 | | | |
| **AQoL-8D (PSD)** | IG | 77.67 | 10.34 | 80.67 | 10.45 | 9.06 | <0.001 | 0.25 (0.01 to 0.48) |
| | CG | 77.67 | 10.81 | 77.9 | 11.77 | | | |
| **AQoL-8D (MSD)** | IG | 62.47 | 10.73 | 68.81 | 11.88 | 11.04 | <0.001 | 0.28 (0.05 to 0.51) |
| | CG | 61.76 | 12.34 | 65.16 | 14.21 | | | |
| **SPE[a] SPE = 0** | IG | 75 | 46.88 | 92 | 57.50 | | 0.51 | |
| SPE = 1 | | 36 | 22.50 | 32 | 20.00 | | | |
| SPE = 2 | | 40 | 25.00 | 28 | 17.50 | | | |
| SPE = 3 | | 9 | 5.63 | 8 | 5.00 | | | |
| **SPE = 0** | CG | 66 | 42.86 | 76 | 49.35 | | | |
| SPE = 1 | | 44 | 28.57 | 38 | 24.68 | | | |
| SPE = 2 | | 33 | 21.43 | 29 | 18.83 | | | |
| SPE = 3 | | 11 | 7.14 | 11 | 7.14 | | | |

IG: Intervention group; CG: Control group; QIDS-SR16: Quick Inventory of Depressive Symptomatology; PSS-10: Perceived Stress Scale; ISI: Insomnia Severity Index; GAD-7: Generalized Anxiety Disorder; PAS: Panic and Agoraphobia Scale; AUDIT-C: Alcohol Use Disorder Identification Test – Consumption Questions; SSS-8: Somatic Symptom Scale; MBI-GS: Maslach-Burnout-Inventory General Survey; EE: Emotional Exhaustion; CY: Cynism; PE: Professional Efficacy; AQoL-8D: Assessment of Quality of Life; PSD: Physical Super Dimension; MSD: Mental Super Dimension; SPE: Subjective Prognosis of Gainful Employment Scale. Data are imputed and based on ITT analyses.
[a]Reported are *N* and %. Tested with Fisher exact test.

95% CI: −0.07 to −0.53), somatic symptoms ($d = -0.39$, 95% CI: −0.16 to −0.63), dimensions of the Maslach burnout inventory (EE: $d = -0.28$, 95% CI: −0.05 to −0.51; PE: $d = 0.4$, 95% CI: 0.64–0.16; CY: $d = -0.24$, 95% CI: 0 to −0.47), and quality of life ($d = 0.29$, 95% CI: 0.52–0.06) (Table 3).

*Negative effects.* Based on the INEP, one-third of IG participants (47/135, 34.8%) reported at least one negative effect attributed to the intervention. In total, 82 negative effects were reported of which most were "intrapersonal changes" ($n = 33$, 40.2%, Supplement 2).

*Use of TAU.* Overall, more participants in IG ($n = 97$, 71.3%) reported use of at least one health service (Table 4) during the last three months compared to CG participants ($n = 63$, 56.8%, $\chi^2(1) = 5.06$, $p = 0.02$).

*Sensitivity analyses.* Analyses of the study completer ($n = 247$, 78.7%) showed that outcomes remained comparable

**Table 4.** Use of psychosocial and (mental) health services in the study population according to corresponding items of Tic-P (study-completer only, $n = 247$ at post-assessment).

| | | Assessment at study begin (T0) | | | Assessment at post-treatment (T1) | | |
|---|---|---|---|---|---|---|---|
| Service | Group | *n* | % | Difference between groups % (95% CI) | *n* | % | Difference between groups % (95% CI) |
| **Use of at least one health service** | IG | 39 | 24.4 | 2.30 (0.64 to 3.95) | 97 | 71.3 | 19.72 (15.31 to 24.12) |
| | CG | 34 | 22.1 | | 63 | 56.8 | |
| **General practitioner** | IG | 111 | 69.4 | 2.05 (0.48 to 3.62) | 87 | 64.0 | 12.62 (8.48 to 16.76) |
| | CG | 110 | 71.4 | | 57 | 51.4 | |
| **Psychotherapist** | IG | 4 | 2.5 | 1.85 (0.36 to 3.34) | 3 | 2.2 | 0.5 (−0.38 to 1.37) |
| | CG | 1 | 0.7 | | 3 | 2.7 | |
| **Specialist for psychiatry, neurology or psychosomatic medicine** | IG | 9 | 5.6 | 3.03 (1.13 to 4.93) | 5 | 3.7 | 0.83 (−0.30 to 1.96) |
| | CG | 4 | 2.6 | | 5 | 4.5 | |
| **Clinic for psychiatry/psychosomatic** | IG | 1 | 0.6 | 0.63 (−0.25 to 1.51) | 2 | 1.5 | 2.13 (0.33 to 3.93) |
| | CG | 0 | 0.0 | | 4 | 3.6 | |
| **Prescribed antidepressants** | IG | 17 | 10.6 | 7.38 (4.49 to 10.27) | 13 | 9.6 | 5.96 (3.00 to 8.91) |
| | CG | 5 | 3.3 | | 4 | 3.6 | |
| **SVLFG psychosocial health services** | IG | 2 | 1.3 | 2.65 (0.87 to 4.42) | 3 | 2.2 | 4.1 (1.63 to 6.57) |
| | CG | 6 | 4.0 | | 7 | 6.3 | |
| **Other psychosocial health services** | IG | 13 | 8.1 | 2.91 (0.81 to 5.01) | 8 | 5.9 | 5.83 (2.91 to 8.75) |
| | CG | 17 | 11.0 | | 13 | 11.7 | |

IG: Intervention group; CG: Control group; TIC-P: Treatment Inventory of Costs in Patients with psychiatric disorders; SVLFG: social insurance for farmers, foresters, and gardeners. Based on study completer answers (T0: $N_{IG} = 160$, $N_{CG} = 154$; T1: $N_{IG} = 136$, $N_{CG} = 111$).

with greater effect sizes compared to the main analyses (Supplement 3). Additionally, small effects of an increased reduction in insomnia ($d = -0.22$ [95% CI: −0.47 to 0.03]) and panic and agoraphobic symptoms ($d = -0.3$ [95% CI: −0.6 to 0.01]) were found in IG compared to CG.

Baseline depressive symptom severity and coaching duration were uncorrelated ($r = 0.12$, $p = 0.11$) but baseline symptom severity predicted depressive symptoms at post-treatment ($\beta = 0.51$, $t(111) = 7.35$, $p < 0.0001$, $R^2 = 0.29$ [95% CI: 0.17–0.42]; Supplement 3).

## Discussion

We reported the post-treatment results from a pragmatic randomized controlled trial comparing a personalized telephone coaching to a control group receiving brief psychoeducational material. Coaching participants showed a significantly greater reduction in depressive symptom severity ($d = -0.39$, 95% CI: −0.64 to −0.15) compared to CG. No significant differences in reliable symptom improvement and depression onset were observed, but coaching led to fewer reliable deteriorations compared to CG. Positive effects in favor of IG were also shown for stress, anxiety, somatic symptoms, burnout risk, and quality of life.

To our knowledge, this is the first trial for personalized preventive telephone coaching targeting depressive symptoms in farmers. The observed reduction in depressive symptoms is comparable to previous findings for indicated depression prevention with effect sizes reported at post-treatment as $d = 0.35$ (95% CI: −0.23 to −0.47) for face-to-face interventions,[49] $d = -0.39$ (95% CI: −0.25 to −0.53) for general online interventions,[60] and $d = -0.28$ (95% CI: −0.50 to −0.07) for guided online-interventions for farmers.[61] The average intervention duration of 13 sessions over six months in this study is higher compared to most online trainings (6–12 sessions, 3–12 weeks)[60] but similar to preventive health coaching (1–20 sessions, 1–24 weeks)[62] and face-to-face interventions in subthreshold (5–16 session)[10] and major depression (3–24 sessions, 3–36 weeks).[63] Studies on telephone coaching based on cognitive behavioral therapy or motivational interviewing for adults with MDD reported slightly higher effect sizes for symptom reduction with $d = -0.76$ (95% CI: −0.40 to −0.80) at four months[21] and $d = -0.45$ (95% CI: −0.14 to −0.75) at 12 months.[22]

Regarding the nonsignificant group differences in MDE onset, three potential reasons should be considered. First, overall onset rates at post-treatment were low in both groups (CG: 12.0%, IG: 4.4% based on CIDI). Onset rates for nonactive control groups are reported with incidence rates of 25%–30% in (indicated) prevention over 12 months.[10,60] As shown by Reins et al.[60] depression onset in subclinical control conditions occurred on average eight months after study enrolment. Thus, our follow-up period might have been too short to observe group differences in depression onset. Second, our control group was not completely inactive but

received psychoeducational material about stress, depression, and healthcare services. Yet, brief psychoeducational interventions can result in depressive symptom reduction.[64] Third, the study was powered to detect statistically significant differences in depressive symptom severity, which might not have been sufficient to detect differences in MDD onset.

The recent Covid-19 pandemic has increased the demand and use of telehealth services rapidly.[65–67] Although the increased exposure to telehealth options might have positively influenced the attitudes of practitioners an users,[68,69] it also revealed that telehealth is still considered a more niche intervention option in many countries resulting in a lack of training of heath care professionals and reimbursement opportunities.[70–72] In line with other (pre-)pandemic studies, our results show that telephone coaching can effectively promote mental health, not only as a necessity in a global health crisis but also in easing access for remotely living communities,[73] lowering socioeconomic costs,[74,75] and offering alternatives for people with restricted (broadband) internet access.[75–77] Implementing a structure for telehealth in routine care can further strengthen and enable the health care system to spontaneously and flexibly react to future large-scale health crisis.[67,71,72]

### *Limitations*

We neither applied a cutoff on self-reported depression nor conducted diagnostic interviews at baseline in order to mimic routine care. Thus, around a third of included individuals were potential MDD cases and about a quarter of participants reported lifetime MDE. The telephone coaching might therefore have functioned as recurrence/relapse prevention or (early) treatment. Future studies should clarify whether telephone coaching is equally effective in the prevention of first depression onset, prevention of recurrence/relapse, and early treatment, respectively.

Although we evaluated the coaching as a depression prevention intervention and included participants with at least mild depressive symptoms, we did not exclude participants with comorbid symptoms. Due to the personalized approach, coaches and participants might have agreed to focus on other symptoms. This might partly explain why only 36% ($n = 113$) of participants showed a treatment response, a finding similar to other studies on online indicated prevention[60] and preventive face-to-face interventions.[49] Treatment response might be especially critical for this target group in order to avoid an even greater reluctance to seek professional help.[78] Nonresponse can lead to (early) dropout.[79] However, early termination of telephone coaching was low in this study ($n = 17/160$, 10.6%), which might be due to the low-threshold telephone delivery[80] or because of the personalized approach.

Nevertheless, the highly personalized intervention and diverse coaches' backgrounds restrict the generalizability of the intervention effects. More research on preventive (telephone) coaching on mental health outcomes is needed, especially to determine effective methods and effects of personalization.[81]

Finally, the assumed advantage of telephone coaching making health care more accessible might have been limited due to study-specific inclusion criteria (e.g. internet access). The multistage inclusion procedure for study participation might have resulted in the inclusion of highly motivated participants.

## Conclusion

Personalized telephone coaching can effectively reduce depression symptoms and improve mental health in farmers. It could play an important role in intervening at an early stage of mental health problems and if needed facilitating access to further support (e.g. psychotherapy, farming specific offers) which is especially crucial in rural areas and populations with low help-seeking behavior. Long-term (cost)-effectiveness analyses and the evaluation of the implementation into routine care will gain more insight into the potential of telephone coaching in this population.

### Acknowledgements

The authors would like to thank Friederike Dietz and Annika Montag for programming the online surveys. Further, the authors thank Sarah Banellis, Merle Bloom, Hanna Böckeler, Albina Chafisuf, Doro Dressel, Johanna Finitzer and Tomris Ohloff for their engagement in enrolling and supporting participants throughout the study. Lukas Fuhrmann and Marvin Franke are thanked for carrying out the randomization and Mathias Harrer for statistical advice. The authors thank the IVPNetworks staff who provided details on their intervention and Madison Ehmann for proofreading the manuscript as well.

### Authors' contributions

Janika Thielecke: Conceptualization, Methodology, Project administration, Investigation, Data Curation, Formal analysis, Visualization, Writing – Original Draft; Claudia Buntrock: Supervision, Conceptualization, Methodology, Writing – Original Draft; Ingrid Titzler: Project administration, Conceptualization, Methodology, Writing – Review & Editing; Lina Braun: Resources, Writing – Review & Editing; Johanna Freund: Resources, Writing – Review & Editing; Harald Baumeister: Resources, Conceptualization, Funding acquisition, Writing – Review & Editing; Matthias Berking: Funding acquisition, Writing – Review & Editing; David Daniel Ebert: Funding acquisition, Conceptualization, Methodology, Writing – Review & Editing. All authors provided critical revision of the article and approved the final manuscript.

## Data availability

Access to the final pseudonomized trial dataset can be provided to fellow researchers upon request, depending on to be specified data security and data exchange regulation agreements.

## Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: DDE has served as a consultant to/on the scientific advisory boards of Sanofi, Novartis, Minddistrict, Lantern, Schoen Kliniken, Ideamed, and German health insurance companies (BARMER, Techniker Krankenkasse) and a number of federal chambers for psychotherapy. DDE and MB are stakeholders of the Institute for health training online (GET.ON/HelloBetter), which aims to implement scientific findings related to digital health interventions into routine care. MB is scientific advisor of mentalis GmbH, a provider for digital aftercare. HB reports having received consultancy fees and fees for lectures/workshops from chambers of psychotherapists and training institutes for psychotherapists in the e-mental-health context. IT reports having received fees for lectures/workshops in the e-mental-health context from training institutes for psychotherapists. She was the research and implementation project lead of the trial site Institute for health training online (GET.ON) for the European implementation research project ImpleMentAll (11/2017-03/2021) funded by the European Commission.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The German insurance company SVLFG provided a financial contribution to the Friedrich-Alexander Universität Erlangen-Nürnberg and Ulm University as expense allowance. SVLFG had no role in study design, data collection, analyses, interpretation or writing the manuscript and the decision to publish it.

## ORCID iDs

Janika Thielecke https://orcid.org/0000-0002-1310-3485
Ingrid Titzler https://orcid.org/0000-0002-8963-2324

## Supplemental material

Supplemental material for this article is available online.

## References

1. Gutiérrez-Rojas L, Porras-Segovia A, Dunne H, et al. Prevalence and correlates of major depressive disorder: a systematic review. *Braz J Psychiatry* 2020; 42: 657–672.
2. Lépine JP and Briley M. The increasing burden of depression. *Neuropsychiatr Dis Treat* 2011; 7: 3–7.
3. Ferrari AJ, Charlson FJ, Norman RE, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med* 2013; 11: 1001547.
4. Greenberg PE, Fournier AA, Sisitsky T, et al. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry* 2015; 76: 155–162.
5. Vos T, Haby MM, Barendregt JJ, et al. The burden of major depression avoidable by longer-term treatment strategies. *Arch Gen Psychiatry* 2004; 61: 1097–1103.
6. Mitchell AJ, Yadegarfar M, Gill J, et al. Case finding and screening clinical utility of the patient health questionnaire (PHQ-9 and PHQ-2) for depression in primary care: a diagnostic meta-analysis of 40 studies. *BJPsych Open* 2016; 2: 127–138.
7. Scott KM, de Jonge P, Stein DJ, et al. *Mental disorders around the world: facts and figures from the WHO world mental health surveys*. Cambridge, UK: Cambridge University Press, 2018. 1–384 p.
8. Andrews G, Issakidis C, Sanderson K, et al. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. *Br J Psychiatry* 2004; 184: 526–533.
9. Chisholm D, Sanderson K, Ayuso-Mateos JL, et al. Reducing the global burden of depression: population-level analysis of intervention cost-effectiveness in 14 world regions. *Br J Psychiatry* 2004; 184: 393–403.
10. Cuijpers P, Pineda BS, Quero S, et al. Psychological interventions to prevent the onset of depressive disorders: A meta-analysis of randomized controlled trials. Vol. 83, Clinical Psychology Review. Elsevier Inc.; 2021. p. 101955.
11. van Zoonen K, Buntrock C, Ebert DD, et al. Preventing the onset of major depressive disorder: a meta-analytic review of psychological interventions. *Int J Epidemiol* 2014; 43: 318–329.
12. Yazd SD, Wheeler SA and Zuo A. Key risk factors affecting farmers' mental health: A systematic review. Vol. 16, International Journal of Environmental Research and Public Health. MDPI AG; 2019.
13. Bundespsychotherapeutenkammer. Ein Jahr nach der Reform der Psychotherapie-Richtlinie. Wartezeiten 2018 [Internet]. 2019. Available from: https://www.bptk.de/publikationen/bptk-studie/
14. Judd F, Jackson H, Fraser C, et al. Understanding suicide in Australian farmers. *Soc Psychiatry Psychiatr Epidemiol* 2006; 41: 1–10.
15. Jackson H, Judd F, Komiti A, et al. Mental health problems in rural contexts: what are the barriers to seeking help from professional providers? *Aust Psychol* 2007; 42: 147–160.
16. Wrigley S, Jackson H, Judd F, et al. Role of stigma and attitudes toward help-seeking from a general practitioner for mental health problems in a rural town. *Aust N Z J Psychiatry* 2005; 39: 514–521.
17. Hutchison AJ and Breckon JD. A review of telephone coaching services for people with long-term conditions. *J Telemed Telecare* 2011; 17: 451–458.
18. Coughtrey AE and Pistrang N. The effectiveness of telephone-delivered psychological therapies for depression and anxiety: a systematic review. *J Telemed Telecare*. 2018; 24: 65–74.
19. Lin T, Heckman TG and Anderson T. The efficacy of synchronous teletherapy versus in-person therapy: a meta-analysis of randomized clinical trials. *Clin Psychol Sci Pract* 2021; 29; 167–178.
20. Krzyzaniak N, Greenwood H, Scott AM, et al. The effectiveness of telehealth versus face-to face interventions for anxiety disorders: a systematic review and meta-analysis. *J Telemed Telecare* 2021.

21. Lerner D, Adler DA, Rogers WH, et al. A randomized clinical trial of a telephone depression intervention to reduce employee presenteeism and absenteeism. *Psychiatr Serv* 2015; 66: 570–577.
22. Schröder S, Fleckenstein J and Wunderlich M. Effektivität eines telefonbasierten Coaching-Programms für Patienten mit einer depressiven Erkrankung. *Monit Versorgungsforsch* 2017; 10: 65–68.
23. Thielecke J, Buntrock C, Titzler I, et al. Clinical and cost-effectiveness of personalized tele-based coaching for farmers, foresters and gardeners to prevent depression: study protocol of an 18–month follow-up pragmatic randomized controlled trial (TEC-A). *Front Psychiatry* 2020; 11.
24. Freund J, Titzler I, Thielecke J, et al. Implementing internet- and tele-based interventions to prevent mental health disorders in farmers, foresters and gardeners (ImplementIT): study protocol for the multi-level evaluation of a nationwide project. *BMC Psychiatry* 2020; 20: 424.
25. Terhorst Y, Braun L, Titzler I, et al. Clinical and cost-effectiveness of a guided internet-based acceptance and commitment therapy to improve chronic pain–related disability in green professions (PACT-A): study protocol of a pragmatic randomised controlled trial. *BMJ Open* 2020; 10: e034271.
26. Braun L, Titzler I, Ebert DD, et al. Clinical and cost-effectiveness of guided internet-based interventions in the indicated prevention of depression in green professions (PROD-A): study protocol of a 36-month follow-up pragmatic randomized controlled trial. *BMC Psychiatry* 2019; 19: 278.
27. Bouwmans C, De Jong K, Timman R, et al. Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). *BMC Health Serv Res* 2013; 13: 127.
28. Hakkaart-van Roijen L, Straten A, Donker M, et al. Manual trimbos/iMTA questionnaire for costs associated with psychiatric illness (TIC-P). *Rotterdam, Netherlands: Erasmus University* 2002.
29. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry* 2003; 54: 573–583.
30. Tavakol M and Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011; 2: 53–55.
31. Auerbach RP, Mortier P, Bruffaerts R, et al. WHO World Mental Health Surveys International College student project: prevalence and distribution of mental disorders. *J Abnorm Psychol* 2018; 127: 623–638.
32. Lamoureux BE, Linardatos E, Fresco DM, et al. Using the QIDS-SR16 to identify Major depressive disorder in primary care medical patients. *Behav Ther* 2010;41: 423–431.
33. Reis D, Lehr D, Heber E, et al. The German version of the perceived stress scale (PSS-10): evaluation of dimensionality, validity, and measurement invariance with exploratory and confirmatory bifactor modeling. *Assessment* 2019;26: 1246–1259.
34. Bastien CH, Vallières A and Morin CM. Validation of the insomnia severity index as an outcome measure for insomnia research. *Sleep Med* 2001; 2: 297–307.
35. Gierk B, Kohlmann S, Kroenke K, et al. The Somatic Symptom Scale-8 (SSS-8): a brief measure of somatic symptom burden. *JAMA Intern Med* 2014;174: 399–407.
36. Löwe B, Decker O, Müller S, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Med Care* 2008: 46: 266–274.
37. Spitzer RL, Kroenke K, Williams JBW, et al. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;106: 1092–1097.
38. Bandelow B. Assessing the efficacy of treatments for panic disorder and agoraphobia. II. The panic and agoraphobia scale. *Int Clin Psychopharmacol* 1995;10: 83–93.
39. Bush K, Kivlahan DR, McDonell MB, et al. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. *Arch Intern Med* 1998;158: 1789–1795.
40. Richardson J, Iezzi A, Khan MA, et al. Validity and reliability of the assessment of quality of life (AQoL)-8D multi-attribute utility instrument. *Patient* 2014; 7: 85–96.
41. Mittag O and Raspe H. Eine kurze Skala zur Messung der subjektiven Prognose der Erwerbstätigkeit: Ergebnisse einer Untersuchung an 4279 Mitgliedern der gesetzlichen Arbeiterrentenversicherung zu Reliabilität (Guttman-Skalierung) und Validität der Skala. *Rehabilitation* 2003; 42: 169–174.
42. Cillien P, Fischbach A, Mörsdorf A, et al. Maslach Burnout Inventory-General Survey German Version 1.0 (MBI-GS-D V1.0.). 2006.
43. Maslach C, Jackson SE and Leiter MP. Maslach Burnout Inventory: Third edition. Evaluating stress: A book of resources. 1997.
44. Attkisson CC and Zwick R. The client satisfaction questionnaire. *Eval Program Plann* 1982; 5: 233–237.
45. Boß L, Lehr D, Schaub MP, et al. Efficacy of a web-based intervention with and without guidance for employees with risky drinking: results of a three-arm randomized controlled trial. *Addiction* 2018; 113: 635–646.
46. Schmidt J and Wittmann WW. Fragebogen zur Messung der Patientenzufriedenheit. In: E Brähler, J Schumacher and B Strauss (eds) *Diagnostische Verfahren in der Psychotherapie*. Göttingen, Germany: Hogrefe, 2002, 392–396.
47. Munder T, Wilmers F, Leonhart R, et al. Working alliance inventory-short revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clin Psychol Psychother* 2010;17(3):231–239.
48. Ladwig I, Rief W and Nestoriuc Y. What are the risks and side effects of psychotherapy? – development of an inventory for the assessment of negative effects of psychotherapy (INEP). *Verhaltenstherapie* 2014;24: 252–263.
49. Cuijpers P, Koole SL, Van Dijke A, et al. Psychotherapy for subclinical depression: meta-analysis. *Br J Psychiatry* 2014;17: 231–239.
50. Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *Br Med J* 2008; 337: a2390.
51. R Core Team. R: A language and environment for statistical computing (version 4.1.0) [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: http://www.r-project.org/
52. van Buuren S and Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. *J Stat Softw* 2011; 45.

53. Graham JW, Olchowski AE and Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev Sci* 2007; 8: 206–213.
54. Rubin DB. Multiple Imputation for Nonresponse in Surveys. Rubin DB, editor. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 1987. 258 p. (Wiley Series in Probability and Statistics).
55. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc* 1996; 91: 473–489.
56. Freeman PR, Hedges L V and Olkin I. *Statistical methods for meta-analysis*. New York, USA: Biometrics, 1986.
57. Jacobson NS and Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991;59: 12–19.
58. Bender R. Calculating confidence intervals for the number needed to treat. *Control Clin Trials* 2001; 22: 102–110.
59. Kraemer HC and Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. *Biol Psychiatry* 2006; 59: 990–996.
60. Reins JA, Buntrock C, Zimmermann J, et al. Efficacy and moderators of internet-based interventions in adults with subthreshold depression: an individual participant data meta-analysis of randomized controlled trials. *Psychother Psychosom* 2021; 90: 94–106.
61. Braun L, Titzler I, Terhorst Y, et al. Effectiveness of guided internet-based interventions in the indicated prevention of depression in green professions (PROD-A): results of a pragmatic randomized controlled trial. *J Affect Disord* 2021; 278: 658–671.
62. Dejonghe LAL, Becker J, Froboese I, et al. Long-term effectiveness of health coaching in rehabilitation and prevention: a systematic review. *Patient Educ Couns* 2017;100: 1643–1653.
63. Cuijpers P, Huibers M, Daniel Ebert D, et al. How much psychotherapy is needed to treat depression? A metaregression analysis. *J Affect Disord* 2013; 149: 1–13.
64. Donker T, Griffiths KM, Cuijpers P, et al. Psychoeducation for depression, anxiety and psychological distress: a meta-analysis. *BMC Med* 2009; 7.
65. Weiner JP, Bandeian S, Hatef E, et al. In-person and telehealth ambulatory contacts and costs in a large US insured cohort before and during the COVID-19 pandemic. *JAMA Netw Open* 2021; 4: 1–13.
66. Zhu D, Paige SR, Slone H, et al. Exploring telemental health practice before, during, and after the COVID-19 pandemic. *J Telemed Telecare* 2021:1357633X211025943.
67. Smith AC, Thomas E, Snoswell CL, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). *J Telemed Telecare* 2020; 26: 309–313.
68. Montoya MI, Kogan CS, Rebello TJ, et al. An international survey examining the impact of the COVID-19 pandemic on telehealth use among mental health professionals. *J Psychiatr Res* 2022; 148: 188–196.
69. Juan NVS, Shah P, Schlief M, et al. Service user experiences and views regarding telemental health during the COVID-19 pandemic: a co-produced framework analysis. *PLoS One* 2021; 16: 1–21.
70. Abraham A, Jithesh A, Doraiswamy S, et al. Telemental health use in the COVID-19 pandemic: a scoping review and evidence gap mapping. *Front Psychiatry* 2021; 12: 1–16.
71. Thomas EE, Haydon HM, Mehrotra A, et al. Building on the momentum: sustaining telehealth beyond COVID-19. *J Telemed Telecare* 2020; 28: 301–308.
72. Giacalone A, Marin L, Febbi M, et al. Ehealth, telehealth, and telemedicine in the management of the COVID-19 pandemic and beyond: lessons learned and future perspectives. *World J Clin Cases* 2022; 10: 2363–2368.
73. Myers CR. Using telehealth to remediate rural mental health and healthcare disparities. *Issues Ment Health Nurs* 2019; 40: 233–239.
74. Jennett PA, Affleck Hall L, Hailey D, et al. The socio-economic impact of telehealth: a systematic review. *J Telemed Telecare* 2003; 9: 311–320.
75. Mahtta D, Daher M, Lee MT, et al. Promise and perils of telehealth in the current era. *Curr Cardiol Rep* 2021; 23: 1–6.
76. Carneiro AC, de Pinho GS, Belo JV, et al. Outcomes of telemedicine care during the COVID-19 pandemic: experience from an intervention program designed for vulnerable population in Brazil. *J Telemed Telecare* 2022. 1357633X2210891.
77. Zhang D, Shi L, Han X, et al. Disparities in telehealth utilization during the COVID-19 pandemic: findings from a nationally representative survey in the United States. *J Telemed Telecare* 2021: 1357633X211051677.
78. Rozental A, Andersson G, Boettcher J, Ebert DD, Cuijpers P, Knaevelsrud C, Ljótsson B, Kaldo V, Titov N, & Carlbring P. Consensus statement on defining and measuring negative effects of Internet interventions. *Internet Interventions* 2014; 1: 12–19.
79. Delgadillo J, McMillan D, Lucock M, et al. Early changes, attrition, and dose-response in low intensity psychological interventions. *Br J Clin Psychol* 2014; 53: 114–130.
80. Mohr DC, Ho J, Duffecy J, et al. Effect of telephone-administered vs face-to-face cognitive behavioral therapy on adherence to therapy and depression outcomes among primary care patients: a randomized trial. *JAMA* 2012; 307: 2278–2285.
81. Hill B, Hons B, Richardson B, et al. Do we know how to design effective health coaching interventions: a systematic review of the state of the literature. *Am J Health Promot*. 2015; 29: 158–169.

**Supplement 1: Additional information on participating coaches**

*Table 1 Sociodemographic and occupational information of participating coaches (self-report, n=17)*

| Variable | | M | SD | Min | Max | n | % |
|---|---|---|---|---|---|---|---|
| **Age** | | 47.87 | 9.66 | 30 | 64 | | |
| **Gender** | Female | | | | | 15 | 88.2 |
| | Male | | | | | 2 | 11.8 |
| **Country of birth** | Germany | | | | | 16 | 94.1 |
| | Russia | | | | | 1 | 5.9 |
| **State** | Hamburg | | | | | 12 | 70.6 |
| | North Rhine-Westphalia | | | | | 2 | 11.8 |
| | Schleswig-Holstein | | | | | 2 | 11.8 |
| | Bavaria | | | | | 1 | 5.9 |
| **Subject of degree** [a] | Psychology | | | | | 17 | 100.0 |
| | Education science | | | | | 1 | 5.9 |
| **State-licenced psychotherapist** | No | | | | | 12 | 70.6 |
| | Yes | | | | | 3 | 17.6 |
| | In training | | | | | 2 | 11.8 |
| **Description of the trainings** [a] | Systemic therapy | | | | | 6 | 35.3 |
| | Hypnotherapy | | | | | 5 | 29.4 |
| | Conflict Counselling/ Moderation | | | | | 4 | 23.5 |
| | Supervision | | | | | 4 | 23.5 |
| | Psychodynamic therapy | | | | | 4 | 23.5 |
| | Traumatherapy/-counselling | | | | | 3 | 17.7 |
| | Gestalt therapy | | | | | 3 | 17.7 |
| | Behavioral therapy | | | | | 3 | 17.7 |
| | Educational guidance | | | | | 2 | 11.8 |
| | Other coaching or counseling training | | | | | 5 | 29.5 |
| **Duration of the trainings** [a] | Indicated in hours | 622.36 | 942.28 | 40 | 3500 | | |
| | Indicated in years | 2.75 | 0.83 | 2 | 4 | | |
| **Type of employment** | Permanent employee | | | | | 12 | 70.6 |
| | Freelancer | | | | | 5 | 29.4 |
| **Clinical experience** | 0-6 months | | | | | 1 | 5.9 |
| | 12-24 months | | | | | 2 | 11.8 |
| | 5-6 years | | | | | 2 | 11.8 |
| | 7-8 years | | | | | 3 | 17.6 |
| | 8-9 years | | | | | 1 | 5.9 |
| | 9-10 years | | | | | 1 | 5.9 |
| | More than 10 years | | | | | 7 | 41.2 |
| **Coaching experience at study begin** | 0-6 months | | | | | 4 | 23.5 |
| | 6-12 months | | | | | 1 | 5.9 |
| | 12-24 months | | | | | 1 | 5.9 |
| | 2-3 years | | | | | 1 | 5.9 |
| | 3-4 years | | | | | 2 | 11.8 |
| | 5-6 years | | | | | 1 | 5.9 |
| | 6-7 years | | | | | 2 | 11.8 |
| | 9-10 years | | | | | 1 | 5.9 |
| | More than 10 years | | | | | 4 | 23.5 |

[a] multiple answers possible

**Supplement 2: Negative effects of the personalized telephone coaching**

*Table 1 Negative effects of the intervention reported by participants (n=135) in the INEP*

| Item | N | % |
|---|---|---|
| **Intrapersonal change** | **33** | **23.9** |
| Longer periods of feeling bad | 13 | 9.6 |
| Depending on coach | 9 | 6.7 |
| Severity of suffering from past experiences/events | 5 | 3.7 |
| Neglect of hobbies and social contacts because of coaching | 5 | 3.7 |
| Worsening of symptoms | 2 | 1.5 |
| Experiencing new thinking and behavior patterns as harmful | 2 | 1.5 |
| Difficulties to make decisions alone | 2 | 1.5 |
| Less trusting others | 1 | 0.7 |
| Feeling more lonely | 0 | 0.0 |
| As a human being changed to the negative | 0 | 0.0 |
| Thoughts/plans to commit suicide for the first time | 0 | 0.0 |
| **Relationship** | **13** | **9.6** |
| Problem with partner's jealousy | 10 | 7.4 |
| More arguments/conflicts in relationship | 5 | 3.7 |
| **Stigma** | **11** | **8.1** |
| Financial worries | 10 | 7.4 |
| Worries about (potentially) increasing insurance fees | 2 | 1.5 |
| Fear of others discovering about the program usage | 2 | 1.5 |
| **Therapeutic malpractice of the coach** | **10** | **7.4** |
| Feeling of being forced to do exercises given by coach | 8 | 5.9 |
| Hurtful statements by coach | 3 | 2.2 |
| Feeling of data security not being ensured during the coaching | 1 | 0.7 |
| Feeling of being made fun of by the coach | 0 | 0.0 |
| **Friends and Family** | **2** | **1.5** |
| Worsened relationship with family | 2 | 1.5 |
| Worsened relationship with friends | 0 | 0.0 |

INEP=Inventory for Assessing Negative Effects of Psychotherapy. Data based on *n*=135 study completer. Only negative effects attributed directly to the intervention are reported.

**Supplement 3: Sensitivity Analyses**

*Table 1 Primary and secondary outcomes at post-treatment (study completer only, n= 168)*

| Outcome | Group | n | Baseline assessment (T0) | | 6m post-treatment (T1) | | F | p | Between-group effect size Cohen's d (95%-CI) |
|---|---|---|---|---|---|---|---|---|---|
| | | | M | SD | M | SD | | | |
| **QIDS-SR16** | IG | 85 | 9.11 | 3.74 | 5.71 | 3.17 | 34.47 | <0.001 | -0.74 (-1.02 - -0.45) |
| | CG | 83 | 10.19 | 4.26 | 8.71 | 4.35 | | | |
| **PSS-10** | IG | 85 | 21.69 | 5.46 | 16.98 | 6.50 | 12.2 | <0.001 | -0.44 (-0.73 - -0.14) |
| | CG | 83 | 21.70 | 6.20 | 20.23 | 7.97 | | | |
| **ISI** | IG | 85 | 10.27 | 4.77 | 8.05 | 5.20 | 4.28 | 0.04 | -0.24 (-0.54 - 0.06) |
| | CG | 83 | 10.89 | 5.64 | 9.30 | 5.36 | | | |
| **GAD-7** | IG | 85 | 8.15 | 3.38 | 6.02 | 3.44 | 8.49 | <0.001 | -0.37 (-0.67 - -0.07) |
| | CG | 83 | 8.92 | 4.24 | 7.41 | 3.96 | | | |
| **PAS** | IG | 85 | 5.18 | 5.21 | 3.19 | 4.91 | 5.46 | 0.02 | -0.3 (-0.6 - 0.01) |
| | CG | 83 | 6.96 | 7.55 | 4.84 | 6.16 | | | |
| **AUDIT-C** | IG | 85 | 3.01 | 1.59 | 2.86 | 1.53 | 0.62 | 0.43 | 0.07 (-0.23 - 0.37) |
| | CG | 83 | 2.87 | 1.83 | 2.75 | 1.74 | | | |
| **SSS-8** | IG | 85 | 10.45 | 4.23 | 8.04 | 4.13 | 26.34 | <0.001 | -0.58 (-0.87 - -0.29) |
| | CG | 83 | 12.24 | 5.76 | 10.94 | 5.46 | | | |
| **MBI (EE)** | IG | 85 | 14.81 | 6.49 | 12.34 | 7.14 | 21.74 | <0.001 | -0.48 (-0.78 - -0.19) |
| | CG | 83 | 16.46 | 7.39 | 15.95 | 7.47 | | | |
| **MBY (CY)** | IG | 85 | 8.36 | 5.55 | 7.84 | 6.08 | 9.03 | <0.001 | -0.36 (-0.66 - -0.06) |
| | CG | 83 | 9.45 | 6.08 | 10.22 | 6.98 | | | |
| **MBI (PE)** | IG | 85 | 27.86 | 5.59 | 27.92 | 6.67 | 8.48 | <0.001 | 0.35 (0.05 - 0.65) |
| | CG | 83 | 24.17 | 7.96 | 25.30 | 7.95 | | | |
| **AQoL-8D (total)** | IG | 85 | 67.64 | 7.59 | 72.99 | 7.86 | 33.41 | <0.001 | 0.51 (0.22 - 0.81) |
| | CG | 83 | 65.70 | 11.71 | 67.50 | 12.39 | | | |
| **AQoL-8D (PSD)** | IG | 85 | 78.45 | 8.78 | 81.61 | 9.54 | 27.79 | <0.001 | 0.45 (0.15 - 0.74) |
| | CG | 83 | 76.37 | 11.38 | 76.61 | 12.09 | | | |
| **AQoL-8D (MSD)** | IG | 85 | 63.21 | 9.12 | 69.46 | 9.21 | 23.85 | <0.001 | 0.48 (0.18 - 0.77) |
| | CG | 83 | 61.33 | 13.19 | 63.76 | 13.76 | | | |
| **SPE** [a] | | | | | | | | | |
| SPE=0 | IG | 85 | 40 | 47.1 | 52 | 61.2 | | 1.0 | |
| SPE=1 | | | 21 | 24.7 | 15 | 17.7 | | | |
| SPE=2 | | | 21 | 24.7 | 17 | 20.0 | | | |
| SPE=3 | | | 3 | 3.5 | 1 | 1.2 | | | |
| SPE=0 | CG | 83 | 32 | 38.6 | 36 | 43.4 | | | |
| SPE=1 | | | 22 | 26.5 | 18 | 21.7 | | | |
| SPE=2 | | | 22 | 26.5 | 20 | 24.1 | | | |
| SPE=3 | | | 7 | 8.4 | 9 | 10.8 | | | |

Data is based on study completer (n=168). IG=Intervention group; CG=Control group; QIDS-SR16=Quick Inventory of Depressive Symptomatology; PSS-10=Perceived Stress Scale; ISI=Insomnia Severity Index; GAD-7=Generalized Anxiety Disorder; PAS=Panic and Agoraphobia Scale; AUDIT-C=Alcohol Use Disorder Identification Test – Consumption Questions; SSS-8=Somatic Symptom Scale; MBI-GS=Maslach-Burnout-Inventory General Survey; EE=Emotional Exhaustion; CY=Cynism; PE=Professional Efficacy; AQoL-8D=Assessment of Quality of Life; PSD=Physical Super Dimension; MSD=Mental Super Dimension; SPE=Subjective Prognosis of Gainful Employment Scale. [a] Reported are N and %. Tested with Fishers exact test.

*Table 2 Linear regression models to assess influence of intervention characteristics on primary outcome (QIDS-SR16) in IG (n=160)*

| Model | Term | β | SE | t | df | p | Nagelkerk $R^2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | $R^2$ | lower 95%-CI | upper 95%-CI |
| **Model 1: symptom severity baseline** | Intercept | 1.40 | 0.73 | 1.92 | 125.39 | 0.06 | 0.29 | 0.17 | 0.42 |
| | QIDS-SR16 (T0) | 0.51 | 0.07 | 7.35 | 110.99 | <.001 | | | |
| **Model 2: symptom severity baseline + coaching durations in hours** | Intercept | 4.10 | 1.75 | 2.34 | 77.03 | 0.02 | 0.31 | 0.19 | 0.44 |
| | QIDS-SR16 (T0) | 0.24 | 0.19 | 1.28 | 48.19 | 0.21 | | | |
| | Duration (h) | -0.25 | 0.14 | -1.80 | 98.70 | 0.08 | | | |
| | QIDS-SR16 (T0) * duration (h) | 0.02 | 0.01 | 1.69 | 62.38 | 0.10 | | | |
| **Model 3: symptom severity baseline + months accompanied by coach** | Intercept | 3.37 | 2.58 | 1.30 | 82.96 | 0.20 | 0.31 | 0.19 | 0.43 |
| | QIDS-SR16 (T0) | 0.25 | 0.27 | 0.95 | 51.40 | 0.34 | | | |
| | duration (month) | -0.32 | 0.40 | -0.81 | 98.39 | 0.42 | | | |
| | QIDS-SR16 (T0) * duration (month) | 0.04 | 0.04 | 1.07 | 62.36 | 0.29 | | | |
| **Model 4: symptom severity baseline + session count** | Intercept | 3.70 | 2.11 | 1.76 | 66.30 | 0.08 | 0.31 | 0.19 | 0.43 |
| | QIDS-SR16 (T0) | 0.26 | 0.23 | 1.16 | 43.06 | 0.25 | | | |
| | Session count | -0.17 | 0.14 | -1.22 | 80.05 | 0.22 | | | |
| | QIDS-SR16 (T0) * session count | 0.02 | 0.01 | 1.25 | 52.23 | 0.22 | | | |
| **Model 5: symptom severity baseline + mean frequency of sessions** | Intercept | 0.91 | 2.31 | 0.40 | 110.80 | 0.69 | 0.30 | 0.18 | 0.43 |
| | QIDS-SR16 (T0) | 0.57 | 0.22 | 2.57 | 102.59 | 0.01 | | | |
| | Frequency | 0.03 | 0.15 | 0.19 | 111.76 | 0.85 | | | |
| | QIDS-SR16 (T0) * frequency | 0.00 | 0.02 | -0.23 | 90.42 | 0.82 | | | |

IG=Intervention group; QIDS-SR16=Quick Inventory of Depressive Symptomology; T0=baseline measurement. Reported are regression analyses based on intervention group data (ITT analyses, n=160).

*Table 3 Linear regression models to assess influence of coaches background on primary outcome (QIDS-SR16) in IG (n = 160)*

| Model | Term | β | SE | t | df | p | Nagelkerk $R^2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | $R^2$ | lower 95%-CI | upper 95%-CI |
| **Model 1: symptom severity baseline** | Intercept | 1.40 | 0.73 | 1.92 | 125.39 | 0.06 | 0.29 | 0.17 | 0.42 |
| | QIDS-SR16 (T0) | 0.51 | 0.07 | 7.35 | 110.99 | <.001 | | | |
| **Model 2: symptom severity baseline + cognitive-behavioral-trained** | Intercept | 1.64 | 0.90 | 1.83 | 109.18 | 0.07 | 0.29 | 0.17 | 0.42 |
| | QIDS-SR16 (T0) | 0.49 | 0.09 | 5.69 | 99.59 | 0.00 | | | |
| | cognitive-behavioral-trained | -0.82 | 1.55 | -0.53 | 136.16 | 0.60 | | | |
| | QIDS-SR16 (T0) * cognitive-behavioral-trained | 0.09 | 0.15 | 0.59 | 129.34 | 0.55 | | | |
| **Model 3: symptom severity baseline + systemic trained coach** | Intercept | 0.85 | 0.93 | 0.92 | 133.85 | 0.36 | 0.30 | 0.17 | 0.42 |
| | QIDS-SR16 (T0) | 0.56 | 0.08 | 6.62 | 129.41 | 0.00 | | | |
| | Systemic-trained | 1.37 | 1.54 | 0.89 | 110.42 | 0.37 | | | |
| | QIDS-SR16 (T0) * systemic-trained | -0.12 | 0.15 | -0.78 | 89.04 | 0.43 | | | |

QIDS-SR16=Quick Inventory of Depressive Symptomology; T0=baseline measurement. Reported are regression analyses based on intervention group data (ITT analyses, n=160)). Reference category is "other coaching or counselling training.

*Table 4 Overview over number of participants treated by coach with specific therapeutic background (n = 160)*

| Coach trained in … | Participants treated | |
|---|---|---|
| | N | % |
| **Syctemic counselling** | 74 | 46.3 |
| **Cognitive behavioral therapy** | 51 | 31.88 |
| **not specified coaching and counselling methods** | 29 | 18.13 |
| **Hypnotherapeuty** | 28 | 17.5 |
| **Analytic therapy** | 16 | 10.0 |
| **Gestalttherapy** | 12 | 7.5 |
| **Not specified additional methods** | 58 | 36.3 |

*Table 5 Linear regression models to assess influence of participant characteristic on primary outcome (QIDS-SR16) (N = 314)*

| Model | Term | β | SE | t | df | p | Nagelkerk $R^2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | $R^2$ | lower 95%-CI | upper 95%-CI |
| **Model 1: group** | Intercept | 8.16 | 0.39 | 21.02 | 160.82 | <.001 | 0.04 | 0.01 | 0.1 |
| | Group: IG | -1.72 | 0.54 | -3.18 | 166 | .001 | | | |
| **Model 2: group + treatment preference**[a] | Intercept | 7.96 | 0.58 | 13.81 | 169.72 | <.001 | 0.05 | 0.01 | 0.11 |
| | Group: IG | -1.73 | 0.78 | -2.22 | 193.21 | 0.03 | | | |
| | Preference for telephone coaching | 0.25 | 0.79 | 0.32 | 182.73 | 0.75 | | | |
| | Preference for information material | 1.21 | 1.51 | 0.8 | 207.89 | 0.42 | | | |
| | Group: IG * Preference for telephone coaching | -0.02 | 1.07 | -0.02 | 226.81 | 0.98 | | | |
| | Group: IG * Preference for information material | 0.33 | 2.06 | 0.16 | 227.19 | 0.87 | | | |
| **Model 3: group + psychotherapy experience**[b] | Intercept | 7.91 | 0.42 | 18.8 | 149.38 | <.001 | 0.06 | 0.02 | 0.13 |
| | Group: IG | -1.77 | 0.58 | -3.07 | 190.09 | <.001 | | | |
| | Experience with psychotherapy | 1.8 | 1.14 | 1.58 | 127.66 | 0.12 | | | |
| | Group: IG * Experience with psychotherapy | -0.22 | 1.43 | -0.15 | 179.76 | 0.88 | | | |

QIDS-SR16=Quick Inventory of Depressive Symptomology. Reported are regression analyses based on imputed data (ITT analyses, n=314)). [a]Reference group model 2: no preference. [b]Reference group model 3: no experience with psychotheray

# 5

## ARTICLE 3 — EFFECTIVENESS OF TELEPHONE COACHING AT 12 & 18 MONTHS (TEC-A)

Title:
Does telephone coaching improve farmers' mental health in the long term? Results of the 12- and 18-month follow-up of a pragmatic randomized controlled trial

Summary:
Farmers and related occupations are subject to a number of risk factors for depression, such as labor-intensive, financially challenging, and seasonal working conditions. They often live and work together with their family in remote and underserved areas and show low help-seeking behavior. Personalized telephone coaching could address the various stressors and needs of the target group and serve as a low-threshold intervention to improve mental health. Thus, the aim of this study was to evaluate the long-term effectiveness of a personalized telephone coaching in reducing depressive symptom severity and other mental health problems in farmers after 12 and 18 months.

Participating farmers with elevated depressive symptoms (Patient Health Questionnaire; PHQ-9 $\geq$ 5) were randomly allocated to receive a personalized telephone coaching (N = 160) or treatment as usual enhanced by psychoeducation and information about routine care (N = 154). The coaching was conducted by psychologists with training in coaching or therapeutic methods. It could be personalized in session count, frequency, duration (at most 850 minutes over 6 months with optional prolonging of 150 minutes /3 months), and topics were based on individual needs. A referral to on-site coaching or other social and health services (e.g. socioeconomic or family counselling for farmers), could be arranged or recommended if indicated. Depressive symptom severity (Quick Inventory of Depressive Symptomatology, QIDS-SR16) as well as other mental health and work-related outcomes were assessed at baseline and at follow-ups 12 and 18 months after baseline and group differences were analyzed using covariance analysis (ANCOVA) controlling for baseline severity.

Participants named on average 5 stressors (SD = 2.00, range: 1 – 11) in everyday life and coaches documented an average of 3 top-

ics per participants (SD = 1.39, range 1 – 7) which were focused on during the coaching. Group differences in favor of the intervention were observed in depressive symptom severity ($d_{12\ months}$ = -0.34 [95%–CI: -0.58, -0.10]; $d_{18\ months}$ = -0.31 [95%–CI: -0.56, -0.05]), perceived stress ($d_{12\ months}$ = -0.42 [95%–CI: -0.66, -0.17], $d_{18\ months}$ = -0.38 [95%–CI: -0.62, -0.13]), generalized anxiety ($d_{12\ months}$ = -0.28 [95%–CI: -0.53, -0.03]), somatic symptoms ($d_{12\ months}$ = -0.32 [95%–CI: -0.57, -0.07], $d_{18\ months}$ = -0.26 [95%–CI: -0.52, -0.01]), and quality of life ($d_{12\ months}$ = 0.30 [95%–CI: 0.06, 0.53]). For work-related characteristics, group differences were observed in favor of the intervention for emotional exhaustion ($d_{12\ months}$ = -0.29 [95%–CI: -0.55, -0.03]; $d_{18\ months}$ = -0.28 [95%–CI: -0.54, -0.03]) and cynicism ($d_{12\ months}$ = -0.33 [95%–CI: -0.57, -0.09]; $d_{18\ months}$ = -0.35 [95%–CI: -0.60, -0.09]). For depression onset, insomnia severity, symptoms of panic and agoraphobia and alcohol consumption no differences were observed at any time as well as for professional efficiency or self-reported prognosis of work (all p >.05).

The follow-up assessments coincided with the COVID-19 pandemic which could potentially influence the results, even though it was tried to account for that. Limitations of the study include that no upper cut-off on depressive symptom severity was applied and thus no distinction can be made between prevention and early treatment. This study is the first RCT to evaluate the effectiveness of a preventive telephone coaching focusing on the needs of the farming population in Germany. Results suggest long-term effectiveness in reducing depressive symptoms and promoting mental health up to a year after coaching, which could potentially reduce depression related burden of disease.

Manuscript in preparation:
Thielecke* J, Titzler* I (* shared first authorship), Braun L, Ebert DD, Freund J, Baumeister H, Berking M, and Buntrock C. Does telephone coaching improve farmers' mental health in the long term? Results of the 12- and 18-month follow-up of a pragmatic randomized controlled trial. In preparation

Contribution:
Janika Thielecke contributed to the study conceptualization, design and methods. She prepared and submitted the research proposal for the local ethics committee. She was responsible for the recruitment, study administration, data curation and analysis. She presented and discussed results with project partners and on international conferences. She drafted the methods and results parts in the original manuscript, revised them after feedback from her co-authors and gave impulses for introduction and discussion parts.

# Does telephone coaching improve farmers' mental health in the long term? Results of the 12- and 18-month follow-up of a pragmatic randomized controlled trial

Janika Thielecke*[1, 2], Ingrid Titzler*[1], Lina Braun[4], David Daniel Ebert[2], Johanna Freund[1, 2], Harald Baumeister[4], Matthias Berking[1], and Claudia Buntrock[1,3]

[1] Department of Clinical Psychology and Psychotherapy, Institute of Psychology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

[2] Professorship of Psychology and Digital Mental Health Care, Technical University of Munich, Munich, Germany

[3] Institute of Social Medicine and Health Systems Research, Faculty of Medicine, Otto von Guericke University Magdeburg, Magdeburg, Germany

[4] Department of Clinical Psychology and Psychotherapy, Institute of Psychology and Education, Ulm University, Ulm, Germany

**Word count:** 4836 / 3500

**Author Note**

* These authors contributed equally to this paper.

Correspondence concerning this Article should be addressed to Janika Thielecke and Ingrid Titzler, Friedrich-Alexander-University Erlangen-Nürnberg, Department of Clinical Psychology and Psychotherapy, Nägelsbachstr. 25a, 91052 Erlangen. E-Mail: janika.thielecke@fau.de and ingrid.titzler@fau.de

**Abstract**

Objectives: Farmers and related professions are at risk of depression due to their labor-intense, economically difficult, and seasonal working conditions, their residence in underserved areas, and their low help-seeking behavior. The aim of this study was to evaluate the long-term effectiveness of personalized telephone coaching in reducing depressive symptom severity and other mental health problems in farmers.

Methods: Participating farmers with elevated depressive symptoms (PHQ-9 ≥ 5) were randomly allocated to receive up to 6 months of personalized telephone coaching (N=160) or enhanced treatment-as-usual (TAU+) with psychoeducational via mail (N=154). The coaching was conducted by psychologists and personalized in terms of session count, frequency, duration, and topics based on individual needs. Depressive symptom severity and other mental health and work-related outcomes were assessed at baseline, after 12 and 18 months. Group differences were determined by covariance analysis controlling for baseline severity.

Results: Group differences were observed in favor of the telephone coaching compared to TAU+ in depressive symptom severity ($d_{12months}$=0.34, $d_{18months}$=0.31), perceived stress ($d_{12\ months}$=0.41, $d_{18\ months}$=0.37), and emotional exhaustion ($d_{12months}$=0.29, $d_{18\ months}$=0.28) and cynism ($d_{12\ months}$=0.32, $d_{18\ months}$=0.34) at both follow-ups. Group differences in anxiety (d=0.28) and quality of life (d=0.30) were only observed at 12 months assessment. No group differences were observed in depression onset, insomnia severity, symptoms of panic and agoraphobia, alcohol consumption, professional efficiency or self-reported prognosis of working capacity.

Conclusions: Personalized telephone coaching as a remote prevention offer for farmers is a promising approach to reduce depression and other mental health problems sustainably up to 18-months.

**Clinical trial registration:** DRKS00015655

**Introduction**

Major depressive disorder (MDD) is associated with functional impairment worldwide and has an average lifetime and 12-month prevalence estimated at 12.6% and 4.8%, respectively, in high-income countries, including Germany [1]. There is ample evidence for the effectiveness of pharmacotherapy and psychotherapy in the treatment of adult MDD, even though the reported higher acceptability of psychotherapy suggests a preference for psychological interventions [2]. Optimal evidence-based treatments with 100% coverage could reduce the disease burden of depression by only one-third [3], but in fact only one-fifth of those affected by MDD receive specialized treatment [1]. This strongly underlines the importance of prevention to further reduce the burden of MDD [4]. Meta-analyses in the field of prevention research demonstrated that psychological interventions can effectively reduce depressive symptom severity at post-treatment (g = 0.35 [95%-CI: 0.23, 0.47]) [5] and the risk of MDD onset one year after the preventive intervention by 19% (RR =0.81 [95% CI: 0.72, 0.91]) [6]. In particular, selective prevention focusing on individuals with high-risk for the onset of MDD, e.g. due to exposure to a combination of risk factors, or indicated prevention focusing on individuals with subthreshold symptoms has proven to be an effective approach [6].

Previous research reported that farming [7–10] and forestry [11] occupations are associated with increased risk for mental health problems including depression, in comparison to non-agricultural setting. Thus, special attention must be paid on the accumulation of risk factors, as individuals from agricultural occupations are exposed to multiple stressors in their work environment, such as pesticide exposure, stress, high work load, working with family, unpredictability in the workplace, an additional off-farm job, poor physical health, previous injury, agricultural policies, the treatment of farmers in society and in the media [12–14]. However, help-seeking behavior for mental health problems tends to be hesitant in rural communities, which is still strongly related to agrarian values, such as concern about stigma, gossip in the community, one's own stoicism, and view that seeking counseling services is a sign of weakness [15]. At the same time,

low mental health care utilization among farmers [7,8] combined with limited treatment availability in rural and remote areas [16], increases barriers to care.

Technological advances in recent decades have encouraged the adoption of digital services, which can bridge gaps in healthcare and make a meaningful contribution to service delivery in prevention and treatment in rural areas [17]. As one of few nationwide prevention initiatives targeting a farming population, the German Social Insurance for Agriculture, Forestry and Horticulture (SVLFG) began a nationwide role-out of digital prevention offers in 2017. The project evaluates and implements a tailored internet-based intervention (IBI) program and personalized telephone coaching (pTC) for depression prevention in policyholders (i.e. farmers, foresters, horticulturists) in several randomized controlled trials on short- and long-term clinical and cost-effectiveness [18–21] and in implementation studies [22,23]. Both technological approaches follow an indicated prevention approach and tailor the psychological interventions to the enhanced depressive symptoms and stress factors that increase the risk of depression onset in participants. pTC can address the wide range of stressors in the farming target group by initiating individual problem- and solution-oriented processes, reactivating and building resources, and using psychoeducational and psychological methods, for example from systemic counseling or cognitive behavioral therapy (CBT), enabling participants to deal with stress, acute problems, and general concerns [21]. Furthermore, it can address the diverse individual and work-related needs, using a telecommunication medium that is widely available and flexibly compatible with professional life.

There is already meta-analytic evidence that standardized evidence-based psychotherapies based on treatment manuals (e.g., CBT, Interpersonal Psychotherapy, Behavioral Activation) can be delivered over the telephone, showing positive effects in treatment of MDD compared to control conditions (TAU, enhanced TAU+, active treatment, waitlist-control, or no treatment) at post-treatment with Cohen's *d* = 0.26 (95%-CI: 0.14, 0.39) [24] and with standardized mean difference [SMD] = -.85 (95%-CI: -1.56, -0.15) [25], respectively, and with *d* ranging from 0.25-1.98 at short-term [26]. Furthermore, synchronous telephone therapy (CBT) resulted in similar or more favorable effects

on depressive symptoms or demonstrated to be non-inferior compared to in-person therapy at post-treatment and 6-month follow-up [27,28]. With regard to the treatment of subthreshold depression, a study on a highly-structured telephone CBT blended with an Employee Assistance Program (EAP) compared to standalone EAP in workers reported positive short-term effectiveness in depression symptom reduction, *d*=0.69 (95%-CI: 0.32, 1.05) [29].

However, for coaching or counselling administered via the telephone (e.g., work-focused CBT, motivational interviewing techniques), only two RCTs in mostly white-collar workers or privately insured individuals with MDD reported moderate effects in reducing depressive symptom severity in 4-month short-term (*d*=-0.60 [95%-CI: -0.39, -0.80) [30] and over 12-month long-term *(d*=-0.45 [95%-CI: -0.14, -0.75], limited by 70% dropout rate in intervention group ) [31].

However, there is a lack of research on whether pTC is effective in the treatment of subthreshold depressive symptoms in blue collar workers (e.g., farmers), leading to long-term positive effects (> 12-month) and preventing the development of MDD.

pTC for depression in the farming population was already shown to effectively reduce depressive symptom severity (*d*= -0.39 [95%-CI: -0.64; -0.15]) and further outcomes, such as stress, anxiety, somatic symptoms and burnout risk at post-treatment (6-month) compared to a control group receiving brief psychoeducational material (TAU+) [21]. The aim of this study is to evaluate the long-term effectiveness of pTC compared to TAU+ in reducing depressive symptom severity and multiple mental health problems in farmers with elevated depressive symptoms over a 12- and 18-month follow-up.

## Methods

### Study design

This two-armed pragmatic randomized controlled trial compared the effectiveness in depressive symptom reduction of a personalized telephone coaching to TAU+ (psychoeducational material). Primary outcome was depressive symptom reduction at post-treatment (6-month, t1) and is reported elsewhere [21]. This article focusses on depression and other secondary mental health outcomes at the 12- (t2) and 18-months follow-up (t3). The trial was approved by the Medical Ethics Committee of the Friedrich-Alexander University Erlangen-Nürnberg (No. 345_18 B) and registered in the German clinical trial register (DRKS00015655). Details can be found in the corresponding study protocol [32].

### Participants

Individuals were included if they were (a) insured by the SVLFG, (b) entrepreneur, assisting family member or pensioned farmer, (c) 18 years or older, (d) showed elevated depressive symptoms (PHQ-9≥5), (e) had internet/telephone access for online assessments/coaching, (f) provided written informed consent, (g) not currently receiving psychotherapy, (h) willing to sign a non-suicide contract in case of suicidal ideation and (i) not living in the German Federal States Bavaria or Schleswig-Holstein, as national pilot roll-out had begun there [33]. Additionally, we controlled that participants did not participate in the parallel trials evaluating a tailored IBI program which recruited in the same population [18–20]. No diagnostic interviews were conducted and no clinical exclusion criteria (e.g. co-morbidity, medication) were established. As described previously [21], female (*n*=149, 47.5%) and male participants (*n*=165, 52.5%) were relatively equally distributed with an average age of 52 years (*SD*=9.98). On average participants reported mild to moderate (*M*=10.03, *SD*=4.26) depressive symptom severity and 88 participants (28.0%) were potential MDD cases (QIDS-SR16≥13). Most participants were Caucasian (*n*=311, 99.0%), lived in partnership (*n*=281, 89.5%), had middle education (*n*=197, 62.7%), and were entrepreneurs (*n*=197, 62.7%).

### Procedure

Recruitment and participant enrollment took place from December 2018 to April 2019. Inclusion criteria were assessed via online-screening. Eligible individuals who provided written informed consent and

completed baseline assessment were randomized either to the intervention (IG) or control group (CG) (see Figure 1).

Permuted block randomization with randomly arranged block sizes (4, 6, 8, 12 via www.sealedenvelope.com) and an allocation ratio of 1:1 was conducted on an individual level by an independent researcher not involved in the study. Coaches and participants could not be blinded but data collectors were blind to group allocation.

Follow-ups after 12-month and 18-month were assessed form October 2019 to August 2020 and from June 2020 to February 2021, respectively. Completion of each follow-up was rewarded with €15 in both groups.

**Study arms**

Participants had unrestricted access to routine care. Actual health service use in both groups was monitored with items of the TiC-P (see intervention-related outcomes) [34].

***Intervention group***

Participants in the intervention group (IG) received a personalized telephone coaching delivered by an independent healthcare provider "IVPNetworks" over up to six months with the option of a three months prolonging after case-based decisions. The study team registered participants on the management platform (IVPnet). Case managers at IVPNetworks assigned participants to coaches. Coaches worked with 1-20 participants (*M*=4.71, *SD*=2.26) during the study time. All 34 coaches were psychologists with a diploma or master's degree in psychology and trained in different psychological methods including systemic counselling (*n*=18, 53.0%), cognitive behavioral methods (*n*=11, 32.0%), hypnotherapy (*n*=8, 24.0%), gestalt- or analytic therapy (*n*=7, 20.5%). Accordingly, no fixed procedures or manuals were applied in the coaching and methods varied depending on the coach's background. Licensed psychotherapists were available for supervision.

Use of the intervention was based on the documentation in the ivp.net and has been reported elsewhere [21]. Main characteristics are summarized as follows: Participants received on average *M*=13.43

(*SD*=6.02) coaching sessions, scheduled roughly biweekly (*M*=14.09, *SD*=5.87 days apart) of *M*=48.57 (*SD*=16.96) minutes each over a course of 6.14 (*SD*=1.99) months.

***Control group***

Participants in the control group (CG) received a pdf via e-mail with brief psychoeducation material on stress and depression combined with information about access to regular care once after randomization.

**Measurements**

All participant measurements were assessed via online-questionnaire (www.unipark.com). Depressive symptom severity was assessed via telephone by the study team or with a shorter version of the online-questionnaire if participants did not complete the full assessment within two months. Information on the intervention (e.g. session count, topics) was documented by coaches on the IVPnet.

***Mental Health Outcomes***

The German Version of the Quick Inventory Depressive Symptomology (QIDS-SR16) was used to assess depressive symptom severity with 16 items and a score range from 027. Scores can be interpreted categorically as none (0-5), mild (6-10), moderate (11-15), severe (16-20) and very severe symptoms (≥21) [35]. Reliability in this study was acceptable with Chronbach's α between .72 and .79.

Self-reported onset of a major depressive episodes (MDE) within the past year, was assessed using adapted items from the web version of the Composite International Diagnosis Interview version 3.0 [CIDI, 36]. Additionally, a sore of 13 or greater on the QIDS-SR16 was applied to identify possible acute MDD cases [37].

Further mental health related outcomes were measured using the Perceived Stress Scale (PSS, range 0-40, α=.88 - .89) [38], the Insomnia Severity Index (ISI, range 0-28, α=.88 - .89) [39], the Somatic Symptom Scale (SSS-8, range 0-32, α=.79 - .81) [40], the Generalizied Anxiety Disorder Scale-7 (GAD-7, range 0-21, α=.86 - .87) [41], the Panic and Agoraphobia Scale (PAS, range 0-48, α=0.91 - 0.92) [42], the consumption questions from Alcohol Use Disorder Identification Test (AUDIT-C, range 0-12, α=0.56 - 0.62) [43] and the Assessment of Quality of Life-8D (AQoL-8D, range 0-100, α=.94 - .95) [44].

***Work-related outcomes***

The Subjective Prognosis of Gainful Employment Scale (SPE, Guttman scale categories from 0=no threat to gainful employment to 3=very high threat to gainful employment , $rep_{T0-3}$=.95) [45] and the Maslach Burnout Inventory (MBI-GS, subscales: "Emotional Exhaustion" , EE, range 0-30; "Cynicism", CY, range 0-30; "Professional Efficacy", PE, range 0-36; α=.77 - .92) [46] were administered to assess impact on work related aspects.

***Intervention-related outcomes***

Intervention related outcomes were assessed at t1 and t2 since 43 participants (32.3%) had not finished their coaching at t1. Context-adapted versions of the Client Satisfaction Questionnaire for internet interventions (CSQ-I, range 8-32, $\alpha_{IG}$=.95 - .96) [47], the Working Alliance Inventory (WAI-SR, range 12-60, $\alpha_{T1/2}$=.94) [48] and the Inventory for the Assessment of Negative Effects of Psychotherapy (INEP) [49] were assed in IG. To monitor TAU use, the Trimbos Institute and Institute of Medical Technology Questionnaire for Costs Associated with Psychiatric Illness (TiC-P) [34] was adapted to the farming context.

***Baseline characteristics, stressors and coaching topics***

Participant characteristics were assessed at baseline and included detail information about the farm (e.g. produce, size), family situation (e.g. child and nursing care), health information (e.g. previous injury or accident [50]) and social surrounding (social status [51], discrimination [52], loneliness [53] and minority status) as they can pose as potential risk for depression.

Everyday stressors were assessed by presenting participants a list of 19 potential stressors (multiple choice) and open end questions for further stressors. Open answers were summarized to form new categories. Coaching topics were documented throughout the coaching on the IVPnet based on a list of 13.

**Adaptations in light of the Covid-19 pandemic**

Since the Covid-19 potential topics (multiple choice) and open answer fields which were summarized as "other topics" due to high heterogeneity.

Adaptations in light of pandemic began during t2 assessment in Germany in March 2020, we decided to add 10 self-developed questions into t3 assessment asking how the pandemic influenced everyday life

and health in the past 6 months (since t2 assessment) in order to approximate effects of the pandemic situation on our results. Questions were included in t3 assessment from September 2020 onwards. Participants who completed t3 assessment before that (*n*=65) were send a separate questionnaire and invited to answer the questions separately (53.8% or 35/65 responded).

Participants were asked to rate the effect the pandemic had on eight different areas of their life (e.g. family life, financial situation) on a 5-point scale from 1 "very negative effects" to 5 "very positive effects". Two items (5-point scale, range 1-5) assessed subjective burden because of the pandemic ("no psychological burden" to "very high psychological burden") and if mental health services were used because of strains resulting from the pandemic ("strongly disagree" to "strongly agree").

**Data Analysis**

The Consolidated Standards of Reporting Trials (CONSORT) 2010 Statement and its extension for reporting pragmatic trials [54] guided the reporting of the study. A needed sample of 312 was calculated with G*Power Version 3.1.9.2 to be able to detect an effect of d=0.35 [5] using an independent t-test with a power 80%, based on an alpha of 0.05 (two-sided) and an attrition rate of 20% [31].

Analyses were performed in R 4.1.0 based on intention-to-treat principles with a two-sided significance level of .05. Missing data was assumed to be missing at random and were imputed across all time points (t0-t3) using the R-package *mice* to apply Multiple Imputation by Chained Equations [55]. A total of 50 imputed datasets were generated. Analyses were run in each set and Rubin's Rule [56] was used to pool estimates. Intervention-related outcomes and health service use were not imputed. Outcome comparison between study groups were performed using analysis of covariance (ANCOVA) to adjust for baseline scores in continues outcomes. Between-group differences and as Cohen's d effect sizes with corresponding 95%-CIs according to Hedges and Olkin [57] are reported. Group differences in categorical outcomes were analyzed using $Chi^2$-test adjusted for multiple imputation [58].

***Treatment response, reliable deterioration & close-to-symptom-free status***

Treatment response in depressive symptom severity at an individual level is examined by a reliable change from baseline to t2 and t3 according to the reliable change index (RCI) of Jacobson and Truax [59].

Taking into account the reliability of the QIDSR-16, reliable improvement is thus defined as a decline by more than 1.96 standard units, equivalent to a point change of at least 6 on the QIDS-SR16. A corresponding increase in depression scores identified by the same method is indicated reliable deterioration.

A score of ≤ 5 on the QIDS-SR16 at follow-ups was defined as close-to-symptom-free status in individuals with at least mild depression symptoms (QIDS-SR16≥6) at baseline.

Group differences at follow-ups in numbers of cases with treatment response, deterioration and close-to-symptom-free status were assessed using $\chi^2$ tests. NNT (Numbers-needed-to-treat; with 95%-CI) were calculated as the inverse of the risk difference to estimate the number of participants needed for one more respective case [60].

***Onset and remission of potential MDD***

In the subsample without self-reported MDE/MDD according to CIDI or QIDS-SR16<13 at baseline, group differences in depression onset at follow-ups were assessed with Poisson regression estimating incidence rate ratios (IRR) with 95%-CIs using time between baseline and assessment as the offset.

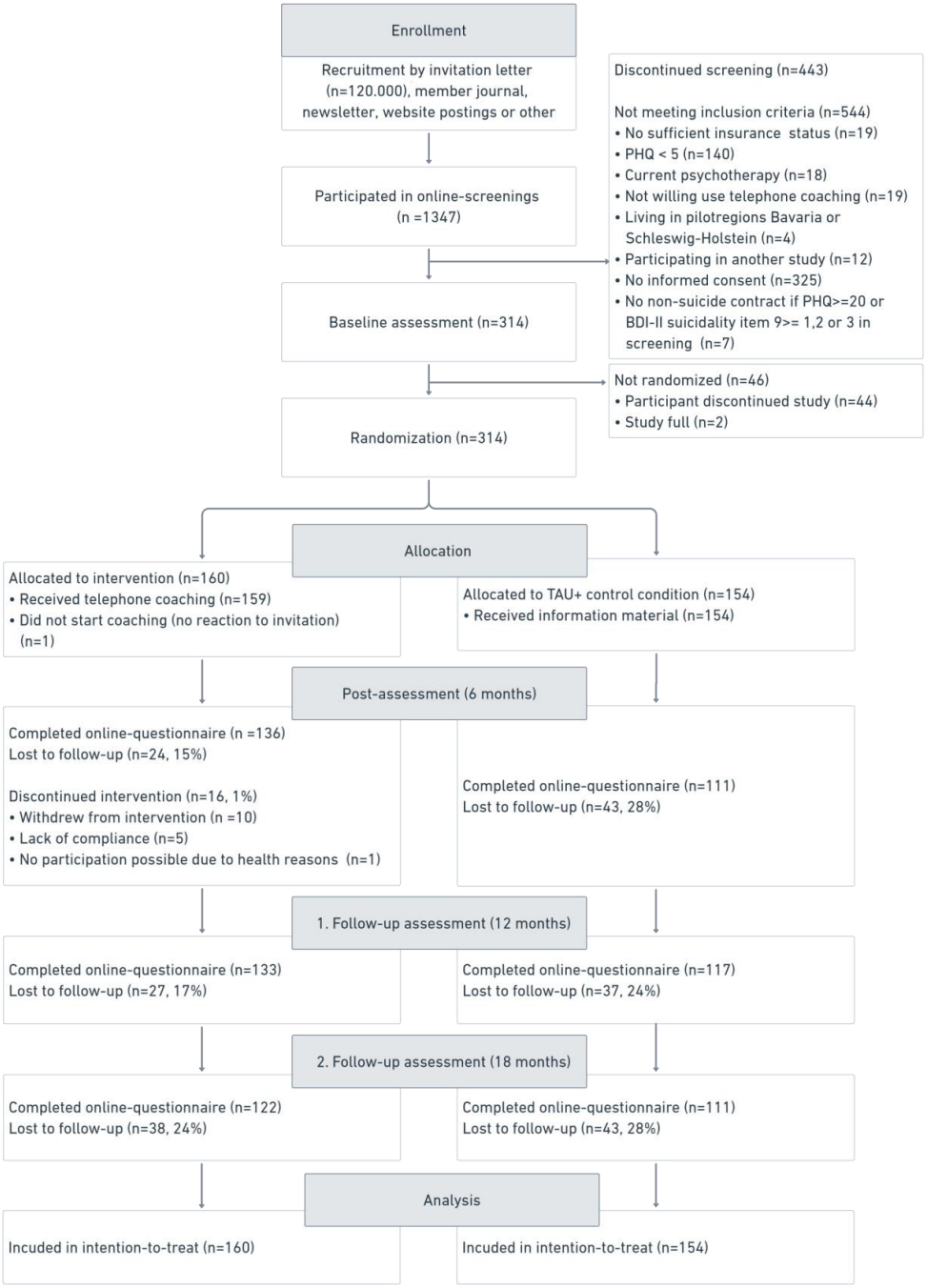
In the subsample of individuals with potential MDD at baseline (QIDS-SR16≥13) group differences in remission at t2 & t3 were assessed using logistic regression. For individuals with MDE according to CIDI at baseline, MDE remission was additionally compared at t2. Odds Ratios (OR) and 95%-CIs were reported.

***Sensitivity and moderation analyses***

Subgroup analyses with study completers (i.e. supplied at least primary outcome data) were conducted to test the robustness of the results. Possible influence of the intervention (i.e. coaching duration, session count) on depressive symptom severity at t2 and t3 was analyzed using linear regression. To estimate impact of the Covid-19 pandemic on our results, linear regression analyses were conducted controlling for baseline severity and including self-reported psychological burden of the pandemic as a potential moderator.

**Figure 1**

*CONSORT-Flowchart depicting study flow*



## Results

### Participants

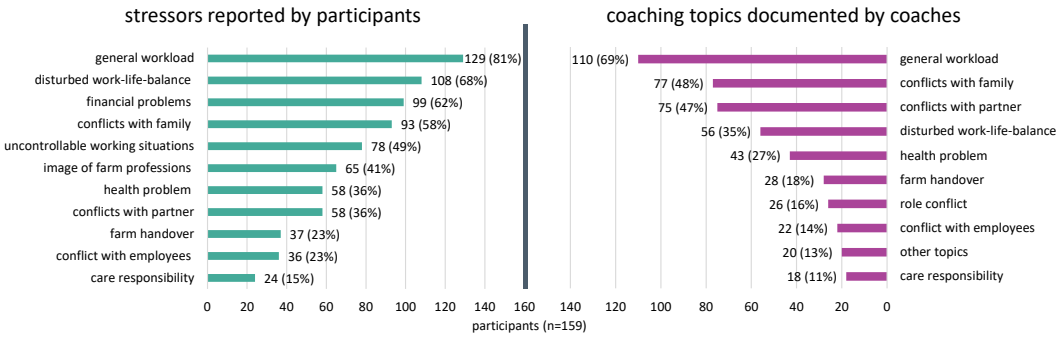
From 1347 individuals who completed the screening, 314 were randomized after baseline assessment. In CG, follow-up data could not be obtained from, 37 (24%) participants after 12 months and 43 (28%) participants after 18 months. In IG, 27 (17%) and 38 (24%) participants were lost to follow-up at t2 and t3, respectively (Figure 1). Study dropout was not associated with any baseline characteristics (p>.05).

### Stressors and coaching topics

Participants named on average 5 stressors in everyday life (*SD*=2.00, range: 1-11) with 100 participants (62.8%) naming ≥5 stressors. On averages, coaches documented 3 topics (*SD*=1.39, range 1-7) per participant (SD=1.39) and in 63.5% (n=101) cases ≥3 topics (Supplement 1). Stressors and coaching topics are compared in Figure 2. A detailed description of the farming situation and risk factors associated with the occupations are given in Table 1.

**Figure 2**

*Stressors reported by coaching participants (n = 159) at baseline and coaching topics as documented by coaches (n= 159)*




*Note.* Stressors reported by participants were assessed in baseline questionnaire and coaching topics documented by coaches on the IVPnet. Response formats (multiple choice) differed in questionnaire and IVPnet. For better comparison due to the different assessments, some categories of stressors (e.g. family conflicts differentiated by involved parties) and coaching topics were combined for presentation.

Participants' problems with under 10 nominations not portrait. One participant reported no problems and did not start the coaching.

***Table 1***

Baseline characteristics of study population (N=314) and description of risk factors associated with the farming occupations

| Characteristics | IG (N=160) | | CG (N=154) | | total (N=314) | |
|---|---|---|---|---|---|---|
| | n | % | n | % | n | % |
| **CLINICAL CHARACTERISTICS** | | | | | | |
| **Physical illness**[a] | | | | | | |
| No | 66 | 41.25 | 69 | 44.81 | 135 | 430 |
| Yes | 94 | 58.75 | 85 | 55.19 | 179 | 57.0 |
| **Suicide in family** | | | | | | |
| No | 135 | 84.3 | 138 | 89.6 | 273 | 86.9 |
| Yes | 21 | 13.1 | 13 | 8.4 | 34 | 10.8 |
| I don't know | 4 | 2.5 | 3 | 1.9 | 7 | 2.2 |
| **Own suicide attempts** | | | | | | |
| No | 156 | 97.5 | 152 | 98.7 | 308 | 98.1 |
| Yes, one time | 3 | 1.9 | 2 | 1.3 | 5 | 1.6 |
| Yes, multiple times | 1 | 0.6 | 0 | 0 | 1 | 0.3 |
| **Life threatening accident** | | | | | | |
| No | 124 | 77.5 | 134 | 87.0 | 258 | 82.2 |
| Yes | 36 | 22.5 | 20 | 13.0 | 56 | 17.8 |
| *If yes: age of accident ( M, SD)* | 27.83 | 15.30 | 26.65 | 17.41 | 27.41 | 15.96 |
| **Own serious illness or injury in past 6 months** | | | | | | |
| No | 148 | 92.5 | 139 | 90.26 | 287 | 91.4 |
| Yes | 12 | 7.5 | 15 | 9.74 | 27 | 8.6 |
| **Family characteristics** | | | | | | |
| **Members in farm household** | | | | | | |
| Entrepreneur | 131 | 89.1 | 129 | 89.0 | 260 | 89.0 |
| Family members | 119 | 81.0 | 111 | 76.6 | 230 | 78.8 |
| Pensioner (*altenteiler*) | 45 | 30.6 | 40 | 27.6 | 85 | 29.1 |
| Seasonal workers | 4 | 2.7 | 8 | 5.5 | 12 | 4.1 |
| Foreign workers | 6 | 4.1 | 4 | 2.8 | 10 | 3.4 |
| Trainees | 8 | 5.4 | 10 | 7.0 | 18 | 6.2 |
| Other | 3 | 2.0 | 1 | 0.7 | 4 | 1.4 |
| **Nursing care in farm household** | 38 | 25.9 | 36 | 24.8 | 74 | 25.3 |
| *If yes: who's responsible for care* | | | | | | |
| Me | 9 | 23.7 | 10 | 27.8 | 19 | 25.7 |
| Someone else | 13 | 34.2 | 15 | 41.7 | 28 | 37.8 |
| Me and someone else equally | 16 | 42.1 | 11 | 30.6 | 27 | 36.5 |

***Table 1 (continued)*** *Baseline characteristics of study population (N=314) and description of risk factors associated with the farming occupations*

| Characteristics | IG (N=160) | | CG (N=154) | | total (N=314) | |
|---|---|---|---|---|---|---|
| | n | % | n | % | n | % |
| **Children in farm household** | | | | | | |
| One or more children in the house | 96 | 60.0 | 96 | 62.3 | 192 | 61.2 |
| Children living with other parent | 1 | 0.6 | 4 | 2.6 | 5 | 1.6 |
| Children left home | 45 | 28.1 | 37 | 24.0 | 82 | 26.1 |
| No children | 17 | 10.6 | 15 | 9.7 | 32 | 10.2 |
| Other | 1 | 0.6 | 2 | 1.3 | 3 | 1.0 |
| *If children in household: who's responsible for child care* | | | | | | |
| Me | 60 | 41.7 | 50 | 35.7 | 110 | 38.7 |
| Someone else | 32 | 22.2 | 34 | 24.3 | 66 | 23.2 |
| Equally shared | 52 | 36.1 | 56 | 40.0 | 108 | 38.0 |
| **Everyday burden (range 0-10)[b]** | | | | | | |
| Family burden (*M, SD*) | 6.60 | 2.14 | 6.43 | 2.40 | 6.51 | 2.27 |
| Business burden (*M, SD*) | 7.10 | 1.93 | 7.10 | 2.10 | 7.10 | 2.01 |
| **Financal burden** | | | | | | |
| Business situation is financially difficult | 37 | 25.2 | 31 | 21.4 | 68 | 23.3 |
| Private situation is financially difficult | 3 | 2.0 | 6 | 4.1 | 9 | 3.1 |
| Both (business and private situation) are financially difficult | 43 | 29.3 | 54 | 37.2 | 97 | 33.2 |
| No financial difficulties | 51 | 34.7 | 44 | 30.3 | 95 | 32.5 |
| I can't/don't want to answer that | 13 | 8.8 | 10 | 6.9 | 23 | 7.9 |
| **FARM CHARACTERISTICS[c]** | | | | | | |
| **Participant's profession status[a]** | | | | | | |
| Entrepreneur | 94 | 58.8 | 103 | 66.9 | 197 | 62.7 |
| Entrepreneurs spouse | 35 | 21.9 | 25 | 16.2 | 60 | 19.1 |
| Family member of entrepreneur | 14 | 8.8 | 12 | 7.8 | 26 | 8.3 |
| Pensioner (*altenteiler*) | 16 | 10.0 | 13 | 8.4 | 29 | 9.2 |
| Incapacitated for work | 1 | 0.6 | 1 | 0.7 | 2 | 0.6 |
| **Second job off the farm[a]** | | | | | | |
| No | 42 | 85.7 | 33 | 89.2 | 75 | 87.2 |
| Yes | 7 | 14.3 | 4 | 10.8 | 11 | 12.8 |
| **Main branch of produce[a]** | | | | | | |
| Animal | 33 | 22.5 | 18 | 12.4 | 51 | 17.5 |
| Milk | 34 | 23.1 | 35 | 24.1 | 69 | 23.6 |
| Field | 27 | 18.4 | 30 | 20.7 | 57 | 19. |
| Fruit | 1 | 0.7 | 5 | 3.5 | 6 | 2.1 |
| Vegetables | 9 | 6.1 | 11 | 7.6 | 20 | 6.9 |
| Wine | 19 | 12.9 | 25 | 17.2 | 44 | 15.1 |
| Fish | 0 | 0 | 0 | 0 | 0 | 0 |

***Table 1 (continued)*** *Baseline characteristics of study population (N=314) and description of risk factors associated with the farming occupations*

| Characteristics | IG (N=160) | | CG (N=154) | | total (N=314) | |
|---|---|---|---|---|---|---|
| | n | % | n | % | n | % |
| Forest | 2 | 1.4 | 0 | 0 | 2 | 0.7 |
| Garden | 11 | 7.5 | 7 | 4.8 | 18 | 6.2 |
| Landscape | 1 | 0.7 | 1 | 0.7 | 2 | 0.7 |
| Gas | 1 | 0.7 | 3 | 2.1 | 4 | 1.4 |
| Service | 0 | 0 | 2 | 1.4 | 2 | 0.7 |
| Special | 2 | 1.4 | 0 | 0 | 2 | 0.7 |
| Other | 1 | 0.7 | 2 | 1.4 | 3 | 1.0 |
| Marketing | 3 | 2.0 | 3 | 2.1 | 6 | 2.1 |
| Horse caring | 2 | 1.4 | 2 | 1.4 | 4 | 1.4 |
| Multiple named | 1 | 0.7 | 1 | 0.7 | 2 | 0.7 |
| **Business size** | | | | | | |
| 1 - 4 persons | 78 | 53.1 | 77 | 53.1 | 155 | 53.1 |
| 5 - 9 persons | 43 | 29.3 | 48 | 33.1 | 91 | 31.2 |
| 10 - 24 persons | 23 | 15.7 | 13 | 9.0 | 36 | 12.3 |
| 25 - 49 persons | 1 | 0.7 | 4 | 2.8 | 5 | 1.7 |
| 50 - 99 persons | 2 | 1.4 | 2 | 1.4 | 4 | 1.4 |
| 100 - 249 persons | 0 | 0 | 1 | 0.7 | 1 | 0.3 |
| 250 or more persons | 0 | 0 | 0 | 0 | 0 | 0 |
| **Farm handover** | | | | | | |
| No handover planned | 82 | 55.8 | 81 | 55.9 | 163 | 55.8 |
| Handover in the past five years completed | 20 | 13.6 | 16 | 11.0 | 36 | 12.3 |
| Handover in the next five years and is settled | 14 | 9.5 | 7 | 4.9 | 21 | 7.2 |
| Handover in the next five years and is not settled | 27 | 18.4 | 34 | 23.4 | 61 | 20.9 |
| Farm will be abandoned in the next five years | 4 | 2.7 | 7 | 4.8 | 11 | 3.8 |
| **SOCIAL CHARACTERISTICS** | | | | | | |
| **Perceived discrimination/unfair treatment** [e] | | | | | | |
| No | 48 | 30 | 50 | 32.5 | 98 | 31.2 |
| Yes | 112 | 70 | 104 | 67.5 | 216 | 68.8 |
| *If yes: reason for discrimination* | | | | | | |
| Occupation | 29 | 18.1 | 31 | 20.1 | 60 | 19.1 |
| Family conflicts | 7 | 4.4 | 1 | 0.7 | 8 | 2.6 |
| Sex | 1 | 0.6 | 5 | 3.3 | 6 | 1.9 |
| Weight | 0 | 0 | 6 | 3.9 | 6 | 1.9 |
| Believes/religion | 3 | 1.9 | 2 | 1.3 | 5 | 1.6 |
| Age | 2 | 1.3 | 3 | 2.0 | 5 | 1.6 |
| Other reasons | 29 | 18.1 | 22 | 14.6 | 51 | 14.9 |

***Table 1 (continued)*** *Baseline characteristics of study population (N=314) and description of risk factors associated with the farming occupations*

| Characteristics | IG (N=160) | | CG (N=154) | | total (N=314) | |
|---|---|---|---|---|---|---|
| | n | % | n | % | n | % |
| **Perceived minority** [e] | | | | | | |
| No | 113 | 70.6 | 101 | 65.6 | 214 | 68.2 |
| Yes | 30 | 18.8 | 40 | 26.0 | 70 | 22.3 |
| I don't know | 17 | 10.6 | 13 | 8.4 | 30 | 9.6 |
| *If yes: perceived minority because of* | | | | | | |
| Farming occupation | 21 | 13.1 | 33 | 21.4 | 54 | 17.2 |
| Other reasons or not stated | 29 | 18.2 | 19 | 12.5 | 48 | 12.6 |
| **Social status (range 0-10)** (*M*, *SD*) [e] | 6.37 | 1.95 | 6.09 | 2.1 | 6.23 | 2.03 |
| **Loneliness (range 3-9)** (*M*, *SD*)[f] | 4.91 | 1.52 | 4.95 | 1.70 | 4.93 | 1.61 |

*Note.* Parts of the description have been previously presented elsewhere [21]. [a] previously reported elsewhere [21]; [b] everyday burden by family and professional situation rated on scale from 1 (no burden / very relaxed) to 10 (high burden /very tensed); [c] only answered by n=292 participants who claimed to be currently active on a farm; [d] reasons named less than five times are summarized under "other reasons"; [e] social status rated on scale 1 (lowest status) to 10 (highest status); [f] loneliness was assessed with the UCLA scale (range 3-9, never experiencing loneliness to often experiencing loneliness) [53]**.** *Abbreviations.* IG=Intervention group; CG=Control group; QIDS-SR16=Quick Inventory of Depressive Symptomatology; CIDI=Composite International Diagnostic Interview; MDE=Major Depressive Episode.

**Mental health outcomes**

An overview on all measurements and time points is given in Table 2.

***Depression***

Both, IG and CG, showed reduced depression symptom severity after 12 months (IG: 3.43 points, T(5654)=- -3.43, $p$<.001, $d$=-0.77 [95%-CI: -0.99, -.55]; CG: 2.47 points, T(6731)=-4.98, $p$<.001, $d$=-0.54 [95%-CI: -0.32, -0.75]) and 18 months (IG: 3.57 points, T(5238)=- -7.62, $p$<.001, $d$=-0.79 [95%-CI: -1.00, -0.59]; CG: 2.75 points, T(1672)=-5.03, $p$<.001, $d$=-0.58 [95%-CI: -0.35, -0.81]) on the QIDS-SR16. Scores of depressive symptom severity were statistically significantly more reduced in IG than in CG while accounting for baseline values (t2: F(1,942)=8.65, $p$=.003; t3: F(1,631)=6.42, $p$=.012). Corresponding effect sizes were small to medium ($d_{t2}$ = -0.34 [95%-CI: -0.58, -0.10]; $d_{t3}$= -0.31 [95%-CI:-0.56, -0.05]).

**Treatment response.** Cases of reliable improvement were similar in IG ($n_{t2}$=56, 36.4%, $n_{t3}$=52, 33.8%) and CG ($n_{t2}$=47, 30.5%, $n_{t3}$=50, 32.5%). In IG, three cases (1.9%) of deterioration were reported at t2 and two cases (1.3%) at t3 while in CG six (3.9%) and three (1.9%) cases were reported. No differences in the distribution of cases of reliable improvement, deterioration and unchanged scores was observed at any time point (12 months: $\chi^2$ (2, 2013)=0.84, $p$=.43; 18 months: $\chi^2$ (2, 759)=0.52, $p$ =.56). NNT for one addition case of reliable improvement therefore was 19 [95%-CI:-19.736, 6.393] at t2 and 204 [95%-CI:-10.00, 9.10] at t3.

**Onset and remission of depression.** There were no group differences in MDE or MDD onset or remission according to CIDI or QIDS-SR16 self-report observed at any time point (Table 3).

Close-to-symptom-free status. While more participants in IG (57/140, 40.7%) than in CG (40/133, 30.1%) were close-to-symptom-free at t2, the difference was not statistically significant ($\chi^2$(3,1310.4)=0.55, $p$=0.65, NNT=9.4 [95%-CI:4.58, 174.87]). Similarly, at t3 more participants in IG (62/140, 44.3%) than in CG (43/133, 32.3%) were close-to-symptom-free but without statistical significance ($\chi^2$(3,730.9)=0.7, $p$=0.55, NNT=8.36 [95%-CI: 4.28, 174.75]).

***Further Mental health Outcomes***

Group comparisons using ANOVAs further showed significant group differences in favor of IG for perceived stress ($d_{t2}$=-0.42 [95%-CI: -0.66, -0.17], $d_{t3}$=-0.38 [95%-CI: -0.62, -0.13]), generalized anxiety ($d_{t2}$=-0.28 [95%-CI: -0.53, -0.03]), somatic symptoms ($d_{t2}$=-0.32 [95%-CI: -0.57, -0.07], $d_{t3}$=-0.26 [95%-CI: -0.52, -0.01]), and quality of life ($d_{t2}$=0.30 [95%-CI: 0.06, 0.53]). For insomnia severity, symptoms of panic and agoraphobia and alcohol consumption no differences were observed at any time ($p$>.05).

**Work-related outcomes**

IG participants reported statistically significantly less emotional exhaustion ($d_{t2}$-0.29 [95%-CI: -0.55, -0.03]; $d_{t3}$=-0.28 [95%-CI: -0.54, -0.03]) and cynicism ($d_{t2}$=-0.33 [95%-CI: -0.57, -0.09]; $d_{t3}$=-0.35 [95%-CI: -0.60, -0.09]) on the respective MBI subscales. No differences in professional efficiency or self-reported prognosis of work capacity was observed ($p$>.05).

**Table 2**

*Group comparisons of continuous outcomes at baseline and follow-ups based on multiple imputation (N=314)*

| Outcome | t | IG | | CG | | F | df | p | Between-group effect size Cohen's d [95% CI] | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | M | SD | M | SD | | | | | |
| **QIDS-SR16** | t0 | 9.81 | 4.13 | 10.26 | 4.4 | | | | | |
| | t2 | 6.38 | 4.24 | 7.79 | 3.92 | 8.65 | 942 | 0.003 | -0.34 | [-0.58, -0.10] |
| | t3 | 6.24 | 3.84 | 7.51 | 4.34 | 6.42 | 631 | 0.012 | -0.31 | [-0.56, -0.05] |
| **PSS-10** | t0 | 21.76 | 5.78 | 21.95 | 6.36 | | | | | |
| | t2 | 15.86 | 7.15 | 18.87 | 6.95 | 12.03 | 649 | 0.001 | -0.42 | [-0.66, -0.17] |
| | t3 | 16.02 | 7.2 | 18.69 | 6.79 | 9.73 | 691 | 0.002 | -0.38 | [-0.62, -0.13] |
| **ISI** | t0 | 10.95 | 4.98 | 10.72 | 5.05 | | | | | |
| | t2 | 7.6 | 5.51 | 8.63 | 5.51 | 2.47 | 556 | 0.117 | -0.19 | [-0.44, 0.07] |
| | t3 | 8.03 | 5.14 | 8.31 | 5.64 | 0.32 | 1421 | 0.574 | -0.05 | [-0.31, 0.21] |
| **GAD-7** | t0 | 8.28 | 3.67 | 8.7 | 4.05 | | | | | |
| | t2 | 5.98 | 3.9 | 7.14 | 4.29 | 5.88 | 662 | 0.016 | -0.28 | [-0.53, -0.03] |
| | t3 | 5.82 | 3.81 | 6.66 | 4.51 | 2.90 | 783 | 0.089 | -0.20 | [-0.45, 0.05] |
| **PAS** | t0 | 6.08 | 6.67 | 6.65 | 7.1 | | | | | |
| | t2 | 4.4 | 6.91 | 4.81 | 6.48 | 0.34 | 1442 | 0.559 | -0.06 | [-0.31, 0.19] |
| | t3 | 4.11 | 6.21 | 4.02 | 5.64 | 0.23 | 4729 | 0.634 | 0.02 | [-0.24, 0.27] |
| **AUDIT-C** | t0 | 3.04 | 1.77 | 3.19 | 2.05 | | | | | |
| | t2 | 2.76 | 1.79 | 2.77 | 1.9 | 0.18 | 3012 | 0.670 | -0.01 | [-0.25, 0.24] |
| | t3 | 2.9 | 1.84 | 2.82 | 1.92 | 0.44 | 1131 | 0.505 | 0.04 | [-0.22, 0.30] |

**Table 2 (continued)**

*Group comparisons of continuous outcomes at baseline and follow-ups based on multiple imputation (N=314)*

| Outcome | t | IG | | CG | | F | df | p | Between-group effect size Cohen's d [95% CI] | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | M | SD | M | SD | | | | | |
| **SSS-8** | t0 | 11.03 | 4.66 | 11.74 | 5.5 | | | | | |
| | t2 | 7.85 | 5.04 | 9.58 | 5.63 | 7.72 | 535 | 0.006 | -0.32 | [-0.57, -0.07] |
| | t3 | 8.28 | 5.15 | 9.76 | 6.14 | 4.84 | 573 | 0.028 | -0.26 | [-0.52, -0.01] |
| **MBI (CY)** | t0 | 8.5 | 5.84 | 9.41 | 5.86 | | | | | |
| | t2 | 6.9 | 5.9 | 8.9 | 6.21 | 8.18 | 854 | 0.004 | -0.33 | [-0.57 - -0.09] |
| | t3 | 6.9 | 5.94 | 9.1 | 6.6 | 7.95 | 629 | 0.005 | -0.35 | [-0.60 - -0.09] |
| **MBI (EE)** | t0 | 15.33 | 7.17 | 16.44 | 7.58 | | | | | |
| | t2 | 11.57 | 7.56 | 13.9 | 8.28 | 6.02 | 433 | 0.015 | -0.29 | [-0.55, -0.03] |
| | t3 | 11.79 | 7.69 | 14.07 | 8.25 | 5.49 | 595 | 0.019 | -0.28 | [-0.54, -0.03] |
| **MBI (PE)** | t0 | 26.69 | 7.19 | 25.12 | 7.58 | | | | | |
| | t2 | 28.58 | 7.02 | 27.43 | 7.17 | 1.96 | 1066 | 0.162 | 0.16 | [-0.08, 0.41] |
| | t3 | 28.51 | 6.72 | 26.76 | 8.32 | 3.27 | 534 | 0.071 | 0.23 | [-0.03, 0.49] |
| **AQoL-8D (total)** | t0 | 66.89 | 9.38 | 66.39 | 10.90 | | | | | |
| | t2 | 73.79 | 9.89 | 70.61 | 11.50 | 8.08 | 909 | 0.005 | 0.30 | [0.06, 0.53] |
| | t3 | 72.87 | 10.32 | 70.84 | 11.88 | 2.56 | 741 | 0.110 | 0.18 | [-0.07, 0.43] |
| **AQoL-8D (MSD)** | t0 | 62.47 | 10.73 | 61.76 | 12.34 | | | | | |
| | t2 | 70.4 | 11.68 | 66.91 | 12.91 | 5.92 | 393 | 0.015 | 0.28 | [0.03, 0.54] |
| | t3 | 69.81 | 11.95 | 67.4 | 13.67 | 2.66 | 798 | 0.103 | 0.19 | [-0.06, 0.43] |
| **AQoL-8D (PSD)** | t0 | 77.67 | 10.34 | 77.67 | 10.81 | | | | | |
| | t2 | 81.87 | 10.24 | 78.84 | 11.51 | 6.45 | 779 | 0.011 | 0.28 | [0.03, 0.52] |
| | t3 | 80.31 | 10.77 | 78.76 | 11.57 | 1.39 | 524 | 0.239 | 0.14 | [-0.12, 0.40] |
| **Outcome** | | **n** | **%** | **n** | **%** | **$X^2$** | **df** | **p** | | |
| **SPE** | | | | | | | | | | |
| SPE=0 | t0 | 75 | 46.88 | 66 | 42.86 | | | | | |
| SPE=1 | | 36 | 22.50 | 44 | 28.57 | | | | | |
| SPE=2 | | 40 | 25.00 | 33 | 21.43 | | | | | |
| SPE=3 | | 9 | 5.63 | 11 | 7.14 | | | | | |
| SPE=0 | T2 | 97 | 60.6 | 80 | 52.0 | | | | | |
| SPE=1 | | 26 | 16.3 | 33 | 21.4 | 0.95 | 847 | 0.418 | | |
| SPE=2 | | 29 | 18.1 | 29 | 18.8 | | | | | |
| SPE=3 | | 8 | 5.0 | 12 | 7.8 | | | | | |
| SPE=0 | T3 | 92 | 57.5 | 76 | 49.4 | | | | | |
| SPE=1 | | 32 | 20.0 | 33 | 21.4 | 0.946 | 451 | 0.418 | | |
| SPE=2 | | 24 | 15.0 | 36 | 21.4 | | | | | |
| SPE=3 | | 13 | 8.1 | 9 | 5.8 | | | | | |

*Note.* Data are imputed and based on intention-to-treat analyses. *Abbreviations.* IG: Intervention group (telephone coaching); CG: Control group (enhanced treatment as usual); t0: Baseline assessment; t1: post-treatment (6-month); t2: 1. Follow-up (12-month); t3: 2. Follow-up (18-month); QIDS-SR16: Quick Inventory of Depressive Symptomatology; PSS-10: Perceived Stress Scale; ISI: Insomnia Severity Index; GAD-7: Generalized Anxiety Disorder; PAS: Panic and Agoraphobia Scale; AUDIT-C: Alcohol Use Disorder Identification Test – Consumption Questions; SSS-8: Somatic Symptom Scale; MBI-GS: Maslach-Burnout-Inventory General Survey; EE: Emotional Exhaustion; CY: Cynism; PE: Professional Efficacy; AQoL-8D: Assessment of Quality of Life; PSD: Physical Super Dimension; MSD: Mental Super Dimension; SPE=Subjective Prognosis of Gainful Employment Scale (0= no threat to gainful employment to 3=very high threat to gainful employment ).

**Table 3**

*Depression status at baseline and overview on onset and remission according to CIDI and QIDS-SR16 at follow-ups (N=314)*

| Onset | Baseline | | 1. Follow-Up (12 months) | | | | | | 2. Follow-Up (18 months) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | % | n onset | % | T | df | p | IRR [95%-CI] | n onset | % | T | df | p | IRR [95%-CI] |
| *Potential MDE (CIDI)*[a] | | | | | | | | | | | | | | |
| **CG** | 125 | 81.2 | 22 | 17.6 | | 638.19 | 0.39 | 0.70 [0.31, 1.59] | 22 | 17.6 | -0.79 | 427.59 | 0.43 | 0.71 [0.31, 1.65] |
| **IG** | 135 | 84.4 | 14 | 10.4 | | | | | 15 | 11.1 | | | | |
| *Potential MDD (QIDS-SR16)* | | | | | | | | | | | | | | |
| **CG** | 106 | 68.8 | 4 | 3.8 | -0.11 | 923.97 | 0.91 | 0.91 [0.18, 4.62] | 8 | 7.5 | -1.12 | 548.99 | 0.27 | 0.41 [-0.09, 1.96] |
| **IG** | 120 | 75.0 | 4 | 3.3 | | | | | 4 | 3.3 | | | | |
| **Remission** | | | | | | | | | | | | | | |
| | n | % | n remissions | % | T | df | p | OR [95%-CI] | n remissions | % | T | df | p | OR [95%-CI] |
| *Potential MDE (CIDI)*[a] | | | | | | | | | | | | | | |
| **CG** | 29 | 18.8 | 14 | 48.3 | -0.37 | 3257.43 | 0.71 | 0.81 [0.25, 2.54] | 18 | 62.1 | -0.13 | 1564.77 | 0.90 | 0.92 [0.27, 3.17] |
| **IG** | 25 | 15.6 | 14 | 56.0 | | | | | 16 | 64.0 | | | | |
| *Potential MDD (QIDS-SR16)* | | | | | | | | | | | | | | |
| **CG** | 48 | 31.2 | 36 | 75.0 | 0.49 | 2258.73 | 0.62 | 1.32 [0.44, 4.01] | 34 | 70.8 | 1.74 | 1596.92 | 0.08 | 3.12 [0.87, 11.19] |
| **IG** | 40 | 25.0 | 32 | 80.0 | | | | | 35 | 87.5 | | | | |

*Note.* Analysis based on imputed data. *Abbreviations.* IG: Intervention group (telephone coaching); CG: Control group (enhanced treatment as usual); QIDS-SR16: Quick Inventory of Depressive Symptomatology; CIDI=Composite International Diagnostic Interview; MDE:Major Depressive Episode; MDD: Major Depressive Dissorder; OR: Odds Ratio; IRR: Incidence Risk Ratio. MDE remission according to CIDI was not calculated at post-assessment due to overlap in assessment point; [a]Controlled for MDE status at post-treatment.

**Intervention related outcomes**

***Satisfaction & Working alliance***

Coaching participants (n=126) reported a satisfaction rating based on CSQ-I scores of M=27.39 (SD=6.06). The working alliance was rated with M=3.98 (SD=0.71) by participants 12 months after randomization.

***Negative effects***

Using the INEP, nearly one-third of IG participants (35/126, 27.8%) reported at least one negative effect. In total, 64 negative effects were attributed to the intervention with "intrapersonal changes" (*n*=19, 15.1%) and "therapeutic malpractice" (*n*=12, 9.5%) named most often (Supplement 2).

***Use of TAU***

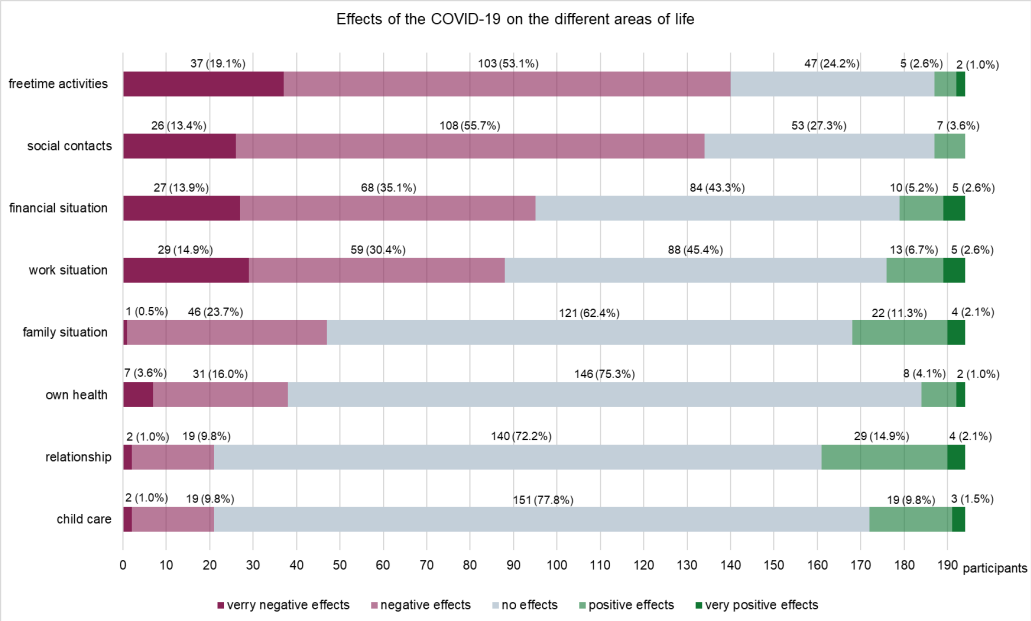
Ues of TAU (Supplement 3) was similar ($\chi^2(1)=1.62$, $p=0.203$) across groups with 82 (71.8%) participants in CG and 88 (63.9%) in IG reporting to have used at least one further health service in the three month leading to t2. At t3, use of at least one health service was reported by slightly more participants in CG (*n*=82, 73.9%) than in IG (*n*=74,59.8%; $\chi^2(1)=4.01$, $p=0.045$).

**Impact of Covid-19 pandemic**

At t3, 91.2% (*n*=177/194) of participants reported to have experienced negative effects of the pandemic situation in at least one area of their life during the past six months and 32.5% (*n*= 63/194) reported at least one positive effect of the pandemic situation (Figure 3). Self-rated psychological strain by pandemic was low (*M*=2.12, *SD*=1.24) and 8.8% (*n*=17/194) at least partly agreed that they used mental health services because of the pandemic situation. No differences were observed between groups (p<.05).

**Figure 3**

*Effects of the COVID-19 pandemic on areas of everyday life assessed at 18-month Follow-up (n=194)*




*Note.* Report based on *n*=194 study completer who answered the additional questions about the impact of the COVID-19 pandemic on their everyday life in the past six months.

***Sensitivity and moderation analyses***

Study completer analyses at t2 (*n*=235) and t3 (*n*=222) revealed comparable results to ITT analyses, with an additional effect on the SPE scale indicating less endangerment of working ability in IG at both time points and no group difference for quality of life at t3. No moderating effects of intervention characteristics, or psychological burden due to the Covid-19 pandemic on mental health outcomes were observed at the two follow-ups (p>.05). For details see Supplements 4-6.

## Discussion

### Summary of main findings

We found small statistically significant effects of reduction in depressive symptom severity in favor of telephone coaching at 12- and 18-month follow-up ($d_{t2}$=-0.34 [95%-CI: -0.58, -0.10]; $d_{t3}$=-0.31 [95%-CI:-0.56, -0.05]). We did not find any effects for reliable symptom improvement or deterioration, depression onset or remission at any follow-up time point.

### Comparison to existing literature

Comparison with previous studies is limited because research on pTC for MDD prevention or long-term follow-up is scarce. The 12- and 18-month effects were comparable, but slightly declining over time to the effect of depressive symptom reduction observed at post-treatment ($d_{t1}$=-0.39 [95%-CI: −0.64, −0.15]) [21].

The observed long-term effects were slightly lower compared to evidence-based brief Interpersonal Psychotherapy or motivational interviewing administered over telephone for MDD treatment [31,61]. Two studies in rural people living with HIV or in policyholders reported slightly higher moderate effect sizes in reducing depressive symptoms compared to control conditions, *d*=0.47 at 8-months follow-up [61] and *d*=-0.45 [95%-CI: -0.14, -0.75] at 12-month follow-up [31].

The observed effects are comparable but slightly higher to previous findings for indicated depression prevention with IBIs reporting a small effect size at 12-month follow-up, *g* = 0.27 [95% -CI: 0.07, 0.47] [62]. Different findings have been reported for tailored IBIs to treat subthreshold depressive symptoms in farmers, which achieved small effects at post-treatment of *d*=-0.28 [95%-CI: −0.50, −0.07] and at 6-month follow-up of *d*=-0.35 [95%-CI: −0.57, −0.14], but not at 12-month follow-up [18]. However, intervention length and follow-up time points varied between the studies (e.g., IBIs: 9-weeks post-treatment, 6-/12-month follow-up; pTC: 24-weeks post-treatment, 12-/18-month follow-up; all assessments post-randomization).

Our study found a non-significant effect of pTC on MDD onset despite a longer 12-/18-month follow-up period. This is consistent with a study on IBIs in farmers which reported no effect on MDD

onset and similarly low onset rates in both groups at 6- and 12-month follow-up (<21% based on CIDI) [18]. However, previous meta-analyses in the field of indicated prevention reported contrasting that psychological interventions (standardized CBT or problem-solving therapy, guided self-help, IBIs) for subthreshold depression could effectively prevent MDD onset at 6- to 12-month [5,62]. A meta-analysis on psychological (mostly CBT-based) prevention found a non-significant effect on incidence up to 8 years follow-up, but few available studies limited statistical power [6]. Therefore, it cannot be conclusively assessed whether pTC can prevent MDD in farmers.

Moreover, the previously found positive effects of the pTC post-treatment [21] regarding stress, somatic symptoms and burnout risk on the MBI subscales emotional exhaustion and cynism sustained over the 12- and 18-months follow-up, whereas anxiety and quality of life sustained until the 12-months follow-up.

Even though no overall effect of psychological burden due to the pandemic was observed, participating farmers reported a negative pandemic impact on different areas of life, which are well-documented as most reported farmer mental health risk factors, such as financial and work situation, social contacts [13]. In contrast to the low rated psychological burden, research reported an impact of COVID-19 on increased mental health problems and higher prevalence of depression in farmer [63,64].

**Implications for research and practice**

Telephone consultations in addition to in-person treatments are recommended by the recent NICE guideline [65] on treatment of depression to increase access and uptake for people who are under-served and under-represented in services. Our study enriches the field of prevention research and guildelines with evidence on long-term effectiveness up to 12-/18-month of pTC on the reduction of depressive outcomes in an agricultural target group and shows that personalized and not just CBT-based remote interventions work in symptom reduction. This pTC is very resource intensive and comparable to short-term therapy for MDD in Germany. Future research should investigate whether fewer sessions achieve comparable results.

This study contributes moreover to the need for research in the field of farmers' mental health interventions which include holistic, multi-component programs and targeted approaches [66]. Participants reported different occupation-specific stressors well-documented as risk factors for MDD in farmers, e.g. general workload, financial problems, conflicts with family/partner/employees, uncontrollable working situations, farm succession, and stress [13,67–70]. The high degree of personalization enabled coaches to pick up on the various stressors present in the individual case and to apply specifically tailored interventions. Thus, pTC is a promising remote prevention approach targeting potential risk factors in a specific occupational group [12] and addressing the need for tailored preventive services to farm-specific needs to increase likelihood of help-seeking behavior [71]. Future research should examine the impact of tailoring strategies on effectiveness and uptake.

If the nationwide implementation of pTC succeed, this remote service could be feasible in the long term and improve access to and uptake of mental health care in rural areas. While the pandemic has accelerated the adoption of digital technologies in the health sector, it now still requires a so-called digital mindset among health workers and affected people in need of prevention to actually use such remote services [72].

**Limitations**

The study results are restricted by several limitations. First, the effectiveness study was conducted to mimic care as usual, so an upper cut-off value of self-reported depressive symptom severity, diagnostic interviews and exclusion criteria, such as comorbid disorders were omitted. This increases the external validity and generalizability of results in the routine setting. As a considerable proportion of the study participants had clinically depressive symptoms at baseline (*n*=88/314, 28%) and MDE in lifetime (*n*=76/314, 24.2%) [21], this heterogeneous study group in terms of symptom burden and intervention need (e.g. prevention, early treatment, relapse prevention) had a limited representativeness and generalizability for the target group of indicated prevention. Future research should exclude MDD cases or plan a-priori sample sizes that allow subgroup-analyses for subclinical

cases. Second, this is a low-standardized intervention without coaching manual or predetermined coaching procedures, which limits internal validity and monitoring of coaching content. However, standardization would counteract the rational of this personalized coaching approach and the diversity of therapeutic backgrounds of university-trained psychologists in routine care. The high degree of personalization enabled a coaching close to everyday life and work by addressing participants' psychosocial needs and exposure to multi-risk situations in the agricultural working environment. Future research should consider the inclusion of coaching, occupational and health psychological frameworks with defined criteria to ensure monitoring and documenting of personalised interventions and coaching fidelity. Third, the COVID-19 pandemic outbreak temporally confounded with the 12-/18-month follow-ups and the impact on mental health in the study population cannot be ruled out, even though it was methodologically addressed. We did not have a reliable and valid pandemic-related item battery. Bias in the results are possible due to the different timing of the follow-up and COVID-19 item surveys in relation to the end of the intervention. Forth, selection bias in the sample cannot be ruled out because intervention access was not possible for insured residents in two pilot implementation states and linked to willingness to study participation, presence of both internet and telephone access and, the multistage study enrollment process via the study team was more complex than in routine care.

**Conclusions**

Personalized telephone coaching shows long-term effectiveness after 12- and 18-months in terms of multifaceted outcomes such as reduction of depressive and other psychological symptoms in farmers. It could serve as accessible early intervention in depression prevention and could ease access to psychological care in telephone delivery format by reaching farmers with low help-seeking behavior with a targeted approach for their needs. Clinical and cost-effectiveness analyses in routine care are further important evaluation objectives.

## Declarations

**Funding**

The German insurance company SVLFG provided a financial contribution to the Friedrich-Alexander Universität Erlangen-Nürnberg and Ulm University as expense allowance. SVLFG had no role in study design, data collection, analyses, interpretation or writing the manuscript and the decision to publish it.

**Conflict of Interest**

DDE has served as a consultant to/on the scientific advisory boards of Sanofi, Novartis, Minddistrict, Lantern, Schoen Kliniken, Ideamed, and German health insurance companies (BARMER, Techniker Krankenkasse) and a number of federal chambers for psychotherapy. DDE and MB are stakeholders of the Institute for health training online (GET.ON/HelloBetter), which aims to implement scientific findings related to digital health interventions into routine care. MB is scientific advisor of mentalis GmbH, a provider for digital aftercare. HB reports having received consultancy fees and fees for lectures/workshops from chambers of psychotherapists and training institutes for psychotherapists in the e-mental-health context. IT reports having received fees for lectures/workshops in the e-mental-health and digital interventions context from professional associations and training institutes for psychotherapists. She was the research and implementation project lead of the trial site Institute for health training online (GET.ON) for the European implementation research project ImpleMentAll (11/2017-03/2021) funded by the European Commission. JT, LB, JF, and CB report no conflict of interest.

**Data and/or Code availability**

[Possibilities to share data and control group material are currently und evaluation.]

**Contributorship statement**

The authors contributed in the following ways to the manuscript:
Janika Thielecke: Conceptualization, Methodology, Project administration, Investigation, Data Curation, Formal analysis, Visualization, Writing - Original Draft (Methods & Results), Writing - Review & Editing; Ingrid Titzler: Conceptualization, Methodology, Supervision, Resources, Project administration, Investigation, Writing - Original Draft (Introduction & Discussion), Writing - Review & Editing; Lina Braun: Resources, Writing - Review & Editing; David Daniel Ebert: Funding acquisition, Conceptualization, Methodology, Writing - Review & Editing; Johanna Freund: Resources, Writing - Review & Editing; Harald Baumeister: Funding acquisition, Resources, Conceptualization, Writing - Review & Editing; Matthias Berking: Funding acquisition, Writing - Review & Editing; Claudia Buntrock: Conceptualization, Methodology, Supervision, Writing - Review & Editing.
Janika Thielecke and Ingrid Titzler contributed equally to this paper. All authors provided critical revision of the article and approved the final manuscript.

**Acknowledgements**

Authors would like to thank Friederike Dietz and Annika Montag for programming the online surveys. Further, authors thank Sarah Banellis, Merle Bloom, Hanna Böckeler, Albina Chafisuf, Doro Dressel, Johanna Finitzer and Tomris Ohloff for their engagement in enrolling and supporting participants throughout the study. Lukas Fuhrmann and Marvin Franke are thanked for carrying out the randomization and Mathias Harrer and Lea Schuurmanns for statistical advice. Authors thank the IVPNetworks staff who provided details on their intervention and [N.N] for proofreading the manuscript as well.

**References**


1 Bromet EJ, Andrade LH, Bruffaerts R, *et al.* Major Depressive Disorder. In: Scott KM, de Jonge P, Stein DJ, *et al.*, eds. *Mental Disorders Around the World: Facts and Figures from the WHO World Mental Health Surveys.* Cambridge: : Cambridge University Press 2018. 41–56. doi:10.1017/9781316336168

2 Cuijpers P, Noma H, Karyotaki E, *et al.* A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. *World Psychiatry* 2020;**19**:92–107. doi:10.1002/wps.20701

3 Andrews G, Issakidis C, Sanderson K, *et al.* Utilising survey data to inform public policy: Comparison of the cost-effectiveness of treatment of ten mental disorders. *Br J Psychiatry* 2004;**184**:526–33. doi:10.1192/bjp.184.6.526

4 Muñoz RF, Cuijpers P, Smit F, *et al.* Prevention of Major Depression. *Annu Rev Clin Psychol* Published Online First: 2010. doi:10.1146/annurev-clinpsy-033109-132040

5 Cuijpers P, Koole SL, van Dijke A, *et al.* Psychotherapy for subclinical depression: meta-analysis. *Br J Psychiatry* 2014;**205**:268–74. doi:10.1192/bjp.bp.113.138784

6 Cuijpers P, Pineda BS, Quero S, *et al.* Psychological interventions to prevent the onset of depressive disorders: A meta-analysis of randomized controlled trials. Clin. Psychol. Rev. 2021;**83**:101955. doi:10.1016/j.cpr.2020.101955

7 Judd F, Jackson H, Fraser C, *et al.* Understanding suicide in Australian farmers. *Soc Psychiatry Psychiatr Epidemiol* Published Online First: 2006. doi:10.1007/s00127-005-0007-1

8 Roy P, Tremblay G, Oliffe JL, *et al.* Male farmers with mental health disorders: A scoping review. *Aust J Rural Health* Published Online First: 2013. doi:10.1111/ajr.12008

9 Sanne B, Mykletun A, Moen BE, *et al.* Farmers are at risk for anxiety and depression: The Hordaland Health Study. *Occup Med (Chic Ill)* 2004;**54**:92–100. doi:10.1093/occmed/kqh007

10 Torske MO, Bjørngaard JH, Hilt B, *et al.* Farmers' mental health: A longitudinal sibling comparison - The HUNT study, Norway. *Scand J Work Environ Heal* Published Online First:

2016. doi:10.5271/sjweh.3595

11 Sanne B, Mykletun A, Dahl AA, *et al.* Occupational differences in levels of anxiety and depression: The Hordaland Health Study. *J Occup Environ Med* Published Online First: 2003. doi:10.1097/01.jom.0000069239.06498.2f

12 Onwuameze OE, Paradiso S, Peek-Asa C, *et al.* Modifiable risk factors for depressed mood among farmers. *Ann Clin Psychiatry* 2013.

13 Yazd SD, Wheeler SA, Zuo A. Key Risk Factors Affecting Farmers' Mental Health: A Systematic Review. *Int J Environ Res Public Heal 2019, Vol 16, Page 4849* 2019;**16**:4849. doi:10.3390/IJERPH16234849

14 Kallioniemi MK, Simola A, Kaseva J, *et al.* Stress and Burnout Among Finnish Dairy Farmers. *J Agromedicine* 2016;**21**:259–68. doi:10.1080/1059924X.2016.1178611

15 Judd F, Jackson H, Komiti A, *et al.* Help-Seeking by Rural Residents for Mental Health Problems: The Importance of Agrarian Values. *Aust New Zeal J Psychiatry* 2006;**40**:769–76. doi:10.1080/j.1440-1614.2006.01882.x

16 Schang L, Schüttig W, Sundmacher L. Unterversorgung im ländlichen Raum–Wahrnehmung der Versicherten und ihre Präferenzen für innovative Versorgungsmodelle. 2016.

17 Benavides-Vaello S, Strode A, Sheeran BC. Using Technology in the Delivery of Mental Health and Substance Abuse Treatment in Rural Communities: A Review. *J Behav Health Serv Res* 2013;**40**:111–20. doi:10.1007/s11414-012-9299-6

18 Braun L, Titzler I, Terhorst Y, *et al.* Are guided internet-based interventions for the indicated prevention of depression in green professions effective in the long run? Longitudinal analysis of the 6- and 12-month follow-up of a pragmatic randomized controlled trial (PROD-A). *Internet Interv* 2021;**26**:100455. doi:10.1016/j.invent.2021.100455

19 Braun L, Titzler I, Terhorst Y, *et al.* Effectiveness of guided internet-based interventions in the indicated prevention of depression in green professions (PROD-A): Results of a pragmatic randomized controlled trial. *J Affect Disord* 2021;**278**:658–71. doi:10.1016/j.jad.2020.09.066

20 Braun L, Terhorst Y, Titzler I, *et al.* Lessons learned from an attempted pragmatic randomized controlled trial for improvement of chronic pain-associated disability in green professions: Long-term effectiveness of a guided online-based Acceptance and Commitment Therapy (PACT-A). *Int J Environ Res Public Health* 2022;**19**:13858.

21 Thielecke J, Buntrock C, Titzler I, *et al.* Telephone coaching for the prevention of depression in depression in farmers: Results from a pragmatic from a pragmatic randomized controlled trial. *J Telemed Telecare* 2022;:1357633X2211060. doi:10.1177/1357633X221106027

22 Freund J, Ebert DD, Thielecke J, *et al.* Using the Consolidated Framework for Implementation Research to evaluate a nationwide depression prevention project (ImplementIT) from the perspective of health care workers and implementers: Results on the implementation of digital interventions for farmers. *Front Digit Heal* 2023;**4**. doi:10.3389/fdgth.2022.1083143

23 Vis C, Schuurmans J, Aouizerate B, *et al.* Effectiveness of self-guided tailored implementation strategies in integrating and embedding Internet-based Cognitive Behavioural Therapy in routine mental health care. Results of the ImpleMentAll project. (Preprint). *J Med Internet Res* 2022;**25**:1–18. doi:10.2196/41532

24 Mohr DC, Vella L, Hart S, *et al.* The effect of telephone-administered psychotherapy on symptoms of depression and attrition: A meta-analysis. Clin. Psychol. Sci. Pract. 2008;**15**:243–53. doi:10.1111/j.1468-2850.2008.00134.x

25 Castro A, Gili M, Ricci-Cabello I, *et al.* Effectiveness and adherence of telephone-administered psychotherapy for depression: A systematic review and meta-analysis. *J Affect Disord* 2020;**260**:514–26. doi:10.1016/j.jad.2019.09.023

26 Coughtrey AE, Pistrang N. The effectiveness of telephone-delivered psychological therapies for depression and anxiety : A systematic review. 2018;**24**:65–74. doi:10.1177/1357633X16686547

27 Bellanti DM, Kelber MS, Workman DE, *et al.* Rapid Review on the Effectiveness of Telehealth Interventions for the Treatment of Behavioral Health Disorders. *Mil Med* 2021;**187**:E577–88.

doi:10.1093/milmed/usab318

28 Lin T, Heckman TG, Anderson T. The efficacy of synchronous teletherapy versus in-person therapy: A meta-analysis of randomized clinical trials. *Clin Psychol Sci Pract* 2022;**29**:167–78. doi:10.1037/cps0000056

29 Furukawa TA, Horikoshi M, Kawakami N, *et al.* Telephone cognitive-behavioral therapy for subthreshold depression and presenteeism in workplace: A randomized controlled trial. *PLoS One* 2012;**7**. doi:10.1371/journal.pone.0035330

30 Lerner D, Adler DA, Rogers WH, *et al.* A Randomized Clinical Trial of a Telephone Depression Intervention to Reduce Employee Presenteeism and Absenteeism. *Psychiatr Serv* 2015;**66**:570–7. doi:10.1176/appi.ps.201400350

31 Schröder S, Fleckenstein J, Wunderlich M. Effektivität eines telefonbasierten Coaching-Programms für Patienten mit einer depressiven Erkrankung. *Monit Versorgungsforsch* 2017;**10**:65–8. doi:10.24945/MVF.04.17.1866-0533.2022

32 Thielecke J, Buntrock C, Titzler I, *et al.* Clinical and Cost-Effectiveness of Personalized Tele-Based Coaching for Farmers, Foresters and Gardeners to Prevent Depression: Study Protocol of an 18-Month Follow-Up Pragmatic Randomized Controlled Trial (TEC-A). *Front Psychiatry* 2020;**11**:125. doi:10.3389/fpsyt.2020.00125

33 Freund J, Titzler I, Thielecke J, *et al.* Implementing internet- and tele-based interventions to prevent mental health disorders in farmers, foresters and gardeners (ImplementIT): study protocol for the multi-level evaluation of a nationwide project. *BMC Psychiatry* 2020;**20**:424. doi:10.1186/s12888-020-02800-z

34 Bouwmans C, De Jong K, Timman R, *et al.* Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). *BMC Health Serv Res* 2013;**13**:217. doi:10.1186/1472-6963-13-217

35 Rush AJ, Trivedi MH, Ibrahim HM, *et al.* The 16-Item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric

evaluation in patients with chronic major depression. *Biol Psychiatry* 2003;**54**:573–83. doi:10.1016/S0006-3223(02)01866-8

36 Auerbach RP, Mortier P, Bruffaerts R, *et al.* WHO world mental health surveys international college student project: Prevalence and distribution of mental disorders. *J Abnorm Psychol* 2018;**127**:623–38. doi:10.1037/abn0000362

37 Lamoureux BE, Linardatos E, Fresco DM, *et al.* Using the QIDS-SR16 to Identify Major Depressive Disorder in Primary Care Medical Patients. *Behav Ther* 2010;**41**:423–31. doi:10.1016/j.beth.2009.12.002

38 Reis D, Lehr D, Heber E, *et al.* The German Version of the Perceived Stress Scale (PSS-10): Evaluation of Dimensionality, Validity, and Measurement Invariance With Exploratory and Confirmatory Bifactor Modeling. *Assessment* 2019;**26**:1246–59. doi:10.1177/1073191117715731

39 Bastien C. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med* 2001;**2**:297–307. doi:10.1016/S1389-9457(00)00065-4

40 Gierk B, Kohlmann S, Kroenke K, *et al.* The Somatic Symptom Scale-8 (SSS-8): A brief measure of somatic symptom burden. *JAMA Intern Med* Published Online First: 2014. doi:10.1001/jamainternmed.2013.12179

41 Spitzer RL, Kroenke K, Williams JBW, *et al.* A brief measure for assessing generalized anxiety disorder: The GAD-7. *Arch Intern Med* Published Online First: 2006. doi:10.1001/archinte.166.10.1092

42 Bandelow B, Hajak G, Holzrichter S, *et al.* Assessing the efficacy of treatments for panic disorder and agoraphobia. I. Methodological problems. *Int Clin Psychopharmacol* 1995;**10**:83–93. doi:http://dx.doi.org/10.1097/00004850-199506000-00004http://dx.doi.org/10.1097/00004850-199506000-00004

43 Bush K. The AUDIT Alcohol Consumption Questions (AUDIT-C)<subtitle>An Effective Brief Screening Test for Problem Drinking</subtitle>. *Arch Intern Med* 1998;**158**:1789–1795.

doi:10.1001/archinte.158.16.1789

44 Richardson J, Iezzi A, Khan MA, *et al.* Validity and reliability of the assessment of quality of life (AQoL)-8D multi-attribute utility instrument. *Patient* 2014;**7**:85–96. doi:10.1007/s40271-013-0036-x

45 Mittag O, Raspe H. Eine kurze Skala zur Messung der subjektiven Prognose der Erwerbstätigkeit: Ergebnisse einer Untersuchung an 4279 Mitgliedern der gesetzlichen Arbeiterrentenversicherung zu Reliabilität (Guttman-Skalierung) und Validität der Skala. *Rehabilitation (Stuttg)* 2003;**42**:169–74. doi:10.1055/s-2003-40095

46 Cillien P, Fischbach A, Mörsdorf A, *et al. Maslach Burnout Inventory-General Survey Deutsche Version 1.0 (MBI-GS-D V1.0.)*. 2006.

47 Boß L, Lehr D, Reis D, *et al.* Reliability and Validity of Assessing User Satisfaction With Web-Based Health Interventions. *J Med Internet Res* 2016;**18**:e234. doi:10.2196/jmir.5952

48 Munder T, Wilmers F, Leonhart R, *et al.* Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clin Psychol Psychother* 2009;:231–9. doi:10.1002/cpp.658

49 Ladwig I, Rief W, Nestoriuc Y. Welche Risiken und Nebenwirkungen hat Psychotherapie? - Entwicklung des Inventars zur Erfassung Negativer Effekte von Psychotherapie (INEP). *Verhaltenstherapie* 2014;**24**:252–63. doi:10.1159/000367928

50 Brugha TS, Cragg D. The List of Threatening Experiences: the reliability and validity of a brief life events questionnaire. *Acta Psychiatr Scand* Published Online First: 1990. doi:10.1111/j.1600-0447.1990.tb01360.x

51 Adler NE, Epel ES, Castellazzo G, *et al.* Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Heal Psychol* Published Online First: 2000. doi:10.1037/0278-6133.19.6.586

52 Janssen I, Hanssen M, Bak M, *et al.* Discrimination and delusional ideation. *Br J Psychiatry* 2003;**182**:71–6. doi:10.1192/bjp.182.1.71

53 Hughes ME, Waite LJ, Hawkley LC, *et al.* A short scale for measuring loneliness in large surveys: Results from two population-based studies. Res. Aging. 2004. doi:10.1177/0164027504268574

54 Zwarenstein M, Treweek S, Gagnier JJ, *et al.* Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;**337**:a2390–a2390. doi:10.1136/bmj.a2390

55 Buuren S van, Groothuis-Oudshoorn K. mice : Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011;**45**. doi:10.18637/jss.v045.i03

56 Rubin DB. Multiple Imputation after 18+ Years. *J Am Stat Assoc* 1996;**91**:473–89. doi:10.1080/01621459.1996.10476908

57 Freeman PR, Hedges L V., Olkin I. Statistical Methods for Meta-Analysis. *Biometrics* Published Online First: 1986. doi:10.2307/2531069

58 Enders CK. *Applied MD Anlysis*. 2010.

59 Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991;**59**:12–9. doi:10.1037/0022-006X.59.1.12

60 Kraemer HC, Kupfer DJ. Size of Treatment Effects and Their Importance to Clinical Research and Practice. *Biol Psychiatry* 2006;**59**:990–6. doi:10.1016/j.biopsych.2005.09.014

61 Heckman TG, Markowitz JC, Heckman BD, *et al.* A Randomized Clinical Trial Showing Persisting Reductions in Depressive Symptoms in HIV-Infected Rural Adults Following Brief Telephone-Administered Interpersonal Psychotherapy. *Ann Behav Med* 2018;**52**:299–308. doi:10.1093/ABM/KAX015

62 Reins JA, Buntrock C, Zimmermann J, *et al.* Efficacy and Moderators of Internet-Based Interventions in Adults with Subthreshold Depression: An Individual Participant Data Meta-Analysis of Randomized Controlled Trials. *Psychother Psychosom* 2021;**90**:94–106. doi:10.1159/000507819

63 Thompson R, Hagen BNM, Lumley MN, *et al.* Mental Health and Substance Use of Farmers in

Canada during COVID-19. *Int J Environ Res Public Health* 2022;**19**. doi:10.3390/ijerph192013566

64 Sapbamrer R, Chittrakul J, Sirikul W, *et al.* Impact of COVID-19 Pandemic on Daily Lives, Agricultural Working Lives, and Mental Health of Farmers in Northern Thailand. *Sustain* 2022;**14**. doi:10.3390/su14031189

65 National Institute for Health and Care Excellence [NICE]. Depression in adults: treatment and management NICE guideline [NG222]. 2022.www.nice.org.uk/guidance/ng222 (accessed 15 Apr 2023).

66 Younker T, Radunovich HL. Farmer mental health interventions: A systematic review. *Int J Environ Res Public Health* 2022;**19**. doi:10.3390/ijerph19010244

67 Logstein B. Work Demands, Independence, Valuation as a Farmer, and Mental Health in Farming. A Study of Mental Health among Dairy Farmers and Vegetable- And Potato Farmers in Norway. *Eur Countrys* 2021;**13**:175–92. doi:10.2478/euco-2021-0010

68 Hagen BNM, Sawatzky A, Harper SL, *et al.* What impacts perceived stress among canadian farmers? A mixed-methods analysis. *Int J Environ Res Public Health* 2021;**18**. doi:10.3390/ijerph18147366

69 Truchot D, Andela M. Burnout and hopelessness among farmers: The Farmers Stressors Inventory. *Soc Psychiatry Psychiatr Epidemiol* 2018;**53**:859–67. doi:10.1007/s00127-018-1528-8

70 Reed DB, Claunch DT. Risk for Depressive Symptoms and Suicide Among U.S. Primary Farmers and Family Members: A Systematic Literature Review. *Work Heal Saf* 2020;**68**:236–48. doi:10.1177/2165079919888940

71 Hagen BNM, Albright A, Sargeant J, *et al.* Research trends in farmers' mental health: A scoping review of mental health outcomes and interventions among farming populations worldwide. *PLoS One* 2019;**14**:1–20. doi:10.1371/journal.pone.0225661

72 Biancuzzi H, Dal Mas F, Bidoli C, *et al.* Economic and Performance Evaluation of E-Health

before and after the Pandemic Era: A Literature Review and Future Perspectives. *Int J Environ Res Public Health* 2023;**20**:4038. doi:10.3390/ijerph20054038

LONG-TERM EFFECTIVENESS TELEPHONE COACHING

**Online Supplement to:**

Thielecke*, J., Titzler*, I., Braun, L., Ebert, D. D., Johanna Freund, Baumeister, H., Berking, M., & Buntrock, C. (2023). Does telephone coaching improve farmers' mental health in the long term? Results of the 12- and 18-month follow-up of a pragmatic randomized controlled trial

**Content:**



Correspondence concerning this Article should be addressed to Janika Thielecke and Ingrid Titzler, Friedrich-Alexander-University Erlangen-Nürnberg, Department of Clinical Psychology and Psychotherapy, Nägelsbachstr. 25a, 91052 Erlangen. E-Mail: janika.thielecke@fau.de and ingrid.titzler@fau.de

LONG-TERM EFFECTIVENESS TELEPHONE COACHING

**Supplement 1: Number of stressors and coaching topics per participant**

**Table 1**

*Number of stressors reported by participants (n=194)*

| n stressors | n TN | % | cumulative % |
|---|---|---|---|
| 1 | 7 | 4.4 | 4.4 |
| 2 | 11 | 6.9 | 11.3 |
| 3 | 16 | 10.1 | 21.4 |
| 4 | 25 | 15.7 | 37.1 |
| 5 | 39 | 24.5 | 61.6 |
| 6 | 24 | 15.1 | 76.7 |
| 7 | 15 | 9.4 | 86.1 |
| 8 | 14 | 8.8 | 94.9 |
| 9 | 6 | 3.8 | 98.7 |
| 10 | 1 | 0.6 | 99.3 |
| 11 | 1 | 0.6 | 99.9 |

**Table 2**

*Number of coaching topics documented by coaches per participant (n=194)*

| n coaching topics | n TN | % | cumulative % |
|---|---|---|---|
| 1 | 19 | 11.9 | 11.9 |
| 2 | 39 | 24.5 | 36.4 |
| 3 | 40 | 25.2 | 61.6 |
| 4 | 37 | 23.3 | 84.9 |
| 5 | 17 | 10.7 | 95.6 |
| 6 | 3 | 1.9 | 97.5 |
| 7 | 4 | 2.5 | 100.0 |

LONG-TERM EFFECTIVENESS TELEPHONE COACHING

**Supplement 2: Negative effects of the personalized telephone coaching (INEP)**

**Table 3**

*Negative effects of the intervention reported by participants (n=126) in the INEP at 12- months Follow-up assessment*

| Item | N | % |
|---|---|---|
| **Intrapersonal change** | **19** | **15.2** |
| Depending on coach | 8 | 6.4 |
| Neglect of hobbies and social contacts because of coaching | 8 | 6.4 |
| Longer periods of feeling bad | 4 | 3.2 |
| Less trusting others | 4 | 3.2 |
| Feeling more lonely | 4 | 3.2 |
| Severity of suffering from past experiences/events | 3 | 2.4 |
| Worsening of symptoms | 2 | 1.6 |
| Experiencing new thinking and behavior patterns as harmful | 2 | 1.6 |
| Difficulties to make decisions alone | 1 | 0.8 |
| As a human being changed to the negative | 1 | 0.8 |
| Thoughts/plans to commit suicide for the first time | 1 | 0.8 |
| **Therapeutic malpractice of the coach** | **12** | **9.5** |
| Feeling of being forced to do exercises given by coach | 7 | 5.6 |
| Feeling of data security not being ensured during the coaching | 4 | 3.2 |
| Hurtful statements by coach | 2 | 1.6 |
| Feeling of being forced to do exercises given by coach | 7 | 5.6 |
| **Relationship** | **8** | **6.4** |
| More arguments/conflicts in relationship | 5 | 4.0 |
| Problem with partner's jealousy | 3 | 2.4 |
| **Stigma** | **3** | **2.4** |
| Worries about (potentially) increasing insurance fees | 7 | 5.6 |
| Financial worries | 4 | 3.2 |
| Fear of others discovering about the program usage | 4 | 3.2 |
| **Friends and Family** | **3** | **2.4** |
| Worsened relationship with family | 3 | 2.4 |
| Worsened relationship with friends | 2 | 1.6 |

*Notes.* Data based on n=126 study completer. Only negative effects attributed directly to the intervention are reported. Bold print represents number of participants reporting at least one negative effect of the intervention. *Abbreviations.* INEP=Inventory for Assessing Negative Effects of Psychotherapy.

LONG-TERM EFFECTIVENESS TELEPHONE COACHING

**Supplement 3:Use of routine care health services**

**Table 4**

*Use of psychosocial and (mental) health services in the study population according to corresponding items of Tic-P at baseline and follow-ups*

| Services | t | CG | | IG | | Difference between groups % [95%-CI] |
|---|---|---|---|---|---|---|
| | | n | % | n | % | |
| **At least one service used** | t0 | 115 | 71.4 | 120 | 73.8 | 2.4 [0.7, 4.1] |
| | t2 | 87 | 71.8 | 88 | 63.9 | 7.9 [4.6, 11.2] |
| | t3 | 82 | 73.9 | 74 | 59.8 | 14.1 [9.6, 18.6] |
| **General practitioner** | t0 | 116 | 68.2 | 117 | 68.8 | 0.6 [-0.3, 1.5] |
| | t2 | 72 | 59.0 | 75 | 52.6 | 6.4 [3.4, 9.4] |
| | t3 | 68 | 61.3 | 63 | 50.0 | 11.3 [7.2, 15.4] |
| **Psychotherapist** | t0 | 1 | 0.6 | 4 | 2.5 | 1.9 [0.4, 3.4] |
| | t2 | 5 | 4.3 | 4 | 3.0 | 1.3 [-0.1, 2.7] |
| | t3 | 4 | 3.6 | 3 | 2.5 | 1.1 [-0.2, 2.4] |
| **Specialist for psychiatry, neurology or psychosomatic medicine** | t0 | 4 | 2.6 | 9 | 5.6 | 3.0 [1.1, 4.9] |
| | t2 | 2 | 1.7 | 5 | 3.8 | 2.1 [0.3, 3.9] |
| | t3 | 2 | 1.8 | 4 | 3.3 | 1.5 [-0.1, 3.1] |
| **Clinic for psychiatry/ psychosomatic** | t0 | 0 | 0.0 | 1 | 0.6 | 0.6 [-0.3, 1.5] |
| | t2 | 1 | 0.9 | 5 | 3.8 | 2.9 [0.8, 5.0] |
| | t3 | 1 | 0.9 | 2 | 1.6 | 0.7 [-0.4, 1.8] |
| **Prescribed antidepressants** | t0 | 5 | 3.2 | 17 | 10.6 | 7.4 [4.5, 10.3] |
| | t2 | 8 | 6.8 | 8 | 6.0 | 0.8 [-0.3, 1.9] |
| | t3 | 4 | 3.6 | 9 | 7.4 | 3.8 [1.3, 6.3] |
| **SVLFG psychosocial health services** | t0 | 6 | 3.9 | 2 | 1.3 | 2.6 [0.8, 4.4] |
| | t2 | 9 | 7.7 | 3 | 2.3 | 5.4 [2.6, 8.2] |
| | t3 | 3 | 2.7 | 2 | 1.6 | 1.1 [-0.2, 2.4] |
| **Other psychosocial health services** | t0 | 17 | 11.0 | 13 | 8.1 | 2.9 [1.0, 4.8] |
| | t2 | 23 | 19.7 | 13 | 9.8 | 9.9 [6.2, 13.6] |
| | t3 | 10 | 9.0 | 8 | 6.6 | 2.4 [0.4, 4.4] |

*Note.* Based on study completer answers (T0: $N_{CG}$=154, $N_{IG}$=160, T2: $N_{CG}$ =117, $N_{IG}$=133, T3: $N_{CG}$ =111, $N_{IG}$=122). *Abbreviations. t*0 = Baseline, t2= 1. Follow-up (12 months), t3 =2. Follow-up (18 months); IG=Intervention group (telephone coaching), CG=Control group (enhanced treatment as usual), TIC-P=Treatment Inventory of Costs in Patients with psychiatric disorders, SVLFG=German Social Insurance for Agriculture, Forestry and Horticulture.

**Supplement 4: Adaptations in light of the Covid-19 pandemic**

**Table 5**

*Moderating effects of psychological burden due to Covid-19 pandemic in ANCOVA controlled for baseline severity in depressive symptom severity (QIDS-SR16)*

| Outcome | t | F | df | p | d [95%-CI] |
|---|---|---|---|---|---|
| **QIDS-SR16** | t2 | 0.33 | 1410.17 | 0.57 | -0.02 [-0.24, 0.20] |
| | t3 | 0.25 | 845.86 | 0.62 | 0.02 [-0.19, 0.24] |
| **PSS-10** | t2 | 0.47 | 720.38 | 0.49 | -0.07 [-0.28, 0.14] |
| | t3 | 0.26 | 1806.6 | 0.61 | -0.03 [-0.23, 0.18] |
| **ISI** | t2 | 0.45 | 1513.19 | 0.5 | -0.09 [-0.31, 0.12] |
| | t3 | 0.52 | 764.93 | 0.47 | -0.1 [-0.32, 0.12] |
| **GAD-7** | t2 | 0.35 | 797.06 | 0.55 | -0.09 [-0.31, 0.12] |
| | t3 | 0.32 | 1036.81 | 0.57 | -0.06 [-0.28, 0.17] |
| **PAS** | t2 | 0.28 | 817.93 | 0.6 | -0.06 [-0.28, 0.16] |
| | t3 | 0.43 | 1290.88 | 0.51 | -0.09 [-0.29, 0.11] |
| **AUDIT-C** | t2 | 0.45 | 775.42 | 0.5 | 0.08 [-0.13, 0.28] |
| | t3 | 0.64 | 440.83 | 0.42 | 0.09 [-0.13, 0.30] |
| **MBI (EE)** | t2 | 1.83 | 397.88 | 0.18 | 0.08 [-0.13, 0.30] |
| | t3 | 0.73 | 637.63 | 0.39 | 0.04 [-0.18, 0.25] |
| **MBI (PE)** | t2 | 0.49 | 691.64 | 0.48 | 0.10 [-0.11, 0.31] |
| | t3 | 0.19 | 2052.18 | 0.66 | 0.02 [-0.19, 0.22] |
| **MBI (CY)** | t2 | 0.66 | 649.87 | 0.42 | 0.05 [-0.16, 0.26] |
| | t3 | 0.26 | 1935.07 | 0.61 | -0.03 [-0.24, 0.18] |
| **SSS-8** | t2 | 0.48 | 1004.87 | 0.49 | -0.03 [-0.26, 0.21] |
| | t3 | 0.89 | 572.81 | 0.35 | 0.06 [-0.14, 0.27] |
| **AQoL-8D (total)** | t2 | 0.4 | 1269.96 | 0.53 | 0.10 [-0.11, 0.32] |
| | t3 | 0.31 | 1361.92 | 0.58 | 0.07 [-0.14, 0.28] |
| **AQoL-8D (MSD)** | t2 | 0.46 | 1949.84 | 0.5 | 0.10 [-0.13, 0.32] |
| | t3 | 0.33 | 1536.34 | 0.57 | 0.08 [-0.13, 0.29] |
| **AQoL-8D (PSD)** | t2 | 0.79 | 516.48 | 0.38 | 0.11 [-0.13v0.34] |
| | t3 | 0.46 | 1603.96 | 0.5 | -0.01 [-0.21, 0.19] |

*Note*. Test statistics for the moderation effect of the item "How much would you attribute the psychological distress you experienced in the last 6 months to the impact of the Corona pandemic?" (range: 0 – no psychological burden to very high psychological burden) are reported. Based on multiple imputation according to intention-to-treat analyses (N=314). *Abbreviations.* IG: Intervention group

LONG-TERM EFFECTIVENESS TELEPHONE COACHING

(telephone coaching); CG: Control group (enhanced treatment as usual); t2: 1. Follow-up (12 months); t3: 2. Follow-up (18 months); QIDS-SR16: Quick Inventory of Depressive Symptomatology; PSS-10: Perceived Stress Scale; ISI: Insomnia Severity Index; GAD-7: Generalized Anxiety Disorder; PAS: Panic and Agoraphobia Scale; AUDIT-C: Alcohol Use Disorder Identification Test – Consumption Questions; SSS-8: Somatic Symptom Scale; MBI-GS: Maslach-Burnout-Inventory General Survey; EE: Emotional Exhaustion; CY: Cynism; PE: Professional Efficacy; AQoL-8D: Assessment of Quality of Life; PSD: Physical Super Dimension; MSD: Mental Super Dimension; Analyses based on study completer.

**Supplement 5: Sensitivity analysis (complete cases)**

**Table 6**

*Group comparisons of continuous outcomes at baseline and follow-ups based on study completer*

| Variable | t | Intervention | | | Control | | | F | p | Between-group effect size Cohen's d [95%-CI] |
|---|---|---|---|---|---|---|---|---|---|---|
| | | N | M | SD | N | M | SD | | | |
| | t0 | 160 | 9.81 | 4.13 | 154 | 10.26 | 4.4 | | | |
| **QIDS-SR16** | t2 | 133 | 6.41 | 4.23 | 117 | 7.76 | 3.89 | 9.466 | 0.002 | -0.33 [-0.57, -0.08] |
| | t3 | 122 | 6.25 | 3.85 | 111 | 7.53 | 4.35 | 7.878 | 0.005 | -0.31 [-0.57, -0.06] |
| | t0 | 160 | 21.76 | 5.78 | 154 | 21.95 | 6.36 | | | |
| **PSS-10** | t2 | 131 | 15.89 | 7.12 | 111 | 18.96 | 6.94 | 14.87 | <0.001 | -0.43 [-0.68, -0.18] |
| | t3 | 119 | 15.98 | 7.17 | 108 | 18.64 | 6.85 | 10.81 | 0.001 | -0.37 [-0.63, -0.12] |
| | t0 | 160 | 10.95 | 4.98 | 154 | 10.72 | 5.05 | | | |
| **ISI** | t2 | 130 | 7.58 | 5.52 | 111 | 8.71 | 5.5 | 3.92 | 0.049 | -0.21 [-0.46, 0.05] |
| | t3 | 119 | 8 | 5.1 | 108 | 8.26 | 5.64 | 0.18 | 0.671 | -0.05 [-0.31, 0.21] |
| | t0 | 160 | 8.28 | 3.67 | 154 | 8.7 | 4.05 | | | |
| **GAD-7** | t2 | 131 | 5.99 | 3.92 | 111 | 7.14 | 4.29 | 7.13 | 0.008 | -0.28 [-0.53, -0.03] |
| | t3 | 119 | 5.76 | 3.79 | 108 | 6.72 | 4.57 | 4.30 | 0.039 | -0.23 [-0.49, 0.03] |
| | t0 | 160 | 6.08 | 6.67 | 154 | 6.65 | 7.1 | | | |
| **PAS** | t2 | 131 | 4.35 | 6.88 | 111 | 4.78 | 6.49 | 0.34 | 0.559 | -0.07 [-0.32, 0.19] |
| | t3 | 119 | 4.03 | 6.19 | 107 | 4.04 | 5.69 | 0.00 | 0.996 | 0.00 [-0.26, 0.26] |
| | t0 | 160 | 3.04 | 1.77 | 154 | 3.19 | 2.05 | | | |
| **AUDIT-C** | t2 | 131 | 2.74 | 1.78 | 111 | 2.73 | 1.85 | 0.01 | 0.935 | 0.01 [-0.25, 0.26] |
| | t3 | 119 | 2.87 | 1.85 | 107 | 2.79 | 1.9 | 0.36 | 0.551 | 0.04 [-0.22, 0.31] |
| | t0 | 160 | 11.03 | 4.66 | 154 | 11.74 | 5.5 | | | |
| **SSS-8** | t2 | 130 | 7.82 | 5.02 | 110 | 9.57 | 5.7 | 11.20 | 0.001 | -0.32 [-0.58, -0.07] |
| | t3 | 117 | 8.3 | 5.18 | 107 | 9.75 | 6.16 | 6.27 | 0.013 | -0.25 [-0.52, 0.01] |
| | t0 | 160 | 15.33 | 7.17 | 154 | 16.44 | 7.58 | | | |
| **MBI (CY)** | t2 | 130 | 11.57 | 7.61 | 110 | 13.85 | 8.29 | 10.28 | 0.002 | -0.33 [-0.59, -0.08] |
| | t3 | 117 | 11.74 | 7.69 | 107 | 14.12 | 8.27 | 9.37 | 0.002 | -0.35 [-0.61, -0.09] |
| | t0 | 160 | 8.5 | 5.84 | 154 | 9.41 | 5.86 | | | |
| **MBI (EE)** | t2 | 130 | 6.86 | 5.91 | 110 | 8.91 | 6.21 | 8.79 | 0.003 | -0.29 [-0.54, -0.03] |
| | t3 | 117 | 6.85 | 5.91 | 107 | 9.08 | 6.63 | 7.25 | 0.008 | -0.30 [-0.56, -0.04] |
| | t0 | 160 | 26.69 | 7.19 | 154 | 25.12 | 7.58 | | | |
| **MBI (PE)** | t2 | 130 | 28.63 | 6.97 | 110 | 27.41 | 7.23 | 2.60 | 0.109 | 0.17 [-0.08, 0.43] |
| | t3 | 117 | 28.55 | 6.6 | 107 | 26.77 | 8.41 | 4.05 | 0.045 | 0.24 [-0.03, 0.5] |
| | t0 | 160 | 66.89 | 9.38 | 154 | 66.39 | 10.9 | | | |
| **AQoL-8D (total)** | t2 | 127 | 73.72 | 10.37 | 110 | 70.38 | 12.17 | 2.615 | 0.107 | 0.17 [-0.1, 0.43] |
| | t3 | 116 | 72.65 | 10.88 | 107 | 70.7 | 12.6 | 10.59 | 0.001 | 0.30 [0.04, 0.55] |
| | t0 | 160 | 62.47 | 10.73 | 154 | 61.76 | 12.34 | | | |
| **AQoL-8D (MSD)** | t2 | 127 | 70.35 | 11.83 | 110 | 66.92 | 13.42 | 8.59 | 0.004 | 0.27 [0.02, 0.52] |
| | t3 | 116 | 69.5 | 12.22 | 107 | 67.47 | 14.14 | 2.17 | 0.142 | 0.15 [-0.11, 0.42] |
| | t0 | 160 | 77.67 | 10.34 | 154 | 77.67 | 10.81 | | | |
| **AQoL-8D (PSD)** | t2 | 127 | 81.95 | 10.19 | 110 | 78.82 | 11.64 | 10.11 | 0.002 | 0.28 [0.03, 0.54] |
| | t3 | 117 | 80.4 | 10.71 | 107 | 78.6 | 11.65 | 2.80 | 0.096 | 0.16 [-0.10, 0.42] |

**Table 6 (continued)**

*Group comparisons of continuous outcomes at baseline and follow-ups based on study completer*

| Outcome | | n | % | n | % | $X^2$ | df | p |
|---|---|---|---|---|---|---|---|---|
| **SPE** | | | | | | | | |
| SPE=0 | t0 | 75 | 46.9 | 66 | 42.9 | | | |
| SPE=1 | | 36 | 22.5 | 44 | 28.6 | | | |
| SPE=2 | | 40 | 25.0 | 33 | 21.4 | | | |
| SPE=3 | | 9 | 5.6 | 11 | 7.1 | | | |
| SPE=0 | t2 | 77 | 61.1 | 57 | 52.3 | | | |
| SPE=1 | | 20 | 15.9 | 24 | 22 | 22.78 | 3 | 0.002 |
| SPE=2 | | 23 | 18.3 | 20 | 18.3 | | | |
| SPE=3 | | 6 | 4.8 | 8 | 7.3 | | | |
| SPE=0 | t3 | 67 | 57.8 | 52 | 49.1 | | | |
| SPE=1 | | 23 | 19.8 | 23 | 21.7 | 22.78 | 3 | 0.002 |
| SPE=2 | | 17 | 14.7 | 25 | 23.6 | | | |
| SPE=3 | | 9 | 7.8 | 6 | 5.7 | | | |

*Note.* Analyses based on study completer. *Abbreviations.* IG: Intervention group (telephone coaching); CG: Control group (enhanced treatment as usual); t0: Baseline assessment; t2:1. Follow-up (12 months); t3: 2. Follow-up (18 months); QIDS-SR16: Quick Inventory of Depressive Symptomatology; PSS-10: Perceived Stress Scale; ISI: Insomnia Severity Index; GAD-7: Generalized Anxiety Disorder; PAS: Panic and Agoraphobia Scale; AUDIT-C: Alcohol Use Disorder Identification Test – Consumption Questions; SSS-8: Somatic Symptom Scale; MBI-GS: Maslach-Burnout-Inventory General Survey; EE: Emotional Exhaustion; CY: Cynism; PE: Professional Efficacy; AQoL-8D: Assessment of Quality of Life; PSD: Physical Super Dimension; MSD: Mental Super Dimension; SPE=Subjective Prognosis of Gainful Employment Scale (0= no threat to gainful employment to 3=very high threat to gainful employment ).

LONG-TERM EFFECTIVENESS TELEPHONE COACHING

**Table 7**

*Overview on onset and remission according to CIDI and QIDS-SR16 across measurements (study completer)*

| | Baseline | | 1. Follow-Up (12 months) | | | | | 2. Follow-Up (18 months) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Onset** | **n cases (%)** | | **n onset (%)** | | **z** | **p** | **IRR [95%-CI]** | **n onset (%)** | | **z** | **p** | **IRR [95%-CI]** |
| *Potential MDE (CIDI)* [a] | | | | | | | | | | | | |
| **CG** | 125 | (81.2) | 13/94 | (13.8) | -0.67 | 1.79 | 0.7 [0.31, 1.79] | 12/88 | (14.8) | -0.82 | 0.41 | 0.7 [0.29, 1.66] |
| **IG** | 135 | (84.4) | 9/111 | (8.1) | | | | 2/101 | (7.9) | | | |
| *Potential MDD (QIDS-SR16)* | | | | | | | | | | | | |
| **CG** | 106 | (68.8) | 2/84 | (2.4) | -0.3 | 5.25 | 0.7 [0.10, 5.25] | 4/76 | (5.3) | -1.04 | 0.3 | 0.4 [0.08, 2.22] |
| **IG** | 120 | (75.0) | 3/102 | (2.9) | | | | 2/97 | (2.1) | | | |
| **Remission** | | | | | | | | | | | | |
| | **n cases (%)** | | **n remis. (%)** | | **z** | **p** | **OR [95%-CI]** | **n remis. (%)** | | **z** | **p** | **OR [95%-CI]** |
| *Potential MDE (CIDI)* | | | | | | | | | | | | |
| **CG** | 29 | (18.8) | 10/23 | (43.5) | -0.33 | 1.95 | 0.9 [0.39, 1.95] | 10/23 | (43.5) | -0.09 | 0.93 | 1 [0.37, 2.48] |
| **IG** | 25 | (15.6) | 11/22 | (50.0) | | | | 9/21 | (42.9) | | | |
| *Potential MDD (QIDS-SR16)* | | | | | | | | | | | | |
| **CG** | 48 | (31.2) | 23/33 | (69.7) | 0.7 | 4.58 | 1.5 [0.49, 4.58] | 23/35 | (65.7) | 1.55 | 0.12 | 2.7 [0.76, 9.82] |
| **IG** | 40 | (25.0) | 24/31 | (77.4) | | | | 21/25 | (84.0) | | | |

*Note.* MDE remission according to CIDI was not calculated at post-assessment due to overlap in assessment point; [a]Controlled for MDE status at post-treatment. *Abbreviations.* IG: Intervention group (telephone coaching); CG: Control group (enhanced treatment as usual); QIDS-SR16: Quick Inventory of Depressive Symptomatology; CIDI=Composite International Diagnostic Interview; MDE: Major Depressive Episode; MDD: Major Depressive Dissorder; OR: Odds Ratio; IRR: Incidence Risk Ratio.

**Supplement 6: Moderation analysis**

**Table 8**

*Moderating effects of coaching characteristics on depressive symptom severity (QIDS-SR16) in IG (n=160)*

| Intervention Characteristic | t | β | SE | t | df | p | d [95%-CI] |
|---|---|---|---|---|---|---|---|
| **Session Count** | t2 | -0.06 | 0.08 | -0.68 | 4760.72 | 0.50 | -0.01 [-0.05, 0.02] |
| | t3 | -0.03 | 0.08 | -0.39 | 2531.15 | 0.69 | -0.01 [-0.04, 0.03] |
| **Total Coaching Time (minutes)** | t2 | 0.00 | 0.00 | -0.94 | 6660.71 | 0.35 | 0 [0, 0] |
| | t3 | 0.00 | 0.00 | -1.01 | 2015.67 | 0.32 | 0 [0, 0] |
| **Time in programm (weeks)** | t2 | -0.01 | 0.06 | -0.22 | 2210.79 | 0.83 | 0 [-0.03, 0.03] |
| | t3 | 0.02 | 0.06 | 0.37 | 2123.89 | 0.72 | 0.01 [-0.02, 0.03] |

*Note.* Based on multiple imputation according to intention-to-treat analyses (n=160).*Abbreviations.* IG: Intervention group (telephone coaching); CG: Control group (enhanced treatment as usual); t2: 1. Follow-up (12 months); t3: 2. Follow-up (18 months); QIDS-SR16: Quick Inventory of Depressive Symptomatology.

# 6

## ARTICLE 4 — INTERVIEW STUDY TO SUPPORT IMPLEMENTATION AND UPTAKE

Title:
How to Promote Usage of Telehealth Interventions for Farmers' Mental Health? A Qualitative Study on Supporting and Hindering Aspects for Acceptance and Satisfaction with a Personalized Telephone Coaching

Summary:
Low-threshold and remotely delivered interventions are important for farmers and related professions, who often experience multiple risk factors for depression, live in underserved areas and show low help-seeking behavior. As such, personalized telephone coaching can be used as a preventive intervention to support farmer's mental health and reduce depressive symptom severity. In order to use the full potential in reducing depression burden, telephone coaching needs to be implemented and actually used in routine care. This qualitative study therefore aimed to identify factors with facilitate or hinder uptake and use in routine care.

Semi-structured interviews were conducted with 20 out of 66 invited participants which had taken part in the telephone coaching program for up to six months as part of an ongoing effectiveness study or implementation project. The interview guide was based on the 'Unified Theory of Acceptance and Use of Technology' for acceptance and the 'Evaluation' and 'Discrepancy' models for satisfaction. The interviews were audio-recorded, transcribed, and analyzed using deductive-inductive qualitative content analysis developing a combined coding system for both constructs. Independent coding by two individuals resulted in good agreement of $\kappa = 0.80$. Participants were asked to rate the identified themes in an online questionnaire for validation.

Using qualitative content analysis, 8 expectations (E), 32 supporting (SF) and 14 hindering factors (HF) that could have influenced acceptance and satisfaction were identified. Themes were organized into the following five main categories: Coaching result (E = 4, SF = 9, HF = 3), coach (E = 4, SF = 9, HF = 1), organization (SF = 5, HF = 2), participant characteristics (SF = 5, HF = 3), and the telephone as de-

livery format (SF = 4, HF = 5). 'Flexible appointment arrangement' (n = 19/95%), 'low effort' (n = 17/85%) and a 'trusting coaching relationship' (n = 16/80%) were the supporting factors most often named. 'Lack of visual cues' (n = 12/60%) and 'social/professional involvement restricting change process' (n = 10/50%) were the hindering aspects most often named. Expectation fulfillment concerning the coaching result (M = 85.3%, SD = 16.2%) and the coach (M = 96.2%, SD = 7.0%) were high.

Limitations of this study include the convenient sample mostly from a study context, which might have been biased in degree of acceptance and satisfaction and might not translate well to routine care use. Adaptation of the theoretical components had to be made to consider both constructs in one analysis.

In conclusion, interviewees revealed that acceptance and satisfaction were influenced by expected and perceived changes attributed to the coaching, the low effort because of the telephone conduct and the personalization of the intervention. These aspects could be further enhanced by recommending and offering more choice in terms of delivery medium (e.g. offering video calls) and training coaches more in farming related topics.

Submitted manuscript:
Thielecke J, Buntrock C, Freund J, Braun L, Ebert DD, Berking M, Baumeister H, and Titzler I. How to Promote Usage of Telehealth Interventions for Farmers' Mental Health? A Qualitative Study on Supporting and Hindering Aspects for Acceptance and Satisfaction with a Personalized Telephone Coaching. Internet Interventions. submitted

Contribution:
Janika Thielecke co-developed the interview guide for telephone coaching participants with focus on theoretical models of acceptance and satisfaction. She was responsible for the recruitment of interview participants, the coordination of the interviews and collection of interview data. She supervised and contributed to the qualitative data analysis and interpretation and lead discussion rounds on results in project team. She was rsponsible for data curation, formal analysis and visualization. She drafted the original manuscript, revised it after feedback from her co-authors and managed the submission process.

**How to promote usage of telehealth interventions for farmers' mental health? A qualitative study on supporting and hindering aspects for acceptance and satisfaction with a personalized telephone coaching**

Janika Thielecke[1;2*], Claudia Buntrock[1;3], Johanna Freund[1;2], Lina Braun[4], David D Ebert[2], Matthias Berking[1], Harald Baumeister[4] and Ingrid Titzler[1]

[1] Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

[2] Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

[3] Institute of Social Medicine and Health Systems Research, Otto von Guericke University Magdeburg, Magdeburg, Germany

[4] Department of Clinical Psychology and Psychotherapy, Ulm University, Ulm, Germany

**Word count:** 8429 words (including Tables and References)

**Word count:** 5167 words (excluding Tables and References)

*Corresponding author: Janika Thielecke, Friedrich-Alexander-University Erlangen-Nürnberg, Department of Clinical Psychology and Psychotherapy, Nägelsbachstr. 25a, D-91052 Erlangen, Germany. Email: janika.thielecke@fau.de

**Abstract**

Low-threshold and remotely delivered preventive interventions, like telephone coaching, are warranted for farmers who experience multiple risk factors for depression, live in underserved areas and show low help-seeking behavior. Factors facilitating uptake and actual use of effective remote interventions are important to reduce depression disease burden. This study aimed at identifying factors that influence acceptance of and satisfaction with a telephone coaching in this occupational group.

Semi-structured interviews were based on the 'Unified Theory of Acceptance and Use of Technology', the 'Evaluation' and 'Discrepancy' models for satisfaction. Interviews were conducted with 20 of 66 invited participants of a 6-months telephone coaching during an effectiveness or implementation study. Audio-recorded interviews were transcribed and analyzed (deductive-inductive qualitative content analysis). Independent coding by two persons resulted in good agreement (K=0.80). Participants validated results via questionnaire.

Overall, 32 supporting (SF) and 14 hindering factors (HF) for acceptance and satisfaction were identified and organized into five categories: Coaching result (SF=9, HF=3), coach (SF=9, HF=1), organization (SF=5, HF=2), the telephone as communication medium (SF=4, HF=5) and participant characteristics (SF=5, HF=3). Most named SFs were 'Flexible appointment arrangement' (n=19/95%) and 'low effort' (n=17/85%) while most reported HFs were 'lack of visual cues' (n=12/60%) and 'social/professional involvement restricts change process' (n=10/50%). Interviewees perceived changes initiated by coaching, low effort through telephone conduct and personalization as important influencing factors on acceptance and satisfaction. Both may be further enhanced by offering choice and advice for delivery formats (e.g. video-calls) and training of coaches in farm-related issues.

**Study registration:** German Clinical Trial Registrations: DRKS00017078 and DRKS00015655.

**Keywords**

Interviews, acceptance, satisfaction, telephone coaching, depression, farmers

## 1 Introduction

Major depression disorder (MDD[1]) is a highly prevalent mental disorder (Gutiérrez-Rojas et al., 2020) and ranks first as single contributor to all years lost to disabilities worldwide (World Health Organization;, 2017). Effective psychological and pharmacological treatments exist but only one out of five individuals with MDD receives adequate care (Bromet et al., 2018). Preventive interventions could target individuals at risk for depression and improve mental health care in an early stage.

Psychological interventions for the prevention of MDD have shown to effectively reduce depression symptoms (Cuijpers et al., 2014) and avoid or at least delay MDD onset (Cuijpers et al., 2021; van Zoonen et al., 2014) in people with known risk factors (selective prevention) or with subthreshold symptoms (indicated prevention). Nevertheless, uptake of indicated preventive interventions is low with estimations that less than one in a hundred individuals with subthreshold depression begins a preventive program (Cuijpers et al., 2010).

A promising approach to address the challenge of low uptake of mental health services could be telehealth interventions which can be used flexible in time and location and thus, facilitate health care access (Ebert et al., 2018; Snoswell et al., 2021). Easily accessible preventive services are especially warranted for farmers and related professions, who experience multiple risk factors for depression (e.g. long working hours, uncontrollable weather conditions)(Daghagh Yazd et al., 2019) while often living in remote, underserved areas (Myers, 2019; Salize et al., 2007; Weinhold and Gurtner, 2014) and showing low help-seeking behavior for mental health problems (Dollman et al., 2021; Judd et al., 2006; Wrigley et al., 2005).

The nationwide pilot project 'With us in balance' by the German Social Insurance for Agriculture, Forestry and Horticulture (SVLFG)[2] is therefore addressing this target group by implementing internet- and tele-based interventions to prevent depression. Two randomized controlled trials (RCTs) have shown that these guided and tailored internet-based interventions (Braun et al., 2021a, 2021b, 2019) and the personalized telephone coaching (Thielecke et al., 2022, 2020) can effectively reduce depressive symptom severity and promote mental health in farmers when compared to a control group receiving psychoeducational material. Actual uptake and use by the target population are evaluated in the implementation study ImplementIT(Freund et al., 2020). The mixed-methods study ImplementIT includes the views of various stakeholders including acceptance and satisfaction with the intervention by participating farmers to adequately address the needs of the population and promote use of the intervention (Damschroder et al., 2009; Proctor et al., 2011).

Acceptance of new digital interventions has been shown to influence actual uptake of such, both in industry (Venkatesh et al., 2003) and in (mental) health care(Harst et al., 2019; Philippi et al., 2021; Sora et al., 2021) settings. Participants' satisfaction is an important criterion for health care quality and supports

[1] MDD: Major Depression Disorder
[2] SVLFG: Social Insurance for Agriculture, Forestry and Horticulture

development of health services (Samartzis and Talias, 2020). For telehealth intervention in patients with mostly somatic conditions, a recent review identified treatment effectiveness, preference, and ease of use as most important contributing factors to satisfaction (Kruse et al., 2017). Reported patient experiences with telephone-delivered cognitive behavioral therapy (CBT[3]) for MDD highlight overall acceptance of the intervention, the convenience of remote delivery, and a good therapeutic relationship (Bee et al., 2010; Haller et al., 2019). Participants in these studies also report an initial doubt about the effectiveness, technical problems, and lack of visual feedback as disadvantages of telephone-administered CBT. To the best of our knowledge, there is no research on users' experiences with preventive telephone interventions. Qualitative studies report therapeutic guidance, target group adaptation, and flexible use as important factors for intervention use for other remotely delivered health services, namely internet-based interventions for depression (Freund et al., n.d.; Holst et al., 2017; Mayer et al., 2019). Low internet skills, lack of individual fit, and data security concerns are on the other hand described as barriers for intervention use (Braun et al., 2022; Freund et al., 2022; Gerhards et al., 2011; Holst et al., 2017; Mayer et al., 2019). When delivering mental health interventions over telephone, some of these factors could be enhanced (e.g. intense guidance) or circumvented (e.g. need for internet skills), but research is missing.

This study addresses the following research question: 'Which factors influence the acceptance of and satisfaction with a personalized telephone coaching in farmers?' in order to best meet the needs of the German farming population with a preventive telephone coaching and optimize uptake in routine care.

## 2 Material and methods

### 2.1 Study setting

The ethics committee of the Friedrich-Alexander-Universität Erlangen-Nürnberg had approved this qualitative interview study as part of the implementation study ImplementIT and was registered in the German Clinical Trial Registration (DRKS00017078). Interviewees participated in the telephone coaching as clients under routine care conditions during the pilot phase of ImplementIT or as study participants in the associated RCT (TEC-A[4], trial registration: DRKS00015655). The Consolidated Criteria for Reporting Qualitative Research (COREQ)(Tong et al., 2007) have guided the reporting of methods and results in this study (see Supplement 1).

### 2.2 Intervention

All study participants took part in a personalized telephone coaching by IVPNetworks, which was financed by their insurance (SVLFG). The study enrollment process differed for TEC-A and ImplementIT

[3] CBT: Cognitive Behavioral Therapy
[44] TEC-A: study acronym, telephone coaching for agriculturists

participants, while coaching conduct was the same. Participants were registered on the IVPNetworks platform (IVPnet) by the university study team after submitting informed consent and being randomized to the intervention group (TEC-A) or by SVLFG call center agents after consultation about different services (ImplementIT[5]). The case managers at IVP networks assigned the coach who then contacted participants.

A coaching volume of 850 minutes over six months was available for each participant with the possibility of additional 150 minutes over three months if approved by the insurance company. As part of the personalization of the coaching, session length and frequency of the coaching as well as topics were oriented on the individual participants needs. If indicated and available, on-site coaching or other support-services (e.g. socioeconomic counseling) could be arranged or recommended.

Coaches were psychologists with master's degree and training in diverse psychological methods (e.g. systemic, cognitive behavioral, hypnotherapeutic). No fixed manual was applied for the coaching. According to the coaches, typical elements of their coaching were psychoeducation, conjoint goal setting, and a three-phase model ('introduction and alliance building,' 'working,' and 'stabilizing' phase). Licensed psychotherapists were available for supervision. Intervention details can be found elsewhere (Freund et al., 2020; Thielecke et al., 2022, 2020).

**2.3 Participants and recruitment**

Based on inclusion criteria for the studies (TEC-A, ImplementIT), all interviewees were 18 years or older, an agricultural entrepreneur or collaborating family member or pensioner and insured at SVLFG with access to internet and telephone. Interviewees recruited within TEC-A were required to have at least subclinical depression (PHQ≥5)[6]. Participants could not currently receive psychotherapy and had to be able to distance from suicidal ideation. All participants provided extra written informed consent for the interviews, data-matching, and audio-recording.

Coaching participants (40/160, 25%) from TEC-A were invited by mail in August and September 2019 who had stated initial interest in participating in the interview, had not withdrawn from the study in the meantime, and had completed coaching or would do so within a month. Out of 20 responders (50%), 17 consented to the interviews, scheduled an appointment, and took part (42%).

At the same time, coaches were asked to inform the 26 current participants of ImplementIT about the interviews and refer interested participants to the study team. Out of 12 responders (46%), three consented to the interviews, scheduled an appointment, and took part (25%).

[5] ImplementIT: study acronym, Implementation of internet- and tele-based prevention programs
[6] PHQ: Patient Health Questionaire

**2.4 Development of interview guide**

Questions for acceptance were based on the Unified Theory of Acceptance and Use of Technology (UTAUT)[7](Venkatesh et al., 2003) which promoted the intention-to-use and actual use of technology. Questions for satisfaction were based on the eight characteristics of health care service described in Evaluation Model (Blum, 1998; Ware et al., 1983) in order to encompass participants' experience with a broad range of aspects. These characteristics comprise and specify the quality dimensions of structure, process, and outcome as suggested in Donadebian's classic Quality of Care Model (Donadebian, 2005). The Discrepancy Theory (Fox and Storms, 1981) informed the question on expectation fulfillment with the coaching, coach, and organization as an indicator for satisfaction (numerical rating form 0%–100%). Table 1 presents the operationalization of all theoretic components and their adaptation to the telephone context in the interview guide.

[7] UTAUT: Unified Theory of Acceptance and Use of Technology

**Table 1.** *Examples from the interview guide representing the theoretical components of acceptance and satisfaction with their adapted definitions used in the study.*

| Component | Definition | Example question |
|---|---|---|
| *Theory of Acceptance and Use of Technology (acceptance):* | | |
| Performance expectancy | Participant's beliefs on how the personalized telephone coaching is favorable for his/her professional or private situation | *When you think back to the beginning of the coaching: What expectations did you have regarding what should change through the coaching for you?* |
| Effort expectancy | Participant's expectations about the associated effort of the coaching | *How much time and work did you imagine you would have to put into the coaching?* |
| Social influence | Degree to which the social environment knows about the coaching and supports participation | *Which role did the support from friends and family play in your participation in the personalized coaching?* |
| Facilitating conditions | Participant's beliefs about what organizational and technical characteristics could support coaching participation | *What are the advantages [disadvantages] of a telephone coaching compared to a traditional on-site preventive service?* |
| Behavioral intention | The intention to use the personalized telephone coaching (again) | *How willing would you be to participate in another coaching?* |
| *Evaluation Model (satisfaction):* | | |
| Technical care quality | Competence of coach or his/her quality of care | *To what extent did you feel understood by your coach in terms of professional e.g. farming-related topics?* |
| Psychosocial care quality | Non-clinical aspect of coach-participant interaction (e.g. friendliness, empathy) | *How would you describe your relationship with your coach?* |
| Accessibility | Organizational conditions that support coaching participation | *Regarding the organization, which aspects were you especially [less] content with?* |
| Spatial and technical equipment | Spatial or technical difficulties during the coaching | *Which technical difficulties occurred during the coaching?* |
| Treatment outcome | Degree to which the coaching improved problematic situation or psychological burden | *What tangible changes came about in your everyday life through the coaching?* |
| Continuity of care | Continuity of coach | *Were you attended by the same coach the whole time?* |
| Financing | Financing of the coaching | *Would you have participated in the personalized telephone coaching if it was associated with costs for you?* |
| Availability | Availability of the personalized telephone coaching in care | *Would you also have chosen the telephone coaching if you had a comparable service available on-site?* |
| *Discrepancy Model (satisfaction):* | | |
| Expectation | Expectation regarding the coaching, the organization and the possible outcome | *When you think back to the beginning of the coaching: What did you expect from your coach?* |
| Expectation fulfillment | Discrepancy between expectations at the beginning of the coaching and real life conditions | *To what extent were your expectations fulfilled? Please name a percentage between 0 and 100%.* |

The semi-structured interview guide contained 25 questions derived from the theoretical components and organized in four blocks: 'coaching process and result', 'coach', 'organization' and 'telephone delivery'. Additionally, 21 memo and inquiry questions were given to ensure complete and detailed answers. The interview guide was developed by psychologists with a master degree (LB, JF, JT) under supervision of a clinical expert and qualitative research specialist (IT).

### 2.5 Data collection

#### *2.5.1 Qualitative data collection*

Interviews were conducted mostly after the end of the coaching via telephone between August and October 2019. Two students with a bachelor degree in psychology (JB, MM) conducted each 10 one-on-one-interviews. Each interviewer conducted a role-play test-interview and received feedback on their first participant-interview by JT. Average interview duration was M=44 minutes (SD=17 minutes). Interviews were audio-recorded using PhonerLite (Sommerfeld, 2019) and audacity (Audacity Team, 2021) and transcribed verbatim according to a transcription protocol. Names were anonymized during transcription.

#### *2.5.2 Quantitative data collection.*

Sociodemographic data were assessed online for interviewees of ImplementIT or were extracted from TEC-A baseline assessment. For later comparison and a triangulation approach (Heale and Forbes, 2013), data on depressive symptom severity (QIDS-SR16[8], Rush et al., 2003), range 0–27), intervention satisfaction (CSQ-I[9], Boß et al., 2016, range 8-32), and working alliance (WAI-SR[10], Munder et al., 2009, range 0-5) were derived for RCT data as well.

After the qualitative analyses of the interview material, interviewees were invited to provide feedback on the identified themes in August to September 2021 via an online validation questionnaire indicating their agreement with 'yes' or 'no'.

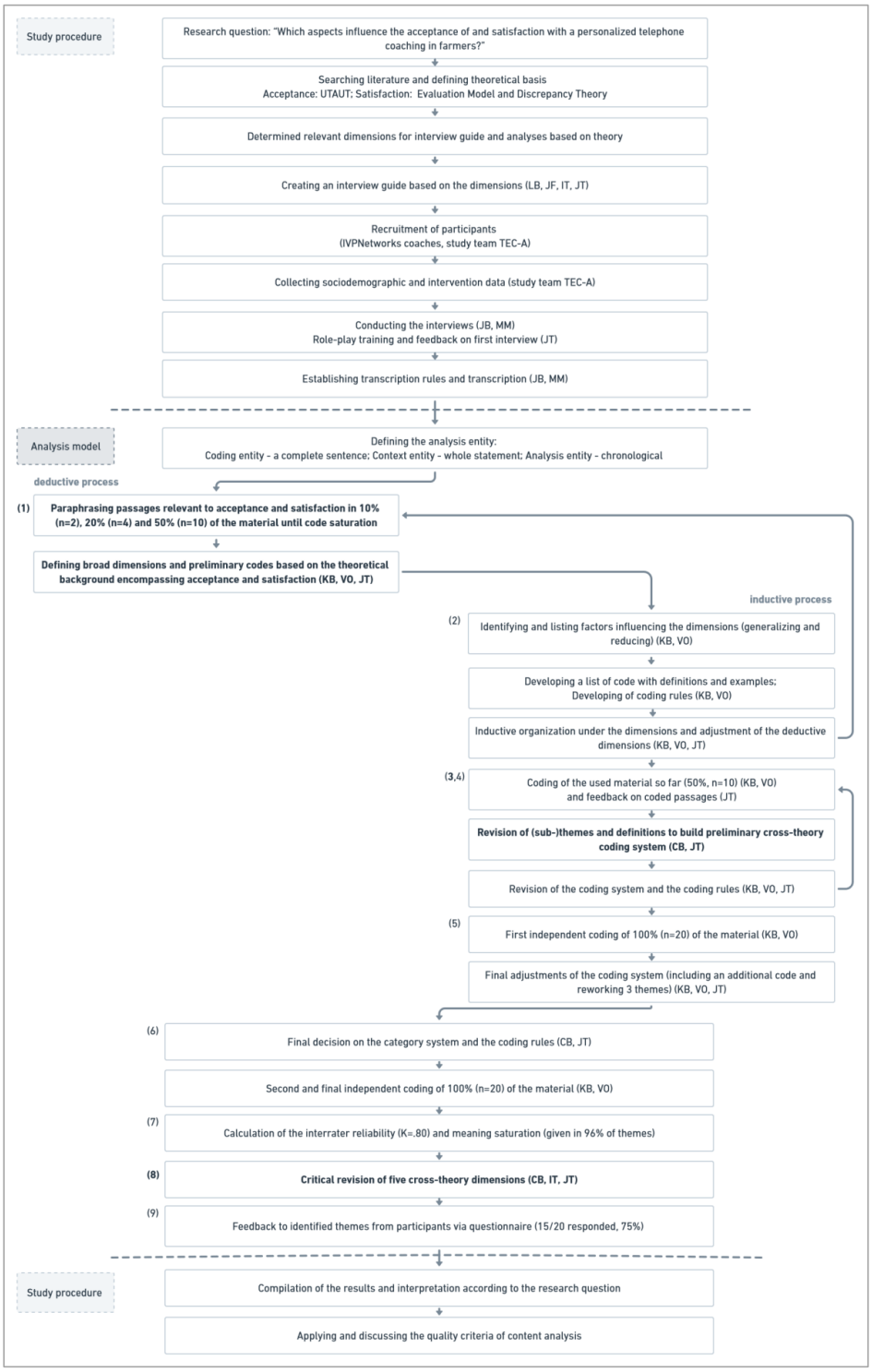
### 2.6 Data analyses

A content analysis was conducted according to standardized steps in qualitative research to structure and summarize interview information and identify themes relevant for the research question (Mayring, 2015; Schreier, 2014). Data collection and analyses procedure are illustrated in Figure 1 with bold printed steps elaborated in the text.

---

[8] QIDS-SR16: Quick Inventory of Depressive Symptomology
[9] CSQ-I: adapted Client Satisfaction Questionnaire for Internet-based Interventions,
[10] WAI-SR: Working Alliance Inventory - Short Revised,

**Figure 1.** Visualized data collection and analyses procedure as suggested by Mayring (Mayring, 2015).




*Note.* Bold printed steps are elaborated in detail in the text.

Adaptations to the representation of the theoretic components had to be made in different steps of the analyses in order to depict both concepts (i.e., acceptance and satisfaction) well. In step (1), combined dimensions were defined for components partly shared by both constructs (e.g. in treatment expectancy and performance expectancy). In step (3), after reviewing the 50% coded material, CB and JT decided not to include the theoretical components 'Continuity of care', 'Financing', and 'Availability' in the analysis, as the participants either answered too short or only described the status quo without evaluating it. Therefore, no insight could be obtained on how these components influenced satisfaction. The preliminary code system was formed to address the two categories, supporting (SF)[11] and hindering factors (HF)[12], which were each organized in cross-theory dimensions including 'Coaching Result', 'Coach', 'Organization', 'Participant Characteristics', and 'Telephone as Communication Medium'. Finally, in step (8), these five cross-theory dimensions were critically revised and arranged with HF and SF for each dimension (CB, IT, JT). Expectations were identified for the dimensions Coaching Result and Coach. No changes to codes or definitions were made. This was done to ensure a higher fit of mapping codes to the theoretical background and stringent communication of results.

MAXQDA 2020 (Sozialforschung GmbH, 2021) was used to conduct qualitative analysis and R 4.1.0 (R Core Team, 2020) and Excel (Microsoft Corporation, 2018) were used for quantitative analysis.

## 3 Results

### 3.1 Participant characteristics and intervention use

Half of the interviewees were male (n=10, 50%), most lived in partnership (n = 18, 90%), had middle education (n=14, 70%), and were, on average, 55 years old (SD=8.62). Interviewees recruited from TEC-A reported high satisfaction with the intervention (M = 30.06, SD = 3.75) and a good working alliance with their coach (M = 4.30, SD = 0.48) six months after study begin. Interview participants represented RCT coaching participants fairly in demographics (Table 2) and intervention use (Table 3). Time between end of coaching and interview varied between two weeks and six months with two interviewees (10%) still in the prolonged coaching at time of the interview.

[11] SF: supporting factor
[12] HF: hindering factor

**Table 2**. Interview participant's characteristics compared to not interviewed coaching participants in the RCT TEC-A.

| Variable | Interview participants (n=20) | | | Not interviewed coaching participants in TEC-A (n=143) | | |
|---|---|---|---|---|---|---|
| | **M** | **SD** | **range** | **M** | **SD** | **range** |
| **Age** | 55.00 | 8.62 | 31-74 | 52.87 | 9.30 | 30-78 |
| **QIDS-SR16 at baseline**[a] | 9.06 | 3.68 | 4-16 | 9.90 | 4.19 | 0-24 |
| **QIDS-SR16 after 6 months**[a] | 5.50 | 3.81 | 1-14 | 6.52 | 3.91 | 0-20 |
| | n | % | | n | % | |
| **Sex** | | | | | | |
| Male | 10 | 50 | | 75 | 52 | |
| Female | 10 | 50 | | 68 | 48 | |
| **Birthplace** | | | | | | |
| Germany | 20 | 100 | | 141 | 99 | |
| Other | 0 | 0 | | 2 | 1 | |
| **Relationship status** | | | | | | |
| No partnership | 2 | 10 | | 10 | 7 | |
| With partnership | 18 | 90 | | 133 | 93 | |
| **Children** [a] | | | | | | |
| Children | 15 | 88 | | 128 | 90 | |
| No children | 2 | 12 | | 15 | 10 | |
| **Education level** | | | | | | |
| Low | 0 | 0 | | 14 | 10 | |
| Middle | 14 | 70 | | 91 | 64 | |
| High | 6 | 30 | | 38 | 26 | |
| **Profession** | | | | | | |
| Entrepreneur | 13 | 65 | | 84 | 59 | |
| Entrepreneurs spouse | 4 | 20 | | 31 | 22 | |
| Pensioner | 2 | 10 | | 14 | 9 | |
| Family member | 1 | 5 | | 13 | 9 | |
| Incapacitated for work | 0 | 0 | | 1 | 1 | |
| **Experience with psychotherapy** [a] | | | | | | |
| No experience with psychotherapy | 14 | 82 | | 115 | 80 | |
| Experience with psychotherapy | 3 | 17 | | 28 | 20 | |

*Note.* [a]Data only available for TEC-A participants (T0: $n_{\text{interviewed RCT participants}}$=17, $n_{\text{not interviewed RCT participants}}$ =143; post: $n_{\text{interviewed RCT participants}}$ =16, $n_{\text{not interviewed RCT participants}}$ =121). Abbreviations: T0: baseline assessment study begin; T1: post assessment 6 months after randomization; QIDS-SR16: Quick Inventory of Depressive Symptomology (range 0-27).

**Table 3.** Intervention related characteristics for interviewees and not interviewed coaching participants in RCT TEC-A six month after study begin.

| Variable | Interview participants (n=20) | | | Not interviewed coaching participants in TEC-A (n=143) | | |
|---|---|---|---|---|---|---|
| | **M** | **SD** | **range** | **M** | **SD** | **range** |
| **Days from enrollment to first contact by coach**[a] | 2.41 | 2.81 | 0-11 | 3.15 | 3.41 | 0-13 |
| **Number of sessions** | 15.40 | 4.59 | 6-24 | 13.18 | 6.10 | 1-32 |
| **Coaching volume in minutes** | 732.30 | 239.33 | 265-1057 | 643.07 | 332.26 | 50-1598 |
| **Duration in weeks** | 25.74 | 6.37 | 10.29-44.43 | 24.39 | 8.09 | 1.43-38.86 |
| **CSQ-I** [a] | 30.06 | 3.75 | 18-32 | 27.92 | 5.75 | 9-32 |
| **WAI-SR total**[a] | 4.30 | 0.48 | 3.5-5.0 | 4.02 | 0.68 | 1.42-5.0 |
| WAI-SR bond[a] | 4.48 | 0.42 | 3.5-5.0 | 4.25 | 0.70 | 1.5-5.0 |
| WAI-SR task[a] | 4.16 | 0.52 | 3.0-5.0 | 3.77 | 0.78 | 1.0-5.0 |
| WAI-SR goal[a] | 4.27 | 0.69 | 3.0-5.0 | 4.05 | 0.75 | 1.0-5.0 |
| **Days between coaching end and interview**[b] | 41.44 | 46.68 | 5-191 | | | |
| | **n** | **%** | | **n** | **%** | |
| **Prolonging of coaching** | | | | | | |
| No | 18 | 90 | | 115 | 81 | |
| Yes | 2 | 10 | | 27 | 19 | |
| **Documented discharge reason** | | | | | | |
| Improved/stabilized quality of life | 14 | 70 | | 84 | 59 | |
| Stabilized and further support recommended/ arranged | 6 | 30 | | 42 | 30 | |
| *Intervention dropout:* | | | | | | |
| Withdrew consent for intervention | 0 | 0 | | 7 | 5 | |
| Lack of compliance | 0 | 0 | | 5 | 3 | |
| Intervention not fitting | 0 | 0 | | 3 | 2 | |
| Other | 0 | 0 | | 1 | 1 | |

*Note.* [a]Data only available for TEC-A participants ( $n_{\text{interviewed RCT participants}}$=17, T0: $n_{\text{not interviewed RCT participants}}$=119); [b]n=18 since two participants were still in coaching, final intervention data reported in other variables. Session information only reported for participants with at least one coaching session ($n_{\text{interviewed partcipants TEC-A and ImpleMentIT}}$ =20, $n_{\text{not interviewed RCT participants}}$=142). Abbreviations: CSQ-I: Client Satisfaction Questionnaire for Internet-based Interventions (range 8-32); WAI-SR: Working Alliance Inventory - Short Revised (range 0-5).

**3.2 Aspects influencing acceptance and satisfaction**

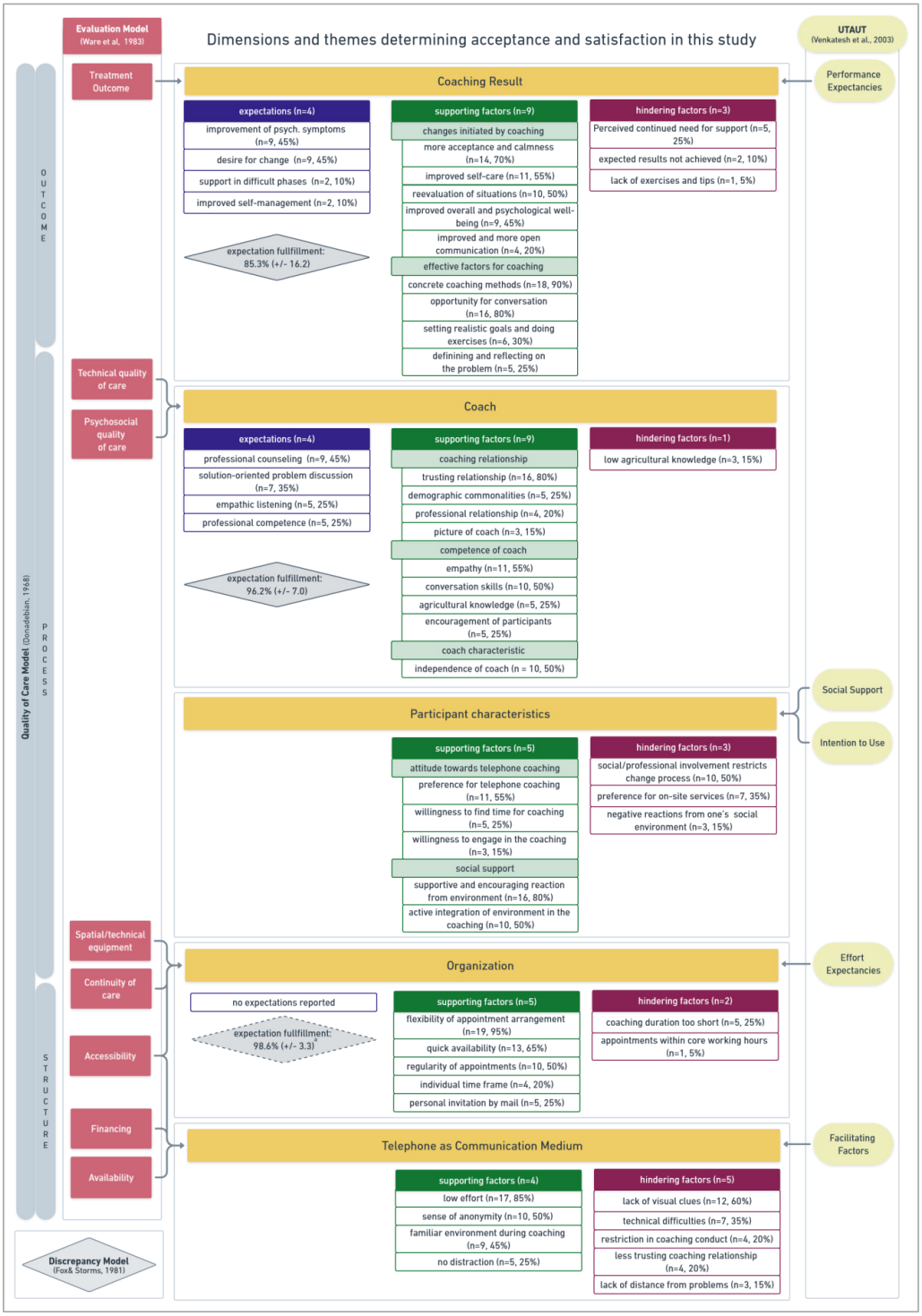
In total, eight expectations (E)[13], 32 SF, and 14 HF that can influence acceptance and satisfaction were identified and organized under five dimensions: Coaching result (E=4, SF=9, HF=3), Coach (E=4, SF=9, HF=1), Organization (SF=5, HF=2), Participant Characteristics (SF=5, HF=3) and Telephone as Communication Medium (SF=4, HF=5).

gives an overview over all identified themes and how the theoretical components are represented in each of the reported dimensions. The most often reported themes for expectations, supporting and hindering factor for each cross-theory dimension are described below and illustrated with a quote. All emerged themes with definitions and quotas are presented in Supplement 2 in the form of the coding system used for the analyses.

[13] E: Expectations

**Figure 2**

*Overview on all identified themes with regard to the theoretic foundation of the study.*



*Note.* Arrows indicate which theoretical component is represented in the dimension of the coding system.

[a]Participants reported experienced instead of expected organization. Expectation fulfillment rating can be interpreted as degree of satisfaction with the addressed organization

3.2.1 Coaching Result

As expectations toward the coaching result, most often interviewees reported either *improvement of psychological symptoms* and/or a *general desire for change* (each n=9, 45%). An interview quote illustrating these themes is as follows:

> *'I think, I generally had the hope that I can find a way out of this dark thinking patterns. So, I did not really have precise expectations in that sense, but I really had this kind of basic need, that this darkness can become lighter.' (p11).*

According to participants, *concrete coaching methods* such as breathing exercises and goal setting exercises (n=18, 90%) were most helpful to reach their goals.

> *'For example, I am looking at a table right now: (…) what I want to do differently, then when I start with it and how I reward myself when I completed it.' (p4).*

Less HFs were identified that could have influenced satisfaction with the coaching result. A quarter of interviewees (n=5, 25%) described *perceived continued need for support* ranging from wanting information about further help services to starting face-to-face therapy and hence indicated that their desired goal was not completely met:

> *'That you could at least have a therapy place once a month or here [on-site].' (p3).*

***3.2.2 Coach*** With regard to their coach, participants most often expected a person that can deliver *professional counseling* and give advice in their situation (n=9, 45%).

> *'That there is somebody, who can give me hints or advice, how I can get out of this situation in which I was in at that point (…).' (p11).*

In terms of SFs, most of the participants (n=16, 80%) described to have experienced a very *trusting relationship* with their coach:

> *'That developed over time and then it was, well, trusting and a very good relationship, but also respectful and friendly.' (p1).*

Only one HF, namely *low agricultural knowledge* of the coach, was named by four participants (20%).

*'When we talked about my role in the family and on the farm (…) it would have been helpful if she had a bit more insigh' (p9).*

***3.2.3*** ***Organization*** The most named (supporting) theme of all (n=19, 95%) was *flexibility of appointment arrangement* which facilitated participation. Interviewees described an easy integration of scheduling into their everyday life:

*'If I was really busy on the field for once, one postponed the appointment. That worked even on relative short notice.' (p5).*

A quarter of participants (n=5) perceived the *coaching duration as too short* to achieve their goal and that possibly diminishing satisfaction:

*'Because with half a year, you are not simply out of the problem like that.' (p15).*

***3.2.4*** ***Participant Characteristics***The most often named SF on participant level was *supportive and encouraging reaction from environment* (n=16, 80%):

*'My wife thought it was really great that I was doing something like that.' (p5).*

Half of the participants described their high *social and professional involvement* as restricting for the change process initiated by the coaching and as hindering for achieving their goals (n=10, 50%):

'*That you're so tied up in the job that there's a lack of opportunity.' (p7).*

***3.2.5*** ***Telephone as the Communication Medium.*** Most controversial topic was the telephone as the communication medium yielding most HFs (n=5) but likewise four SFs. Most named supporting factor was the perceived *low effort* through telephone delivery (n=17, 85%):

*'That was crucial for me – that it was over the phone. I would not have gotten into a car and driven somewhere for a quarter of an hour. (…)' (p1).*

Over half the participants reported the *lack of visual cues* via telephone conversation as the most common hindering factor (n=12, 60%):

> *'Except that you can't see whether the coach is listening or sometimes a bit of facial expression or certain feedback (…). That is missing a bit.' (p20).*

### 3.3 Degree of satisfaction

Satisfaction with the intervention was indicated by the numerical ratings on expectation fulfillment. The above reported expectations toward the coaching result (n=20, M=85%, SD=16%) and the coach (n=20, M=96%, SD=7%) were met to a high degree. Even though no expectation but actual experiences were named with regard to the organization, fulfillment were rated high as well (n=18, M=99%, SD=3%).

### 3.4 Results of the validation questionnaire

All interviewees were invited and 17/20 (85%) responded. In the validation questionnaire, on average, 66% (range: 0%-100%) of participants agreed to the identified themes: Agreement rates were higher for expectations (M=87%, range 56%-100%) and SFs (M = 81%, range: 44%–100%) than for HFs (M=20%, range 0%-50%). Supplement 3 provides a comparison of qualitative coding results and validation.

## 4 Discussion

In this interview study, farmers participating in a personalized telephone coaching to prevent depression were asked for their experiences with the intervention in order to identify factors that influence their acceptance and satisfaction.

### 4.1 Main findings

The influencing factors identified were organized under the dimensions 'Coaching Result', 'Coach', 'Organization', 'Participant Characteristics', and 'Telephone as Communication Medium'. More than half of interviewees stated to have preferred the telephone offer over a comparable on-site intervention from the beginning, and all participants used the telephone coaching for at least 10 weeks, which indicated acceptance for the intervention. The interviewees reported an almost complete fulfillment of their expectations for the coaching, which indicated their satisfaction with the intervention. Correspondingly, twice more supporting than hindering factors to satisfaction and acceptance were identified, indicating positive experiences by all interviewees.

With regard to the theoretic components, codings for the coaching result, the organization, and the telephone as communication medium revealed SFs and HFs to both acceptance and satisfaction. Participants'

experiences with respect to the coach and especially the good coach-participant relationship primarily supported satisfaction with the intervention, while reported participant characteristics (having a positive attitude and social support) mainly facilitated acceptance.

The remote delivery format via telephone was viewed divisive by interviewees. Almost all participants highlighted the perceived low effort of the telephone coaching (e.g. no need to change clothes or travel) and half of the interviewees valued the feeling of (visual) anonymity and staying in a familiar environment during the telephone coaching. On the other hand, telephone-related disadvantages (e.g. lack of visual clues, technical problems) were the most reported hindering aspects and deemed to diminish the coaching conduct or effectiveness in participants' views.

The results of the validation questionnaire supported the interpretation that the HFs named only by a few interviewees (e.g. coaching being not practical enough or too short) were indeed rare as average agreement rates for these were low.

### 4.2 Comparison with literature

The described participants' experiences with personalized telephone coaching are very similar to earlier studies reporting experiences by MDD patients with more standardized CBT over telephone. These reflect the advantages of the telephone use (e.g. convenience, anonymity) as well as potential problems of it (e.g. technical problem, lack of visual feedback) (Bee et al., 2010; Haller et al., 2019). The farming population in the current study added the flexibility and adaptation to occupational background as part of the personalized intervention as facilitating intervention use. Despite sporicidal skepticism, the good therapeutic relationship reported in our and earlier studies (Bee et al., 2010; Haller et al., 2019) is in line with studies reporting no difference in therapeutic relationship in telephone- or face-to-face therapy (Irvine et al., 2020; Stiles-Shields et al., 2014).

The findings in this qualitative study are coherent with the quantitative results observed in the associated RCT (Thielecke et al., 2022) that suggested good acceptance (e.g. only 11% intervention dropout) and high satisfaction with the intervention (CSQ-I: M=28.17, SD=5.58) in farmers. Aligned with the experienced changes initiated by the coaching (e.g. more acceptance, improved self-care) reported by the interviewees, the RCT showed a reduction in depressive symptom severity and further mental health outcomes including stress and anxiety. Similarly, the stated supporting aspects of a good coaching relationship and a realistic goal setting in the interviews corresponded to the observed ratings on the subscales of the working alliance questionnaire in the RCT.

Both the UTAUT model and the theories for patient satisfaction could be applied to our remote health care setting with slight adaptations. Some aspects seem to be especially worth looking at with regard to our farming sample:

In terms of acceptance, perceived low effort and quick availability when compared to on-site psychotherapy (Bundespsychotherapeutenkammer, 2019) were often named by interviewees, suggesting those factors might support the uptake of mental health care. Researchers on telehealth interventions have argued that remotely delivered health service might offer low-threshold access to (mental) health care (Ebert et al., 2018; Snoswell et al., 2021) and improves health care in underserved areas (Myers, 2019; Salize et al., 2007), which is supported for farming communities by our interviewees.

For the experienced satisfaction with the intervention, the personalization of the intervention seems to be a promoting factor. Multiple themes under different dimensions were named by participants that indicate overall personalization:

(1) Individual time composition: Alongside the reduced effort by eliminating traveling times, the flexible arrangement of appointments was highlighted that allows for an easy integration into everyday life. Given time restrictions by often long and unpredictable working hours in this occupational group (Braun et al., 2022; Daghagh Yazd et al., 2019; Vayro et al., 2020), individual time composition appears to be especially important for adherence and might help to dissolve the conflict of prioritizing work over help-seeking reported in other studies (Vayro et al., 2020).

(2) Appreciation for occupational background: The study invitation, addressing common stressors in farming occupations, was perceived as very personal by interviewees. Coaches' agricultural knowledge, if present, was perceived as supporting. Open address of farming-specific topics and expression of appreciation in the communication about mental health care might therefore be important for usage, which are highlighted by studies focusing on help-seeking behavior and intervention design in this population (Lotfi et al., 2010; Vayro et al., 2020).

(3) Adaptation of coaching methods: The reported realistic goal setting and exercises allowing for constrains by interviewees' work life reflect a personalized procedure in this intervention. Knowing this occupational background allows coaches to adapt their approach better to the participants' need (Cole and Bondy, 2020; Stumpp and Sauer-Zavala, 2021).

The mentioned supporting aspects of intervention design being adapted to participants' background, therapeutic relationship, anonymity, and easy accessibility as drivers for intervention use were also found in interviews with farmers participating in an online intervention to prevent depression (Freund et al., 2022). However, as expected in the beginning, some of the typical barriers for internet-based interventions such as data security worries, lack of personalization, insufficient computer skills, or internet availability (Braun et al., 2022; Freund et al., 2022; Gerhards et al., 2011; Holst et al., 2017; Mayer et al., 2019) were not mentioned by our interviews with regard to the telephone coaching. Telephone interventions therefore enrich the possibilities in the realm of remotely delivered health intervention especially for individuals with restricted internet use and

in need of a flexible, highly personalized intervention in which they can participate from their familiar environment.

### 4.3 Implication for practice and research

Some direct implications from the interviews for coaching in farmers can be derived. Personalization as a key aspect for satisfaction could be further enhanced by offering additional videoconferencing, which has been accepted and effective for psychotherapy (Berryhill et al., 2019). Videoconferencing could serve those who are missing visual clues during telephone coaching. It should, however, not be forced on those enjoying the visual anonymity as results from a survey study suggest that telephone interventions might overcome fear of being judged by visual appearance and promote disclosure to mental health experts (Lingley-Pottie et al., 2013). The adaptation of intervention content to the farming background as a further mean of personalization should be kept up and enhanced by training coaches in farming-specific topic as already suggested in other qualitative studies on farmers' mental health (Cole and Bondy, 2020; Vayro et al., 2021).

This study results support the ongoing implementation in routine care by stressing that telehealth can play an important role in the promotion of mental health in underserved communities by eliminating travel times, providing anonymity and a low-threshold access to mental health care. However, besides the overall positive attitudes to telephone coaching and the reported advantages, almost half of the participants stated that it would not have been their first choice given alternatives. Personalization of health care should therefore begin with offering different delivery formats (face-to-face, internet- and telephone-based, blended approaches) and settings (group vs. individual programs) and supporting participants in accessing the health care service that best fits their clinical need and individual preference (Smit and Van, 2013; Stumpp and Sauer-Zavala, 2021).

In routine care, clients are counseled by the SVLFG call center about available services and can choose between different face-to-face, telephone-, and internet-based interventions in the preventive setting. However, in the present study it became obvious that participants could not always clearly name their expectations toward the intervention. Introducing standardized assessments of preferences (e.g. on-site, remote), expectations (e.g. amount of guidance, time investment), and personal characteristics (e.g. treatment experience, computer skills) could help call center agents to guide the decision process to what best serves the individual's needs. Besides individual preferences, this process should also take into account the evidence for the interventions in question and current treatment guidelines as well as considering the associated costs of the different delivery methods.

Scalability is one of the main arguments for implementation of telehealth compared to face-to-face services, but different delivery modes are also associated with varying costs and time investment from health care professionals (e.g. one-on-one coaching vs. unguided internet-based intervention) posing a potential

bottleneck in care distribution (Berryhill et al., 2019; Nobis et al., 2018). More research is needed to determine how the best match between participant characteristics, health care resources, and health outcome can be achieved and if combinations of different approaches might be beneficial.

### 4.4 Strengths and limitations

While the theory-based approach, the documented standardized steps, and the triangulation with quantitative results in the same population should be highlighted as strengths of this qualitative study, some limitations should be addressed.

This study provides insights in how the different elements of the UTAUT and Evaluation Model contribute to the acceptance of and the satisfaction with the telephone coaching from a farmer's perspective even though the two constructs could not be distinguished clearly from each other. A conjoint category system was built for both constructs after defining deductive cross-theory dimensions, which solved the construct overlap but reduced the specificity of the study. These kind of changes are to be expected in qualitative research and should help future researchers in preparing their studies (Schreier, 2014).

More adaption of the analysis process was necessary with regard to the structural elements in the Evaluation Model. The interviews yielded no deeper understanding of three of the proposed characteristics and their influence on satisfaction than the model originally suggested and were thus excluded from the analyses. In both routine care and study setting, participants could take part with no extra costs involved and were coached continuously by the same person, thus the theoretical components of 'Financing' and 'Continuity of Care' probably rather promoted satisfaction than hindering it. In terms of 'Availability', most participants were recruited from a study setting (TEC-A). Therefore, availability was dependent on study participation and randomization luck. Only three participants were recruited under routine care conditions (ImplementIT) resulting in a different enrollment process. The coaching began quickly after enrollment and therefore could pose as a supporting factor in both settings. While routine care experience might have been preferable in order to evaluate in a more detail manner on factors like accessibility and availability, in this study, TEC-A participants could be recruited earlier, thus interview results could already be used in the early stages of the implementation process (Glasgow et al., 2003).

For the expectations toward the intervention, it was only feasible to assess them after participants experienced the intervention, thus making reported (fulfilled) expectations prone to hindsight bias (Groß et al., 2017). A priori acceptability studies are warranted for our intervention and target population (Gunn et al., 2022, 2021). We cannot foreclose that our sample might be especially selective and highly satisfied or motivated because of convenient and small sampling. Because of the limited number of unsatisfied participants and intervention dropouts, purposefully sampling those might have increased meaning saturation or identified

additional themes but could also have screwed representativeness. Because of the individual time frame of the coaching, it could not be avoided that interviews were conducted with different intervals between coaching and interview, possibly changing perception, for example, on achieved changes initiated by coaching which only become noticeable after time (Reio et al., 2017; Smidt et al., 2009).

Overall, high satisfaction and the sampling method also could explain why we did not reach meaning saturation for all identified hindering themes and why HFs were often not supported by many participants in the validation questionnaire. Recruiting more coaching participants might have improved meaning saturation, however, because of the heterogeneity of the intervention and the deductive-inductive analysis approach, achieving full saturation might not be reasonably assumed for our research question (O'Reilly and Parker, 2013). Some HFs were only named rarely, however, the personalized nature of the intervention should be able to address individual treatment barriers during coaching.

Finally, even though we identified the personalization of the intervention as a key factor for the coaching practice, heterogeneity in coaching conduct and adjustment to the individual challenges generalizability of our research results (Stumpp and Sauer-Zavala, 2021). The next step to evidence-based personalized health interventions should be systems for monitoring and evaluating personalization in telehealth as suggested by Aswad and Lessard (Aswad and Lessard, 2021).

## 5 Conclusion

Theory-based insights were gained on how dimensions of acceptance and satisfactions influenced interviewed farmer's evaluation of a personalized telephone coaching to prevent depression. Perceived changes initiated by coaching, low effort through telephone conduct, and the personalization in terms of time management and farming specificity were perceived key for this overall accepted and satisfactory intervention. While these results support ongoing implementation, more research is needed to optimize personalization and resource input in telehealth care to support sustainable implementation into routine care and reduce depression burden.

**Final declarations**

**Conflict of Interest**

IT reports to have received fees for lectures/workshops in the e-mental-health context from training institutes and congresses for psychotherapists. She was the scientific project leader for the research project ImpleMentAll (11/2017 -03/2021, funded by the European Commission) at the Institute for health training online (GET.ON) which aimed to investigate the effectiveness of tailored implementation strategies compared to implementation as usual. DDE has served as a consultant to/on the scientific advisory boards of Sanofi, Novartis, Minddistrict, Lantern, Schoen Kliniken, Ideamed and German health insurance companies (BARMER, Techniker Krankenkasse) and a number of federal chambers for psychotherapy. DDE and MB are current or former stakeholders of the GET.ON Institute for health training online, which aims to implement scientific findings related to digital health interventions into routine care. MB is scientific advisor of mentalis GmbH, a provider for digital aftercare. HB reports to have received consultancy fees, fees for lectures or workshops from chambers of psychotherapists and training institutes for psychotherapists and license fees for an Internet-intervention. JF, CB, JT, LB report no conflicts of interest.

**Funding**

The German insurance company SVLFG provided a financial contribution to the Friedrich-Alexander University Erlangen-Nürnberg and Ulm University as expense allowance. SVLFG had no role in study design, data collection, analysis, interpretation and decision to prepare or publish this manuscript.

**Ethical approval**

The ethics committee of the Friedrich-Alexander-Universität Erlangen-Nürnberg had approved this qualitative interview study as part of the implementation study ImplementIT and was registered in the German Clinical Trial Registration (DRKS00017078). Interviewees participated in the telephone coaching as clients under routine care conditions during the pilot phase of ImplementIT or as study participants in the associated RCT (TEC-A, trial registration: DRKS00015655). All participants provided extra written informed consent for the interviews, data-matching with RCT data (if applicable), and audio-recording.

**Contributorship**

DDE, HB and MB obtained funding for the SVLFG evaluation project. JT, IT, JF and LB developed the study design and interview guide. JT was responsible for recruitment of interview participants, coordination, collection of interview data, development of code system and analysis. CB and IT provided feedback on the analytically derived code system. IT supervised and contributed to the trial management, data collection and analyses as

operational lead of the project and qualitative research expert. JT drafted the manuscript. IT supervised and contributed to the further writing of the manuscript. All authors provided critical revision of the article and approved the final manuscript.

**Acknowledgements**

Authors would like to thank Jana Burger (JB) and Melanie Martin (MM) for the conduction and transcription of the interviews and Katja Boué (KB) and Vanessa Oth (OV) who were involved in the development of the code system and as independent rater in the analysis as part of their theses. Further, authors thank Sarah Banellis and Johanna Finitzer for participant their engagement in enrolling the participants throughout the study. Authors thank the participants for their willingness to take part in the interview, IVPNetworks coaches for helping recruit interview partners and delivering the intervention and the SVLFG as implementing organization. Authors further thank Jennifer Kulke, Johanna Finitzer and Maria Bachmakova for the forward-backward translation of the code system and quotas. Finally, authors would like to thank enago Academy for proofreading the manuscript and TUM Graduate School for covering the costs for this service as part of JT's PhD program.

**References**

Aswad, O., Lessard, L., 2021. Toward a service design method for telehealth personalization. AMCIS 2021 Proc. 1, 1007.

Audacity Team, 2021. Audacity 2.3.3.

Bee, P.E., Lovell, K., Lidbetter, N., Easton, K., Gask, L., 2010. You can't get anything perfect: "User perspectives on the delivery of cognitive behavioural therapy by telephone." Soc. Sci. Med. 71, 1308–1315.

Berryhill, M.B., Culmer, N., Williams, N., Halli-Tierney, A., Betancourt, A., Roberts, H., King, M., 2019. Videoconferencing Psychotherapy and Depression: A Systematic Review. Telemed. e-Health 25, 435–446.

Blum, K., 1998. Patientenzufriedenheit bei ambulanten Operationen. Juventa, Weinheim.

Boß, L., Lehr, D., Reis, D., Vis, C., Riper, H., Berking, M., Ebert, D.D., 2016. Reliability and Validity of Assessing User Satisfaction With Web-Based Health Interventions. J. Med. Internet Res. 18, e234.

Braun, L., Freund, J., Thielecke, J., Baumeister, H., Ebert, D.D., Titzler, I., 2022. Barriers to and Facilitators of Engaging With and Adhering to Guided Internet-Based Interventions for Depression Prevention and Reduction of Pain-Related Disability in Green Professions: Mixed Methods Study. JMIR Ment. Heal. 9, e39122.

Braun, L., Titzler, I., Ebert, D.D., Buntrock, C., Terhorst, Y., Freund, J., Thielecke, J., Baumeister, H., 2019. Clinical and cost-effectiveness of guided internet-based interventions in the indicated prevention of depression in green professions (PROD-A): study protocol of a 36-month follow-up pragmatic randomized controlled trial. BMC Psychiatry 19, 278.

Braun, L., Titzler, I., Terhorst, Y., Freund, J., Thielecke, J., Ebert, D.D., Baumeister, H., 2021a. Are guided internet-based interventions for the indicated prevention of depression in green professions effective in the long run? Longitudinal analysis of the 6- and 12-month follow-up of a pragmatic randomized controlled trial (PROD-A). Internet Interv. 26, 100455.

Braun, L., Titzler, I., Terhorst, Y., Freund, J., Thielecke, J., Ebert, D.D., Baumeister, H., 2021b. Effectiveness of guided internet-based interventions in the indicated prevention of depression in green professions (PROD-A): Results of a pragmatic randomized controlled trial. J. Affect. Disord. 278, 658–671.

Bromet, E.J., Andrade, L.H., Bruffaerts, R., Williams, D.R., 2018. Major Depressive Disorder. In: Scott, K.M., de Jonge, P., Stein, D.J., Kessler, R.C. (Eds.), Mental Disorders Around the World: Facts and Figures from the WHO World Mental Health Surveys. Cambridge University Press, Cambridge, pp. 41–56.

Bundespsychotherapeutenkammer, 2019. Ein Jahr nach der Reform der Psychotherapie-Richtlinie. Wartezeiten 2018 [WWW Document]. URL https://www.bptk.de/publikationen/bptk-studie/

Cole, D.C., Bondy, M.C., 2020. Meeting Farmers Where They Are–Rural Clinicians' Views on Farmers' Mental Health. J. Agromedicine 25, 126–134.

Cuijpers, P., Koole, S.L., van Dijke, A., Roca, M., Li, J., Reynolds, C.F., 2014. Psychotherapy for subclinical depression: meta-analysis. Br. J. Psychiatry 205, 268–274.

Cuijpers, P., Pineda, B.S., Quero, S., Karyotaki, E., Struijs, S.Y., Figueroa, C.A., Llamas, J.A., Furukawa, T.A., Muñoz, R.F., 2021. Psychological interventions to prevent the onset of depressive disorders: A meta-analysis of randomized controlled trials. Clin. Psychol. Rev. 83, 101955.

Cuijpers, P., van Straten, A., Warmerdam, L., van Rooy, M.J., 2010. Recruiting participants for interventions to prevent the onset of depressive disorders: Possibile ways to increase participation rates. BMC Health Serv. Res. 10, 181.

Daghagh Yazd, Wheeler, Zuo, 2019. Key Risk Factors Affecting Farmers' Mental Health: A Systematic Review. Int. J. Environ. Res. Public Health 16, 4849.

Damschroder, L.J., Aron, D.C., Keith, R.E., Kirsh, S.R., Alexander, J.A., Lowery, J.C., 2009. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement. Sci. 4, 50.

Dollman, J., Gunn, K.M., Hull, M.J., 2021. Sociodemographic Predictors of Attitudes to Support Seeking From a Medical Doctor or Other Health Provider Among Rural Australians. Int. J. Behav. Med. 28, 616–626.

Donadebian, A., 2005. Evaluating the Quality of Medical Care. Milbank Q. 83, 691–729.

Ebert, D.D., Van Daele, T., Nordgreen, T., Karekla, M., Compare, A., Zarbo, C., Brugnera, A., Øverland, S., Trebbi, G., Jensen, K.L., Kaehlke, F., Baumeister, H., 2018. Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health. Eur. Psychol. 23, 167–187.

Fox, J.G., Storms, D.M., 1981. A different approach to sociodemographic predictors of satisfaction with health care. Soc. Sci. Med. Part A Med. Psychol. Med. 15, 557–564.

Freund, J., Buntrock, C., Braun, L., Thielecke, J., Baumeister, H., Berking, M., Ebert, D.D., Titzler, I., 2022. Digital prevention of depression for farmers? A qualitative study on participants' experiences regarding determinants of acceptance and satisfaction with a tailored guided internet intervention program. Internet Interv. 29, 100566.

Freund, J., Buntrock, C., Braun, L., Thielecke, J., Baumeister, H., Berking, M., Ebert, D.D., Titzler, I., n.d. Digital Prevention of Depression for Farmers? A Qualitative Study on Participants' Experiences regarding Determinants of Acceptance and Satisfaction with a Tailored Guided Internet Intervention Program. Internet Interv. (under review).

Freund, J., Titzler, I., Thielecke, J., Braun, L., Baumeister, H., Berking, M., Ebert, D.D., 2020. Implementing internet- and tele-based interventions to prevent mental health disorders in farmers, foresters and gardeners (ImplementIT): study protocol for the multi-level evaluation of a nationwide project. BMC Psychiatry 20, 424.

Gerhards, S.A.H., Abma, T.A., Arntz, A., de Graaf, L.E., Evers, S.M.A.A., Huibers, M.J.H., Widdershoven, G.A.M., 2011. Improving adherence and effectiveness of computerised cognitive behavioural therapy without support for depression: A qualitative study on patient experiences. J. Affect. Disord. 129, 117–125.

Glasgow, R.E., Lichtenstein, E., Marcus, A.C., 2003. Why Don't We See More Translation of Health Promotion Research to Practice? Rethinking the Efficacy-to-Effectiveness Transition. Am. J. Public Health 93, 1261–1267.

Groß, J., Blank, H., Bayen, U.J., 2017. Hindsight Bias in Depression. Clin. Psychol. Sci. 5, 771–788.

Gunn, K.M., Barrett, A., Hughes-Barton, D., Turnbull, D., Short, C.E., Brumby, S., Skaczkowski, G., Dollman, J., 2021. What farmers want from mental health and wellbeing-focused websites and online interventions. J. Rural Stud. 86, 298–308.

Gunn, K.M., Skaczkowski, G., Dollman, J., Vincent, A.D., Short, C.E., Brumby, S., Barrett, A., Harrison, N., Turnbull, D., 2022. Combining Farmers' Preferences With Evidence-Based Strategies to Prevent and Lower Farmers' Distress: Co-design and Acceptability Testing of ifarmwell. JMIR Hum. Factors 9, e27631.

Gutiérrez-Rojas, L., Porras-Segovia, A., Dunne, H., Andrade-González, N., Cervilla, J.A., 2020. Prevalence and correlates of major depressive disorder: A systematic review. Brazilian J. Psychiatry 42, 657–672.

Haller, E., Besson, N., Watzke, B., 2019. "unrigging the support wheels" - A qualitative study on patients' experiences with and perspectives on low-intensity CBT. BMC Health Serv. Res. 19, 1–13.

Harst, L., Lantzsch, H., Scheibe, M., 2019. Theories predicting end-user acceptance of telemedicine use: Systematic review. J. Med. Internet Res.

Heale, R., Forbes, D., 2013. Understanding triangulation in research. Evid. Based Nurs. 16, 98–98.

Holst, A., Nejati, S., Björkelund, C., Eriksson, M.C.M., Hange, D., Kivi, M., Wikberg, C., Petersson, E.-L., 2017. Patients' experiences of a computerised self-help program for treating depression – a qualitative study of Internet mediated cognitive behavioural therapy in primary care. Scand. J. Prim. Health Care 35, 46–53.

Irvine, A., Drew, P., Bower, P., Brooks, H., Gellatly, J., Armitage, C.J., Barkham, M., McMillan, D., Bee, P., 2020. Are there interactional differences between telephone and face-to-face psychological therapy? A systematic review of comparative studies. J. Affect. Disord. 265, 120–131.

Judd, F., Jackson, H., Fraser, C., Murray, G., Robins, G., Komiti, A., 2006. Understanding suicide in Australian farmers. Soc. Psychiatry Psychiatr. Epidemiol. 41, 1–10.

Kruse, C.S., Krowski, N., Rodriguez, B., Tran, L., Vela, J., Brooks, M., 2017. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 7, e016242.

Lingley-Pottie, P., McGrath, P.J., Andreou, P., 2013. Barriers to mental health care: Perceived delivery system differences. Adv. Nurs. Sci. 36, 51–61.

Lotfi, L., Flyckt, L., Krakau, I., Mårtensson, B., Nilsson, G.H., 2010. Undetected depression in primary healthcare:

Occurrence, severity and co-morbidity in a two-stage procedure of opportunistic screening. Nord. J. Psychiatry 64, 421–427.

Mayer, G., Gronewold, N., Alvarez, S., Bruns, B., Hilbel, T., Schultz, J.H., 2019. Acceptance and expectations of medical experts, students, and patients toward electronic mental health apps: Cross-sectional quantitative and qualitative survey study. JMIR Ment. Heal. 6, 1–15.

Mayring, P., 2015. Qualitative Inhaltsanalyse, Qualitative Inhaltsanalyse. Beltz Verlag, Weinheim.

Microsoft Corporation, 2018. Microsoft Excel.

Munder, T., Wilmers, F., Leonhart, R., Linster, H.W., Barth, J., 2009. Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. Clin. Psychol. Psychother. 231–239.

Myers, C.R., 2019. Using Telehealth to Remediate Rural Mental Health and Healthcare Disparities. Issues Ment. Health Nurs. 40, 233–239.

Nobis, S., Ebert, D.D., Lehr, D., Smit, F., Buntrock, C., Berking, M., Baumeister, H., Snoek, F., Funk, B., Riper, H., 2018. Web-based intervention for depressive symptoms in adults with types 1 and 2 diabetes mellitus: a health economic evaluation. Br. J. Psychiatry 212, 199–206.

O'Reilly, M., Parker, N., 2013. "Unsatisfactory Saturation": A critical exploration of the notion of saturated sample sizes in qualitative research. Qual. Res. 13, 190–197.

Philippi, P., Baumeister, H., Apolinário-Hagen, J., Ebert, D.D., Hennemann, S., Kott, L., Lin, J., Messner, E.-M., Terhorst, Y., 2021. Acceptance towards digital health interventions – Model validation and further development of the Unified Theory of Acceptance and Use of Technology. Internet Interv. 26, 100459.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., Hensley, M., 2011. Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda. Adm. Policy Ment. Heal. Ment. Heal. Serv. Res. 38, 65–76.

R Core Team, 2020. R: A language and environment for statistical computing (version 4.1.0).

Reio, T.G., Rocco, T.S., Smith, D.H., Chang, E., 2017. A Critique of Kirkpatrick's Evaluation Model. New Horizons Adult Educ. Hum. Resour. Dev. 29, 35–53.

Rush, A.J., Trivedi, M.H., Ibrahim, H.M., Carmody, T.J., Arnow, B., Klein, D.N., Markowitz, J.C., Ninan, P.T., Kornstein, S., Manber, R., Thase, M.E., Kocsis, J.H., Keller, M.B., 2003. The 16-Item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol. Psychiatry 54, 573–583.

Salize, H.J., Rössler, W., Becker, T., 2007. Mental health care in Germany. Eur. Arch. Psychiatry Clin. Neurosci. 257, 92–103.

Samartzis, L., Talias, M.A., 2020. Assessing and improving the quality in mental health services. Int. J. Environ. Res. Public Health 17, 1–31.

Schreier, M., 2014. Varianten qualitativer Inhaltsanalyse: ein Wegweiser im Dickicht der Begrifflichkeiten. Forum Qual. Soc. Res. 15, 18.

Smidt, A., Balandin, S., Sigafoos, J., Reed, V.A., 2009. The Kirkpatrick model: A useful tool for evaluating training outcomes. J. Intellect. Dev. Disabil. 34, 266–274.

Smit, E.S., Van, J.C.M., 2013. Taking online computer-tailoring forward. Eur. Heal. Psychol. 17, 25–31.

Snoswell, C.L., Chelberg, G., De Guzman, K.R., Haydon, H.H., Thomas, E.E., Caffery, L.J., Smith, A.C., 2021. The clinical effectiveness of telehealth: A systematic review of meta-analyses from 2010 to 2019. J. Telemed. Telecare 1357633X2110229.

Sommerfeld, H., 2019. PhonerLite.

Sora, B., Nieto, R., Montesano del Campo, A., Armayones, M., 2021. Acceptance and Use of Telepsychology From the Clients' Perspective: Questionnaire Study to Document Perceived Advantages and Barriers. JMIR Ment. Heal. 8, e22199.

Sozialforschung GmbH, 2021. MAXQDA 2020, Software for qualatative data analyses.

Stiles-Shields, C., Kwasny, M.J., Cai, X., Mohr, D.C., 2014. Therapeutic alliance in face-to-face and telephone-administered cognitive behavioral therapy. J. Consult. Clin. Psychol. 82, 349–354.

Stumpp, N.E., Sauer-Zavala, S., 2021. Evidence-Based Strategies for Treatment Personalization: A Review. Cogn. Behav. Pract.

Thielecke, J., Buntrock, C., Titzler, I., Braun, L., Freund, J., Berking, M., Baumeister, H., Ebert, D.D., 2020. Clinical and Cost-Effectiveness of Personalized Tele-Based Coaching for Farmers, Foresters and Gardeners to Prevent Depression: Study Protocol of an 18-Month Follow-Up Pragmatic Randomized Controlled Trial (TEC-A). Front. Psychiatry 11, 125.

Thielecke, J., Buntrock, C., Titzler, I., Braun, L., Freund, J., Berking, M., Baumeister, H., Ebert, D.D., 2022. Telephone coaching for the prevention of depression in depression in farmers: Results from a pragmatic from a pragmatic randomized controlled trial. J. Telemed. Telecare 1357633X2211060.

Tong, A., Sainsbury, P., Craig, J., 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int. J. Qual. Heal. Care 19, 349–357.

van Zoonen, K., Buntrock, C., Ebert, D.D., Smit, F., Reynolds, C.F., Beekman, A.T.F., Cuijpers, P., 2014. Preventing the onset of major depressive disorder: A meta-analytic review of psychological interventions. Int. J. Epidemiol. 43, 318–329.

Vayro, C., Brownlow, C., Ireland, M., March, S., 2020. 'Farming is not Just an Occupation [but] a Whole Lifestyle': A Qualitative Examination of Lifestyle and Cultural Factors Affecting Mental Health Help-Seeking in Australian Farmers. Sociol. Ruralis 60, 151–173.

Vayro, C., Brownlow, C., Ireland, M., March, S., 2021. "Don't … Break Down on Tuesday Because the Mental

Health Services are Only in Town on Thursday": A Qualitative Study of Service Provision Related Barriers to, and Facilitators of Farmers' Mental Health Help-Seeking. Adm. Policy Ment. Heal. Ment. Heal. Serv. Res. 48, 514–527.

Venkatesh, Morris, Davis, Davis, 2003. User Acceptance of Information Technology: Toward a Unified View. MIS Q. 27, 425–478.

Ware, J.E., Snyder, M.K., Wright, W.R., Davies, A.R., 1983. Defining and measuring patient satisfaction with medical care. Eval. Program Plann. 6, 247–263.

Weinhold, I., Gurtner, S., 2014. Understanding shortages of sufficient health care in rural areas. Health Policy (New. York). 118, 201–214.

World Health Organization;, 2017. Depression and Other Common Mental Disorders: Global Health Estimates. Geneva.

Wrigley, S., Jackson, H., Judd, F., Komiti, A., 2005. Role of stigma and attitudes toward help-seeking from a general practitioner for mental health problems in a rural town. Aust. N. Z. J. Psychiatry 39, 514–521.

**Online Supplement to:**

Thielecke, J., Buntrock, C., Johanna Freund, Braun, L., Ebert, D. D., Berking, M., Baumeister, H., & Titzler, I. (2023). How to promote usage of telehealth interventions for farmers' mental health? A qualitative study on supporting and hindering aspects for acceptance and satisfaction with a personalized telephone coaching

**Content:**



Corresponding author: Janika Thielecke, Friedrich-Alexander-University Erlangen-Nürnberg, Department of Clinical Psychology and Psychotherapy, Germany. Mail: janika.thielecke@fau.de

Supplement 1: COREQ checklist

**Supplement Table 1.** COnsolidated criteria for REporting Qualitative research (COREQ): 32-item checklist.

| No. Item | Guide questions/description | Reported |
|---|---|---|
| Domain 1: research team and reflexivity | | |
| *Personal characteristics* | | |
| 1. Interviewer/-facilitator | Which people conducted the interview or focus group? | Jana Burger (JB)<br>Melanie Martin (MM) |
| 2. Credentials | What were the researcher's credentials? E.g. PhD, MD | Bachelor of science, psychology |
| 3. Occupation | What was their occupation at the time of the study? | Master students |
| 4. Gender | Was the researcher male or female? | Female |
| 5. Experience and training | What experience or training did the researcher have? | Role-play training, feedback, exchange, and supervision by authors |
| *Relationship with participants* | | |
| 6. Relationship established | Was a relationship established prior to study commencement? | TEC-A RCT participants had already been in contact with the study team prior to the interviews but not to the interviewers. Pilot study participants had no prior contact |
| 7. Participant knowledge of the interviewer | What did the participants know about the researcher? (e.g. personal goals, reasons for doing the research) | The interviewers were not prior known to the participants. Participants were informed in the beginning of the interviews about the background of this study and that it was part of the model project "With us in balance." Research question was described as "interest in positive as well as negative experiences with the coaching" |
| 8. Interviewer characteristics | What characteristics were reported about the interviewer/-facilitator? (e.g. bias, assumptions, reasons, and interests in the research topic) | Interviewers conducted the interviews as part of their master theses and could disclose that to participants |
| Domain 2: study design | | |
| *Theoretical framework* | | |
| 9. Methodological orientation and theory | What methodological orientation was stated to underpin the study? (e.g. grounded theory, discourse analysis, ethnography, phenomenology, or content analysis) | Inductive-deductive approach (e.g. theory-based analysis and content analysis). The interview guide and interpretation was based on the adapted version of:<br>Theory of Acceptance and Use of Technology (UTAUT) [25]<br>Discrepancy Theory of Satisfaction (Fox & Storms, 1981)<br>Evaluation categories for satisfaction with health care service (Blum, 1998)<br>Quality of Care model (Donadebian, 1968) |

**Supplement Table 1 (continued).** COnsolidated criteria for REporting Qualitative research (COREQ): 32-item checklist.

| No. Item | Guide questions/description | Reported |
|---|---|---|
| *Participant selection* | | |
| 10. Sampling | How were participants selected? (e.g. purposive, convenience, consecutive, or snowball) | Convenience |
| 11. Method of approach | How were participants approached? (e.g. face-to-face, telephone, mail, or e-mail) | RCT: Via mail through the study team<br>Implementation study: via coaches |
| 12. Sample size | How many participants were in the study? | *N* = 20 |
| 13. Non-participation | How many people refused to participate or dropped out? Reasons? | Potentially 66 people were contacted and invited,<br>32 indicated initial interest, 20 signed inform consent, scheduled an interview, and participated |
| *Setting* | | |
| 14. Setting of data collection | Where was the data collected? (e.g. home, clinic, or workplace) | At home or workplace, via telephone interviews |
| 15. Presence of non-participants | Was anyone else present besides the participants and researchers? | No |
| 16. Description of sample | What are the important characteristics of the sample? (e.g. demographic data, date) | People with an occupational background in farming. Sociodemographic characteristics are reported (see Manuscript Table 2) |
| *Data collection* | | |
| 17. Interview guide | Were questions, prompts, guides provided by the authors? Was it pilot tested? | Theory-based questions, prompts, and guides were provided. Pilot testing and individual feedback was conducted with the first interview of each interviewer |
| 19. Audio/visual recording | Did the research use audio or visual recording to collect the data? | Interviews were audio-recorded (sipgate & audacity) |
| 20. Field notes | Were field notes made during and/or after the interview or focus group? | No additional notes were made |
| 21. Duration | What was the duration of the interviews or focus group? | Average interview duration was *M* = 44 min (*SD* = 17 min, *Min* = 19 min, *Max* = 78 min) |
| 23. Transcripts returned | Were transcripts returned to participants for comment and/or correction? | No |
| Domain 3: analysis and findings | | |
| *Data analysis* | | |
| 24. Number of data coders | How many data coders coded the data? | Katja Bouje<br>Vanessa Oth, B.Sc. |
| 25. Description of the coding tree | Did authors provide a description of the coding tree? | Yes |
| 26. Derivation of themes | Were themes identified in advance or derived from the data? | Main categories were identified in advance as part of the dimensions of the theories; themes were derived from the data based on an inductive approach in order to identify what led to support or hindering influence in the dimensions |

**Supplement Table 1 (continued).** COnsolidated criteria for REporting Qualitative research (COREQ): 32-item checklist.

| No. Item | Guide questions/description | Reported |
|---|---|---|
| 27. Software | What software, if applicable, was used to manage the data? | MAXQDA (VERBI software, 2020) |
| 28. Participant checking | Did participants provide feedback on the findings? | Yes, a validation survey was carried out with 15 (75%) of participants responding |
| *Reporting* | | |
| 29. Quotations presented | Were participant quotations presented to illustrate the themes/findings? Was each quotation identified (e.g. ID)? | Yes, quotations including the ID of the participants are reported |
| 30. Data and findings consistent | Was there consistency between the data presented and the findings? | Yes |
| 31. Clarity of major themes | Were major themes clearly presented in the findings? | Yes, all dimensions and themes are reported (see Figure 2 and Supplement 2) and the most often named are elaborated in detail in the main text |
| 32. Clarity of minor themes | Is there a description of diverse cases or discussion of minor themes? | Yes, diverse cases are discussed |

Table based on: Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care. 2007. Volume 19, Number 6: pp. 349–357.

**Supplement Table 2.** Overview of all identified themes, frequencies, definition, and example statement as used in the coding system.

| Theme | n | % | k | Definition | Supporting quote |
|---|---|---|---|---|---|
| ***COCHING RESULT*** | | | | | |
| ***Expectations (Q=4, k=27)*** | | | | | |
| Improvement of psychological symptoms | 9 | 45 | 11 | Participants expected an improvement of clearly named psychological symptoms (e.g. depression, anxiety, psychosomatic or neurological complaints, restlessness, stress) through the coaching. | *"The second reason was that I have been stuck in a burnout for years or, how should I say it, that I can't get back on my feet."* (p7) |
| Desire for change | 9 | 45 | 11 | Participants express a nonspecific desire for change, overall improved well-being or (re-)gain of balance. | "*But regaining balance and that was, at the start, kind of my motivation or demand and hope that I had in it.*" (p10) |
| Support in difficult phases | 2 | 10 | 3 | Participants expected that by participating in the coaching, they would receive support in difficult phases in order to prevent even worse consequences. | "*Well, what was important to me is that I am somehow guided through this extremely exhausting phase because in spring it's [harvest] season and I was thinking "I can't do all this!*" (p1) |
| Improved self-management | 2 | 10 | 2 | Participants expected to achieve better self-management or self-control through the coaching. | "*My goal or my wish was: (…) I want to get broad shoulders. That I don't always cave in and start crying, sometimes it worked but sometimes it's still like that.*" (p1) |
| ***Supporting Factors (Q=9, k=156)*** | | | | | |
| *Effective factors for coaching* | | | | | |
| Concrete coaching methods | 18 | 90 | 31 | Participants name or describe concrete coaching methods such as breathing exercises, cognitive restructuring, take-home exercises or journaling as helpful for achieving results. | "*She [the coach] also breathed with me. When I started to fall back into old behavior patterns, during the conversation as well. That honestly helped me the most.*" (p6) |
| Opportunity for conversation | 16 | 80 | 23 | Participants rate the opportunity to have a conversation in itself or talking about problems/topics as a conducive factor to achieving the coaching result. | "*Well, definitely the conversation. That contributed to it, simply because it gave me suggestions (…) and a positive feeling.*" (p2) |
| Setting realistic goals and doing exercises | 6 | 30 | 6 | Participants name realistic goal setting as well as exercises during the coaching as helpful. | "*But it wasn't like there were any ideas, let's say (…) 'This can't go on like this, you have to take a 4-week vacation. Do that immediately.' I wouldn't have been able to execute (…). Everything we talked about was doable.*" (p8) |
| Defining and reflecting on the problem | 5 | 25 | 8 | Participants name describing the problem, gaining awareness for the problem, and consciously dealing with one's own situation as helpful for achieving the coaching result. | "*What always helped me a lot was that you could describe the problem precisely and find approaches for how to escape the vicious circle and that helped me a lot.*" (p10) |
| *Changes initiated by coaching* | | | | | |
| More acceptance and calmness | 14 | 70 | 24 | Participants report having developed more acceptance for certain aspects in life and more calmness or self-control as a result of the coaching. | "*Well additionally, the main change was actually making me more relaxed, gaining more distance between me and my company, that I can evaluate things better*" (p4) |

**Supplement 2: Coding system representing all identified themes**

**Supplement Table 2 (continued).** Overview of all identified themes, frequencies, definition, and example statement as used in the coding system.

| Theme | n | % | k | Definition | Supporting quote |
|---|---|---|---|---|---|
| Improved self-care | 11 | 55 | 26 | Participants report an improved self-management as a result of the coaching, which includes improved self-care as well as increased self-confidence. | "*Just that I'm more calm, more selfish, that I can lead conversations with my coworkers differently, that I am more understanding in certain situations with my coworkers.*" (p4) |
| Reevaluation of situations | 10 | 50 | 17 | Participants report actually evaluating situations differently as a result of the coaching instead of merely being shown a different view by the coach. | "*In my daily life - In principle nothing changed. But my attitude toward things has changed.*" (p6) |
| Improved overall and psychological well-being | 9 | 45 | 15 | Participants report having improved their overall or psychological well-being through the coaching (specifically mental health problems such as sleeplessness, stress, or depressive symptoms and nonspecific improvements of well-being). | "*I would say that you gain a certain distance to your job or company, that you also have periods of rest, your own breaks, that you sleep through the night again and your thoughts don't continue rattling.*" (p10) |
| Improved and more open communication | 4 | 20 | 6 | Participants report having improved communication with their social environment and talking more openly about private and health-related topics as a result of the coaching. | "*As I said, overall, that I became more calm, that I can talk more easily about the illness.*" (p3) |
| **Hindering Factors (Q=3, k=10)** | | | | | |
| Perceived continued need for support | 5 | 25 | 5 | Participants still see a need for support after the coaching and wish for specialized counseling centers, aftercare, and outpatient or inpatient psychotherapy or group services. | "*It could get more concrete, that you would, for example, get recommended two or three points of contact that have already proven themselves in the field. (…)*" (p17) |
| Expected results not achieved | 2 | 10 | 2 | As a negative aspect, participants name not achieving the expected coaching results. | "*For me, it was primarily about finding out the reasons for my depressive phase. And for that, I believe, the amount of time was too short.*" (p5) |
| Lack of exercises and tips | 1 | 5 | 3 | Participants describe a lack of exercises and tips for implementation in daily life. | "*Yes I actually wished I would have gotten a few practical little instructions. Along the lines of, (…) Let's say, when I'm totally stressed, take 5 deep breaths or something like that, these practical instructions, [were missing]*" (p8) |
| **COACH** | | | | | |
| ***Expectations (Q=4, k=32)*** | | | | | |
| Professional counseling | 9 | 45 | 12 | Participants expected that the coach would support them through professional counseling and assessment of their situation and that they would give them practical tips, advice, and exercise suggestions. | "*(…) I have a case of mental illness in my family. My brother suffers from severe depression and that burdens me greatly and I hoped to get advice from the coaching on how I should act in front of my brother.*" (p9) |
| Solution-oriented problem discussion | 7 | 35 | 8 | Participants expected from the coach that they would discuss different topics/problems with them (in a solution-oriented manner). | " *(…) That I could talk about it, that I could get new ideas and support so that I could categorize it.*" (p10) |
| Empathetic listening | 5 | 25 | 7 | Participants expected from the coach that they would actively listen and respond to the participants in an empathetic way. | "*The expectation that he would take what I say seriously, that we can talk on eye level. And (…) that he would respond to what I'm telling him.*" (p2) |

**Supplement Table 2 (continued).** Overview of all identified themes, frequencies, definition, and example statement as used in the coding system.

| Theme | n | % | k | Definition | Supporting quote |
|---|---|---|---|---|---|
| Professional competence | 5 | 25 | 5 | Participants expected that coach would be competent in their sphere. | "*That he [the coach] understands his job. (…) I actually assumed that someone would be chosen who knows exactly what is expected from him.*" (p4) |
| ***Supporting Factors (Q=9, k=115)*** | | | | | |
| Competence of coach | | | | | |
| Empathy | 11 | 55 | 24 | Participants perceive the coach's empathy which includes aspects that take place on an emotional level (e.g. understanding the other person and being genuinely interested in the individual's situation) as positive. | "*I mean, I have a hands-on profession and I am also not the only, let's say patient, she had similar conversation before and after me, and that's unbelievable how someone can respond to a person in such a way and give them the feeling, you're not one of many, but I know exactly how you mean it and I'm going to help you solve this.*" (p4) |
| Conversation skills | 10 | 50 | 22 | Participants perceive the way the coach led conversations (i.e., asking targeted questions, activating and steering the conversation in a solution-oriented direction) as positive without describing concrete coaching methods. | "*I was not told, you have to do it a certain way, not that, but she did make sure that I develop [my own thoughts] and then she dug deeper in the right places.*" (p1) |
| Encouragement of participants | 5 | 25 | 11 | Participants perceive the coach's positive reinforcement in the form of praise, encouragement, and validation of one's own statements as positive. | "*Just through the conversation and that she encouraged what I said myself (…).*" (p1) |
| Agricultural knowledge | 5 | 25 | 7 | Participants perceive the coach's agricultural knowledge as agreeable. | "*Well, it wasn't overly important, but it was quite nice that there was certain background knowledge and the coach was also interested in things I told him about agriculture.*" (p2) |
| *Coaching relationship* | | | | | |
| Trusting relationship | 16 | 80 | 23 | Participants describe their relationship with the coach as trusting/secure or friendly/comradely. | "*The relationship was relaxed, trusting.*" (p2) |
| Demographic commonalities | 5 | 25 | 6 | As a conducive factor in relation to the coach, commonalities in demographic data (e.g. sex, ethnicity, age, etc.) are named. | "*I also thought about whether I would be able to talk to my husband as well as with her, I probably would, but it was actually nice being able to have a woman-to-woman talk.*" (p4) |
| Professional relationship | 4 | 20 | 4 | The relationship with the coach is described as professional/respectful. | "*Well, it was a good relationship but not in the sense that I was dependent on her. I did think, well you have to make sure that you remain in control of yourself and don't make her responsible for your happiness. There was a clear separation between what we both did professionally and privately.*" (p4) |
| Picture of coach | 3 | 15 | 3 | Participants rate the coach's picture that was previously sent to them per e-mail as positive. | "*(…) Maybe I still belong to the older [generation] who sometimes still want to know who's on the other side of the phone if you're calling almost every week (inaudible). Therefore I thought the picture was quite nice (…)*" (p10) |

**Supplement Table 2 (continued).** Overview of all identified themes, frequencies, definition, and example statement as used in the coding system.

| Theme | n | % | k | Definition | Supporting quote |
|---|---|---|---|---|---|
| *Coach characteristic* | | | | | |
| Independence of coach | 10 | 14 | 15 | Participants rate the coach's independence/objectivity as positive. | "*And it's always better to talk to someone independent.*" (p3) |
| ***Hindering Factors (Q=1, k = 9)*** | | | | | |
| Low agricultural knowledge | 4 | 20 | 9 | Participants rate the lack of knowledge in the agricultural sphere on the part of the coach as negative. | "*It was maybe just the technical knowledge. He [the coach] took a lot of what I gave him from the professional sphere. There was maybe this one percent left, because he didn't operate that job-specific.*" (p2) |
| **ORGANIZATION** | | | | | |
| ***Supporting Factors (Q=5, k =85)*** | | | | | |
| Flexibility of appointment arrangement | 19 | 95 | 35 | Participants describe the compatibility with their professional and private duties, e.g. thanks to a flexible arrangement and postponement of appointments. This way the coaching can be successfully integrated in daily life. | "*It all went very well. Well, we arranged appointments and if you couldn't make them for once, then you quickly got in contact with each other. It all went very smoothly, well I was very satisfied*." (p5) |
| Quick availability | 13 | 65 | 23 | Participants rate the quick availability and the short waiting times as positive. | "*In terms of processing speed, starting with my first call and then how quickly the whole thing was organized, I can only say that I was surprised how quick it went.*" (p6) |
| Regularity of appointments | 10 | 50 | 17 | Participants rate the reliable appointments and the regularity of appointments as positive. | "*Well, that we almost always talked on the same day and time, with a few exceptions. I could rely on that.*" (p3) |
| Personal invitation by mail | 5 | 25 | 6 | Participants rate the personal invitation to the coaching by postal letter from the social insurance company (at the right time) as positive. | "*So I was contacted by the agricultural social security that there is this study and that they offer me this and because I have just been under extreme stress at the moment because my husband passed away in October and I had to take care of the gardening myself, this came at the perfect time.*" (p1) |
| Individual time frame | 4 | 20 | 4 | Participants rate the individual time frame (e.g. orientation toward individual need and overall length) of the coaching as positive. | "*And that it went on for a longer period of time, not only three or four calls, but that it went on for a longer period of several months.*" (p2) |
| ***Hindering Factors (Q=2, k=12)*** | | | | | |
| Coaching duration too short | 5 | 25 | 10 | Participants describe the overly short duration of coaching as negative. | "*It's just that this whole thing was a little too short because these things take time, now I can take care of it a little more (...)*" (p10) |
| Appointments within core working hours | 1 | 5 | 2 | Participants describe appointments within core working hours as an aggravating factor. | "*As I touched on just now, sometimes appointments after work would be more relaxed because I don't have anyone behind me wanting something from me.*" (p10) |

**Supplement Table 2 (continued).** Overview of all identified themes, frequencies, definition, and example statement as used in the coding system.

| Theme | n | % | k | Definition | Supporting quote |
|---|---|---|---|---|---|
| **PARTICIPANT CHARACTERISTICS** | | | | | |
| ***Supporing Factors (Q=5, k = 52)*** | | | | | |
| *Social support* | | | | | |
| Supportive and encouraging reaction from environment | 16 | 80 | 19 | Participants report positive or supportive and encouraging reactions from their environment. | "*Well my family had a positive view. When we talked about it, they definitely supported me and said, this is a good thing, you should continue if it makes you feel good.*" (p2) |
| Active integration of environment in the coaching | 10 | 50 | 10 | Participants report an active integration of their social environment in the coaching, or the coaching opened possibilities for conversations. | "*The support was just that I repeated what Ms. [anonymized] and I discussed during the coaching when I was my family.*" (p6) |
| *Attitude toward telephone coaching* | | | | | |
| Preference for telephone coaching instead of on-site service | 11 | 55 | 11 | Participants would have chosen the telephone coaching even if there had been a comparable service on site. | "*At the time, I would have always preferred the telephone coaching because of the time aspect. The other wouldn't have been viable at that point in time. And nowadays I would say the same*." (p10) |
| Willingness to find time for coaching | 5 | 25 | 8 | Participants explain that a willingness to take time for the coaching is important. | "*This was possible because I said: "I need this time now, I have to take time off, I'm gone now*." (p4) |
| Willingness to engage in the coaching | 3 | 15 | 4 | Participants talk about how it is necessary to open up to a stranger and an unfamiliar situation and have to admit to your own problems. | "*But when I got accepted to this program of individual coaching sessions, in that moment, it was clear to me that I have to and want to open up to a complete stranger on the phone.*" (p6) |
| ***Hindering Factors (Q=3, k=32)*** | | | | | |
| social/professional involvement restricts change process | 10 | 50 | 16 | Participants describe being subject to their work life, their social environment, and the lack of possibilities as hindering to achieving results. | "*In the professional sphere, it's difficult because I'm trapped in my role as a wife and company manager and I can't leave just like that.*" (p9) |
| Preference for on-site services | 7 | 35 | 13 | Participants would have preferred a comparable on-site service over telephone coaching. | "*If the comparable on-site service had been a personal conversation, I would have definitely chosen that.*" (p7) |
| Negative reactions from one's social environment | 3 | 15 | 3 | Participants report negative reactions from their social environment in response to their participation in the coaching. | "*I didn't look at it in that way, but there were also my husband's preconceptions. (…) 'That they [the insurance] will use it against you afterwards?'*" (p18) |
| **TELEPHONE AS COMMUNICATION MEDIUM** | | | | | |
| ***Supporting Factors (Q=5, k = 70)*** | | | | | |
| Low effort | 17 | 85 | 42 | The perceived advantage of telephone coaching was low time, organizational and financial effort (compared to on-site services). | "*You don't have to drive anywhere, even if you were at the shed earlier, you don't have to shower or change, but instead you can just sit in your office in work clothes and you can make the phone call. This isn't possible with a personal coaching.*" (p5) |

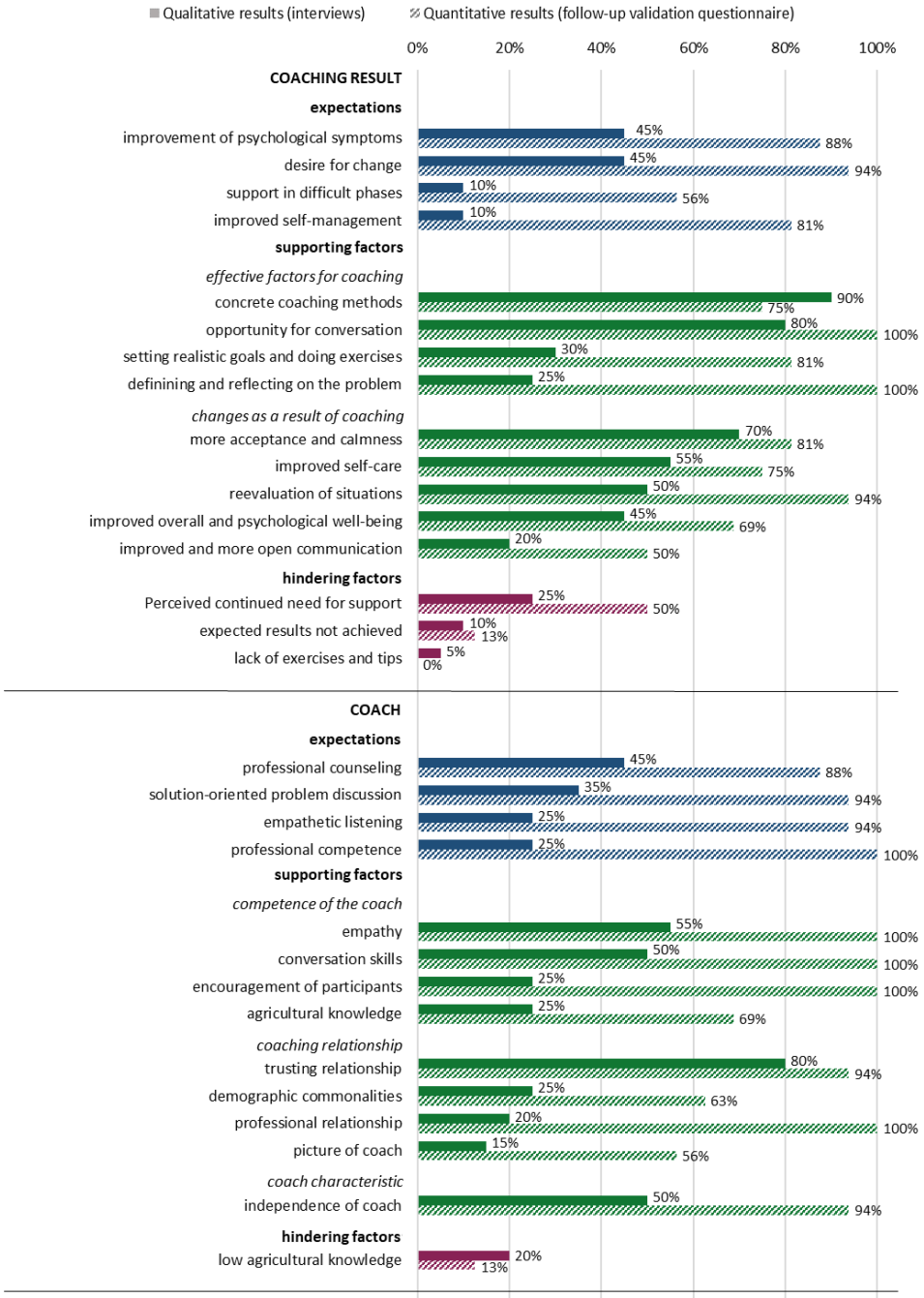
**Supplement Table 2 (continued).** Overview of all identified themes, frequencies, definition, and example statement as used in the coding system.

| Theme | n | % | k | Definition | Supporting quote |
|---|---|---|---|---|---|
| Feeling of anonymity | 10 | 50 | 10 | Participants name the feeling of anonymity and low inhibition threshold as an advantage of telephone coaching. | "*Maybe there are people who are better at talking over the phone, who can open up more easily. Well for those people that would clearly be an advantage, right?*" (p7) |
| Familiar environment during coaching | 9 | 45 | 11 | Participants perceive the familiar environment during telephone coaching as positive. | "*There was the time aspect and actually the familiar environment which made me feel safe to some extent.*" (p11) |
| No distraction | 5 | 25 | 7 | Participants perceive the absence of distracting external factors (e.g. external appearances, stimuli in the treatment room) as conducive to focused conversation, concentration on the essentials, and unbiased participation in the conversation. | "*Well it was definitely good that it was neutral, because I think you subconsciously judge people by their appearance, and this way that [aspect] was eliminated and you could only hear your counterpart but you couldn't see them.*" (p4) |
| ***Hindering Factors (Q=5, k = 42)*** | | | | | |
| Lack of visual clues | 12 | 60 | 20 | Participants name the loss of visual clues (e.g. absence of eye contact and nonverbal communication) as a disadvantage of telephone coaching. | "*Well the disadvantages are clearly that you can't see your counterpart. I do find it different. For some people, the inhibition threshold to talk about something might be lower, but I would find it comfortable if you could sit across each other and talk properly.*" (p8) |
| Technical difficulties | 7 | 35 | 7 | Participants report technical difficulties during the coaching. | *There were indeed problems with the connection being gone, sometime we have that.*" (p8) |
| Less trusting coaching-relationship | 4 | 20 | 5 | Participants mention the difficulty of opening up and building trust with the coach as a potential disadvantage of telephone coaching. | "*I was a little concerned whether you could establish a trusting relationship over the phone as opposed to actually sitting or standing across from each other, eye to eye so to speak, but it actually turned out to be unproblematic.*" (p17) |
| Restrictions in coaching conduct | 4 | 20 | 6 | Participants suspect that not all techniques and approaches that are applicable to telephone coaching are applicable on site. | "*That you can maybe not use certain techniques, I wouldn't know about that. (..) That there are certain techniques that you can't communicate over the phone.*" (p5) |
| Lack of distance from problems | 3 | 15 | 4 | Participants describe remaining in a familiar environment and the lack of distance to the problem during the coaching as hindering to the coaching process. | "*That it takes place in the usual space. That you're not seperated. (..). Well personally, I'm never seperated when I'm in a private, personal space, in my usual environment. But I can seperate myself within seconds when I'm in an unfamiliar atmosphere.*" (p7) |

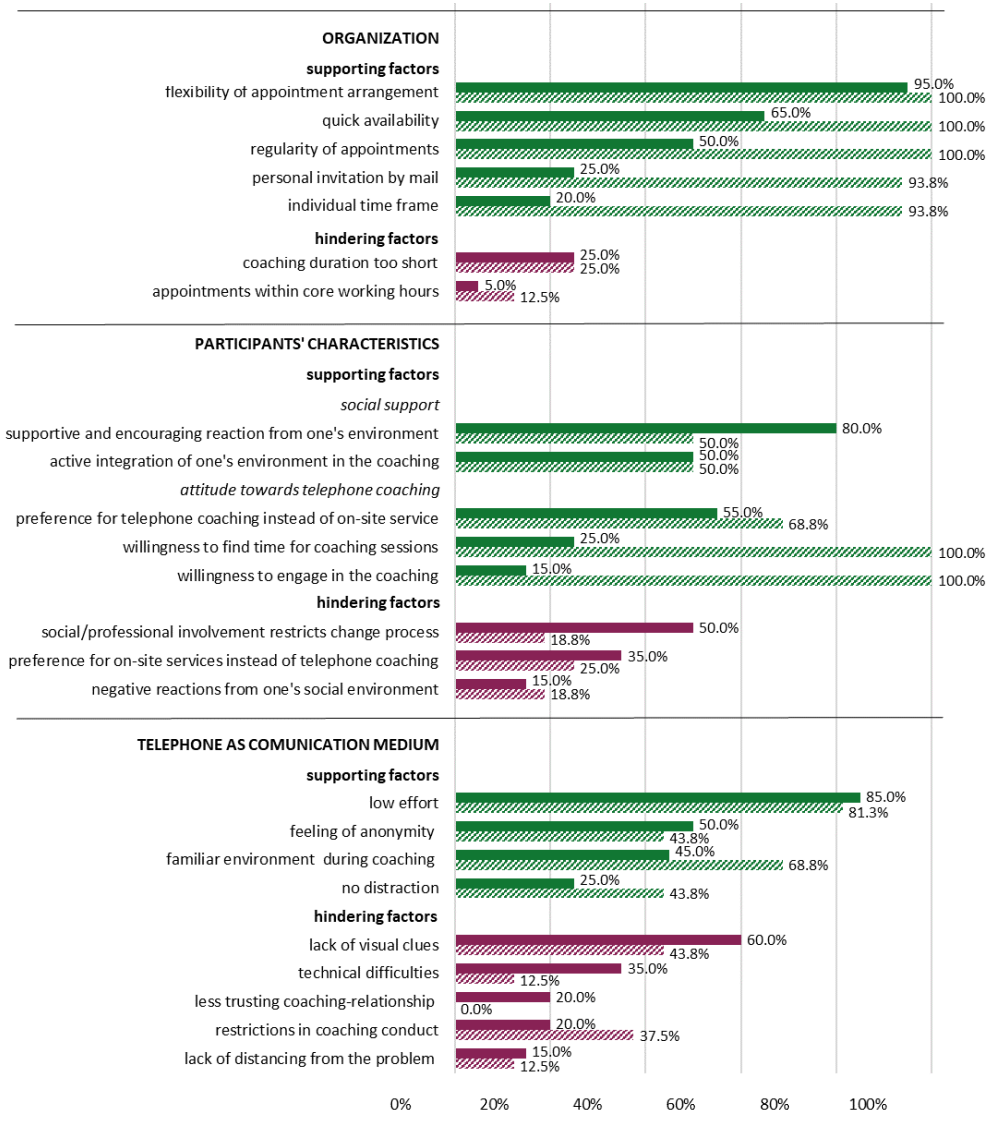
In supporting quotes, pauses and comments removed from quotes for better readability. (...) indicate left out parts without changing the content of the sentence. [..] indicate complements and explanations derived from the context of the quote. Abbreviations: n=number of participants naming the theme; % = percentage of participants referring to the N=20 interviewed persons; k= number of coded elements in the interview data; Q=number of themes identified on the level of a dimensio

Supplement 3: Comparison of identified themes in interview and validation questionnaire

**Supplement Figure 1.** Frequency of identified themes in the interviews (n=20) and the validation questionnaire (n=15) influencing acceptance of and satisfaction with the telephone coaching.



**Supplement Figure 1 (continued).** Frequency of identified themes in the interviews (n=20) and the validation questionnaire (n=15) influencing acceptance of and satisfaction with the telephone coaching.



# 7

## ARTICLE 5 — OUTCOME EXPECTANCY AS A PREDICTOR OF TREATMENT OUTCOME

Title:
Does Outcome Expectancy Predict Outcomes in Online Depression Prevention? Secondary Analysis of Randomized Controlled Trials

Summary:
Web-based interventions have been shown to be efficacious in reducing depressive symptoms and preventing depression onset. However, less is known about what influences treatment outcome. In psychotherapy research, outcome expectancy is considered a common factor across therapeutic approaches that influences depression outcomes. The role of outcome expectancy in prevention and in web-based interventions is less clear. This secondary data analysis aims to explore the role of outcome expectancy in participants who received the preventive web-based training for depression *GET.ON Mood Enhancer*.

For the analysis, data from the intervention groups of two randomized controlled trials were combined (N = 304). Outcome expectancy was assessed with the expectancy subscale from the Credibility and Expectancy Questionnaire (CEQ) in both studies. Multilevel modeling, controlling for age, sex and baseline symptom severity was used to explore the effect of outcome expectancy on depressive symptom severity (Center for Epidemiological Studies Depression Scale, CES-D) and close-to-symptom-free status (CES-D < 16) at post-treatment (6 – 7 weeks) and follow-up (3 – 6 months). Cox regression was applied to assess the effect of outcome expectancy on time to depression onset within 12 months in a subsample of n = 102 participants who were assessed with diagnostic interviews. Moderation analysis explored whether age, sex or baseline symptom severity influenced the relationship of outcome expectancy and depressive symptom severity.

Outcome expectancy did not predict depressive symptom severity and close-to-symptom-free status at post-treatment nor on depression onset within a year. At the follow-up assessment, small effects of outcome expectancy were observed on depressive symptom severity ($\beta$ = -0.39, [95%–CI: -0.75, -0.03], p = 0.032, $p_{adjusted}$ = 0.130) and close-to-symptom-free status (RR = 1.06, [95%–CI: 1.01, 1.11], p = 0.013,

$p_{adjusted}$ = 0.064), but those results were not statistically significant when controlled for multiple testing. Moderator analyses suggested that the relation of outcome expectancy and post-treatment depression might be stronger in females and individuals with higher baseline depressive symptom severity. Results on depression onset are especially limited because of the small sample and reduced power. Results might be influenced by using a unvalidated translation of the CEQ and assessing it before participants had access to the intervention.

Overall, the findings suggest that outcome expectancy could have some influence on long-term depressive outcomes and given the easily assessable and potentially modifiable nature of the construct, it remains an interesting field of research within the preventive context. Methodologically sound studies are warranted to establish these potential effects before investigating methods to enhance outcome expectancy and improve the effectiveness of web-based interventions for depression.

Submitted manuscript:
Thielecke J, Kuper P, Ebert DD, Cuijpers P, Smit F, Riper H, Lehr D, and Buntrock C. Does Outcome Expectancy Predict Outcomes in Online Depression Prevention? Secondary Analysis of Randomized Controlled Trials. Prevention Science. submitted

Contribution:
Janika Thielecke contributed substantially to the study conceptualization, design and methods and prepared and published the preregistration on the Open Sience Framework. She organized the secondary data analysis and shared responsibility for data curation, formal analysis and visualization with a co-author. She interpreted the results, drafted the original manuscript, revised it after feedback from her co-authors and managed the submission process.

# Does Outcome Expectancy Predict Outcomes in Online Depression Prevention? Secondary Analysis of Randomized Controlled Trials

Janika Thielecke[1,2], Paula Kuper[1,3], David Ebert[2], Pim Cuijpers[4,5], Filip Smit [4,6,7], Heleen Riper [4,5,8], Dirk Lehr[9] & Claudia Buntrock[3]

[1] Professorship of Psychology and Digital Mental Health Care, Department of Sports and Health Sciences, Technical University of Munich, Germany

[2] Department of Clinical Psychology and Psychotherapy, Institute of Psychology, Friedrich-Alexander - University Erlangen-Nürnberg, Germany

[3] Institute of Social Medicine and Health Systems Research, Faculty of Medicine, Otto-von-Guericke University Magdeburg, Germany

[4] Department of Clinical, Neuro and Developmental Psychology, VU University, Netherlands

[5] Amsterdam Public Health, Amsterdam University Medical Centers, Netherlands [6]Department of Mental Health & Prevention, Trimbos Institute: Netherlands Institute of Mental Health and Addiction, Netherlands

[7]Department of Epidemiology and Biostatistics, University Medical Center Amsterdam, Netherlands

[8]Department of Psychiatry, VU University Medical Center, Netherlands

[9]Department of Health Psychology and Applied Biological Psychology, Institute for Sustainability, Education & Psychology, Leuphana University Luneburg, Germany

**Author Note**

Correspondence concerning this Article should be addressed to Janika Thielecke, Technical University Munich (TUM), Department of Sports and Health Sciences, Georg-Brauchle-Ring 60/62, 80992 Munich.

Mail: janika.thielecke@tum.de

**Abstract**

Evidence shows that online interventions could prevent depression. However, to improve the effectiveness of preventive online interventions in individuals with subthreshold depression it is worthwhile to study factors influencing intervention outcomes. Outcome expectancy has been shown to predict treatment outcome in psychotherapy for depression. However, little is known about whether this also applies to depression prevention. The aim of this study was to investigate the role of participants' outcome expectancy in an online depression prevention intervention. A secondary data analysis was conducted using data from two randomized controlled trials (N = 304). Multilevel modelling was used to explore the effect of outcome expectancy on depressive symptoms and close-to-symptom-free status at post (6-7 weeks) and follow-up (3-6 months). In a subsample (n = 102), Cox regression was applied to assess the effect on depression onset within 12 months. Explorative analyses included baseline characteristics as possible moderators. Outcome expectancy did not predict post-treatment outcomes or the onset of depression. Small effects were observed at follow-up for depressive symptoms (β = -0.39, 95% CI [-0.75, -0.03], $p = 0.032$, $p_{adjusted} = 0.130$) and close-to-symptom-free status (RR = 1.06, 95% CI [1.01, 1.11], $p = 0.013$, $p_{adjusted} = 0.064$), but statistical significance was not maintained when controlling for multiple testing. Moderator analyses indicated that expectancy could be more influential for females and individuals with higher initial symptom severity. More thoroughly designed, predictive studies targeting outcome expectancy are necessary to assess the full impact of the construct for effective depression prevention.

*Keywords:* depression, prevention, secondary analyses, prediction, expectancy, online intervention, CBT

*Registration:* original studies: DRKS00004709, DRKS00005973; secondary analysis: osf.io/9xj6a

**Introduction**

Subthreshold depression (sD) is highly prevalent (Cuijpers et al., 2004) and associated with poor quality of life (Rucci et al., 2003; Volz et al., 2022), a high risk of developing major depressive disorder (MDD) (Cuijpers & Smit, 2004; Lee et al., 2019), increased mortality (Cuijpers et al., 2013), and use of health care services (Wagner et al., 2000), as well as substantial economic costs (Cuijpers et al., 2007). Individuals are considered to suffer from sD if they show clinically relevant depressive symptoms but standard diagnostic criteria for a MDD are not yet met (Volz et al., 2022). sD can be defined categorically by meeting some but not all criteria for a MDD diagnosis or dimensionally by scoring above a certain cut-off level on a validated self-rated depression scales (Cuijpers et al., 2014).

The importance of preventive interventions aimed at the treatment of sD and prevention of MDD is emphasized by the fact that nearly all individuals who develop a MDD are assumed to have first gone through a phase of sD (Eaton et al., 1995). Meta-analytic evidence shows that face-to-face preventive psychological interventions can reduce the incidence of depression by about 20% (Relative Risk = 0.81, 95% CI [0.72, 0.91]; Cuijpers et al., 2021). Though clearly effective, psychological face-to-face interventions do not reach the majority of people who could benefit from them (Muñoz et al., 2021). Online interventions have the potential to increase access to preventive services.

Online interventions target cognitive, affective, and behavioral changes; are typically based on evidence-based face-to-face interventions; and require active engagement from participants through the completion of online and offline assignments (Beatty & Binnion, 2016). Discussed benefits of online interventions include a flexible use, comparably low costs, and the ability to reach a wide range of users (Ebert et al., 2017; Ebert, Van Daele, et al., 2018; Muñoz et al., 2021). A recent individual participant data meta-analysis on guided and unguided online interventions showed that they can be effective in improving depressive symptom severity (*d* = -0.39, 95% CI [-0.25, -0.53]) and in reducing the incidence of major depressive disorder in individuals with sD (Hazard Ratio = 0.72, 95% CI [0.58–0.89]; Reins et al.,

2021). However, to further increase the effectiveness of preventive online interventions for depression, it is important to investigate factors that predict differential treatment outcome (Ebert & Cuijpers, 2018).

Patient expectancies of treatment outcome—that is, their beliefs of whether treatment will lead to an improvement in health status (Constantino et al., 2011, 2018)—is discussed as a common factor in psychotherapy (Cuijpers et al., 2019; Greenberg et al., 2006) and is meta-analytically associated with psychotherapy outcomes ($r$ = 0.12 - 0.18; Constantino et al., 2011, 2018). Depression-specific studies have shown the predictive value of positive outcome expectancy in individual face-to-face treatment of β = -0.35 on 16 weeks follow-up depression scores (Thiruchselvam et al., 2019). In group CBT an indirect effect of outcome expectancy of β = -1.29 [95%-CI:-2.93, -0.16] mediated by mid-treatment working alliance was observed on depression scores at post-treatment after 10 weeks (Vîslă et al., 2018). Outcome expectancy has attracted attention as a way to maximize treatment outcomes (Beshai et al., 2019; Braun-Koch & Rief, 2022; Mitchell & Gordon, 2007; Stalujanis et al., 2021).

In online interventions, outcome expectancy has mainly been investigated in terms of the acceptability and uptake of various internet health services (Beatty & Binnion, 2016; Musiat et al., 2014; Philippi et al., 2021) but less in its persisting effects on the outcome. Evidence for an association with depression outcomes in guided and unguided online interventions is inconclusive, with three studies supporting outcome expectancy as being predictive for treatment outcome (Alaoui et al., 2016; de Graaf et al., 2009; Myklebost et al., 2022), whereas four studies did not find that association (Cavanagh et al., 2009; Høifødt et al., 2015; Lüdtke et al., 2018; Pearson et al., 2019). Another study reported an effect of outcome expectancy fully mediated by working alliance as reported by patients (Zagorscak et al., 2020).

For preventive online interventions, evidence is scant and indecisive, with outcome expectancy showing correlations with reduction in anxiety symptoms (Kenardy et al., 2003) but not with post-treatment obsessive compulsive disorder symptoms (Boisseau et al., 2017). To the best of our knowledge, no study has investigated outcome expectancy for a preventive online intervention for depression. The

aim of this study was thus to explore the predictive role of outcome expectancy in a preventive online intervention for adults with sD. We hypothesized that outcome expectancy predicted depressive symptom severity (Research question 1, RQ1) and close-to-symptom-free status (RQ2) at post-treatment and follow-up and depression onset within 12 months (RQ3). In exploratory analyses, we examined moderating effects of baseline symptom severity, age, and sex (Exploratory RQ1-3).

## Methods

### Study Design and Participants

Secondary analyses were conducted based on combined data from the intervention groups of two randomized controlled trials (*N* = 304) that evaluated the effectiveness of the same online intervention (GET.ON Mood Enhancer Prevention). Earlier publications from these trials reported effects of the intervention on depressive symptoms (Buntrock et al., 2015; Ebert, Buntrock, et al., 2018) and the progression to major depressive episode over a 12-month follow-up period (Buntrock et al., 2016). The first study (PREV-DEP I, *N* = 406) compared the online intervention with enhanced treatment-as-usual (i.e., online psychoeducation) within a 12-month follow-up period (Buntrock et al., 2014), whereas the second study (PREV-DEP II, *N* = 204) compared the intervention with a wait-list control condition and used a 3-month follow-up period (Ebert, Buntrock, et al., 2018). The studies were approved by the ethics committee of the Philipps University Marburg (2012-35K, PREV-DEP I) and the Leuphana University Lueneburg (Ebert201404_Depr, PREV-DEP II), respectively, and registered in the German Clinical Trial Register under DRKS00004709, and DRKS00005973. These secondary analyses of pseudonymized data was registered on OSF (https://doi.org/10.17605/OSF.IO/9XJ6A) prior to data analyses.

Participants were mainly recruited via a large German statutory health insurance company (BARMER) by announcing the studies in the members' magazine. The studies were also announced in newspaper articles, on-air media, and related websites. Applicants self-identifying with a depressed mood and who (a) screened positive for subthreshold depressive symptoms (Centre for Epidemiological

Studies Depression Scale [CES-D] ≥ 16; Lewinsohn et al., 1997), (b) were aged 18 and above, (c) with Internet access, (d) neither currently in psychotherapy (e) nor in the past 6 months (f) or on a waiting list for psychotherapy, and (g) with no notable suicidal risk (Becks Depression Inventory [BDI item 9] > 1; Hautzinger et al., 2006) were scheduled for a semi-structured clinical interview (SCID) conducted via telephone by trainees in psychotherapy to assess final eligibility. Those meeting DSM-IV criteria for (a) a major depressive episode, (b) bipolar disorder, or (c) psychotic disorder, and (d) having a history of a major depressive disorder in the past 6 months were excluded. As trials were pragmatic, the use of antidepressant medication was allowed as part of treatment-as-usual if participants took a stable dose for at least four weeks prior to study participation.

**Intervention**

The online intervention consisted of six 30-minute modules. The session duration could vary between users. In PREV-DEP II, participants were offered an optional seventh module as a booster session four weeks after completion of the intervention. Participants in both studies were advised to complete two modules per week if possible, but at least one. The intervention was based on psychoeducation, behavioural activation, and problem-solving therapy. Further details on the content of the intervention can be found elsewhere (Buntrock et al., 2014). While participants in the PREV-DEP I were supported by an online coach who provided written individual feedback after each completed module, participants in PREV-DEP II received feedback only upon request. In both studies, feedback focused on helping participants complete the exercises and no therapeutic advice was provided.

**Meassurements**

***Outcome expectancy***

Outcome expectancy was assessed at baseline prior to the start of the intervention with the respective items from the Credibility and Expectancy Questionnaire (CEQ) (Devilly & Borkovec, 2000). The CEQ version used in PREV-DEP I & II was self-translated to German and not validated in this form. The

wording was adapted to specify "online-training" as the intervention and "depressive symptoms" as the outcome. The CEQ expectancy sub-scale included one item on how participants think and two items on how they *feel* about the effect the intervention will have on their depressive symptoms. Items were rated on a scale from 1 to 9, leading to a composite score for expectancy ranging from 3 (low expectancy) to 27 (high expectancy). The original CEQ demonstrated good psychometric properties, with high internal consistency and high test-retest reliability (Devilly & Borkovec, 2000). Cronbach's alpha in the combined data was $\alpha = 0.87$.

***Depressive symptom severity***

Depressive symptom severity was measured with the German version of the Center for Epidemiological Studies Depression Scale (CES-D, Hautzinger et al., 2012). The CES-D is a self-reporting scale consisting of 20 items, each scored 0–3 resulting in a total score from 0 to 60, with a higher score indicating more severe depressive symptoms. Psychometric properties of the CES-D are well established (Hautzinger et al., 2012). Cronbach's alpha in combined data was $\alpha = 0.82$ at baseline, $\alpha = 0.89$ at post and $\alpha = 0.90$ at follow-up, respectively. Close to symptom-free status was defined by a CES-D score < 16.

***Onset of major depressive episode***

Time to onset of a major depressive episode over a 12-month follow-up was only assessed in PREV-DEP I. DSM-IV criteria were assessed via telephone-administered SCID at 6- and 12-month follow-up, covering the period from the previous assessment (Crippa et al., 2008; Lobbestael et al., 2011). To reduce potential recall bias, time to onset of MDD was assessed as accurately as possible using the Life Chart method (Lyketsos et al., 1994). Diagnostic interviews were conducted by psychologists trained in delivering the SCID. The κ coefficient for inter-rater agreement for a diagnosis of a depressive episode was 0.77 (based on data from 12% of the participants), indicating substantial agreement (McHugh, 2012).

**Prognostic factors**

Age, sex, and baseline depressive symptom severity have been repeatedly identified as

predictors of outcome in online interventions for depression treatment (Alaoui et al., 2016; Donker et al., 2013; Høifødt et al., 2015; Pearson et al., 2019). These baseline characteristics were included in the analyses to assess an adjusted effect of expectancy.

**Data analysis**

For combined data, "post-treatment" refers to the first assessment after intervention completion and "follow-up" defines the second assessment after three (PREV-DEP II) and six months (PREV-DEP I), respectively. For each hypothesis, separate regression models were specified in R (R Core Team, 2022). Significant levels for the five effect estimates of expectancy on depressive outcomes were adjusted for multiple testing using the Bonferroni-Holm method (Holm, 1979). Effects of expectancy were reported with the appropriate estimates and confidence intervals and $R^2$ was reported as a measure of overall model fit. The full model specifications are given in Online Resource 1.

***Multilevel-Models***

To account for the nesting of participants in trials, RQ1 and RQ3 are answered using multilevel models. First, a one-stage individual participant data (IPD) approach (Riley et al., 2020) was used to investigate the predictive effect of expectancy on depressive symptom severity. The general recommendations for one-step IPD analyses from Riley et al. (2020) were followed, i.e specifying a stratified intercept, a random slope of expectancy, and stratified prognostic variables (age, sex, baseline CES-D), each centered by their trial means. Second, given the small number of included studies ($k$ = 2), we followed the proposal by Chung et al. (2012) to use a "pseudo-Bayesian" approach in the multilevel-models. Between-study heterogeneity was highly plausible, given the differences in session count, guidance, and assessment points in the trials, but the small number of random-effects levels (trials) may have led to improperly estimated heterogeneity variances at zero. Using a weakly informative gamma prior with a shape parameter of 1.5 and a rate parameter of 0.05 allowed for an approximate maximum a posteriori (MAP) estimate (Chung et al., 2015). The prior, implemented with the 'blme' package

(Chung et al., 2013), helped to avoid boundary-fit-problems while remaining largely uninformative itself.

Models 1 and 2 were defined as linear mixed-effects models predicting depressive symptoms at post (H1) and follow-up (H2) based on expectancy. Close-to-symptom-free status was predicted from expectancy in models 4 (post-treatment) and 5 (follow-up) using a generalized linear effects models with a clog-log link function to retrieve risk ratios (*RR*). Participants with CES-D<16 at baseline despite inclusion criteria were excluded from this analysis. Model 5 included close-to-symptom-free status at post as an additional stratified and centered covariate. The prior rate was adjusted to 0.01 in model 5 to reduce convergence problems. However, models 4 and 5 both showed convergence problems in one of the 50 imputed sets.

***Single-level model***

Model 3 was specified in PREV-DEP I data as a right-censored Cox-regression using the 'survival' (Therneau, 2022) and 'survminer' package (Kassambara et al., 2021). Continuous baseline characteristics were mean centered. Hazard ratios (HR) with their 95% confidence intervals (CI) were reported. In addition to Nagelkerke's pseudo $R^2$ calculated with the 'psfmi' (Heymans, 2022) package, concordance as a more adequate goodness-of-fit measure was given (Therneau, 2022).

***Explorative Analyses***

As exploratory analyses, an interaction term of expectancy and each baseline characteristic (age, sex, baseline symptom severity) was entered into separate linear mixed-effects models examining differential depressive symptoms while additionally controlling for baseline characteristics (comparable to H1). Continuous candidate moderators were centered by trial-specific means. Contrary to the preregistration, we did not include trial means of candidate moderators as a level 2 predictor as recommended by Riley et al (2020) because of multicollinearity with the trial specific intercepts. However, as suggested, we stratified and centered all parameters outside the interaction term including expectancy in order to avoid amalgamation of within and across trial information (Riley et al., 2020).

***Missing data***

Analyses were conducted according to the intention-to-treat principle (ITT). Multiple imputation by chained equations (MICE) was applied by using the R packages 'mice' (van Buuren & Groothuis-Oudshoorn, 2011) and 'miceadds' (Robitzsch & Grund, 2022) assuming data was missing at random (MAR). A total of m = 50 imputation sets were generated. Model parameters were estimated in all datasets and combined according to Rubin's rules (Little & Rubin, 2002; Rubin, 1987).

A random-effects model with fixed effects for level 1 (participant) and random effects level 2 (trial) variables was used to account for the nested structure of the data. Continuous data were imputed using '2l.pan' (Schafer & Yucel, 2002), a special case of a multivariate linear mixed effects model for panel data already included in 'mice'. Due to convergence problems, cluster means of the covariates could not be included in the prediction matrix. For congeniality of imputation and analysis models, the 'blme' (Chung et al., 2013) functionality was called up within 'mice' to apply a weakly informative gamma-prior with a shape parameter of 1.5 and a rate of 0.05.

For RQ 3 no imputation was needed since baseline data was complete and data for depression onset was right-censored, with observation time set to 0 or 26 weeks in cases of dropout occurred after baseline or 6-month follow-up, respectively (*n* = 40).

## Results

### Participants

In PREV-DEP I, participant data was missing in 12% of cases (25/202) at post-treatment, and in 25% (51/202) and 36% (72/202) at 6- and 12-months follow-up due to study dropout (Buntrock et al., 2015, 2016). In PREV-DEP II, dropout was 21% (21/102) and 29% (30/102) respectively at post-treatment and 3-month follow-up (Ebert, Buntrock, et al., 2018). Study dropout ($r$ = 0.03, $T(302)$ = 0.7, $p$ = 0.5) and number of completed sessions ( $r$= 0.04, $T(302)$ = 0.8, $p$ = 0.4) were not related to baseline expectancy. Participants across both trials (Table 1) were predominantly female (*n* = 230, 76%), highly educated (*n* =

**Table 1**

*Baseline characteristics of participants in both trials and the combined sample*

| | total (N = 304) [a] | | PREV-DEP I[b] (N = 202) | | PREV-DEP II[c] (N = 102) | |
|---|---|---|---|---|---|---|
| | N | % | N | % | N | % |
| CES-D sum score (M, SD) | 26.4 | 7.4 | 26.3 | 7.9 | 26.7 | 6.5 |
| CEQ expectancy (M, SD) [a] | 16.7 | 5.0 | 16.8 | 5.1 | 16.4 | 4.8 |
| Age (M, SD) | 45.4 | 11.8 | 45.7 | 11.9 | 44.7 | 11.7 |
| Gender | | | | | | |
| Male | 73 | 24 | 53 | 26 | 20 | 20 |
| Female | 231 | 76 | 149 | 74 | 82 | 80 |
| Relationship | | | | | | |
| Single | 89 | 29 | 62 | 31 | 27 | 27 |
| Married or cohabiting | 167 | 55 | 102 | 50 | 65 | 64 |
| Divorced or separated | 46 | 15 | 37 | 18 | 9 | 9 |
| Widowed | 3 | 1 | 2 | 1 | 1 | 1 |
| Ethnicity | | | | | | |
| White | 244 | 80 | 165 | 81 | 79 | 78 |
| Black | 1 | 0 | 1 | 1 | 0 | 0 |
| Not reported | 60 | 20 | 37 | 18 | 23 | 23 |
| Level of education | | | | | | |
| Low (primary) [d] | 7 | 2 | 5 | 3 | 2 | 2 |
| Middle (secondary) [e] | 49 | 16 | 33 | 16 | 16 | 16 |
| High (A-level or higher) [f] | 248 | 82 | 164 | 81 | 84 | 82 |
| Employment status | | | | | | |
| Employed | 260 | 86 | 170 | 84 | 90 | 88 |
| Unemployed or seeking work | 6 | 2 | 4 | 2 | 2 | 2 |
| On sick-leave | 3 | 1 | 3 | 2 | 0 | 0 |
| Non-working | 36 | 12 | 26 | 12 | 10 | 10 |
| Income in Euro[g] | | | | | | |
| Low (< 10,000) | 25 | 8 | 16 | 8 | 9 | 9 |
| Middle (10 – 60,000) | 215 | 71 | 145 | 72 | 70 | 69 |
| High (> 60,000) | 40 | 13 | 26 | 13 | 14 | 14 |
| Not reported | 27 | 9 | 18 | 9 | 9 | 9 |
| Previous | | | | | | |
| Psychotherapy | 170 | 56 | 88 | 44 | 82 | 40 |
| health training[h] | 74 | 24 | 51 | 25 | 23 [a] | 23 |
| Use of antidepressants | 57 | 19 | 50 | 25 | 7 | 7 |

*Note.* [a]New data added in the current work, [b]Adapted from Buntrock et al. (2015), [c]Adapted from Ebert et al. (2018), [d]Primary education indicates elementary school, [e]Secondary education indicates high school. [f]Indicates A-level examinations ("Abitur") or above (university degree). [g]Yearly gross income (€1.00≈ US $1.13 at time of original study) [h]Preventive interventions as offered by German statutory health insurance companies (eg, stress management, smoking cessation, healthy diet). *Abbreviations*. CES-D, Center for Epidemiologic Studies Depression Scale; CEQ, Credibility and Expectancy Questionnaire

195, 64%), in a relationship (*n* = 166, 55%) and on average 45 years old (*SD* = 11.8). Expectancy ranged from low (min = 4) to high (max = 27) with *M* = 16.7 (*SD* = 5.0).

**Depressive symptom severity (H1 & H2)**

No predictive effect of expectancy on depressive symptom severity at post-treatment was observed (β = -0.30, 95% CI [-0.73, 0.13], $p_{adjusted}$ = 0.352, $R^2$ = 0.21). Higher expectancy at baseline indicated lower depressive symptom scores at follow-up (β = -0.39, 95% CI [-0.75, -0.03]). This effect did not remain significant when controlling for multiple testing (*p* = 0.032, $p_{adjusted}$ = 0.130, $R^2$ = 0.26).

**Depression onset (H3)**

As reported elsewhere (Buntrock et al., 2016), from the *n* = 202 participants in PREV-DEP I, *n* = 55 (27.2%) individuals experiencing onset of depression within the follow-up period. No predictive effect of expectancy on depression onset was found (HR = 0.97, 95% CI [0.93, 1.02], $p_{adjusted}$ = 0.346). The proportional hazard assumptions were met according to the scaled Schoenfeld residual test (global $\chi^2(4)$ = 0.10 *p* = 1.00). The model was overall significant based on the Likelihood ratio test ($\chi^2(4)$ = 16.25, *p* = 0.003) but did not explain the data well (Concordance = 0.66, *SE* = 0.04, Nagelkerke's pseudo $R^2$ = 0.08).

**Close-to-symptom-free status (H4 & H5)**

For analyses of close-to-symptom-free status, participants already close-to-symptom-free at baseline were excluded (*n* = 16), resulting in *n* = 290. No predictive effect of expectancy on close-to-symptom-free status at post-treatment was found (RR = 1.04, 95% CI [0.99, 1.08], $p_{adjusted}$ = 0.326, $R^2$ = 0.11). Comparably to H1/H2, a positive effect of expectancy on reaching close-to-symptom-free status at follow-up was observed (RR = 1.06, 95% CI [1.01, 1.11]) but statistical significance did not remain after adjusting for multiple testing (*p* = 0.013, $p_{adjusted}$ = 0.064, $R^2$ = 0.29).

**Moderation analyses (Exploratory H1-H3)**

While higher baseline expectancy was potentially associated with lower depressive symptoms at post-treatment, sex (β = 0.46, 95% CI [0.03, 0.90], *p* = 0.038, $R^2$ = 0.22) and baseline severity (β = -0.03,

95% CI [-0.05, 0.001], $p = 0.041$, $R^2 = 0.22$) significantly moderated this association but not age (β = 0.01, 95% CI [-0.01, 0.02], $p = 0.270$, $R^2 = 0.21$). With each one-point increase from the trial specific mean in expectancy, depressive symptom severity at post-treatment was additionally reduced by 0.46 points by being female. Each additional one-point increase from the trial specific mean in initial depressive symptom severity increased the effect of expectancy on depressive symptom severity by 0.03 points. No moderating effect was observed at follow-up assessment (Online Resource 2).

**Discussion**

In this secondary analysis, we explored the effect of expectancy in an online intervention for depression prevention on different depressive outcomes at post-treatment and follow-up. No predictive effects of expectancy on post-treatment depressive symptom severity, on close-to-symptom-free status, or on time to depression onset within 12 months could be observed. At follow-up, small effects were found on depressive symptoms (β = -0.39, 95% CI [-0.75, -0.03], $p = 0.032$, $p_{adjusted} = 0.130$) and close-to-symptom-free status (RR = 1.06, 95% CI [1.01 - 1.11], $p = 0.013$, $p_{adjusted} = 0.064$), but significance was lost after adjusting for multiple testing. Exploratory analyses suggested that female sex and higher depressive symptom severity increased the effect of expectancy on symptom severity at post-treatment.

Our findings are in line with other studies that did not find an effect of expectancy on depression outcomes directly after use of an online intervention (Cavanagh et al., 2009; Høifødt et al., 2015; Lüdtke et al., 2018; Pearson et al., 2019). However, given the possible effects at follow-up, overall evidence remains inconclusive. Our observation that expectancy might be more relevant in later follow-ups than in post treatment is similar to findings by de Graaf et al. (de Graaf et al., 2009), who observed a predictive effect of expectancy on reliably changed depression scores after 9 months, but not after 3 months. Indeed, studies that did not find an effect of expectancy all had observation times under 10 weeks (Cavanagh et al., 2009; Høifødt et al., 2015; Lüdtke et al., 2018; Pearson et al., 2019). Zagorscak et al. (2020) suggest that early expectancy predicts mid treatment task and goal agreement, which then leads

to symptom improvement (Zagorscak et al., 2020). This might explain a delayed effect on depressive outcomes in self-help oriented interventions, where identification with tasks and goals might be crucial for the transfer into everyday life.

Studies that found a predictive effect of expectancy shared common methods: namely, that they used a validated questionnaire (i.e., CEQ) and either did not include a randomization process and assessed expectancy after session one or two (Alaoui et al., 2016; Myklebost et al., 2022) or they assessed expectancy after randomization (de Graaf et al., 2009). We also used the CEQ expectancy scale; however, we assessed expectancy before randomization. Not knowing whether they got immediate access to the online intervention might have influenced participant's expectancy ratings.

To our knowledge, our exploratory moderation analyses are first of their kind in online interventions. However, these findings are contrary to what has been found in in-person therapy, where age was observed to be a moderator but not sex (Constantino et al., 2018). Higher baseline depressive symptom severity has previously been reported to correlate with lower expectancy (Constantino et al., 2018; Tsai et al., 2014; Vîslă et al., 2019). This would indicate that individuals with more severe symptoms at baseline are a prime target group when trying to enhance expectancy before preventive interventions and that possible sex-specific responses should be considered.

The large heterogeneity in existing studies' definitions, instruments, measurement times of expectancy and outcome assessments, and intervention characteristics restricts the comparability of study results in all healthcare fields (Laferton et al., 2017; Shedden-Mora et al., 2023). Outcome expectancy has been transferred as a common factor from in-person psychotherapy to online intervention but little research is done on how comparable treatment mechanisms are across prevention vs. treatment or in-person vs. online interventions (Domhardt et al., 2021). The baseline value of expectancy being only slightly above the middle of the scale (range 3-27, *M* = 16.67, *SD* = 5.00) raises a question as to whether participants had distinct expectations about changes in depressive symptoms or if

the tendency towards the midpoint represents uncertainty of what to expect from an (online) prevention intervention (Nadler et al., 2015). This could be specific to online interventions, assuming that individuals generally have an idea of the psychotherapy effects but might be less familiar with the potential of online interventions. More (qualitative) research is needed to understand what individuals expect when signing up for a preventive online intervention for depression.

**Implications for research and practice**

Some evidence already exists that outcome expectancy can increase the intention to use online interventions (Ebert et al., 2015; Philippi et al., 2021). The results of our study suggest that it could be worthwhile to assess and foster outcome expectancy before the start of a preventive online intervention for depression, considering that is as an easily assessable, influencible and amenable characteristic before and during an intervention (Molloy et al., 2021; Stalujanis et al., 2021). However, more thoroughly designed, predictive studies targeting expectancy are highly necessary to ascertain the full impact of expectancy on the effectiveness of depression prevention.

Such studies should, first and foremost, apply validated measurement instruments of expectancy and control for confounding effects (e.g. credibility or working alliance; Constantino et al., 2011, 2018; Devilly & Borkovec, 2000). To the best of our knowledge, no comparative studies have been conducted, to assess how expectancy is influenced by the underlying disorder. This might, however, be crucial when applying the concept to depression prevention and treatment, given that depression is associated with generally more pessimistic expectations (Rief & Joormann, 2019). In addition, as a common factor in in-person psychotherapy, expectancy is discussed as being (partly) mediated through therapeutic alliance, a mechanism also found in guided online interventions (Zagorscak et al., 2020).

Expectancy should therefore be discussed with regard to the amount of guidance provided. For studies examining social anxiety interventions, Nordgreen et al. summarized that expectancy was a predictor for symptom reduction in the unguided but not in the guided intervention arm (Nordgreen et

al., 2012). However, similar conclusions could not be drawn for depression interventions (Zagorscak et al., 2020). Studying outcome expectancy in relation to the intensity of guidance (e.g. unguided, adherence-focused, guidance on demand, guided) might help to maximize individual benefits.

Future studies should also systematically assess when outcome expectancy is best assessed to be able to draw conclusions about its impact. Constantino et al (2018) combined studies that used pre-treatment and early treatment expectancies in their meta-analysis, assuming there was no relevant difference in expectancy. This assumption is challenged by the findings in online interventions that only studies assessing expectancy after exposure to the intervention observed an association with the outcome. Formerly, having some experience with the intervention was thought only important for reliable credibility measurements, but not for outcome expectancy (Constantino et al., 2011).

Finally, it would be prudent to conduct longitudinal studies to understand how initial outcome expectancy emerges and influences the interaction with the intervention while taking into account previous experience (Millard et al., 2021) and information available prior to treatment decision (Morrison et al., 2021) and level of human support. To test wheather expectancy plays a specific role in online interventions, studies directly comparing online and face-to-face preventive offers are warranted. This information could assist in designing interventions or components to foster expectancy starting from the first help-seeking impulse and continuing throughout the intervention use.

If methodologically sound studies establish outcome expectancy as a predictor of depressive outcomes in online interventions, further research should investigate whether manipulating outcome expectancy before and during an intervention (e.g., providing a strong intervention rationale, managing unrealistic expectations, providing a nontechnical overview of efficacy of online interventions) indeed results in greater effects (e.g., reduced risk for depression onset), taking into account participants' characteristics (e.g., initial symptom severity, sex).

**Limitations**

Our study contains secondary analyses of RCT data that was not originally designed to examine expectancy, thus some limitations need to be considered when interpreting the results: First, the German version of the CEQ used in the studies was self-translated, not validated in a German sample, and adapted in wording. This might have influenced the validity of the instrument (Flake & Fried, 2020).

Second, even though we combined two trials, power might have been insufficient to detect small predictive effects of expectancy, especially because the question focusing on depression onset was present in only one study. The sample size also did not allow us to consider study-level characteristics such as guidance as potential moderators, more complex associations like non-linear trends or moderating effects by multiple variables, which should be considered in future research. The sample size was partly reduced as we did not include participants in the control conditions, given the different operationalization and given that expectancy was only assessed with regard to the online intervention.

Third, although age, sex, and baseline depressive symptom severity have been included in the analyses to assess an adjusted effect of expectancy, other prognostic indicators for MDD could not be included (e.g., history of MDD, chronic medical conditions).

Fourth, previous experiences in health care are of interest to better understand outcome expectancies. Even though data on experience with previous health trainings and psychotherapy was assessed in the original studies, this data could not be used because we lacked information on whether these experiences were perceived as positive or negative (Constantino et al., 2018).

**Conclusion**

In this secondary analysis of two RCTs for indicated online depression prevention, we could not find a predictive effect of outcome expectancy on depressive symptoms at post-treatments or on depression onset. Models predicting follow-up depression scores and moderation analyses appear promising, but more research is needed to assess the potential impact of expectancy to enhance effectiveness in preventive online interventions for depression.

## Declarations

### Compliance with Ethical Standards

#### *Funding*

Funding for the original studies was received from the European Union (project number: EFRE: CCI 2007DE161PR001) and the BARMER GEK (German statutory health insurance company). The funders did not have a role in study design; data collection, analysis the interpretation of results or the decision to publish the study results. This secondary analysises did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

#### *Ethics approval*

The original studies were performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of of the Philipps University Marburg (2012-35K, PREV-DEP I) and the Leuphana University Lueneburg (Ebert201404_Depr, PREV-DEP II), respectively. No extra approval for the secondary analysis of the pseudonomized data was needed.

#### *Consent to participate*

All participants signed a written informed consent before being enrolled in the studies.

#### *Conflict of Interest*

DDE is stakeholder of the GET.ON Institute/HelloBetter, which aims to implement scientific findings related to digital health interventions into routine care. DDE has served as a consultant to/on the scientific advisory boards of Sanofi, Novartis, Minddistrict, Lantern, Schoen Kliniken, Ideamed and German health insurance companies (BARMER, Techniker Krankenkasse) and a number of federal chambers for psychotherapy. JT, PK, PC, FS, HR, DL & CB report no conflict of interest.

**Data and/or Code availability**

Access to the final pseudonomized trial dataset can be provided to fellow researchers upon request from the senior author, depending on to be specified data security and data exchange regulation agreements. The analyses scripts can be assessed via OSF: https://osf.io/mjh3d/files/osfstorage

**Authors' contribution statements**

Janika Thielecke: Conceptualization, Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing; Paula Kuper: Conceptualization, Methodology, Formal analysis, Writing - Review & Editing; David Ebert: Conceptualization, Writing - Review & Editing; Pim Cuijpers: Conceptualization, Writing - Review & Editing; Filip Smit: Conceptualization, Writing - Review & Editing; Heleen Riper: Conceptualization, Writing - Review & Editing, Dirk Lehr: Conceptualization, Writing - Review & Editing; Claudia Buntrock: Supervision, Conceptualization, Methodology, Project administration, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing;

**Acknowledgements**

We would like to thank Avery Veldhouse for proofreading the manuscript.

**References**

Alaoui, S. E., Ljótsson, B., Hedman, E., Svanborg, C., Kaldo, V., & Lindefors, N. (2016). Predicting outcome in internet-based cognitive behaviour therapy for major depression: A large cohort study of adult patients in routine psychiatric care. *PLoS ONE*, *11*(9), 1–16. https://doi.org/10.1371/journal.pone.0161191

Beatty, L., & Binnion, C. (2016). A Systematic Review of Predictors of, and Reasons for, Adherence to Online Psychological Interventions. *International Journal of Behavioral Medicine*, *23*(6), 776–794. https://doi.org/10.1007/s12529-016-9556-9

Beshai, S., Watson, L. M., Meadows, T. J. S., & Soucy, J. N. (2019). Perceptions of Cognitive-Behavioral Therapy and Antidepressant Medication for Depression After Brief Psychoeducation: Examining Shifts in Attitudes. *Behavior Therapy*, *50*(5), 851–863. https://doi.org/10.1016/j.beth.2019.01.001

Boisseau, C. L., Schwartzman, C. M., Lawton, J., & Mancebo, M. C. (2017). App-guided exposure and response prevention for obsessive compulsive disorder: An open pilot trial. *Cognitive Behaviour Therapy*, *46*(6), 447–458. https://doi.org/10.1080/16506073.2017.1321683

Braun-Koch, K., & Rief, W. (2022). Maintenance vs. Change of Negative Therapy Expectation: An Experimental Investigation Using Video Samples. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/fpsyt.2022.836227

Buntrock, C., Ebert, D. D., Lehr, D., Cuijpers, P., Riper, H., Smit, F., & Berking, M. (2014). Evaluating the efficacy and cost-effectiveness of web-based indicated prevention of major depression: Design of a randomised controlled trial. *BMC Psychiatry*. https://doi.org/10.1186/1471-244X-14-25

Buntrock, C., Ebert, D. D., Lehr, D., Smit, F., Riper, H., Berking, M., & Cuijpers, P. (2016). Effect of a web-based guided self-help intervention for prevention of major depression in adults with subthreshold

depression a randomized clinical trial. *JAMA - Journal of the American Medical Association*, *315*(17), 1854–1863. https://doi.org/10.1001/jama.2016.4326

Buntrock, C., Ebert, D., Lehr, D., Riper, H., Smit, F., Cuijpers, P., & Berking, M. (2015). Effectiveness of a Web-Based Cognitive Behavioural Intervention for Subthreshold Depression: Pragmatic Randomised Controlled Trial. *Psychotherapy and Psychosomatics*. https://doi.org/10.1159/000438673

Cavanagh, K., Shapiro, D. A., van Den Berg, S., Swain, S., Barkham, M., & Proudfoot, J. (2009). The acceptability of computer-aided cognitive behavioural therapy: A pragmatic study. *Cognitive Behaviour Therapy*, *38*(4), 235–246. https://doi.org/10.1080/16506070802561256

Chung, Y., Gelman, A., Rabe-Hesketh, S., Liu, J., & Dorie, V. (2015). Weakly Informative Prior for Point Estimation of Covariance Matrices in Hierarchical Models. *Journal of Educational and Behavioral Statistics*, *40*(2), 136–157. https://doi.org/10.3102/1076998615570945

Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A., & Liu, J. (2013). A Nondegenerate Penalized Likelihood Estimator for Variance Parameters in Multilevel Models. *Psychometrika*, *78*(4), 685–709. https://doi.org/10.1007/s11336-013-9328-2

Chung, Y., Rabe-hesketh, S., Gelman, A., Dorie, V., & Liu, J. (2012). Avoiding Boundary Estimates in Linear Mixed Models Through Weakly Informative Priors. *U.C. Berkeley Division of Biostatistics Working Paper Series*, *Working Pa*, 1–25.

Constantino, M. J., Ametrano, R. M., & Greenberg, R. P. (2012). Clinician interventions and participant characteristics that foster adaptive patient expectations for psychotherapy and psychotherapeutic change. *Psychotherapy (Chicago, Ill.)*, *49*(4), 557–569. https://doi.org/10.1037/a0029440

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. A. Z. (2011). Expectations. *Journal of Clinical Psychology*, *67*(2), 184–192. https://doi.org/10.1002/jclp.20754

Constantino, M. J., Coyne, A. E., Boswell, J. F., Iles, B. R., & Visla, A. (2018). A Meta-Analysis of the Association between Patients' Early Perception of Treatment Credibility and Their Posttreatment Outcomes. *Psychotherapy (Chicago, Ill.)*, *55*(4), 486–495. https://doi.org/10.1037/pst0000168

Crippa, J. A. S., de Lima Osório, F., Del-Ben, C. M., Filho, A. S., da Silva Freitas, M. C., & Loureiro, S. R. (2008). Comparability Between Telephone and Face-to-Face Structured Clinical Interview for DSM-IV in Assessing Social Anxiety Disorder. *Perspectives in Psychiatric Care*, *44*(4), 241–247. https://doi.org/10.1111/j.1744-6163.2008.00183.x

Cuijpers, P., De Graaf, R., & Van Dorsselaer, S. (2004). Minor depression: Risk profiles, functional disability, health care use and risk of developing major depression. *Journal of Affective Disorders*, *79*(1–3), 71–79. https://doi.org/10.1016/S0165-0327(02)00348-8

Cuijpers, P., Koole, S. L., Van Dijke, A., Roca, M., Li, J., & Reynolds, C. F. (2014). Psychotherapy for subclinical depression: Meta-analysis. *British Journal of Psychiatry*. https://doi.org/10.1192/bjp.bp.113.138784

Cuijpers, P., Pineda, B. S., Quero, S., Karyotaki, E., Struijs, S. Y., Figueroa, C. A., Llamas, J. A., Furukawa, T. A., & Muñoz, R. F. (2021). Psychological interventions to prevent the onset of depressive disorders: A meta-analysis of randomized controlled trials. *Clinical Psychology Review*, *83*, 101955. https://doi.org/10.1016/j.cpr.2020.101955

Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The Role of Common Factors in Psychotherapy Outcomes. *Annual Review of Clinical Psychology*, *15*, 207–231. https://doi.org/10.1146/annurev-clinpsy-050718-095424

Cuijpers, P., & Smit, F. (2004). Subthreshold depression as a risk indicator for major depressive disorder: A systematic review of prospective studies. *Acta Psychiatrica Scandinavica*, *109*(5), 325–331. https://doi.org/10.1111/j.1600-0447.2004.00301.x

Cuijpers, P., Smit, F., Oostenbrink, J., De Graaf, R., Ten Have, M., & Beekman, A. (2007). Economic costs of minor depression: A population-based study. *Acta Psychiatrica Scandinavica*, *115*(3), 229–236. https://doi.org/10.1111/j.1600-0447.2006.00851.x

Cuijpers, P., Vogelzangs, N., Twisk, J., Kleiboer, A., Li, J., & Penninx, B. W. (2013). Differential mortality rates in major and subthreshold depression: Meta-analysis of studies that measured both. *British Journal of Psychiatry*, *202*(1), 22–27. https://doi.org/10.1192/bjp.bp.112.112169

de Graaf, L. E., Huibers, M. J. H., Riper, H., Gerhards, S. A. H., & Arntz, A. (2009). Use and acceptability of unsupported online computerized cognitive behavioral therapy for depression and associations with clinical outcome. *Journal of Affective Disorders*, *116*(3), 227–231. https://doi.org/10.1016/j.jad.2008.12.009

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the Credibility/Expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*. https://doi.org/10.1016/S0005-7916(00)00012-4

Domhardt, M., Cuijpers, P., Ebert, D. D., & Baumeister, H. (2021). More Light? Opportunities and Pitfalls in Digitalized Psychotherapy Process Research. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.544129

Donker, T., Batterham, P. J., Warmerdam, L., Bennett, K., Bennett, A., Cuijpers, P., Griffiths, K. M., & Christensen, H. (2013). Predictors and moderators of response to internet-delivered Interpersonal Psychotherapy and Cognitive Behavior Therapy for depression. *Journal of Affective Disorders*, *151*(1), 343–351. https://doi.org/10.1016/j.jad.2013.06.020

Eaton, W. W., Badawi, M., & Melton, B. (1995). Prodromes and precursors: Epidemiologic data for primary prevention of disorders with slow onset. *American Journal of Psychiatry*, *152*(7), 967–972. https://doi.org/10.1176/ajp.152.7.967

Ebert, D. D., Berking, M., Cuijpers, P., Lehr, D., Pörtner, M., & Baumeister, H. (2015). Increasing the acceptance of internet-based mental health interventions in primary care patients with depressive symptoms. A randomized controlled trial. *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2015.01.056

Ebert, D. D., Buntrock, C., Lehr, D., Smit, F., Riper, H., Baumeister, H., Cuijpers, P., & Berking, M. (2018). Effectiveness of Web- and Mobile-Based Treatment of Subthreshold Depression With Adherence-Focused Guidance: A Single-Blind Randomized Controlled Trial. *Behavior Therapy*, *49*(1), 71–83. https://doi.org/10.1016/j.beth.2017.05.004

Ebert, D. D., & Cuijpers, P. (2018). It Is Time to Invest in the Prevention of Depression. *JAMA Network Open*, *1*(2), e180335. https://doi.org/10.1001/jamanetworkopen.2018.0335

Ebert, D. D., Cuijpers, P., Muñoz, R. F., & Baumeister, H. (2017). Prevention of Mental Health Disorders Using Internet- and Mobile-Based Interventions: A Narrative Review and Recommendations for Future Research. *Frontiers in Psychiatry*, *8*. https://doi.org/10.3389/fpsyt.2017.00116

Ebert, D. D., Van Daele, T., Nordgreen, T., Karekla, M., Compare, A., Zarbo, C., Brugnera, A., Øverland, S., Trebbi, G., Jensen, K. L., Kaehlke, F., & Baumeister, H. (2018). Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health. *European Psychologist*, *23*(2), 167–187. https://doi.org/10.1027/1016-9040/a000318

Flake, J. K., & Fried, E. I. (2020). Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them. *Advances in Methods and Practices in Psychological Science*, *3*(4), 456–465. https://doi.org/10.1177/2515245920952393

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical Psychology Review*, *26*(6), 657–678. https://doi.org/10.1016/j.cpr.2005.03.002

Hautzinger, M., Bailer, M., Hofmeister, D., & Keller, F. (2012). *Allgemeine Depressionsskala* (2nd editio). Hogrefe.

Hautzinger, M., Keller, F., & Kühner, C. (2006). *Das Beck Depressionsinventar II. Deutsche Bearbeitung und Handbuch zum BDI-II.*

Heymans, M. (2022). *psfmi: Prediction model pooling, selection and performance evaluation across multiply imputed datasets* [Manual]. https://CRAN.R-project.org/package=psfmi

Høifødt, R. S., Mittner, M., Lillevoll, K., Katla, S. K., Kolstrup, N., Eisemann, M., Friborg, O., & Waterloo, K. (2015). Predictors of response to web-based cognitive behavioral therapy with high-intensity face-to-face therapist guidance for depression: A Bayesian analysis. *Journal of Medical Internet Research*, *17*(9), e197. https://doi.org/10.2196/jmir.4351

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Kassambara, A., Kosinski, M., & Biecek, P. (2021). *survminer: Drawing survival curves using "ggplot2"* [Manual]. https://CRAN.R-project.org/package=survminer

Kenardy, J., McCafferty, K., & Rosa, V. (2003). Internet-delivered indicated prevention for anxiety disorders: A randomized controlled trial. *Behavioural and Cognitive Psychotherapy*, *31*(3), 279–289. https://doi.org/10.1017/S1352465803003047

Laferton, J. A. C., Kube, T., Salzmann, S., Auer, C. J., & Shedden-Mora, M. C. (2017). Patients' Expectations Regarding Medical Treatment: A Critical Review of Concepts and Their Assessment. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00233

Lee, Y. Y., Stockings, E. A., Harris, M. G., Doi, S. A. R., Page, I. S., Davidson, S. K., & Barendregt, J. J. (2019). The risk of developing major depression among individuals with subthreshold depression: A systematic review and meta-analysis of longitudinal cohort studies. *Psychological Medicine*, *49*(1), 92–102. https://doi.org/10.1017/s0033291718000557

Lewinsohn, P. M., Seeley, J. R., Roberts, R. E., & Allen, N. B. (1997). Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychology and Aging*, *12*(2), 277–287. https://doi.org/10.1037/0882-7974.12.2.277

Little, R. J. A., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. John Wiley & Sons, Inc. https://doi.org/10.1002/9781119013563

Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). *Clinical Psychology & Psychotherapy*, *18*(1), 75–79. https://doi.org/10.1002/cpp.693

Lüdtke, T., Westermann, S., Pult, L. K., Schneider, B. C., Pfuhl, G., & Moritz, S. (2018). Evaluation of a brief unguided psychological online intervention for depression: A controlled trial including exploratory moderator analyses. *Internet Interventions*, *13*(June), 73–81. https://doi.org/10.1016/j.invent.2018.06.004

Lyketsos, C. G., Nestadt, G., Cwi, J., Heithoff, K., & et al. (1994). The Life Chart Interview: A standardized method to describe the course of psychopathology. *International Journal of Methods in Psychiatric Research*, *4*, 143–155.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

Millard, E., Cardona, J., Fernandes, J., Priebe, S., & Carr, C. (2021). I know what I like, and I like what I know: Patient preferences and expectations when choosing an arts therapies group. *The Arts in Psychotherapy*, *75*, 101829. https://doi.org/10.1016/j.aip.2021.101829

Mitchell, N., & Gordon, P. K. (2007). Attitudes Towards Computerized CBT for Depression Amongst a Student Population. *Behavioural and Cognitive Psychotherapy*, *35*(4), 421–430. https://doi.org/10.1017/S1352465807003700

Molloy, A., Ellis, D. M., Su, L., & Anderson, P. L. (2021). Improving Acceptability and Uptake Behavior for Internet-Based Cognitive-Behavioral Therapy. *Frontiers in Digital Health*, *3*(March), 1–11. https://doi.org/10.3389/fdgth.2021.653686

Morrison, N. R., Constantino, M. J., Santorelli, G. D., McVicar, E. L., & Ametrano, R. M. (2021). Determinants of psychotherapy outcome expectations: Qualitative accounts from socially anxious college students. *Counselling and Psychotherapy Research*, *21*(4), 781–791. https://doi.org/10.1002/capr.12411

Muñoz, R. F., Pineda, B. S., Barrera, A. Z., Bunge, E., & Leykin, Y. (2021). Digital tools for prevention and treatment of depression: Lessons from the institute for international internet interventions for health. *Clinica y Salud*, *32*(1), 37–40. https://doi.org/10.5093/CLYSA2021A2

Musiat, P., Goldstone, P., & Tarrier, N. (2014). Understanding the acceptability of E-mental health—Attitudes and expectations towards computerised self-help treatments for mental health problems. *BMC Psychiatry*, *14*(1), 1–8. https://doi.org/10.1186/1471-244X-14-109

Myklebost, S. B., Gjestad, R., Inal, Y., Hammar, A., & Nordgreen, T. (2022). Predictors of Treatment Response to an Internet-Delivered Intervention Targeting Residual Cognitive Symptoms After Major Depressive Disorder. *Frontiers in Psychiatry*, *13*(March). https://doi.org/10.3389/fpsyt.2022.795698

Nadler, J. T., Weston, R., & Voyles, E. C. (2015). Stuck in the Middle: The Use and Interpretation of Mid-Points in Items on Questionnaires. *The Journal of General Psychology*, *142*(2), 71–89. https://doi.org/10.1080/00221309.2014.994590

Nordgreen, T., Havik, O. E., Öst, L. G., Furmark, T., Carlbring, P., & Andersson, G. (2012). Outcome predictors in guided and unguided self-help for social anxiety disorder. *Behaviour Research and Therapy*, *50*(1), 13–21. https://doi.org/10.1016/j.brat.2011.10.009

Pearson, R., Pisner, D., Meyer, B., Shumake, J., & Beevers, C. G. (2019). A machine learning ensemble to predict treatment outcomes following an Internet intervention for depression. *Psychological Medicine*, *49*(14), 2330–2341. https://doi.org/10.1017/S003329171800315X

Philippi, P., Baumeister, H., Apolinário-Hagen, J., Ebert, D. D., Hennemann, S., Kott, L., Lin, J., Messner, E. M., & Terhorst, Y. (2021). Acceptance towards digital health interventions – Model validation and further development of the Unified Theory of Acceptance and Use of Technology. *Internet Interventions*, *26*. https://doi.org/10.1016/j.invent.2021.100459

R Core Team. (2022). *R: A language and environment for statistical computing* [Manual]. https://www.R-project.org/

Reins, J. A., Buntrock, C., Zimmermann, J., Grund, S., Harrer, M., Lehr, D., Baumeister, H., Weisel, K., Domhardt, M., Imamura, K., Kawakami, N., Spek, V., Nobis, S., Snoek, F., Cuijpers, P., Klein, J. P., Moritz, S., & Ebert, D. D. (2021). Efficacy and Moderators of Internet-Based Interventions in Adults with Subthreshold Depression: An Individual Participant Data Meta-Analysis of Randomized Controlled Trials. *Psychotherapy and Psychosomatics*, *90*(2), 94–106. https://doi.org/10.1159/000507819

Rief, W., & Joormann, J. (2019). Revisiting the cognitive model of depression: The role of expectations. *Clinical Psychology in Europe*, *1*(1). https://doi.org/10.32872/cpe.v1i1.32605

Riley, R. D., Debray, T. P. A., Fisher, D., Hattle, M., Marlin, N., Hoogland, J., Gueyffier, F., Staessen, J. A., Wang, J., Moons, K. G. M., Reitsma, J. B., & Ensor, J. (2020). Individual participant data Meta-analysis to examine interactions between treatment effect and Participant-level covariates: Statistical recommendations for conduct and planning. *Statistics in Medicine*, *39*(15), 2115–2137. https://doi.org/10.1002/sim.8516

Robitzsch, A., & Grund, S. (2022). *miceadds: Some additional multiple imputation functions, especially for "mice"* [Manual]. https://CRAN.R-project.org/package=miceadds

Rubin, D. B. (Ed.). (1987). *Multiple Imputation for Nonresponse in Surveys* (Vol. 81). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470316696

Rucci, P., Gherardi, S., Tansella, M., Piccinelli, M., Berardi, D., Bisoffi, G., Corsino, M. A., & Pini, S. (2003). Subthreshold psychiatric disorders in primary care: Prevalence and associated characteristics. *Journal of Affective Disorders*, *76*(1–3), 171–181. https://doi.org/10.1016/S0165-0327(02)00087-3

Schafer, J. L., & Yucel, R. M. (2002). Computational Strategies for Multivariate Linear Mixed-Effects Models With Missing Values. *Journal of Computational and Graphical Statistics*, *11*(2), 437–457. https://doi.org/10.1198/106186002760180608

Shedden-Mora, M. C., Alberts, J., Petrie, K. J., Laferton, J. A. C., von Blanckenburg, P., Kohlmann, S., Nestoriuc, Y., & Löwe, B. (2023). The Treatment Expectation Questionnaire (TEX-Q): Validation of a generic multidimensional scale measuring patients' treatment expectations. *PLOS ONE*, *18*(1), e0280472. https://doi.org/10.1371/journal.pone.0280472

Stalujanis, E., Neufeld, J., Glaus Stalder, M., Belardi, A., Tegethoff, M., & Meinlschmidt, G. (2021). Induction of efficacy expectancies in an ambulatory smartphone-based digital placebo mental health intervention: Randomized controlled trial. *JMIR MHealth and UHealth*, *9*(2). https://doi.org/10.2196/20329

Therneau, T. M. (2022). *A package for survival analysis in R* [Manual]. https://CRAN.R-project.org/package=survival

Thiruchselvam, T., Dozois, D. J. A., Bagby, R. M., Lobo, D. S. S., Ravindran, L. N., & Quilty, L. C. (2019). The role of outcome expectancy in therapeutic change across psychotherapy versus pharmacotherapy for depression. *Journal of Affective Disorders*, *251*(August 2018), 121–129. https://doi.org/10.1016/j.jad.2019.01.046

Tsai, M., Ogrodniczuk, J. S., Sochting, I., & Mirmiran, J. (2014). Forecasting Success: Patients' Expectations for Improvement and Their Relations to Baseline, Process and Outcome Variables in Group

Cognitive-Behavioural Therapy for Depression: Forecasting Success. *Clinical Psychology & Psychotherapy*, *21*(2), 97–107. https://doi.org/10.1002/cpp.1831

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Vîslă, A., Constantino, M. J., Newkirk, K., Ogrodniczuk, J. S., & Söchting, I. (2018). The relation between outcome expectation, therapeutic alliance, and outcome among depressed patients in group cognitive-behavioral therapy. *Psychotherapy Research*, *28*(3), 446–456. https://doi.org/10.1080/10503307.2016.1218089

Vîslă, A., Flückiger, C., Constantino, M. J., Krieger, T., & Grosse Holtforth, M. (2019). Patient characteristics and the therapist as predictors of depressed patients' outcome expectation over time: A multilevel analysis. *Psychotherapy Research*, *29*(6), 709–722. https://doi.org/10.1080/10503307.2018.1428379

Volz, H.-P., Stirnweiß, J., Kasper, S., Möller, H.-J., & Seifritz, E. (2022). Subthreshold depression – concept, operationalisation and epidemiological data. A scoping review. *International Journal of Psychiatry in Clinical Practice*, *0*(0), 1–15. https://doi.org/10.1080/13651501.2022.2087530

Wagner, H. R., Burns, B. J., Broadhead, W. E., Yarnall, K. S. H., Sigmon, A., & Gaynes, B. N. (2000). Minor depression in family practice: Functional morbidity, co-morbidity, service utilization and outcomes. *Psychological Medicine*, *30*(6), 1377–1390. https://doi.org/10.1017/S0033291799002998

Zagorscak, P., Heinrich, M., Schulze, J., Böttcher, J., & Knaevelsrud, C. (2020). Factors contributing to symptom change in standardized and individualized Internet-based interventions for depression: A randomized-controlled trial. *Psychotherapy (Chicago, Ill.)*, *57*(2), 237–251. https://doi.org/10.1037/pst0000276

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

**Online Resource 1 – Full model specifications**

***RQ1: Are outcome expectancy and depression symptom severity at post and follow-up related?***

***Hypothesis 1***

H0(1) = Outcome expectancy at baseline does not predict depressive symptomatology at post-treatment.
H1(1) = Outcome expectancy at baseline does predict depressive symptomatology at post-treatment.

*Model specification:*

```
with(implist,expr= blmer(cesd_1 ~ 0+ factor(trial)+ trial1_cesd_0_c+trial2_cesd_0_c
+trial1_age_c+trial2_age_c+ trial1_sex_c +  trial2_sex_c +ceq_exp + (0 + ceq_exp | trial),
control = lmerControl(optimize= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.05) ))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | lower 95%-CI | upper 95%-CI |
|---|---|---|---|---|---|---|---|
| factor(trial)1 | 24.6310 | 3.1526 | 7.8130 | 160.0 | 0.0000 | 18.4049 | 30.8570 |
| factor(trial)2 | 21.0965 | 2.0241 | 10.4230 | 244.0 | 0.0000 | 17.1096 | 25.0830 |
| trial1_cesd_0_c | 0.5350 | 0.1420 | 3.7680 | 186.0 | 0.0002 | 0.2549 | 0.8150 |
| trial2_cesd_0_c | 0.4669 | 0.0782 | 5.9710 | 235.0 | 0.0000 | 0.3129 | 0.6210 |
| trial1_age_c | 0.0105 | 0.0781 | 0.1340 | 225.0 | 0.8934 | -0.1435 | 0.1640 |
| trial2_age_c | 0.0888 | 0.0530 | 1.6740 | 208.0 | 0.0956 | -0.0158 | 0.1930 |
| trial1_sex_c | -1.9171 | 2.2667 | -0.8460 | 205.0 | 0.3987 | -6.3862 | 2.5520 |
| trial2_sex_c | 2.5911 | 1.4078 | 1.8400 | 226.0 | 0.0670 | -0.1831 | 5.3650 |
| ceq_exp | -0.2981 | 0.2197 | -1.3570 | 272.0 | 0.1760 | -0.7307 | 0.1340 |

Corresponding author: Janika Thielecke, Technical University Munich (TUM), Department of Sports and Health Sciences, Georg-Brauchle-Ring 60/62, 80992 Munich. Mail: janika.thielecke@tum.de

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

***Hypothesis 2***

H0(2) = Outcome expectancy at baseline does not predict depressive symptomatology at follow-up assessment.
H1(2) = Outcome expectancy at baseline does predict depressive symptomatology at follow-up assessment.

*Model specification:*

```
with(implist,expr= blmer(cesd_2 ~ 0+ factor(trial)+ trial1_cesd_0_c+trial2_cesd_0_c
+trial1_age_c+trial2_age_c+ trial1_sex_c +  trial2_sex_c +ceq_exp + (0 + ceq_exp | trial),
control = lmerControl(optimize= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.05) ))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | lower 95%-CI | upper 95%-CI |
|---|---|---|---|---|---|---|---|
| factor(trial)1 | 24.3340 | 3.1270 | 7.7820 | 141.0 | 0.0000 | 18.1523 | 30.5158 |
| factor(trial)2 | 23.8453 | 2.1099 | 11.3010 | 202.0 | 0.0000 | 19.6850 | 28.0057 |
| trial1_cesd_0_c | 0.6002 | 0.1500 | 4.0010 | 156.0 | 0.0001 | 0.3039 | 0.8965 |
| trial2_cesd_0_c | 0.5146 | 0.0828 | 6.2170 | 196.0 | 0.0000 | 0.3514 | 0.6779 |
| trial1_age_c | 0.0255 | 0.0880 | 0.2900 | 143.0 | 0.7723 | -0.1484 | 0.1994 |
| trial2_age_c | 0.1643 | 0.0579 | 2.8350 | 152.0 | 0.0052 | 0.0498 | 0.2787 |
| trial1_sex_c | -0.9581 | 2.3131 | -0.4140 | 199.0 | 0.6792 | -5.5195 | 3.6033 |
| trial2_sex_c | 3.8596 | 1.4880 | 2.5940 | 189.0 | 0.0102 | 0.9244 | 6.7948 |
| ceq_exp | -0.3913 | 0.1819 | -2.1520 | 244.0 | 0.0324 | -0.7495 | -0.0331 |

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

***RQ2: Are outcome expectancy and time to depression onset within a 12-month follow-up period related?***

***Hypothesis 3***

H0(3) = In the subset of participants with clinical interviews, outcome expectancy at baseline does not predict time to depression onset at 12-months follow-up assessment.
H1(3) = In the subset of participants with clinical interviews, outcome expectancy at baseline does predict time to depression onset at 12-months follow-up assessment

*Model specification:*

```
Surv(fu.time, dep.status, type = "right") ~ ceq.exp + sex + age.c + cesd.0.c
```

*Full model:*

| | coef | exp(coef) | se(coef) | z | p.value | lower 95%-CI | upper 95%-CI |
|---|---|---|---|---|---|---|---|
| ceq.exp | -0.028576 | 0.9718 | 0.024588 | -1.162 | 0.2452 | 0.9261 | 1.02 |
| sex1 | 0.155361 | 1.1681 | 0.302467 | 0.514 | 0.6075 | 0.6457 | 2.113 |
| age.c | 0.007878 | 1.0079 | 0.010941 | 0.72 | 0.4715 | 0.9865 | 1.03 |
| cesd.0.c | 0.068768 | 1.0712 | 0.017678 | 3.89 | 0.0001 | 1.0347 | 1.109 |

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

***RQ3: Are outcome expectancy and "close-to-symptom-free"-status at post and follow-up related?***

***Hypothesis 4***

H0(4) = Outcome expectancy at baseline does not predict "close-to-symptom-free"-status at post-treatment.
H1(4) = Outcome expectancy at baseline does predict "close-to-symptom-free"-status at post-treatment.

*Model specification:*

```
with(implist, expr= bglmer(symptom_free_1 ~ 0 + factor(trial+trial1_cesd_0_c+
trial2_cesd_0_c+trial1_age_c+trial2_age_c+ + trial1_sex_c + trial2_sex_c + ceq_exp + (0 +
ceq_exp | trial),
family=binomial(link= "cloglog"),
control = glmerControl(optimizer= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.05)))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | RR | RR CI.lower | RR CI.upper |
|---|---|---|---|---|---|---|---|---|
| factor(trial)1 | -1.4615 | 0.4311 | -3.3903 | 219.9 | 0.0008 | 0.2319 | 0.0992 | 0.5423 |
| factor(trial)2 | -1.2606 | 0.4109 | -3.0681 | 225.9 | 0.0024 | 0.2835 | 0.1262 | 0.6370 |
| trial1_cesd_0_c | -0.0904 | 0.0336 | -2.6937 | 202.4 | 0.0077 | 0.9136 | 0.8551 | 0.9761 |
| trial2_cesd_0_c | -0.0661 | 0.0189 | -3.4976 | 254.8 | 0.0006 | 0.9360 | 0.9018 | 0.9715 |
| trial1_age_c | -0.0009 | 0.0152 | -0.0571 | 248.2 | 0.9545 | 0.9991 | 0.9697 | 1.0295 |
| trial2_age_c | -0.0154 | 0.0109 | -1.4191 | 233 | 0.1572 | 0.9847 | 0.9638 | 1.0060 |
| trial1_sex_c | 0.3696 | 0.4136 | 0.8935 | 218.3 | 0.3726 | 1.4471 | 0.6404 | 3.2698 |
| trial2_sex_c | -0.5263 | 0.3175 | -1.6579 | 213.7 | 0.0988 | 0.5908 | 0.3160 | 1.1045 |
| ceq_exp | 0.0363 | 0.0225 | 1.6113 | 225.3 | 0.1085 | 1.0369 | 0.9920 | 1.0840 |

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

***Hypothesis 5***

H0(5) = Outcome expectancy at baseline does not predict "close-to-symptom-free"-status at follow-up assessment.
H1(5) = Outcome expectancy at baseline does predict "close-to-symptom-free"-status at follow-up assessment.

*Model specification:*

```
fit.blmer <- with(implist,expr= bglmer(symptom_free_2 ~ 0 + trial + trial1_cesd_0_c
+trial2_cesd_0_c +trial1_age_c+trial2_age_c+ + trial1_sex_c +  trial2_sex_c + ceq_exp +
trial1_symptom_free_1_c +  trial2_symptom_free_1_c + (0 + ceq_exp | trial),
family=binomial(link= "cloglog"),
control = glmerControl(optimize= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.01)))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | RR | RR CI.lower | RR CI.upper |
|---|---|---|---|---|---|---|---|---|
| trial1 | -1.3466 | 0.4490 | -2.9992 | 157.0 | 0.0031 | 0.2601 | 0.1072 | 0.6314 |
| trial2 | -1.7490 | 0.4514 | -3.8747 | 159.7 | 0.0002 | 0.1740 | 0.0713 | 0.4242 |
| trial1_cesd_0_c | -0.0435 | 0.0297 | -1.4644 | 182.3 | 0.1448 | 0.9575 | 0.9030 | 1.0152 |
| trial2_cesd_0_c | -0.0550 | 0.0233 | -2.3584 | 181.1 | 0.0194 | 0.9464 | 0.9038 | 0.9910 |
| trial1_age_c | -0.0069 | 0.0139 | -0.4970 | 171.7 | 0.6198 | 0.9931 | 0.9661 | 1.0208 |
| trial2_age_c | -0.0145 | 0.0114 | -1.2722 | 182.6 | 0.2049 | 0.9856 | 0.9638 | 1.0080 |
| trial1_sex_c | -0.1083 | 0.4119 | -0.2629 | 184.3 | 0.7929 | 0.8974 | 0.3981 | 2.0225 |
| trial2_sex_c | -0.2840 | 0.3396 | -0.8362 | 186.1 | 0.4041 | 0.7528 | 0.3852 | 1.4711 |
| ceq_exp | 0.0599 | 0.0238 | 2.5178 | 164.9 | 0.0128 | 1.0618 | 1.0130 | 1.1128 |
| trial1_symptom_free_1_c | -1.0023 | 0.3616 | -2.7720 | 182.5 | 0.0061 | 0.3670 | 0.1798 | 0.7491 |
| trial2_symptom_free_1_c | -1.4666 | 0.2976 | -4.9277 | 169.0 | 0.0000 | 0.2307 | 0.1282 | 0.4152 |

**Supplement Table 1**

*p-values adjustment for multiple testing in hypotheses 1-5 using the Bonferroni-Holm method*

| hypothesis | $p_{unadjusted}$ | $p_{adjusted}$ |
|---|---|---|
| H5: CEQ -> Close-to-symptom-free-status follow-up | 0.0128 | 0.0638 |
| H2: CEQ -> CES-D follow-up | 0.0324 | 0.1296 |
| H4: CEQ -> Close-to-symptom-free-status post | 0.1085 | 0.3255 |
| H1: CEQ -> CES-D post | 0.1760 | 0.3520 |
| H3: CEQ -> Depression onset | 0.2452 | 0.3520 |

*Note.* Table sorted according to original unadjusted p-values.

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

***Exploratory RQ1: Is the relation between outcome expectancy and depression outcome at post-treatment moderated by baseline characteristics?***

***Exploratory Hypothesis 1***

Exploratory H0(1): Age does not moderate the effect of outcome expectancy on depressive symptoms at post-treatment.
Exploratory H1(1): Age moderates the effect of outcome expectancy on depressive symptoms at post-treatment.

*Model specification:*

```
ws.interaction.age <-ceq_exp_c*age_c

with(implist, expr= blmer(cesd_1 ~ 0 +  trial + trial1_cesd_0_c +trial2_cesd_0_c
+trial1_age_c+trial2_age_c + trial1_sex_c+trial2_sex_c+ trial1_ceq_exp_c +
trial2_ceq_exp_c + ws.interaction.age + (0 + ceq_exp |trial),
control = lmerControl(optimize= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.05)))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | lower 95%-CI | upper 95%-CI |
|---|---|---|---|---|---|---|---|
| trial1 | 25.1966 | 3.4203 | 7.3667 | 172.6 | 0.0000 | 18.4455 | 31.9477 |
| trial2 | 20.8763 | 2.0971 | 9.9546 | 246.7 | 0.0000 | 16.7457 | 25.0069 |
| trial1_cesd_0_c | 0.5317 | 0.1422 | 3.7394 | 184.7 | 0.0002 | 0.2512 | 0.8123 |
| trial2_cesd_0_c | 0.4643 | 0.0784 | 5.9250 | 232.8 | 0.0000 | 0.3099 | 0.6186 |
| trial1_age_c | -0.1216 | 0.1489 | -0.8169 | 236.7 | 0.4148 | -0.4150 | 0.1717 |
| trial2_age_c | -0.0523 | 0.1396 | -0.3746 | 239.6 | 0.7083 | -0.3273 | 0.2227 |
| trial1_sex_c | -1.9389 | 2.2687 | -0.8546 | 203.7 | 0.3938 | -6.4120 | 2.5342 |
| trial2_sex_c | 2.6589 | 1.4100 | 1.8857 | 224.3 | 0.0606 | -0.1197 | 5.4375 |
| trial1_ceq_exp_c | -0.4296 | 80.9893 | -0.0053 | 289.0 | 0.9958 | -159.8333 | 158.9741 |
| trial2_ceq_exp_c | -0.1998 | 80.9891 | -0.0025 | 289.0 | 0.9980 | -159.6031 | 159.2036 |
| ws.interaction.age | 0.0083 | 0.0075 | 1.1068 | 260.9 | 0.2694 | -0.0065 | 0.0231 |

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

***Exploratory Hypothesis 2***

Exploratory H0(2): Sex does not moderate the effect of outcome expectancy on depressive symptoms at post-treatment.
Exploratory H1(2): Sex moderates the effect of outcome expectancy on depressive symptoms at post-treatment.

*Model specification:*

```
ws.interaction.sex <-ceq_exp_c*sex

with(implist, expr= blmer(cesd_1 ~ 0 +  trial + trial1_cesd_0_c +trial2_cesd_0_c
+trial1_age_c+trial2_age_c + trial1_sex_c+trial2_sex_c + trial1_ceq_exp_c +
trial2_ceq_exp_c + ws.interaction.sex + (0 + ceq_exp |trial),
control = lmerControl(optimize= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.05)))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | lower 95%-CI | upper 95%-CI |
|---|---|---|---|---|---|---|---|
| trial1 | 25.0347 | 3.4051 | 7.3521 | 171.9 | 0.0000 | 18.3135 | 31.7559 |
| trial2 | 21.3247 | 2.1051 | 10.1299 | 243.7 | 0.0000 | 17.1781 | 25.4712 |
| trial1_cesd_0_c | 0.5336 | 0.1414 | 3.7746 | 184.4 | 0.0002 | 0.2547 | 0.8125 |
| trial2_cesd_0_c | 0.4556 | 0.0779 | 5.8456 | 234.2 | 0.0000 | 0.3020 | 0.6092 |
| trial1_age_c | 0.0121 | 0.0784 | 0.1544 | 223.1 | 0.8775 | -0.1425 | 0.1667 |
| trial2_age_c | 0.0938 | 0.0529 | 1.7711 | 205.2 | 0.0780 | -0.0106 | 0.1981 |
| trial1_sex_c | -9.6631 | 4.4398 | -2.1765 | 223.2 | 0.0306 | -18.4124 | -0.9138 |
| trial2_sex_c | -5.0249 | 3.9122 | -1.2844 | 251.8 | 0.2002 | -12.7296 | 2.6798 |
| trial1_ceq_exp_c | -0.9658 | 80.5190 | -0.0120 | 289.0 | 0.9904 | -159.4437 | 157.5122 |
| trial2_ceq_exp_c | -0.8011 | 80.5189 | -0.0099 | 289.0 | 0.9921 | -159.2790 | 157.6767 |
| ws.interaction.sex | 0.4608 | 0.2212 | 2.0832 | 253.4 | 0.0382 | 0.0252 | 0.8965 |

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

***Exploratory Hypothesis 3***

Exploratory H0(3): Baseline depressive symptoms does not moderate the effect of outcome expectancy on depressive symptoms at post-treatment.
Exploratory H1(3): Baseline depressive symptoms moderates the effect of outcome expectancy on depressive symptoms at post-treatment.

*Model specification:*

```
ws.interaction.cesd_0 <-ceq_exp_c*cesd_0_c

with(implist, expr= blmer(cesd_1 ~ 0 +  trial + trial1_cesd_0_c +trial2_cesd_0_c
+trial1_age_c+trial2_age_c + trial1_sex_c+trial2_sex_c + trial1_ceq_exp_c
+trial2_ceq_exp_c + ws.interaction.cesd_0 + (0 + ceq_exp |trial),
control = lmerControl(optimize= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.05)))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | lower 95%-CI | upper 95%-CI |
|---|---|---|---|---|---|---|---|
| trial1 | 25.2781 | 3.4021 | 7.4302 | 171.9 | 0.0000 | 18.5629 | 31.9932 |
| trial2 | 20.8803 | 2.0879 | 10.0005 | 245.0 | 0.0000 | 16.7677 | 24.9929 |
| trial1_cesd_0_c | 0.9434 | 0.2425 | 3.8909 | 216.4 | 0.0001 | 0.4655 | 1.4214 |
| trial2_cesd_0_c | 0.8994 | 0.2286 | 3.9350 | 211.4 | 0.0001 | 0.4488 | 1.3500 |
| trial1_age_c | 0.0188 | 0.0785 | 0.2391 | 222.6 | 0.8112 | -0.1359 | 0.1734 |
| trial2_age_c | 0.0920 | 0.0530 | 1.7359 | 203.3 | 0.0841 | -0.0125 | 0.1965 |
| trial1_sex_c | -1.9212 | 2.2574 | -0.8511 | 202.5 | 0.3957 | -6.3723 | 2.5299 |
| trial2_sex_c | 2.8279 | 1.4029 | 2.0157 | 225.5 | 0.0450 | 0.0633 | 5.5924 |
| trial1_ceq_exp_c | -0.4256 | 80.4910 | -0.0053 | 289.0 | 0.9958 | -158.8486 | 157.9973 |
| trial2_ceq_exp_c | -0.1962 | 80.4909 | -0.0024 | 289.0 | 0.9981 | -158.6188 | 158.2265 |
| ws.interaction.cesd_0 | -0.0252 | 0.0123 | -2.0575 | 228.2 | 0.0408 | -0.0494 | -0.0011 |

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

**Online Resource 2 – Additional analysis: Moderation at Follow-Up**

***Additional Exploratory RQ: Is the relation between outcome expectancy and depression outcome at follow-up moderated by baseline characteristics?***

***Exploratory Hypothesis 4***

Exploratory H0(4): Age does not moderate the effect of outcome expectancy on depressive symptoms at follow-up assessment..
Exploratory H1(4): Age moderates the effect of outcome expectancy on depressive symptoms at follow-up assessment..

*Model specification:*

```
ws.interaction.age <-ceq_exp_c*age_c

with(implist, expr= blmer(cesd_2 ~ 0 +  trial + trial1_cesd_0_c +trial2_cesd_0_c
+trial1_age_c+trial2_age_c + trial1_sex_c+trial2_sex_c+ trial1_ceq_exp_c +
trial2_ceq_exp_c + ws.interaction.age + (0 + ceq_exp |trial),
control = lmerControl(optimize= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.05)))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | lower 95%-CI | upper 95%-CI |
|---|---|---|---|---|---|---|---|
| trial1 | 24.6997 | 3.6450 | 6.7764 | 141.9 | 0.0000 | 17.4942 | 31.9051 |
| trial2 | 23.6968 | 2.2172 | 10.6877 | 210.1 | 0.0000 | 19.3260 | 28.0677 |
| trial1_cesd_0_c | 0.5992 | 0.1505 | 3.9824 | 155.7 | 0.0001 | 0.3020 | 0.8964 |
| trial2_cesd_0_c | 0.5136 | 0.0830 | 6.1858 | 195.3 | 0.0000 | 0.3499 | 0.6774 |
| trial1_age_c | -0.0238 | 0.1678 | -0.1417 | 153.8 | 0.8875 | -0.3553 | 0.3077 |
| trial2_age_c | 0.1102 | 0.1497 | 0.7362 | 191.4 | 0.4625 | -0.1851 | 0.4055 |
| trial1_sex_c | -0.9614 | 2.3210 | -0.4142 | 197.4 | 0.6792 | -5.5384 | 3.6157 |
| trial2_sex_c | 3.8885 | 1.4923 | 2.6057 | 188.8 | 0.0099 | 0.9448 | 6.8322 |
| trial1_ceq_exp_c | -0.4509 | 82.2459 | -0.0055 | 289.0 | 0.9956 | -162.3279 | 161.4262 |
| trial2_ceq_exp_c | -0.3498 | 82.2458 | -0.0043 | 289.0 | 0.9966 | -162.2264 | 161.5269 |
| ws.interaction.age | 0.0032 | 0.0081 | 0.3921 | 207.1 | 0.6954 | -0.0128 | 0.0191 |

Corresponding author: Janika Thielecke, Corresponding author: Janika Thielecke, Technical University Munich (TUM), Department of Sports and Health Sciences, Georg-Brauchle-Ring 60/62, 80992 Munich. Mail: janika.thielecke@tum.de

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

***Exploratory Hypothesis 5***

Exploratory H0(5): Sex does not moderate the effect of outcome expectancy on depressive symptoms at follow-up assessment.
Exploratory H1(5): Sex moderates the effect of outcome expectancy on depressive symptoms at follow-up assessment.

*Model specification:*

```
ws.interaction.sex <-ceq_exp_c*sex

with(implist, expr= blmer(cesd_2 ~ 0 +  trial + trial1_cesd_0_c +trial2_cesd_0_c
+trial1_age_c+trial2_age_c + trial1_sex_c+trial2_sex_c + trial1_ceq_exp_c +
trial2_ceq_exp_c + ws.interaction.sex + (0 + ceq_exp |trial),
control = lmerControl(optimize= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.05)))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | lower 95%-CI | upper 95%-CI |
|---|---|---|---|---|---|---|---|
| trial1 | 24.6771 | 3.6373 | 6.7845 | 143.2 | 0.0000 | 17.4875 | 31.8668 |
| trial2 | 23.7766 | 2.2463 | 10.5847 | 204.0 | 0.0000 | 19.3477 | 28.2056 |
| trial1_cesd_0_c | 0.6003 | 0.1503 | 3.9941 | 156.0 | 0.0001 | 0.3034 | 0.8972 |
| trial2_cesd_0_c | 0.5124 | 0.0832 | 6.1561 | 194.8 | 0.0000 | 0.3482 | 0.6765 |
| trial1_age_c | 0.0279 | 0.0897 | 0.3108 | 140.2 | 0.7564 | -0.1494 | 0.2051 |
| trial2_age_c | 0.1649 | 0.0582 | 2.8342 | 151.0 | 0.0052 | 0.0500 | 0.2799 |
| trial1_sex_c | -2.4711 | 4.9809 | -0.4961 | 150.9 | 0.6205 | -12.3123 | 7.3702 |
| trial2_sex_c | 2.3686 | 4.4227 | 0.5355 | 166.1 | 0.5930 | -6.3633 | 11.1004 |
| trial1_ceq_exp_c | -0.5551 | 82.2290 | -0.0068 | 289.0 | 0.9946 | -162.3988 | 161.2885 |
| trial2_ceq_exp_c | -0.4669 | 82.2289 | -0.0057 | 289.0 | 0.9955 | -162.3103 | 161.3766 |
| ws.interaction.sex | 0.0905 | 0.2512 | 0.3604 | 164.4 | 0.7190 | -0.4054 | 0.5865 |

Online Resource to: Thielecke, J., Kuper, P., Ebert, D.D., Cuijpers, P., Smit, F., Riper, H., Lehr, D., & Buntrock, C. (2023). Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomized controlled trials

***Exploratory Hypothesis 6***

Exploratory H0(6): Baseline depressive symptoms does not moderate the effect of outcome expectancy on depressive symptoms at follow-up assessment.
Exploratory H1(6): Baseline depressive symptoms moderates the effect of outcome expectancy on depressive symptoms at follow-up assessment.

*Model specification:*

```
ws.interaction.cesd_0 <-ceq_exp_c*cesd_0_c

with(implist, expr= blmer(cesd_2 ~ 0 +  trial + trial1_cesd_0_c +trial2_cesd_0_c
+trial1_age_c+trial2_age_c + trial1_sex_c+trial2_sex_c + trial1_ceq_exp_c
+trial2_ceq_exp_c + ws.interaction.cesd_0 + (0 + ceq_exp |trial),
control = lmerControl(optimize= "Nelder_Mead"),
cov.prior = trial~"gamma"(shape=1.5, rate=0.05)))
```

*Full model:*

| term | estimate | std.error | statistic | df | p.value | lower 95%-CI | upper 95%-CI |
|---|---|---|---|---|---|---|---|
| trial1 | 24.7355 | 3.6373 | 6.8006 | 141.9 | 0.0000 | 17.5452 | 31.9257 |
| trial2 | 23.7056 | 2.2155 | 10.7001 | 208.9 | 0.0000 | 19.3381 | 28.0732 |
| trial1_cesd_0_c | 0.8174 | 0.2581 | 3.1675 | 180.9 | 0.0018 | 0.3082 | 1.3266 |
| trial2_cesd_0_c | 0.7445 | 0.2428 | 3.0660 | 177.9 | 0.0025 | 0.2653 | 1.2236 |
| trial1_age_c | 0.0303 | 0.0898 | 0.3379 | 138.2 | 0.7360 | -0.1473 | 0.2080 |
| trial2_age_c | 0.1659 | 0.0581 | 2.8541 | 150.1 | 0.0049 | 0.0510 | 0.2807 |
| trial1_sex_c | -0.9583 | 2.3153 | -0.4139 | 197.7 | 0.6794 | -5.5241 | 3.6075 |
| trial2_sex_c | 3.9866 | 1.4886 | 2.6781 | 191.1 | 0.0080 | 1.0504 | 6.9228 |
| trial1_ceq_exp_c | -0.4496 | 82.0792 | -0.0055 | 289.0 | 0.9956 | -161.9983 | 161.0992 |
| trial2_ceq_exp_c | -0.3487 | 82.0790 | -0.0042 | 289.0 | 0.9966 | -161.8971 | 161.1997 |
| ws.interaction.cesd_0 | -0.0134 | 0.0131 | -1.0258 | 189.8 | 0.3063 | -0.0392 | 0.0124 |

# 8

## ARTICLE 6 — DIFFERENTIAL EFFICACY OF WEB-BASED SLEEP INTERVENTIONS ON DEPRESSIVE SYMPTOMS

Title:
Who benefits from indirect prevention and treatment of depression using online interventions for insomnia? Results from an individual-participant-data meta-analysis

Summary:
Psychological treatments both in person and delivered online are effective in reducing symptoms of major depressive disorder (MDD) and even preventing onset of MDD. However, uptake of these interventions is low in routine care. A new approach to overcome structural (e.g. waiting times, travel costs) and attitudinal barriers (e.g. self-reliance, stigma) might be web-based indirect prevention and treatment, i.e. focusing on less stigmatized problems like stress, procrastination, or sleep. Insomnia might be a prime target for depression prevention and treatment, given that insomnia predicts depression, is often comorbid with it and insomnia problems often remain after depression treatment. This individual-participant-data (IPD) meta-analysis aimed to evaluate the efficacy of an indirect approach to reduce depressive symptom severity and onset of potentially clinically relevant cases by targeting insomnia. The study focused on (adapted versions of) the web-based intervention *GET.ON Recovery* across different populations and intended to establish recommendations for whom this approach can be used best.

Seven eligible studies were identified and IPD was obtained from four of them (N = 563). The efficacy of depressive outcomes (based on Center for Epidemiological Studies Depression Scale, CES-D) at post-treatment (8 weeks) and follow-up (6 months) were evaluated using a one-stage IPD meta-analysis approach. Generalized Linear Mixed Models were estimated separately for depression symptom severity, onset of potentially clinically relevant cases, close-to-symptom-free-status, reliable change index and anchor-based clinically relevant change while controlling for baseline depressive symptom severity. Candidate moderators for the multivariable moderation analysis were

sociodemographic, clinical and work-related characteristics and were chosen based on 1) significant univariable moderation or 2) a positive variable importance in random-forest analysis. Multivariable decision-trees were built for post-treatment and follow-up using a recursive partitioning algorithm.

The IPD meta-analysis indicated significant group differences in favor of the intervention for all depressive outcomes except depression onset, which was evaluated in the small subsample (n = 121) without clinical levels of depressive symptom severity at baseline. Group differences in depressive symptom severity showed medium to large effects at post-treatment (d = -0.71 [95%–CI: -0.92, -0.51]) and follow-up (d = -0.84 [95%–CI: -1.11, -0.57]). The decision tree split the sample according to baseline severity into four groups at post-treatment with effects being largest in participants with baseline CES-D scores > 28 (d = -0.87 [95%–CI: -1.25, -0.49]) and into three groups at follow up with highest effects in the group with baseline CES-D scores of > 19 but $\leq$ 24 (d = -0.99 [95%–CI: -1.33, -0.65]). No other clinical, sociodemographic, or work-related characteristic besides baseline depression symptom severity moderated the effect.

Even though the study aimed to assess the potential of an indirect approach across different population, this was not possible due to constraints in which IPD could be obtained so that results are restricted to employees with insomnia and high work stress. However, for this group, results suggest that web-based interventions for insomnia can be offered to individuals with all levels of baseline severity across a wide range of sociodemographic, clinical and work-related characteristics. The indirect approach thus seems promising to reduce depressive symptom severity to a comparable degree to what could be expected from web-based interventiond for depression.

Manuscript in Preparation:
Thielecke J, Kuper P, Schuurmans L, Harrer M, Ebert D D, Lehr D, Behrend D, Brückner H, Horvath H, Riper H, Cuijpers P, and Buntrock C. Who benefits from indirect prevention and treatment of depression using online interventions for insomnia? Results from an individual-participant-data meta-analysis? Results from an individual-participant-data meta-analysis. In preparation

Contribution:
Janika Thielecke contributed substantially to the study conceptualization, design and methods and prepared and published the preregistration on Open Sience Framework. She organized the secondary data analysis and shared responsibility for data curation, formal analysis and visualization with the second author. She interpreted the results, drafted the original manuscript and revised it after feedback from her co-authors.

# Who benefits from indirect prevention and treatment of depression using online interventions for insomnia? Results from an individual-participant data meta-analysis

Janika Thielecke[1, 2], Paula Kuper[1,3], Lea Schuurmans[1], Mathias Harrer[1,4], David D. Ebert[1], Dirk Lehr[5], Pim Cuijpers[6,7], Dörte Behrendt[5], Hanna Brückner[5], Hanne Horvath[4], Heleen Riper[6,7,8], & Claudia Buntrock[3]

[1]Professorship of Psychology and Digital Mental Health Care, Department of Sports and Health Sciences, Technical University of Munich, Germany

[2]Department of Clinical Psychology and Psychotherapy, Institute of Psychology, Friedrich-Alexander-University Erlangen-Nürnberg, Germany

[3]Institute of Social Medicine and Health Systems Research, Faculty of Medicine, Otto von Guericke University Magdeburg, Germany

[4]GET.ON Institute for Online Health Trainings GmbH, Berlin, Germany

[5]Department of Health Psychology and Applied Biological Psychology, Institute for Sustainability, Education & Psychology, Leuphana University Luneburg, Germany

[6]Department of Clinical, Neuro and Developmental Psychology, VU University, Netherlands

[7]Amsterdam Public Health, Amsterdam University Medical Centers, Netherlands

[8]Department of Psychiatry, VU University Medical Center, Netherlands

**Words:** 4844 /4500

**Author Note**

Correspondence concerning this Article should be addressed to Janika Thielecke, Technical University Munich, Department of Sports and Health Sciences, Georg-Brauchle-Ring 60/62, 80992 München. Mail: janika.thielecke@tum.de

**Abstract**

**Background.** Major depressive disorder (MDD) is highly prevalent and burdensome for individuals and society. While there are psychological interventions able to prevent and treat MDD, uptake remains low. To overcome structural and attitudinal barriers, an indirect approach of using online insomnia interventions seems promising because insomnia is less stigmatized, predicts MDD onset, is often comorbid and can outlast MDD treatment. This individual-participant-data meta-analysis evaluated the potential of the online insomnia intervention *GET.ON Recovery* as an indirect treatment to reduce depressive symptom severity and potential MDD onset across a range of participant characteristics.

**Methods.** Efficacy on depressive symptom outcomes was evaluated using multilevel regression models controlling for baseline severity. To identify potential effect moderators, clinical, sociodemographic, and work-related variables were preselected using univariable moderation and random-forest methodology before developing a multivariable decision tree.

**Results.** IPD were obtained from four of seven eligible studies (N=561). Depressive symptom severity was significantly lower in the intervention group both at post-assessment (d=−0.71 [95%-CI: −0.92, −0.51]) and at follow-up (d=−0.84 [95%-CI: −1.11, −0.57]). In the subsample (n=121) without potential MDD at baseline, there was no differences in the onset of potential MDD between the groups. Moderation analyses revealed that effects on depressive symptom severity differed across baseline severity groups with effect sizes between d=−0.48 and −0.87 (post) and d= −0.66 to −0.99 (follow-up), while no other sociodemographic, clinical, or work-related characteristics were significant moderators.

**Conclusions.** Online insomnia interventions are a promising approach to reduce depressive symptom severity across demographic groups.

Words: 250/250

**Preregistration**: https://osf.io/xcus5

**Introduction**

Major depressive disorder (MDD) is a highly prevalent disorder (Gutiérrez-Rojas et al., 2020) associated with great individual (Ferrari et al., 2013) and societal burden (Vos et al., 2004). Psychological treatments such as cognitive behavioral therapy (CBT) are the first line treatments for depression (National institute for Health and Care Excellence, 2022) and have the potential to prevent MDD onset (Cuijpers, Pineda, et al., 2021). However, the uptake of psychological interventions remains low, even in high-income countries, where only 28% of individuals in need of treatment receiving it (Chisholm et al., 2016) and less than 1% use indicated preventive interventions (Cuijpers et al., 2010). New approaches to increase the uptake of mental health interventions are required to reduce the overall depression burden (Cuijpers, 2021).

Structural barriers to healthcare access can be addressed using the internet (Ebert et al., 2018), but especially in high-income countries, attitudinal barriers, including a common preference to solve one's own problems and the perceived stigma of mental illness, are important barriers for treatment uptake (Andrade et al., 2014; Clement et al., 2015). An indirect approach to depression prevention and treatment (Cuijpers, 2021) using (guided) online self-help interventions may be a promising alternative to overcome these barriers. Instead of focusing on depression, the idea of an indirect approach is to target problems contributing to depression or frequently comorbid, such as low self-esteem, procrastination (Cuijpers, Smit, et al., 2021), and stress (Harrer et al., 2021; Weisel et al., 2018), or that are less stigmatized, such as insomnia (van der Zweerde et al., 2019). By addressing these problems, depressive symptom severity (DSS) may be reduced but interventions might face more acceptance and match participants' perceived needs better. Insomnia is a particularly promising target as it is an impairing and burdensome disorder even in the absence of depressive symptoms (Morin et al., 2015; Roach et al., 2021; Wade, 2010) and can be effectively treated with specialized CBT for insomnia in person or via the internet (henceforth termed iCBT-I) (Feng et al., 2020; Simon et al., 2023; Ye et al., 2016). Insomnia is also a predictor of new and recurrent depressive episodes (Baglioni et al., 2011; Li et al., 2016), is comorbid in 38%– 83% of depression cases (Bjorvatn et al., 2019; Staner, 2010; R. Stewart et al., 2006), and often persists after depression treatment (Vargas & Perlis, 2020).

Evidence that iCBT-I can effectively reduce DSS in subthreshold cases (Batterham et al., 2017; Cheng, Luik, et al., 2019; Christensen et al., 2016; van der Zweerde et al., 2019) and in patients diagnosed with clinical depression (Blom et al., 2015, 2017) is promising. However, evidence that iCBT-I can prevent onset of new depressive episodes is divided. One study found a preventive effect of iCBT-I on self-reported depression onset after 12 months compared to an active control group (Cheng, Kalmbach, et

al., 2019), but another did not find a preventive effect at 6-month follow-up compared to an active control group using diagnostic interviews (Christensen et al., 2016). Only one study investigated treatment moderators and identified baseline depression severity but no sociodemographic variables as moderators (Cheng, Luik, et al., 2019).

More insight of the potential of this indirect approach can be gained from focusing on one specific intervention used in different populations to allow for a greater precision and guide recommendations for researchers and clinicians by revealing who might profit most from it. Therefore, we performed an individual-participant-data (IPD) meta-analysis that focusses on the web-based insomnia intervention *GET.ON Recovery*, which is based on classic CBT-I components (e.g. sleep hygiene and sleep restriction) and enhanced by behavioral activation and gratitude aspects. This program emphasizes detachment from work-related thoughts by including technique to counter worry and rumination (Thiart et al., 2013). It was originally developed and evaluated in teachers (Ebert et al., 2015; Thiart et al., 2013, 2015) but has since been adapted and evaluated in the general employee population (Behrendt et al., 2020, tba FOD-study, 2023), where it has been shown to reduce insomnia complaints (Behrendt et al., 2020; Ebert et al., 2015; Thiart et al., 2015). Further adaptations and (pilot) tests have been conducted among farmers (Braun et al., 2019), international students (Spanhel et al., 2021), and refugees (Spanhel et al., 2021) in Germany. All of these groups might profit from an indirect treatment approach since stigma of mental health problems are associated with fears specific to context, such as assumed workplace difficulties among employees (Brohan & Thornicroft, 2010), loss of community support among refuges (Satinsky et al., 2019; Shannon et al., 2015), and academic performance, finances, and career anxiety among college students (Cooper et al., 2020; Ebert et al., 2019).

The aim of the current analysis is to 1) evaluate the efficacy of GET.ON Recovery on DSS reduction in individuals with subclinical or clinical depressive symptoms across different populations and to 2) identify the possible moderating effects of various participant and intervention characteristics.

## Methods

This study was designed as an IPD meta-analysis to investigate the efficacy of *GET.ON Recovery* training or an adapted version thereof on depressive symptom outcomes. The study was preregistered on OSF (https://osf.io/xcus5) and follows the Preferred Reporting Items for Systematic Review and Meta-Analyses of IPD (PRISMA-IPD) statement (L. A. Stewart et al., 2015) where applicable. For details and rationale for all deviations from the registration, see Supplement 1.

### Identification and selection of studies

Randomized controlled trials investigating a version of *the GET.ON Recovery* training in comparison to any kind of control group among an adult population, which assessed DSS at post-treatment and/or follow-up were eligible for inclusion. Studies were identified through the scientific advisors at GET.ON institute (DDE) and by searching the German Clinical Trial Registry (DRKS) using the keyword "*GET.ON Recovery*" in November 2021. The authors of the eligible studies were contacted and invited to provide IPD.

### Risk-of-bias assessment

The revised version of the Cochrane risk-of-bias tool for randomized trials (RoB2, Sterne et al., 2019) and the related excel tool (Higgins et al., 2019) were used to assess the quality of included studies, focusing on the intention-to-treat data available for DSS at post-treatment and/or follow-up. The RoB2 assesses possible bias in five domains: "randomization process," "deviations from interventions," "missing data," "outcome measurement," and "selective reporting." Each domain is rated as either "low risk," "some concern," or "high risk." We followed the proposed algorithm to reach an overall judgment, which reflected at least the lowest assessment of an individual domain. Published papers and/or the clinical trial registrations were used for the assessments which were conducted independently by two researchers (PK & JT) who were not involved in the original studies. Disagreements were resolved by discussion.

### Depressive outcomes

All depressive symptom outcomes were based on the German version of the Center for Epidemiological Studies Depression Scale (CES-D, Hautzinger et al., 2012). This self-reporting scale consists of 20 items, each rated 0–3, yielding a total score from 0 to 60 with higher scores indicating more severe depressive symptoms. Psychometric properties of the CES-D are well established with a Cronbach's $\alpha$=0.89 (Hautzinger et al., 2012). As a primary objective, we focused on DSS at post-treatment and follow-up. Additionally, we examined the following secondary outcomes at post-treatment and follow-up assessments: (1) reliable improvement and deterioration according to the

**reliable change index (RCI)** by Jacobson and Truax (1991), (2) **anchor-based clinically relevant change** reflecting a 33% change in CES-D score (Bundesärztekammer et al., 2022), (3) **close-to-symptom-free status** defined as a CES-D score <16, and (4) **onset of potential MDD** based on participant self-report (CES-D ≥ 16).

**Potential moderators of the intervention effect**

Given the limited knowledge of potential moderators, we included a wide range of variables in multivariable analyses. For sociodemographic variables sex, age, relationship status, ethnicity, children, education, and employment, were sought from the original studies. For clinical characteristics, baseline DSS (CES-D), insomnia severity (Insomnia Severity Index, ISI, range: 0–28; Dieck et al., 2018), and previous experience with psychotherapy and/or health training were selected. After obtaining IPD, the following work-related variables were also included: the Effort-Reward Imbalance Scale – Short form (ERI-S; Siegrist et al., 2009) with subscales effort (range 3–15) and reward (range 7–35) used to calculate an effort-reward ratio (>1 indicating imbalance) and work engagement (Utrecht Work Engagement Scale, UWES; Schaufeli & Bakker, 2004) with subscales vigor, dedication, and absorption (score range for each: 0–6).

**Statistical analyses**

The obtained IPD were harmonized by trained personal according to established coding guidelines (https://protectr.netlify.app/coding-guide.html). For all analyses, the significance level was set to $p < 0.05$ (two-sided) and adjusted for multiple comparisons using the Bonferroni method (Emerson, 2020).

This study followed an intention-to-treat approach. Missing post-treatment and follow-up data were estimated separately using the *mice* package (van Buuren & Groothuis-Oudshoorn, 2011) for multiple imputation by chained equations in R (R Core Team, 2022) under a missing at random assumption. The multilevel structure of the data was reflected by defining random-effect models. Stratification by treatment was implemented using the *bygroup* function in *miceadds* (Robitzsch & Grund, 2022), which generates imputations separately for intervention and control conditions. Two-level predictive mean matching (2l.pmm) from the *miceadds* package was used to account for data clustering. Trial means of baseline DSS were used in the prediction of post-treatment and follow-up symptom severity outcomes (Supplement 2). A total of 50 imputed datasets were created for each time point. All analysis models were estimated in each corresponding imputed data set and parameters were combined using Rubin's rules (Little & Rubin, 2002; Rubin, 1987).

To evaluate depression outcomes, we used a one-step IPD approach to better account for the small number of participants/events in the included studies (Riley et al., 2020). Separate generalized linear

mixed models (GLMMs) were specified for each outcome. All models included a random intercept for trial and random slope for the treatment effect, and were adjusted for baseline DSS. Fixed model parameters are reported with 95% confidence intervals (CIs). GLMMs predicting continuous outcomes are reported with model-based Cohen's d values directly estimated by standardizing the outcome using the pooled standard deviation. Models for categorical outcomes were estimated as GLMMs from the combined data with a logit link to retrieve Odds Ratios (ORs). Numbers-Needed-to-Treat (NNTs) were calculated as the inverted absolute risk difference and absolute numbers are given for all dichotomous depressive outcomes. Since not all original studies reported treatment effects on DSS, effects in the individual studies were estimated by separate linear models of relevant imputed data subsets.

A multistep approach was used to investigate potential univariable and multivariable moderating effects. First, potential moderators were separately included in the "unconditional" GLMMs. Following the recommendations of Riley et al. (2020), we centered all covariates by their trial-specific mean and included the mean as a level-two predictor to avoid amalgamation of within- and across-trial information. Second, we investigated multivariable treatment-by-moderator interactions. Putative moderators were first ranked for their relevance by calculating the variable permutation importance using the random-forest method (Garge et al., 2013) in an aggregated dataset as Rubin's rules are not directly applicable for nonparametric approaches. With this method, DSS at post-treatment and follow-up times were regressed on the treatment indicator using 300 bootstrapped samples. All potential moderators were introduced as partitioning variables on their raw scale using the *mobforest* package (Garge et al., 2013). Variables were ranked by relative importance according to the frequency with which they served as a splitting variable in the trees (termed the "permutation accuracy method").

In the final model-based tree analysis to evaluate possible multivariable moderation, all variables with significant interaction with the treatment effect in univariable models and/or that yielded variable importance >0 in the random-forest model were included as partitioning variables. Model-based recursive partitioning allows incorporation of machine learning approaches, specifically recursive partitioning, into a parametric model. The result is an easy to interpret decision tree describing subgroups with differential treatment effects. The tree was operationalized using the R package *glmertree* (Zeileis et al., 2008). In the models for post-treatment and follow-up, DSS was regressed on the treatment indicator in the aggregated data. The nested structure of the patients within the studies was accounted for by specifying a random trial intercept and random treatment effect. Effects were estimated separately in the subgroups to receive model-based Cohen's d as described above. Since the

model-based tree analysis approach is prone to overfitting, we reported an optimism-adjusted $R^2$ (Harrell et al., 1996; Smith et al., 2014).

Between-trial heterogeneity was deemed highly plausible, but the small number of included studies led to improperly estimated heterogeneity variances of zero. Therefore, a "quasi-Bayesian" approach was applied using the functionality of the *blme* package (Chung et al., 2013) throughout imputation, one-stage IPD analyses, and decision-tree building. A weakly informative Wishart prior with *df*=4 and a scale matrix multiplied by 0.05 (adapted to 0.01 or 0.075 in case of convergence problems) was used. The prior helped to avoid boundary-fit issues while remaining largely uninformative itself.

**Sensitivity analysis**

Analysis was repeated in a complete case subsample and in the total sample while excluding the sleep item from the CES-D scores to evaluate the robustness of our results. We decided not to conduct moderation analysis using the complete case sample due to the reduced sample size and power, which would increase the chance for spurious effects in multiple testing.

## Results

**Study selection and IPD obtained**

A total of eight studies evaluating *GET.ON Recovery* (or a modified version) were identified, of which seven were deemed eligible and the authors of six were asked to contribute IPD. One study was deemed ineligible because it did not assess DSS (unpublished, trial registration: DRKS00017737), while the authors of one eligible study were not asked to contribute IPD because the online training was used only in a small subsample (15/150) of participants who could choose from a portfolio of online training programs (Braun et al., 2021a, 2021b). The IPD from two studies were unavailable due to administrative reasons (Spanhel et al., 2021, 2022). Ultimately, the IPD from four studies were included in the meta-analysis, and no integrity concerns were raised (for Flowchart see Supplement 3). Since the study by Ebert et al. (2015) did not assess follow-up data in the control group due to a shorter waiting-list time, this study was excluded from analysis at follow-up. In total, data from 561 participants in four trials were analyzed for DSS at post-treatment (8 weeks post-randomization) and 433 participants from three trials were analyzed at follow-up (24 weeks post-randomization).

**Study and participant characteristics**

All four included studies assessed the effects of *GET.ON Recovery* training on DSS among employees with high work-related rumination or without clear separation of work and private life and used a waiting-list control group design (Table 1). Most participants were female (68%, n=381/561), in a relationship (70%, n=395/561), and had achieved more than high school education (76%, n=429/561). Mean age was 47 years (SD=9.73). The majority had clinically relevant insomnia (ISI≥15, 76%, n=425/563) and 78% (n=440/563) reported clinically relevant levels of depression (CES-D≥16). Work engagement was considered average (subscale means 3.02–3.38), and the effort-reward ratio suggested an imbalance (M=1.41, SD=0.40). For more details, see Supplement 4.

**Table 1**

*Characteristics of the included studies*

| Study | Intervention | Target group and dysfunction | Assessments | | | TN | | Missing data (%) | | Risk-of-bias assessment | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | T0 | T1 | T2 | IG | CG | IG | CG | Rand | Int dev | Out mis | Out meas | Sel rep |
| Thiart et al. (2015) | IG: GET.ON Recovery, guided<br>CG: WLC (6 months) | teachers<br>work-related rumination (IS-CI≥15)<br>clinically relevant insomnia (ISI ≥15) | 0 | 8 | 24 | 64 | 64 | T1: 3.1<br>T2: 6.3 | T1: 12.5<br>T2: 15.6 | + | + | + | X | - |
| Ebert et al. (2015) | IG: GET.ON Recovery, unguided<br>CG: WLC (8 weeks) | teachers<br>work-related rumination (IS-CI≥15)<br>clinically relevant insomnia (ISI ≥15) | 0 | 8 | 24* | 64 | 64 | T1: 23.4<br>T2: 37.5 | T1: 17.2 | + | + | + | X | + |
| Behrendt et al. (2020) | IG: GET.ON Recovery, unguided<br>CG: WLC (6 months) | working adults<br>no clear distinction of work/private life (segmentation supplies <2.25) | 0 | 8 | 24 | 81 | 89 | T1: 40.7<br>T2: 51.9 | T1: 13.5<br>T2: 13.5 | + | + | + | X | + |
| tba (202X) | IG: GET.ON Recovery, feedback on demand<br>CG: WLC (6 months) | working adults<br>no clear distinction of work/private life (segmentation supplies <2.25)<br>clinically relevant insomnia (ISI≥15) | 0 | 8 | 24 | 64 | 64 | T1: 21.9<br>T2: 32.8 | T1: 17.2<br>T2: 26.6 | ? | ? | ? | X | ? |

*Abbreviations.* IG: intervention group, CG: control group, WLC: waiting-list control, IS-CI: Cognitive Irritation Scale (Mohr et al., 2005), ISI: Insomnia Severity Index (Dieck et al., 2018), T0: baseline assessment, T1: post-treatment assessment, T2: follow-up assessment, segmentation supplies: Subscale from the workplace segmentation preferences and supplies (Kreiner, 2006); tba: to be announced (original study in preperation, trial registration: DRKS00006223);*in IG only, Risk-of-bias assessment: "+" low–risk of bias, "–" some concern, "x" high risk, "?" not enough information to assess, Rand: randomization process, Int Dev: deviations from interventions, Out Mis: missing data, Out Meas: outcome measurement, Sel Rep: selective reporting.

**Risk-of-bias assessment**

All included studies were judged to have a high risk-of-bias, mainly due to unblinded participants reporting on the outcome by self-report.

**Effects on depressive symptom outcomes**

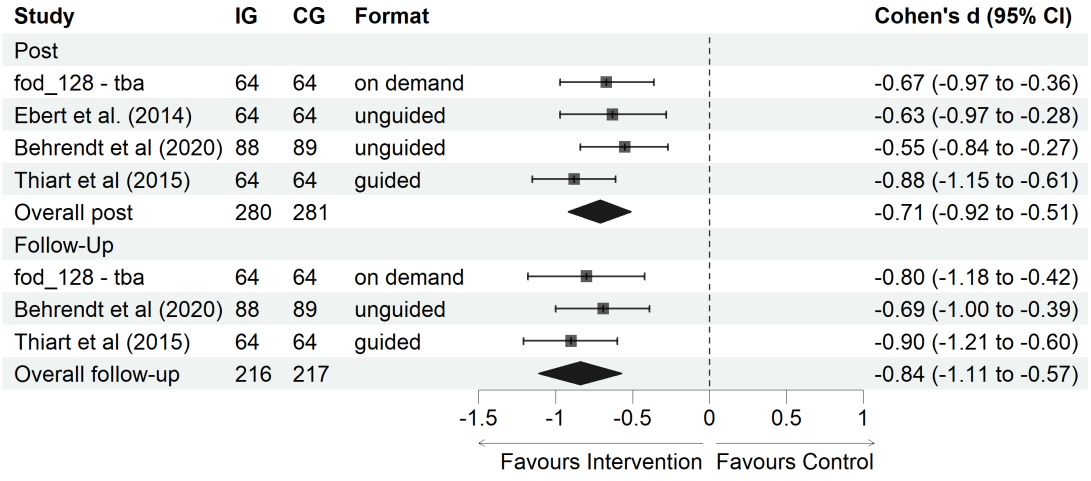
All depression-related outcomes are presented in Table 2. Sensitivity analyses confirmed the robustness of the results (Supplements 5 and 6).

***DSS***

Depression symptom severity was significantly reduced in the intervention group compared to the control group both at post-treatment (ß=−5.99 [95%-CI: −7.68, −4.29], T(419.9)=−6.93, $p_{adjusted} < 0.0001$) and at follow-up (ß=−7.28 [95%-CI: −9.61, 4.95], T(229.2)= −6.93, $p_{adjusted} < 0.0001$). The effects estimated in the individual studies and the overall pooled average effects on DSS at post-treatment (d=−0.71 [95%-CI: −0.92, −0.51]) and at follow-up (d=−0.84 [95%-CI: −1.11, −0.57]) are presented in Figure 1.

**Figure 1**

*Forest plot summarizing the estimated effects estimated in individuals studies (based on multiple imputation) and the average pooled effect from IPD analysis*




**Reliable Change Index**

A statistically significant greater proportion of participants in IG than in CG exhibited reliable symptom improvement at post-treatment (OR=0.19 [95%-CI: 0.12, 0.29], T(411)=−7.39, $p_{adjusted} < 0.001$;

NNT=3 [95%-CI: 3.8, 2.4]) and at follow-up (OR=0.11 [95%-CI: 0.05, 0.22], T(175.6)=−6.09, $p_{adjusted} < 0.001$; NNT=3.2 [95%-CI: 4.1, 2.6]). Additionally, fewer participants in IG than the control group demonstrated reliable deterioration at post-treatment (OR=2.83 [95%-CI: 1.15, 6.95], T(308.6)=2.27, $p_{adjusted=0}$.284; NNT=20.1 [95%-CI: 11.7, 71.6]) and at follow-up (OR=2.95 [95%-CI: 1.15, 7.57], T(194.3)=2.26, $p_{adjusted}$=0.297; NNT=16.6 [95%-CI: −10.2, −43.6]) but without statistical significance.

**Anchor-based clinically relevant change**

A 33% reduction in CES-D score was associated with an average point decrease of 7.30 (SD=2.64). A statistically significant greater proportion of participants in IG than in CG reported anchor-based clinically relevant improvement at post-treatment (OR=0.17 [95%-CI: 0.11, 0.27], T(396.1)=−7.54, $p_{adjusted} < 0.001$) and at follow-up (OR=0.13 [95%-CI: 0.07, 0.23], T(196.6)=−6.51, $p_{adjusted} < 0.001$). NNTs indicated that 2 to 3 more participants would need iCBT-I to achieve one more person with clinically relevant improvement in DSS ($NNT_{post}$=2.3 [95%-CI: 1.9, 2.7]; $NNT_{follow-up}$=2.6 [95%-CI: 2.2, 3.1]).

**Close-to-symptom-free status**

A statistically significant greater proportion of participants in IG than in CG attained close-to-symptom-free status postassessment (152/224, 67.9% vs. 41/216, 19.0%; OR=0.16 [95%-CI: 0.09, 0.28], T(306.1)=−6.72, $p_{adjusted} < 0.001$) and at follow-up (118/173, 68.2% vs. 21/163, 12.9%; OR=0.13 [95%-CI: 0.06, 0.25], T(158.6)=−5.86, $p_{adjusted} < 0.001$). According to NNT analysis, two more participants would need to receive the intervention to produce one additional participant with close-to-symptom-free status ($NNT_{post}$=2.0 [95%-CI: 1.8, 2.4]; $NNT_{follow-up}$=1.8 [95%-CI: 1.6, 2.1]).

***Potential onset of depression***

Within the subsample without clinically relevant depressive symptoms at baseline (n=121), a lower proportion of participants in IG (7/56, 12.5%) than in CG (20/65, 30.8%) exhibited potential onset of MDD at post-treatment, but the difference did not reach statistical significance (OR=2.13 [95%-CI: 0.72, 6.32], T(80.0)=1.38, p=0.17, $p_{adjusted}$=1.00). Similarly, a smaller proportion of participants in IG than in CG exhibited potential MDD onset at follow-up (n=8/43, 18.6% vs. n=22/54, 40.47%), but the difference did not reach statistical significance (OR=2.00 [95%-CI: 0.68, 5.85], T(67.9)=1.28, p=0.20, $p_{adjusted}$=1.00). NNT to achieve one more case without possible MDD onset was 5.5 [95%-CI: 3.1, 24.4] at post-treatment and 4.5 [95%-CI: 2.5,21.7] at follow-up.

**Table 2**

*Overview of depression outcomes at post-treatment (8 weeks post-randomization) and at follow-up (24 weeks post-randomization) based on multiple imputation*

| Variable | Group | $n_{T0}$ | M | SD | ß [95%-CI] | $p_{adjusted}$ | d [95%-CI] | $\tau_{int}$ | $\tau_{group}$ | $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Depressive symptom severity (CES-D) | | | | | | | | | | |
| *Post-treatment* | IG | *280* | 13.86 | 7.12 | −5.99 [−7.68, −4.29] | <0.0001 | −0.71 [−0.51, −0.92] | 0.69 | 1.13 | 0.35 |
| | CG | *281* | 19.7 | 8.19 | | | | | | |
| *Follow-up* | IG | *216* | 14.23 | 6.73 | −7.28 [−9.61, −4.95] | <0.0001 | −0.84 [−0.57, −1.11] | 2.12 | 1.49 | 0.31 |
| | CG | *217* | 21.13 | 8.47 | | | | | | |
| **Variable** | **Group** | $n_{T0}$ | $n_{event}$ | % | **OR [95%-CI]** | $p_{adjusted}$ | **NNT [95%-CI]** | $\tau_{int}$ | $\tau_{group}$ | $R^2$ |
| RCI improvement | | | | | | | | | | |
| *Post-treatment* | IG | *280* | 159 | 56.79 | 0.19 [0.12, 0.29] | <0.0001 | 3.00 [3.8, 2.4] | 0.01 | 0.01 | 0.22 |
| | CG | *281* | 65 | 23.13 | | | | | | |
| *Follow-up* | IG | *216* | 121 | 56.02 | 0.11 [0.05, 0.22] | <0.0001 | 3.20 [4.1, 2.6] | 0.05 | 0.04 | 0.32 |
| | CG | *217* | 34 | 15.67 | | | | | | |
| RCI deterioration | | | | | | | | | | |
| *Post-treatment* | IG | *280* | 7 | 2.5 | 2.83 [1.15, 6.95] | 0.28 | −20.10 [−11.7, −71.6] | 0.04 | 0.05 | 0.03 |
| | CG | *281* | 21 | 7.47 | | | | | | |
| *Follow-up* | IG | *216* | 7 | 3.24 | 2.95 [1.15, 7.57] | 0.30 | −16.60 [−10.2, −43.6] | 0.06 | 0.04 | 0.05 |
| | CG | *217* | 24 | 11.06 | | | | | | |
| Anchor-based clinically relevant change | | | | | | | | | | |
| *Post-treatment* | IG | *280* | 180 | 64.29 | 0.17 [0.11, 0.27] | <0.0001 | 2.30 [2.7, 1.9] | 0.03 | 0.04 | 0.18 |
| | CG | *281* | 56 | 19.93 | | | | | | |
| *Follow-up* | IG | *216* | 139 | 64.35 | 0.13 [0.07, 0.23] | <0.0001 | 2.60 [3.1, 2.2] | 0.03 | 0.05 | 0.23 |
| | CG | *217* | 30 | 13.82 | | | | | | |

**Table 2 (continued)**

*Overview of depression outcomes post-treatment (8 weeks post-randomization) and at follow-up (24 weeks post-randomization) based on multiple imputation*

| Variable | Group | $n_{T0}$ | $n_{event}$ | % | OR [95%-CI] | $p_{adjusted}$ | NNT [95%-CI] | $\tau_{int}$ | $\tau_{group}$ | $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| close-to-symptom-free status[a] | | | | | | | | | | |
| *post-treatment* | IG | *224* | 152 | 67.86 | 0.17 [0.1, 0.28] | <0.0001 | 2.00 [2.4, 1.8] | 0.02 | 0.04 | 0.23 |
| | CG | *216* | 41 | 18.98 | | | | | | |
| *follow-up* | IG | *173* | 118 | 68.21 | 0.13 [0.06, 0.25] | <0.0001 | 1.80 [2.1, 1.6] | 0.03 | 0.05 | 0.23 |
| | CG | 163 | 21 | 12.88 | | | | | | |
| clinically relevant depression (MDD onset)[b] | | | | | | | | | | |
| *post-treatment* | IG | *56* | 7 | 12.5 | 2.13 [0.72, 6.32] | 1.00 | −5.50 [−3.1, −24.4] | 0.03 | 0.04 | 0.06 |
| | CG | *65* | 20 | 30.77 | | | | | | |
| *follow-up* | IG | *43* | 8 | 18.6 | 2.00 [0.68, 5.85] | 1.00 | −4.50 [−2.5, −21.7] | 0.03 | 0.04 | 0.07 |
| | CG | *54* | 22 | 40.74 | | | | | | |

*Note.* Analysis based on multiple imputation. [a]Subgroup exceeding the cut-off for clinically relevant depressive symptoms at baseline (CES-D≥16); [b]Subgroup considered close-to-symptom-free at baseline (CES-D < 16). *Abbreviations.* CES-D: Center for Epidemiological Studies Depression Scale, MDD: Major depressive disorder, RCI: Reliable Change Index, IG: intervention group, CG: control group, NNT: Number-Needed-to-treat, $n_{T0}$: case number at baseline assessment, $n_{T0}$: case number with outcome post-treatment or at follow-up, $\tau_{int}$: intercept variance, $\tau_{group}$: slope variance for the treatment effect.

**Moderation of the treatment effect**

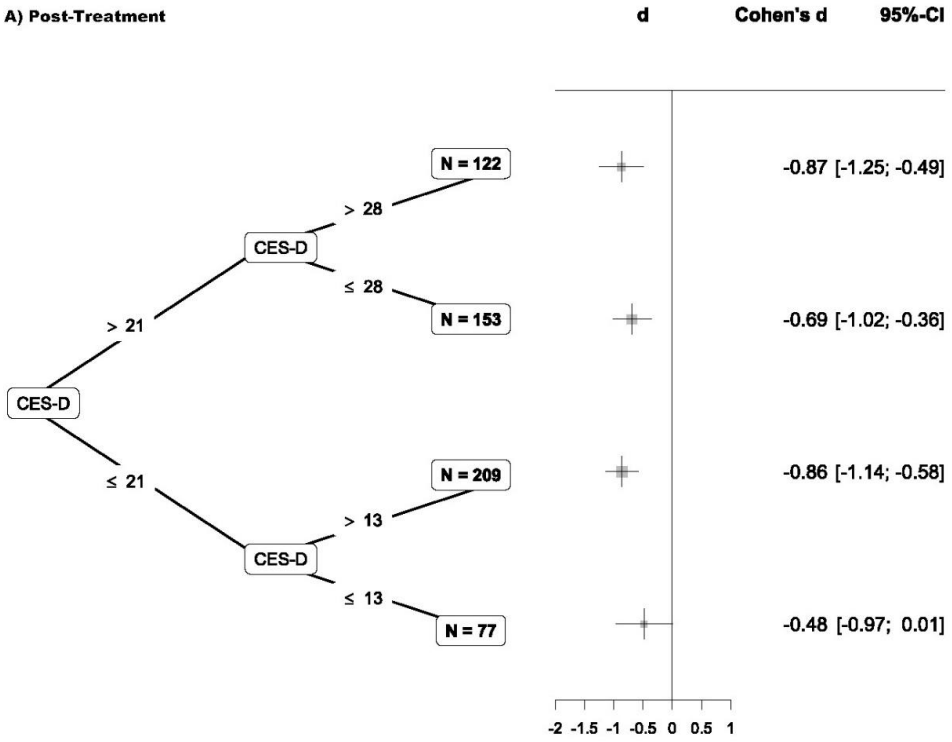
Based on the available IPD, ethnicity and employment were excluded as potential moderators due to a lack of variance, while effort-reward imbalance and work engagement were included in addition to sociodemographic and clinical variables. Univariable moderation analysis (Table 3) identified baseline symptom severity as the only significant moderator of follow-up symptom severity (ß =−0.30 [−0.56, −0.03], p=0.02). Additionally, based on random-forest analysis (Supplement 7), the following variables were included as partitioning variables in the final tree-models at both post-treatment and follow-up: baseline symptoms of depression and insomnia, previous psychotherapy, vigor, dedication, and reward. Relationship status and effort were included in the post-treatment model and age, absorption, and effort-reward ratio were included in the model for follow-up.

In the final tree-based models, only baseline DSS predicted heterogeneous treatment responses. For post-treatment, the first split divided the sample at 21 points with a second spilt occurring at 13 and 28 points in the two branches, respectively (Figure 2A). Optimism-corrected $R^2$ was reduced by 0.11 to $R^2_{adjusted}$=0.30. Statistically significant treatment effects were observed in three of the four terminal nodes with differences in the effect magnitude between subgroups based on partitioning the dataset by baseline CES-D scores of ≤13, >13 but ≤21, >21 but ≤28, and >28. Effects were highest for participants with a baseline score >28 (d=−0.87 [95%-CI: −1.25, −0.48], n=122) and smallest without statistical significance in the small group of participants with baseline scores ≤13 (d=−0.48 [95%-CI: −0.97, 0.01], n=77)

Similarly, in the follow-up model, only baseline DSS explained the heterogeneity in treatment effect (Figure 2B). Two splits were identified, the first split at 24 points on the CES-D and the second at 19 points in the subgroup with baseline CES-D scores ≤24. Optimism-corrected $R^2$ was reduced by 0.11 to $R^2_{adjusted}$=0.27. Treatment effects were significant in all terminal node models, with the biggest effect size in participants with baseline CES-D scores >19 but ≤24 (d=−0.99 [95%-CI: −1.33, −0.64]) and lowest in participants with baseline scores >24 (d=−0.66 [95%-CI: −0.95, −0.36]).

**Figure 2**

*Tree model for depressive symptoms (CES-D) post-treatment (A) and at follow-up (B) derived from model-based recursive partitioning in aggregated data*



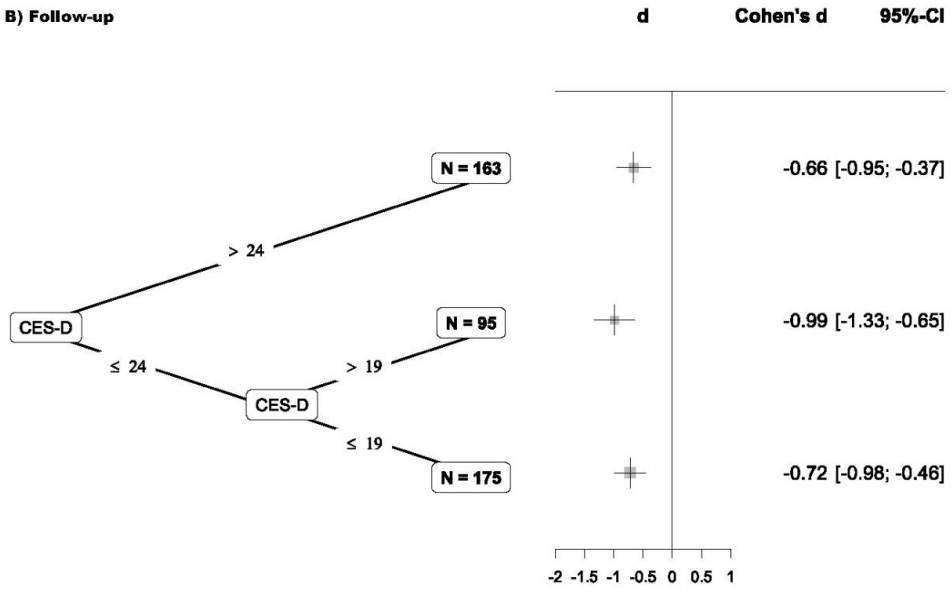



**Table 3**

*Univariable moderation analyses of sociodemographic, clinical, and work-related variables post-treatment (8 weeks post-randomization) and at follow-up (24 weeks post-randomization)*

| Variable | Interaction: trial mean-centered baseline characteristic × treatment group | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Post-treatment (N=563) | | | | | | | | Follow-up (N=433) | | | | | | |
| | Estimate | [95%-CI] | SE | T | df | p | $\tau^2_{int}$ | $\tau^2_{group}$ | Estimate | [95%-CI] | SE | T | df | p | $\tau^2_{int}$ | $\tau^2_{group}$ |
| **Sociodemographic variables** | | | | | | | | | | | | | | | | |
| Sex | 2.20 | [−0.73, 5.12] | 1.49 | 1.48 | 1071.1 | 0.140 | 1.37 | 1.1 | 2.09 | [−1.61, 5.79] | 1.88 | 1.11 | 1.35 | 0.266 | 1.2 | 1.35 |
| Age | 0.01 | [−0.17, 0.19] | 0.09 | 0.12 | 188.0 | 0.906 | 1.55 | 1.12 | 0.13 | [−0.04, 0.30] | 0.09 | 1.49 | 1.32 | 0.137 | 1.84 | 1.32 |
| Relationship status[a] | | | | | | | | | | | | | | | | |
| *Single* | −2.04 | [−5.70, 1.61] | 1.85 | −1.10 | 368.4 | 0.271 | 1.63 | 1.1 | −3.26 | [−7.30, 0.79] | 2.05 | −1.59 | 1.33 | 0.113 | 3.08 | 1.33 |
| *Married/in a relationship* | 1.39 | [−1.91, 4.69] | 1.67 | 0.83 | 392.3 | 0.407 | 1.54 | 1.04 | 1.75 | [−1.87, 5.38] | 1.84 | 0.95 | 1.34 | 0.340 | 3.56 | 1.34 |
| *Divorced/seperated* | 0.74 | [−4.61, 6.10] | 2.72 | 0.27 | 739.7 | 0.785 | 1.54 | 1.07 | 2.12 | [−3.57, 7.80] | 2.89 | 0.73 | 1.34 | 0.463 | 1.35 | 1.34 |
| *Widowed* | 5.02 | [−22.50, 32.54] | 13.91 | 0.36 | 235.1 | 0.718 | 1.17 | 1.13 | 3.29 | [−26.86, 33.44] | 15.22 | 0.22 | 1.48 | 0.829 | 3.83 | 1.48 |
| Education[a] | | | | | | | | | | | | | | | | |
| *Up to high school only* | 0.01 | [−7.46, 7.48] | 3.80 | 0.00 | 929.5 | 0.999 | 0.91 | 1.11 | 2.89 | [−5.21, 11.00] | 4.11 | 0.70 | 1.53 | 0.482 | 1.54 | 1.53 |
| *High school* | 2.64 | [−1.60, 6.89] | 2.16 | 1.22 | 978.8 | 0.221 | 1.57 | 1.1 | 0.18 | [−4.57, 4.94] | 2.41 | 0.08 | 1.31 | 0.939 | 2.02 | 1.31 |
| *After high school* | −0.60 | [−4.62, 3.41] | 2.04 | −0.30 | 740.3 | 0.767 | 1.38 | 1.1 | −0.94 | [−5.26, 3.38] | 2.19 | −0.43 | 1.35 | 0.668 | 1.18 | 1.35 |
| *Postgraduate* | −3.79 | [−9.63, 2.05] | 2.97 | −1.27 | 2461.1 | 0.203 | 1.35 | 1.1 | 0.46 | [−6.08, 7.00] | 3.32 | 0.14 | 1.38 | 0.889 | 1.13 | 1.38 |
| **Clinical variables** | | | | | | | | | | | | | | | | |
| Baseline depression (CES-D) | −0.15 | [−0.36, 0.06] | 0.10 | −1.44 | 229.7 | 0.151 | 0.67 | 1.13 | −0.30 | [−0.56, −0.03] | 0.13 | −2.25 | 1.55 | **0.026** | 2.40 | 1.55 |
| Baseline insomnia (ISI) | −0.40 | [−0.81, 0.00] | 0.21 | −1.95 | 637.1 | 0.052 | 1.45 | 1.09 | −0.43 | [−0.96, 0.09] | 0.26 | −1.65 | 1.34 | 0.100 | 1.32 | 1.34 |
| Previous psychotherapy | −1.54 | [−4.45, 1.37] | 1.48 | −1.04 | 702.9 | 0.297 | 0.75 | 1.07 | 0.14 | [−3.56, 3.84] | 1.87 | 0.08 | 1.51 | 0.940 | 2.62 | 1.51 |
| Previous health training | −0.94 | [−5.01, 3.14] | 2.07 | −0.45 | 827.8 | 0.651 | 1.17 | 1.13 | −2.10 | [−6.94, 2.74] | 2.46 | −0.85 | 1.4 | 0.393 | 3.89 | 1.4 |
| **Work-related variables** | | | | | | | | | | | | | | | | |
| Effort-reward-ratio (ERI-S) | | | | | | | | | | | | | | | | |
| *Effort* | −0.38 | [−1.14, 0.38] | 0.39 | −0.99 | 1679.8 | 0.324 | 0.70 | 1.18 | −0.11 | [−1.06, 0.84] | 0.48 | −0.22 | 1.52 | 0.822 | 3.67 | 1.52 |
| *Reward* | 0.31 | [−0.17, 0.79] | 0.25 | 1.27 | 365.6 | 0.207 | 1.25 | 1.03 | −0.07 | [−0.63, 0.49] | 0.28 | −0.25 | 1.55 | 0.806 | 2.40 | 1.55 |
| *Effort-reward ratio* | −3.37 | [−7.13, 0.38] | 1.90 | −1.77 | 449.1 | 0.077 | 1.00 | 1.09 | −0.33 | [−4.92, 4.26] | 2.32 | −0.14 | 1.55 | 0.887 | 2.81 | 1.55 |
| work engagement (UWES) | | | | | | | | | | | | | | | | |
| *Vigor* | 0.21 | [−0.87, 1.29] | 0.55 | 0.38 | 1115.2 | 0.702 | 1.54 | 1.02 | −0.01 | [−1.42, 1.39] | 0.71 | −0.02 | 1.31 | 0.987 | 2.14 | 1.31 |
| *Absorption* | 0.27 | [−0.64, 1.18] | 0.46 | 0.58 | 1199.6 | 0.559 | 1.60 | 1.06 | −0.64 | [−1.85, 0.58] | 0.61 | −1.03 | 1.31 | 0.302 | 3.27 | 1.31 |
| *Dedication* | 0.59 | [−0.40, 1.59] | 0.51 | 1.18 | 1219.1 | 0.240 | 1.52 | 1.02 | 0.41 | [−0.89, 1.71] | 0.66 | 0.62 | 1.25 | 0.533 | 2.40 | 1.25 |

*Note.* Significant interactions are printed bold. [a] Variable levels were tested individually against all other levels. *Abbreviations.* CES-D: Center for Epidemiological Studies Depression Scale, ISI: Insomnia Severity Index, ERI-S: Effort-Reward Imbalance Scale – Short form, UWES: Utrecht Work Engagement Scale, $\tau^2_{int}$: intercept variance, $\tau^2_{group}$: slope variance of the treatment effect.

**Discussion**

Using data from N= 561 individuals from four studies, we found an average pooled intervention effect of iCBT-I on DSS of d=-0.71 [95%-CI: -0.92, -0.51] at post-treatment and d=-.84 [95%-CI: -1.11, -0.57] at follow-up. The superiority of the intervention compared to waiting-list controls (WLC) group also reflected significant differences in the numbers of participants showing reliable improvement, anchor-based clinically relevant change, and close-to-symptom-free status. In contrast, iCBT-I demonstrated no significant effect on possible depression onset among the subsample without clinical levels of depression at baseline. The model-based decision tree revealed four (three) groups defined by their baseline DSS with differential treatment effects at post-treatment (follow-up) reaching from d=-0.48 to d=-0.87 (d=-0.66 to d=-0.99) with no other sociodemographic, clinical or work-related characteristic moderating effects.

The effect size for DSS found at post-treatment is comparable to what has been reported previously for unguided iCBT-I in mostly subthreshold depression cases compared to an active control group (Cohen's d between 0.60 and 0.64) (Batterham et al., 2017; Cheng, Luik, et al., 2019; Christensen et al., 2016). Few studies have reported longer follow-up times but those that did reported smaller effects at 6 months (d=0.40, Batterham et al., 2017; d=0.48, Christensen et al., 2016), while we report a slightly larger pooled effect (d=0.81) at 24 weeks post-randomization. However, all studies in the analysis included more severely depressed individuals and compared the effect to a WLC. Further, 2 of 4 studies reporting post-treatment effects and two of the three reporting effects at follow-up were guided, which are all factors increasing that may have increased the effect sizes (Furukawa et al., 2014; Werntz et al., 2023). Guidance might be an important factor to consider, given that effects at post-treatment reported here were smaller compared to a study using only guided iCBT-I and an active control for individuals with at least mild depressive symptoms (d=1.05 based on the Patient Health Questionnaire [PHQ] with sleep item) or d=0.76 (PHQ without sleep item) (van der Zweerde et al., 2019). Removing the sleep items from the CES-D in this study reduced the group differences at post-treatment (d=0.65 [−0.44, −0.86]), highlighting the differential sensitivities of various depression measurements to sleep changes. The observed effects were also comparable to those expected for online CBT of mild-to-moderate depression both at post-treatment (d=0.74 [95%-CI: 0.62, 0.86]) and at 3–6 months follow-up (d=0.83 [95%-CI: 0.69, 0.99]) according to a previous meta-analysis, and were also in accord with a small-scale study directly comparing iCBT-I to online CBT for depression in participants with clinical insomnia. While iCBT-I was superior for reducing sleep problems, both iCBT-I and depression CBT were equally effective for reducing

depressive symptoms up to 3 years postintervention (Blom et al., 2015, 2017). Thus, iCBT-I may be a better treatment option for individuals with both depression and insomnia.

We did not find a significant group difference in potential depression onset, likely due to the small sample of participants with no depressive symptoms or only mild symptoms at baseline and the relatively short observation period of 6 months. Similarly, Christensen et al. (2016) found no differences in depression onset at 6 months after iCBT-I using diagnostic interviews for both inclusion and evaluation. In contrast, Cheng et al. (2019) reported that the risk for depression onset was halved in the intervention group (relative risk ratio=0.51 [95%-CI:.26, 0.81]) at one year postintervention based on cut-off scores for self-reported measurements.

While there were no differences in depression onset between intervention and control groups, moderation analysis indicated that iCBT-I had positive effects on DSS in both subthreshold and clinically depressed individuals. The first split at both postintervention and follow-up time points occurred above the international CES-D cut-off of ≥ 16 (Radloff, 1977; Vilagut et al., 2016) as well as above the recommended cut-off in German depression guidelines (>22) for clinically significant cases and the >27 cut-off for "highly likely MDD cases" at post-treatment (Bundesärztekammer et al., 2022). The largest effect sizes were observed in individuals scoring around the cut-off for clinically relevant depression. The results are thus in line with Cheng et al. (2019), who also identified baseline DSS as the sole moderator, with participants in the upper tertile of a mild-moderate depressed sample showing the greatest improvements. Looking more closely on individualized symptoms instead of general symptom severity at baseline may be a promising approach to enhance treatment efficacy, especially given that multiple studies have identified problems initiating sleep as a predictor of later depression onset (Bjorøy et al., 2020; Blanken et al., 2020; Leerssen et al., 2021).

**Implications for research and practice**

Multivariable moderation analyses suggested that severely depressed individuals will also benefit from iCBT-I and do not need to be excluded from it, as was the case in most prior studies (Cheng, Kalmbach, et al., 2019; Christensen et al., 2016; van der Zweerde et al., 2019). In fact, moderate-to-severely depressed individuals with comorbid sleep problems may benefit more from iCBT-I than from an exclusively depression-focused intervention because of the simultaneous improvements in their sleep problems. While the studies included in this IPD meta-analysis did not include suicidal participants but with an appropriate safety protocol, other studies suggest that this indirect approach (targeting insomnia) may also be an opportunity for suicide prevention (Christensen et al., 2016; Kalmbach et al., 2022; Torok et al., 2020).

The easy to understand presentation of the results in a decision tree may facilitate decision making in a clinical context, when individuals present with symptoms of both disorders. This could be important since even though the idea that CBT-I improves depression has been around for decades, no treatment recommendation were established as consequence (Asarnow & Manber, 2019) and over 40% of US physicians surveyed stated that they still focus primarily on depression or anxiety if presented with comorbid insomnia (Morin et al., 2023). Kraepelien et al. (2022) also report no changes in treatment practice even after iCBT-I was demonstrated to be more effective in simultaneously reducing both depression and insomnia symptoms (Blom et al., 2015) and implemented in the same clinic's routine care, no changes in treatment practice occurred. Given that the effectiveness of iCBT-I on DSS did not vary across demographics in this IPD, this approach should be more highlighted in practice and included in the associated guidelines.

The study by Kraepelien et al. (2022) also brought up the topic of patient preference for treatment, because almost all participants receiving depression treatment while having elevated insomnia symptoms actively sought *depression* treatment instead of treatment for insomnia, probably seeing depression as their main problem. However, post-hoc interviews indicated that patients in the iCBT-I group of the initial trial (Blom et al., 2016) were more positive about the treatment than in the depression group. Therefore, patients' attitude towards an indirect approach and their naïve perception of how their symptoms relate to each other should be considered in future research.

Studies directly comparing depression to insomnia treatment (as in Blom et al., 2015) in a preventive setting with subthreshold insomnia and depression are still warranted to inform further personalization of preventive offers. Focusing on insomnia may encourage greater and more timely treatment participation by individuals, when insomnia still presents a predictor for depression but comorbidity is not yet fully developed. Together with an easy accessible, anonymous and self-help oriented intervention, this indirect online approach might overcome more attitudinal barriers than traditional interventions (Cuijpers, 2021; Cuijpers et al., 2010). However, it is still important to test the differences in uptake of insomnia and depression treatment in a more naturalistic setting and in a sample less confounded by high work stress, which is associated with depression risk independent of sleep pathology (Siegrist, 2008). However, not much is known about the long-term consequences of addressing mental health complaints by targeting more socially acceptable problems. With regard to maininingt treatment effects and relapse prevention it would be positive if treatment experience encourage future help-seeking intentions if needed, but it is unclear to what extent an indirect approach can support this. Studies thus far have mainly assessed help-seeking intention as a means to estimate further need after

the intervention (Blom et al., 2015; Christensen et al., 2016). Moreover. promotion of help-seeking and de-stigmatization of mental health problems should not stop at an individual level but should be seen as a societal effort (Clement et al., 2015).

Finally, Asarnow and Manber (2019) found inconclusive evidence for the greater efficacy of combined over sequential insomnia and depression treatment and also suggested that comorbidity may influence adherence and dropout. Internet interventions could serve as an ideal testing ground to address these questions (Domhardt et al., 2021). Module-based online trainings could be used to explore if the order of components, for example behavioral activations and sleep restriction, interact with each other and the treatment outcomes over time or if individual modules targeting potential mediators like rumination (Behrendt et al., 2020; Cheng et al., 2020) are especially crucial for combined treatment. Internet interventions combining aspects of insomnia and depression treatment as a predefined module, on demand additional modules chosen by the user, or recommendations by the program/guiding coach could be used to adapt the intervention to the individual's needs and preferences.

**Limitations**

The current results should be interpreted in light of several limitations. First, depending on the obtained IPD, our analysis included studies from a very homogenous group of with similar interventions and homogenous study groups. For instance, all participants were employed and were predominantly female and highly educated. The intervention itself focused on work detachment in addition to the classic CBT-I components such as sleep hygiene and sleep restriction. While this focus on the work context supports the ecological validity and highlights the potential of an indirect approach for occupational health, it limits transportability to other contexts. Second, we based our analysis on the German version of the CES-D, which does not have one uniform cut-off for depression onset or different categorical levels of DSS, limiting comparability with other studies in the field. We retained the pre-specified cut-off of ≥ 16 but included other cut-offs suggested from more recent depression guidelines, but future research should also consider clinical assessments. Third, the reported effect sizes may be higher than would be expected in routine care given the WLC and high risk-of-bias in outcome-measures. The reported OR were controlled for baseline CES-D to report conditional effects. However, the reported effects may be higher than at the population-level due to the non-collapsibility of ORs (Groenwold et al., 2011). Fourth, our results on subthreshold depression and potential depression onset should be interpreted with special caution because we (a) used a self-reported cut-off to identify potential MDD cases and (b) had a very reduced sample and power to examine the effects on onset in a population with

insomnia/depressive symptom comorbidity. Finally, the univariable and multivariable moderation analyses were the first in the field of indirect prevention and treatment of depression, and due to the relatively small sample size, should be considered exploratory and in need of validation across different samples. The random-forest method is only feasible in an aggregated dataset and when not accounting for the multilevel structure; thus the analysis did not consider imputation insecurity or heterogeneity among included studies. Similarly, the model-based recursive partitioning was also run in an aggregated dataset but could consider the multilevel structure. Further, for ease of interpretation, we did not center variables in the tree, which could introduce ecological bias. Finally, due to the relatively small sample size, the tree analysis was prone to overfitting, so we adjusted the $R^2$.

**Conclusion**

The findings of the current study provide evidence that iCBT-I can effectively reduce DSS across a wide range of demographic and work-related variables in working adults experiencing sleep problems and high work stress. Multivariable moderation analyses suggested that the effect size magnitude of iCBT-I varied according to baseline DSS but is promising intervention approach for both prevention and treatment of comorbid insomniac and depressive symptoms.

**Final disclosure**

**Acknowledgements**

Authors would like to thank enago Academy for proofreading the manuscript and TUM Graduate School for covering the costs for this service as part of JT's PhD program.

**Funding Statement**

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Ethical standards**

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All included studies individually underwent approval by the ethics committee of the responsible institutions as described in the original articles.

**Competing Interests**

MH is a part-time employee at Institute for Health Trainings Online (GET.ON), which aims to implement scientific findings related to digital health interventions into routine care.

DDE reports to have received consultancy fees or served in the scientific advisory board from several companies such as Novartis, Sanofi, Lantern, Schön Kliniken, Minddistrict, and German health insurance companies (BARMER, Techniker Krankenkasse). DDE is a stakeholder of GET.ON. HH is Founder and Chief Commercial Officer at GET.ON. JT, PK, LS, DL, HB, HR,PC, CB report no competing interests.

**Authorship and contributorship**

Janika Thielecke: Conzeptionalization, Methodology, Formal Analysis, Data Curation, Visualization, Project administration, Writing – Original Draft, Writing - Review & Editing. Paula Kuper: Conzeptionalization, Methodology, Formal Analysis, Data Curation, Visualization, Writing - Review & Editing. Lea Schuurmans: Conzeptionalization, Methodology, Formal Analysis, Data Curation, Writing - Review & Editing. Mathias Harrer: Conzeptionalization, Methodology, Data Curation, Formal Analysis, Writing - Review & Editing. David D. Ebert: Conzeptionalization, Resources, Supervision, Writing - Review & Editing. Dirk Lehr: Conzeptionalization, Resources, Writing - Review & Editing. Dörte Behrendt: Resources, Writing - Review & Editing. Hanna Brückner: Resources, Writing - Review & Editing. Hanne

Horvath: Resources, Writing - Review & Editing. Heleen Riper: Resources, Writing - Review & Editing. Pim Cuijpers: Resources, Writing - Review & Editing. Claudia Buntrock: Conzeptionalization, Methodology, Project administration, Supervision, Writing - Review & Editing.

**Data and/or Code availability**

The data are not publicly available. The corresponding author is a data processor, not a data owner and thus cannot provide data access upon request. Access to the data must be sought from authors of original studies and might depend on to be specified data security and data exchange regulation agreements. The analyses scripts can be assessed via OSF: XXXXXXXXXXX

**References**

Asarnow, L. D., & Manber, R. (2019). Cognitive Behavioral Therapy for Insomnia in Depression. *Sleep Medicine Clinics*, *14*(2), 177–184. https://doi.org/10.1016/j.jsmc.2019.01.009

Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., Lombardo, C., & Riemann, D. (2011). Insomnia as a predictor of depression: A meta-analytic evaluation of longitudinal epidemiological studies. *Journal of Affective Disorders*, *135*(1–3), 10–19. https://doi.org/10.1016/j.jad.2011.01.011

Batterham, P. J., Christensen, H., Mackinnon, A. J., Gosling, J. A., Thorndike, F. P., Ritterband, L. M., Glozier, N., & Griffiths, K. M. (2017). Trajectories of change and long-term outcomes in a randomised controlled trial of internet-based insomnia treatment to prevent depression. *BJPsych Open*, *3*(5), 228–235. https://doi.org/10.1192/bjpo.bp.117.005231

Behrendt, D., Ebert, D. D., Spiegelhalder, K., & Lehr, D. (2020). Efficacy of a Self-Help Web-Based Recovery Training in Improving Sleep in Workers: Randomized Controlled Trial in the General Working Population. *Journal of Medical Internet Research*, *22*(1), e13346. https://doi.org/10.2196/13346

Bjorøy, I., Jørgensen, V. A., Pallesen, S., & Bjorvatn, B. (2020). The Prevalence of Insomnia Subtypes in Relation to Demographic Characteristics, Anxiety, Depression, Alcohol Consumption and Use of Hypnotics. *Frontiers in Psychology*, *11*, 527. https://doi.org/10.3389/fpsyg.2020.00527

Bjorvatn, B., Olufsen, I. S., & Sørensen, M. E. (2019). The ICSD-3/DSM-5 diagnostic criteria for insomnia reinforce the association between insomnia, anxiety and depression. *Sleep Medicine*, *64*, S41. https://doi.org/10.1016/j.sleep.2019.11.114

Blanken, T. F., Borsboom, D., Penninx, B. W., & Van Someren, E. J. (2020). Network outcome analysis identifies difficulty initiating sleep as a primary target for prevention of depression: A 6-year prospective study. *Sleep*, *43*(5), 1–6. https://doi.org/10.1093/sleep/zsz288

Blom, K., Jernelöv, S., Kraepelien, M., Bergdahl, M. O., Jungmarker, K., Ankartjärn, L., Lindefors, N., & Kaldo, V. (2015). Internet Treatment Addressing either Insomnia or Depression, for Patients with both Diagnoses: A Randomized Trial. *Sleep*, *38*(2), 267–277. https://doi.org/10.5665/sleep.4412

Blom, K., Jernelöv, S., Lindefors, N., & Kaldo, V. (2016). Facilitating and hindering factors in Internet-delivered treatment for insomnia and depression. *Internet Interventions*, *4*, 51–60. https://doi.org/10.1016/j.invent.2016.03.004

Blom, K., Jernelöv, S., Rück, C., Lindefors, N., & Kaldo, V. (2017). Three-year follow-up comparing cognitive behavioral therapy for depression to cognitive behavioral therapy for insomnia, for patients with both diagnoses. *Sleep*, *40*(8). https://doi.org/10.1093/sleep/zsx108

Braun, L., Titzler, I., Ebert, D. D., Buntrock, C., Terhorst, Y., Freund, J., Thielecke, J., & Baumeister, H. (2019). Clinical and cost-effectiveness of guided internet-based interventions in the indicated prevention of depression in green professions (PROD-A): Study protocol of a 36-Month follow-up pragmatic randomized controlled trial. *BMC Psychiatry*, *19*(1), 278. https://doi.org/10.1186/s12888-019-2244-y

Braun, L., Titzler, I., Terhorst, Y., Freund, J., Thielecke, J., Ebert, D. D., & Baumeister, H. (2021a). Are guided internet-based interventions for the indicated prevention of depression in green professions effective in the long run? Longitudinal analysis of the 6- and 12-Month follow-up of a pragmatic randomized controlled trial (PROD-A). *Internet Interventions*, *26*, 100455. https://doi.org/10.1016/j.invent.2021.100455

Braun, L., Titzler, I., Terhorst, Y., Freund, J., Thielecke, J., Ebert, D. D., & Baumeister, H. (2021b). Effectiveness of guided internet-based interventions in the indicated prevention of depression in green professions (PROD-A): Results of a pragmatic randomized controlled trial. *Journal of Affective Disorders*, *278*, 658–671. https://doi.org/10.1016/j.jad.2020.09.066

Brohan, E., & Thornicroft, G. (2010). Stigma and discrimination of mental health problems: Workplace implications. *Occupational Medicine*, *60*(6), 414–415. https://doi.org/10.1093/occmed/kqq048

Bundesärztekammer, B. Ä. K., Kassenärztliche Bundesvereinigung, K. B. V., & Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften, A. W. M. F. (2022). *Nationale VersorgungsLeitlinie Unipolare Depression—Langfassung* (3.0) [Text/pdf]. https://doi.org/10.6101/AZQ/000493

Cheng, P., Kalmbach, D. A., Castelan, A. C., Murugan, N., & Drake, C. L. (2020). Depression prevention in digital cognitive behavioral therapy for insomnia: Is rumination a mediator? *Journal of Affective Disorders*, *273*(August 2019), 434–441. https://doi.org/10.1016/j.jad.2020.03.184

Cheng, P., Kalmbach, D. A., David A. Kalmbach, Tallent, G., Joseph, C. L. M., Espie, C. A., & Drake, C. L. (2019). Depression prevention via digital cognitive behavioral therapy for insomnia: A randomized controlled trial. *Sleep*, *42*(10). https://doi.org/10.1093/sleep/zsz150

Cheng, P., Luik, A. I., Fellman-Couture, C., Peterson, E., Joseph, C. L. M., Tallent, G., Tran, K. M., Ahmedani, B. K., Roehrs, T., Roth, T., & Drake, C. L. (2019). Efficacy of digital CBT for insomnia to reduce depression across demographic groups: A randomized trial. *Psychological Medicine*, *49*(3), 491–500. https://doi.org/10.1017/S0033291718001113

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: A global return on investment analysis. *The Lancet Psychiatry*, *3*(5), 415–424. https://doi.org/10.1016/S2215-0366(16)30024-4

Christensen, H., Batterham, P. J., Gosling, J. A., Ritterband, L. M., Griffiths, K. M., Thorndike, F. P., Glozier, N., O'Dea, B., Hickie, I. B., & Mackinnon, A. J. (2016). Effectiveness of an online insomnia program (SHUTi) for prevention of depressive episodes (the GoodNight Study): A randomised controlled trial. *The Lancet Psychiatry*, *3*(4), 333–341. https://doi.org/10.1016/S2215-0366(15)00536-2

Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A., & Liu, J. (2013). A Nondegenerate Penalized Likelihood Estimator for Variance Parameters in Multilevel Models. *Psychometrika*, *78*(4), 685–709. https://doi.org/10.1007/s11336-013-9328-2

Clement, S., Schauman, O., Graham, T., Maggioni, F., Evans-Lacko, S., Bezborodovs, N., Morgan, C., Rüsch, N., Brown, J. S. L., & Thornicroft, G. (2015). What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. *Psychological Medicine*, *45*(1), 11–27. https://doi.org/10.1017/S0033291714000129

Cooper, K. M., Gin, L. E., & Brownell, S. E. (2020). Depression as a concealable stigmatized identity: What influences whether students conceal or reveal their depression in undergraduate research experiences? *International Journal of STEM Education*, *7*(1), 27. https://doi.org/10.1186/s40594-020-00216-5

Cuijpers, P. (2021). Indirect prevention and treatment of depression: An emerging paradigm? *Clinical Psychology in Europe*, *3*(4). https://doi.org/10.32872/cpe.6847

Cuijpers, P., Pineda, B. S., Quero, S., Karyotaki, E., Struijs, S. Y., Figueroa, C. A., Llamas, J. A., Furukawa, T. A., & Muñoz, R. F. (2021). Psychological interventions to prevent the onset of depressive disorders: A meta-analysis of randomized controlled trials. *Clinical Psychology Review*, *83*, 101955. https://doi.org/10.1016/j.cpr.2020.101955

Cuijpers, P., Smit, F., Aalten, P., Batelaan, N., Klein, A., Salemink, E., Spinhoven, P., Struijs, S., Vonk, P., Wiers, R. W., de Wit, L., Gentili, C., Ebert, D. D., Bruffaerts, R., Kessler, R. C., & Karyotaki, E. (2021). The Associations of Common Psychological Problems With Mental Disorders Among College Students. *Frontiers in Psychiatry*, *12*(September), 1–9. https://doi.org/10.3389/fpsyt.2021.573637

Cuijpers, P., van Straten, A., Warmerdam, L., & van Rooy, M. J. (2010). Recruiting participants for interventions to prevent the onset of depressive disorders: Possibile ways to increase participation rates. *BMC Health Services Research*, *10*, 181. https://doi.org/10.1186/1472-6963-10-181

Dieck, A., Morin, C. M., & Backhaus, J. (2018). Deutsche Version des Insomnia Severity Index: Validierung und Bestimmung eines Cut-off-Punkts für Insomnie. *Somnologie*, *22*(1), 27–35. https://doi.org/10.1007/s11818-017-0147-z

Domhardt, M., Cuijpers, P., Ebert, D. D., & Baumeister, H. (2021). More Light? Opportunities and Pitfalls in Digitalized Psychotherapy Process Research. *Frontiers in Psychology*, *12*, 544129. https://doi.org/10.3389/fpsyg.2021.544129

Ebert, D. D., Mortier, P., Kaehlke, F., Bruffaerts, R., Baumeister, H., Auerbach, R. P., Alonso, J., Vilagut, G., Martínez, K. U., Lochner, C., Cuijpers, P., Kuechler, A., Green, J., Hasking, P., Lapsley, C., Sampson, N. A., Kessler, R. C., & On behalf of the WHO World Mental Health—International College Student Initiative collaborators. (2019). Barriers of mental health treatment utilization among first-year college students: First cross-national results from the WHO World Mental Health International College Student Initiative. *International Journal of Methods in Psychiatric Research*, *28*(2), e1782. https://doi.org/10.1002/mpr.1782

Ebert, D. D., Thiart, H., Laferton, J. A. C., Berking, M., Riper, H., Cuijpers, P., Sieland, B., & Lehr, D. (2015). Restoring depleted resources: Efficacy and mechanisms of change of an internet-based unguided recovery training for better sleep and psychological detachment from work. *Health Psychology*, *34*(December), 1240–1251. https://doi.org/10.1037/hea0000277

Ebert, D. D., Van Daele, T., Nordgreen, T., Karekla, M., Compare, A., Zarbo, C., Brugnera, A., Øverland, S., Trebbi, G., Jensen, K. L., Kaehlke, F., & Baumeister, H. (2018). Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health. *European Psychologist*, *23*(2), 167–187. https://doi.org/10.1027/1016-9040/a000318

Emerson, R. W. (2020). Bonferroni Correction and Type I Error. *Journal of Visual Impairment & Blindness*, *114*(1), 77–78. https://doi.org/10.1177/0145482X20901378

Feng, G., Han, M., Li, X., Geng, L., & Miao, Y. (2020). The Clinical Effectiveness of Cognitive Behavioral Therapy for Patients with Insomnia and Depression: A Systematic Review and Meta-Analysis. *Evidence-Based Complementary and Alternative Medicine*, *2020*, 1–14. https://doi.org/10.1155/2020/8071821

Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J. L., Vos, T., & Whiteford, H. A. (2013). Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. *PLoS Medicine*, *10*, e1001547. https://doi.org/10.1371/journal.pmed.1001547

Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H., Chen, P., Hunot, V., & Churchill, R. (2014). Waiting list may be a nocebo condition in psychotherapy trials: A contribution from network meta-analysis. *Acta Psychiatrica Scandinavica*, *130*(3), 181–192. https://doi.org/10.1111/acps.12275

Garge, N. R., Bobashev, G., & Eggleston, B. (2013). Random forest methodology for model-based recursive partitioning: The mobForest package for R. *BMC Bioinformatics*, *14*(1), 125. https://doi.org/10.1186/1471-2105-14-125

Groenwold, R. H. H., Moons, K. G. M., Peelen, L. M., Knol, M. J., & Hoes, A. W. (2011). Reporting of treatment effects from randomized trials: A plea for multivariable risk ratios. *Contemporary Clinical Trials*, *32*(3), 399–402. https://doi.org/10.1016/j.cct.2010.12.011

Gutiérrez-Rojas, L., Porras-Segovia, A., Dunne, H., Andrade-González, N., & Cervilla, J. A. (2020). Prevalence and correlates of major depressive disorder: A systematic review. *Brazilian Journal of Psychiatry*, *42*(6), 657–672. https://doi.org/10.1590/1516-4446-2020-0650

Harrell, F. E., Lee, K. L., & Mark, D. B. (1996). Multivariable Prognostic Models: Issues In Developing Models, Evaluating Assumptions And Adequacy, And Measuring And Reducing Errors. *Statistics in Medicine*, *15*(4), 361–387. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361

Harrer, M., Apolinário-Hagen, J., Fritsche, L., Salewski, C., Zarski, A. C., Lehr, D., Baumeister, H., Cuijpers, P., & Ebert, D. D. (2021). Effect of an internet- and app-based stress intervention compared to online psychoeducation in university students with depressive symptoms: Results of a randomized controlled trial. *Internet Interventions*, *24*(February). https://doi.org/10.1016/j.invent.2021.100374

Hautzinger, M., Bailer, M., Hofmeister, D., & Keller, F. (2012). *Allgemeine Depressionsskala* (2nd ed.). Hogrefe.

Higgins, J. P. T., Savović, J., Page, M. J., & Sterne, J. A. (2019). *Revised Cochrane risk-of-bias tool for randomized trials (RoB 2)*. riskofbias.info

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Kalmbach, D. A., Cheng, P., Ahmedani, B. K., Peterson, E. L., Reffi, A. N., Sagong, C., Seymour, G. M., Ruprich, M. K., & Drake, C. L. (2022). Cognitive-behavioral therapy for insomnia prevents and alleviates suicidal ideation: Insomnia remission is a suicidolytic mechanism. *Sleep*, *45*(12), zsac251. https://doi.org/10.1093/sleep/zsac251

Kraepelien, M., Forsell, E., & Blom, K. (2022). Large-scale implementation of insomnia treatment in routine psychiatric care: Patient characteristics and insomnia-depression comorbidity. *Journal of Sleep Research*, *31*(1). https://doi.org/10.1111/jsr.13448

Kreiner, G. E. (2006). Consequences of work-home segmentation or integration: A person-environment fit perspective. *Journal of Organizational Behavior*, *27*(4), 485–507. https://doi.org/10.1002/job.386

Leerssen, J., Lakbila-Kamal, O., Dekkers, L. M. S., Ikelaar, S. L. C., Albers, A. C. W., Blanken, T. F., Lancee, J., Van Der Lande, G. J. M., Maksimovic, T., Mastenbroek, S. E., Reesen, J. E., Van De Ven, S., Van Der Zweerde, T., Foster-Dingley, J. C., & Van Someren, E. J. W. (2021). Treating Insomnia with High Risk of Depression Using Therapist-Guided Digital Cognitive, Behavioral, and Circadian Rhythm Support Interventions to Prevent Worsening of Depressive Symptoms: A Randomized Controlled Trial. *Psychotherapy and Psychosomatics*, *91*, 168–179. https://doi.org/10.1159/000520282

Li, L., Wu, C., Gan, Y., Qu, X., & Lu, Z. (2016). Insomnia and the risk of depression: A meta-analysis of prospective cohort studies. *BMC Psychiatry*, *16*(1). https://doi.org/10.1186/s12888-016-1075-3

Little, R. J. A., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. John Wiley & Sons, Inc. https://doi.org/10.1002/9781119013563

Mohr, G., Rigotti, T., & Müller, A. (2005). Irritation—Ein Instrument zur Erfassung psychischer Beanspruchung im Arbeitskontext. Skalen- und Itemparameter aus 15 Studien. *Zeitschrift für Arbeits- und Organisationspsychologie A&O*, *49*(1), 44–48. https://doi.org/10.1026/0932-4089.49.1.44

Morin, C. M., Bertisch, S. M., Pelayo, R., Watson, N. F., Winkelman, J. W., Zee, P. C., & Krystal, A. D. (2023). What Should Be the Focus of Treatment When Insomnia Disorder Is Comorbid with Depression or Anxiety Disorder? *Journal of Clinical Medicine*, *12*(5), 1975. https://doi.org/10.3390/jcm12051975

Morin, C. M., Drake, C. L., Harvey, A. G., Krystal, A. D., Manber, R., Riemann, D., & Spiegelhalder, K. (2015). Insomnia disorder. *Nature Reviews Disease Primers*, *1*(1), 15026. https://doi.org/10.1038/nrdp.2015.26

National Institute for Health and Care Excellence, N. I. C. E. (2022). *Depression in adults: Treatment and management*. www.nice.org.uk/guidance/ng222

R Core Team. (2022). *R: A language and environment for statistical computing* [Manual]. https://www.R-project.org/

Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Riley, R. D., Debray, T. P. A., Fisher, D., Hattle, M., Marlin, N., Hoogland, J., Gueyffier, F., Staessen, J. A., Wang, J., Moons, K. G. M., Reitsma, J. B., & Ensor, J. (2020). Individual participant data Meta-analysis to examine interactions between treatment effect and Participant-level covariates: Statistical recommendations for conduct and planning. *Statistics in Medicine*, *39*(15), 2115–2137. https://doi.org/10.1002/sim.8516

Roach, M., Juday, T., Tuly, R., Chou, J. W., Jena, A. B., & Doghramji, P. P. (2021). Challenges and opportunities in insomnia disorder. *International Journal of Neuroscience*, *131*(11), 1058–1065. https://doi.org/10.1080/00207454.2020.1773460

Robitzsch, A., & Grund, S. (2022). *miceadds: Some additional multiple imputation functions, especially for "mice"* [Manual]. https://CRAN.R-project.org/package=miceadds

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470316696

Satinsky, E., Fuhr, D. C., Woodward, A., Sondorp, E., & Roberts, B. (2019). Mental health care utilisation and access among refugees and asylum seekers in Europe: A systematic review. *Health Policy*, *123*(9), 851–863. https://doi.org/10.1016/j.healthpol.2019.02.007

Schaufeli, W., & Bakker, A. (2004). *Utrecht Work Engagement Scale (UWES)—Preliminary Manual*. https://www.wilmarschaufeli.nl/tests/#engagement

Shannon, P. J., Wieling, E., Simmelink-McCleary, J., & Becher, E. (2015). Beyond Stigma: Barriers to Discussing Mental Health in Refugee Populations. *Journal of Loss and Trauma*, *20*(3), 281–296. https://doi.org/10.1080/15325024.2014.934629

Siegrist, J. (2008). Chronic psychosocial stress at work and risk of depression: Evidence from prospective studies. *European Archives of Psychiatry and Clinical Neuroscience*, *258*(S5), 115–119. https://doi.org/10.1007/s00406-008-5024-0

Siegrist, J., Wege, N., Pühlhofer, F., & Wahrendorf, M. (2009). A short generic measure of work stress in the era of globalization: Effort–reward imbalance. *International Archives of Occupational and Environmental Health*, *82*(8), 1005–1013. https://doi.org/10.1007/s00420-008-0384-3

Simon, L., Steinmetz, L., Feige, B., Benz, F., Spiegelhalder, K., & Baumeister, H. (2023). Comparative efficacy of onsite, digital, and other settings for cognitive behavioral therapy for insomnia: A systematic review and network meta-analysis. *Scientific Reports*, *13*(1), 1929. https://doi.org/10.1038/s41598-023-28853-0

Smith, G. C. S., Seaman, S. R., Wood, A. M., Royston, P., & White, I. R. (2014). Correcting for Optimistic Prediction in Small Data Sets. *American Journal of Epidemiology*, *180*(3), 318–324. https://doi.org/10.1093/aje/kwu140

Spanhel, K., Burdach, D., Pfeiffer, T., Lehr, D., Spiegelhalder, K., Ebert, D. D., Baumeister, H., Bengel, J., & Sander, L. B. (2021). Effectiveness of an internet-based intervention to improve sleep difficulties in a culturally diverse sample of international students: A randomised controlled pilot study. *Journal of Sleep Research*, *September*, 1–12. https://doi.org/10.1111/jsr.13493

Spanhel, K., Hovestadt, E., Lehr, D., Spiegelhalder, K., Baumeister, H., Bengel, J., & Sander, L. B. (2022). Engaging Refugees With a Culturally Adapted Digital Intervention to Improve Sleep: A Randomized Controlled Pilot Trial. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/fpsyt.2022.832196

Staner, L. (2010). Comorbidity of insomnia and depression. *Sleep Medicine Reviews*, *14*(1), 35–46. https://doi.org/10.1016/j.smrv.2009.09.003

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., … Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*, *366*, l4898. https://doi.org/10.1136/bmj.l4898

Stewart, L. A., Clarke, M., Rovers, M., Riley, R. D., Simmonds, M., Stewart, G., & Tierney, J. F. (2015). Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. *JAMA*, *313*(16), 1657. https://doi.org/10.1001/jama.2015.3656

Stewart, R., Besset, A., Bebbington, P., Brugha, T., Lindesay, J., Jenkins, R., Singleton, N., & Meltzer, H. (2006). Insomnia Comorbidity and Impact and Hypnotic Use by Age Group in a National Survey Population Aged 16 to 74 Years. *Sleep*, *29*(11), 1391–1397. https://doi.org/10.1093/sleep/29.11.1391

Thiart, H., Lehr, D., Ebert, D. D., Berking, M., & Riper, H. (2015). Log in and breathe out: Internet-based recovery training for sleepless employees with work-related strain – results of a randomized controlled trial. *Scandinavian Journal of Work, Environment and Health*, *41*(2), 164–174. https://doi.org/10.5271/sjweh.3478

Thiart, H., Lehr, D., Ebert, D. D., Sieland, B., Berking, M., & Riper, H. (2013). Log in and breathe out: Efficacy and cost-effectiveness of an online sleep training for teachers affected by work-related strain—Study protocol for a randomized controlled trial. *Trials*. https://doi.org/10.1186/1745-6215-14-169

Torok, M., Han, J., Baker, S., Werner-Seidler, A., Wong, I., Larsen, M. E., & Christensen, H. (2020). Suicide prevention using self-guided digital interventions: A systematic review and meta-analysis of randomised controlled trials. *The Lancet Digital Health*, *2*(1), e25–e36. https://doi.org/10.1016/S2589-7500(19)30199-2

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

van der Zweerde, T., van Straten, A., Effting, M., Kyle, S. D., & Lancee, J. (2019). Does online insomnia treatment reduce depressive symptoms? A randomized controlled trial in individuals with both insomnia and depressive symptoms. *Psychological Medicine*, *49*(3), 501–509. https://doi.org/10.1017/S0033291718001149

Vargas, I., & Perlis, M. L. (2020). Insomnia and depression: Clinical associations and possible mechanistic links. *Current Opinion in Psychology*, *34*, 95–99. https://doi.org/10.1016/j.copsyc.2019.11.004

Vilagut, G., Forero, C. G., Barbaglia, G., & Alonso, J. (2016). Screening for Depression in the General Population with the Center for Epidemiologic Studies Depression (CES-D): A Systematic Review with Meta-Analysis. *PLoS ONE*, *11*(5), e0155431. https://doi.org/10.1371/journal.pone.0155431

Vos, T., Haby, M. M., Barendregt, J. J., Kruijshaar, M., Corry, J., & Andrews, G. (2004). The burden of major depression avoidable by longer-term treatment strategies. *Archives of General Psychiatry*, *61*, 1097–1103. https://doi.org/10.1001/archpsyc.61.11.1097

Wade, A. (2010). The societal costs of insomnia. *Neuropsychiatric Disease and Treatment*, 1. https://doi.org/10.2147/NDT.S15123

Weisel, K. K., Lehr, D., Heber, E., Zarski, A. C., Berking, M., Riper, H., & Ebert, D. D. (2018). Severely burdened individuals do not need to be excluded from internet-based and mobile-based stress management: Effect modifiers of treatment outcomes from three randomized controlled trials. *Journal of Medical Internet Research*, *20*(6), e211. https://doi.org/10.2196/jmir.9387

Werntz, A., Amado, S., Jasman, M., Ervin, A., & Rhodes, J. E. (2023). Providing Human Support for the Use of Digital Mental Health Interventions: Systematic Meta-review. *Journal of Medical Internet Research*, *25*, e42864. https://doi.org/10.2196/42864

Ye, Y., Chen, N., Chen, J., Liu, J., Lin, L., Liu, Y., Lang, Y., Li, X., Yang, X., & Jiang, X. (2016). Internet-based cognitive–behavioural therapy for insomnia (ICBT-i): A meta-analysis of randomised controlled trials. *BMJ Open*, *6*(11), e010707. https://doi.org/10.1136/bmjopen-2015-010707

Zeileis, A., Hothorn, T., & Hornik, K. (2008). Model-Based Recursive Partitioning. *Journal of Computational and Graphical Statistics*, *17*(2), 492–514. https://doi.org/10.1198/106186008X319331

**Online Supplement to:**

Thielecke, J., Paula Kuper, Schuurmans, L., Harrer, M., Ebert, D. D., Lehr, D., Behrendt, D., Brückner, H., Horvath, H., Riper, H., Cuijpers, P., & Buntrock, C. (in preparation). Who benefits from indirect prevention and treatment of depression using online interventions for insomnia? Results from an individual-participant-data meta-analysis

**Content:**



Corresponding author: Janika Thielecke, Technical University Munich (TUM), Department of Sports and Health Sciences, Georg-Brauchle-Ring 60/62, 80992 Munich. Mail: janika.thielecke@tum.de

**Supplement 1: Methodological Changes to the preregistration**

**Table 1**

*Methodological Changes to the preregistration, the rational for the changes and the expected consequences*

| Change | Rational & Consequence |
|---|---|
| Exclusion of two additional studies, originally promising their IPD | The preregistration was written with all identified studies in mind. Unfortunately, administrative reasons let to the exclusion of the two studies by Spanhel et al (2021, 2022), which would have contributed a lot to the heterogeneity between studies, given that the interventions had crucial adaptations in content, duration and language due to the different target groups (international students and refugees). Due to the low number of participants in these pilot trials, we deemed a combined approach of IPD and aggregated data not sensible. This led to two major changes in the methodology:<br>1) We excluded level two (study level) variables as potential moderators, because the remaining studies were too similar in their characteristics, which limits our results to a work-focused online CBT-I intervention.<br>2) We reconsidered the potential variables to seek from the original study authors, because the sample now consisted of solely working participants, which had not been studied as an individual group before. The addition of engagement and effort-reward-ration as potential moderators allows us better consider differential effects of the intervention in the working population. Ethnicity and employment were excluded as potential moderators due to a lack of variance. |
| Use of CES-D scores instead of common metrics as the main outcome | Because all included studies assessed the CES-D there was no need to harmonize the outcome by using common metrics. We decided to stay with the original measure in order to not introduce unnecessary insecurity through the score transformation and remain an easier interpretation of the results. |
| separate imputation models for post- and follow-up assessment | In the preregistration we did not consider that one of the potential studies had no follow-up assessment in the control group, making it impossible to estimate a treatment effect. Therefore, that study was excluded from the follow-up analysis. In order to avoid imputation of data for the not-assessed follow-up data, separate imputation models had to be built for the two time points including only the studies, which assessed the time point.<br>Consequently, analysis at post-treatment were done with the full sample of N= 563 from four studies and analysis at follow-up included N=433 individuals from three studies. |

**Table 1 (continued)**

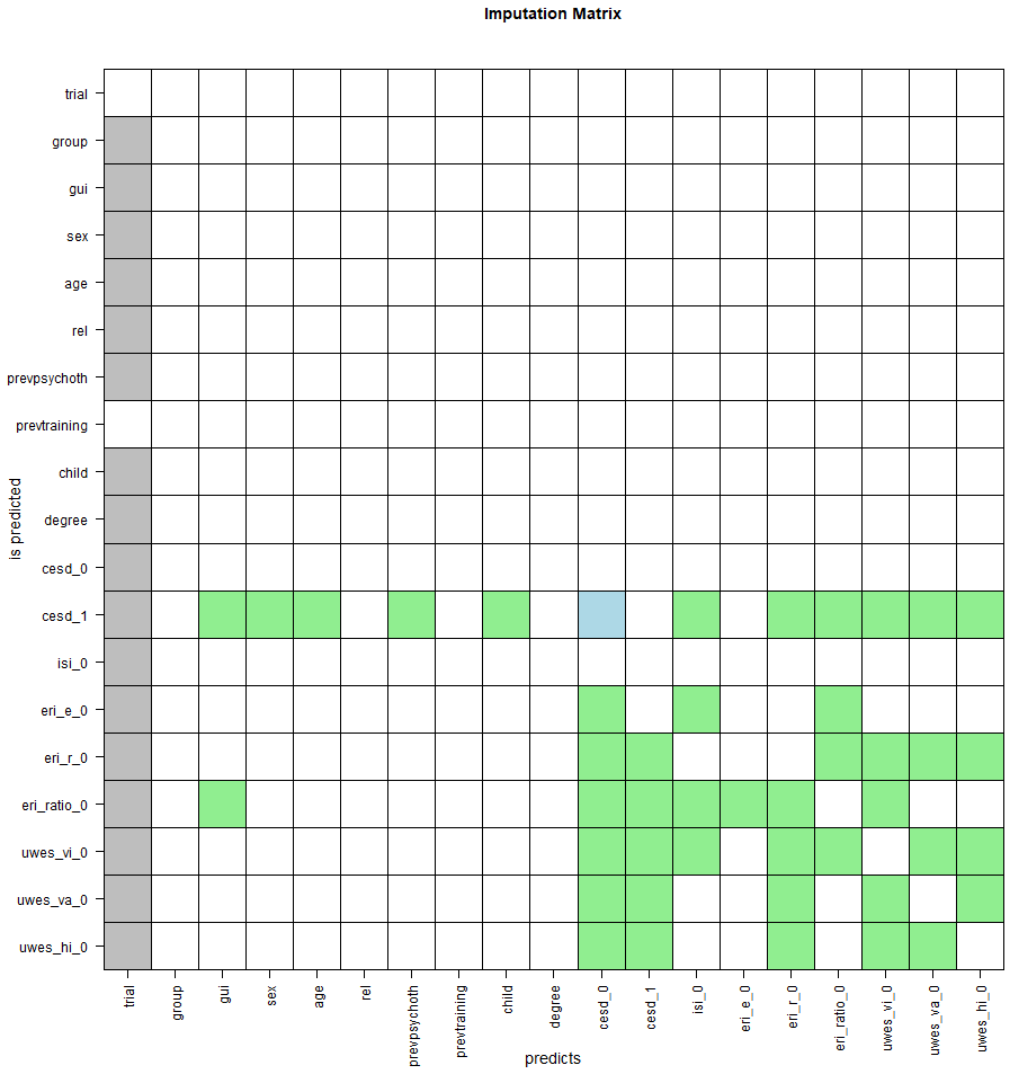
*Methodological Changes to the preregistration, the rational for the changes and the expected consequences*

| Change | Rational & Consequence |
|---|---|
| Imputation method changed from 2l.pan to 2l.pmm | We inspected our imputed data using stripplots, densityplots and traceplots following the diagnostic procedure described by van Buren (2018). Using 2l.pan as the imputation method the density plots showed that density estimates for the marginal distributions imputed data did not fit the observed data very well. Switching the imputation model to 2l.pmm increased the fit while still accounting for the data being nested within studies. |
| Additional sensitivity analysis without the sleep item in CES-D | In preparation of the manuscript, we decided to run an additional sensitivity analysis in which we excluded the sleep-item from the CES-D total to confirm the robustness of our results and conclude that changes in depression symptom severity does not solely derive from changes in the sleep item. |
| $R^2$ reported | We did not pre-specify to report $R^2$ but thought it was helpful to interpret the overall model fit, so we included that information in addition to the pre-specified effect estimates and model parameters. |
| Note to research question 3 | Research question 3 (mechanisms and mediators of the treatment approach) is planned as a separate analysis with a Bayesian methodology and thus was considered beyond the scope of this article. |

**Supplement 2: Imputation models**

**Figure 1**

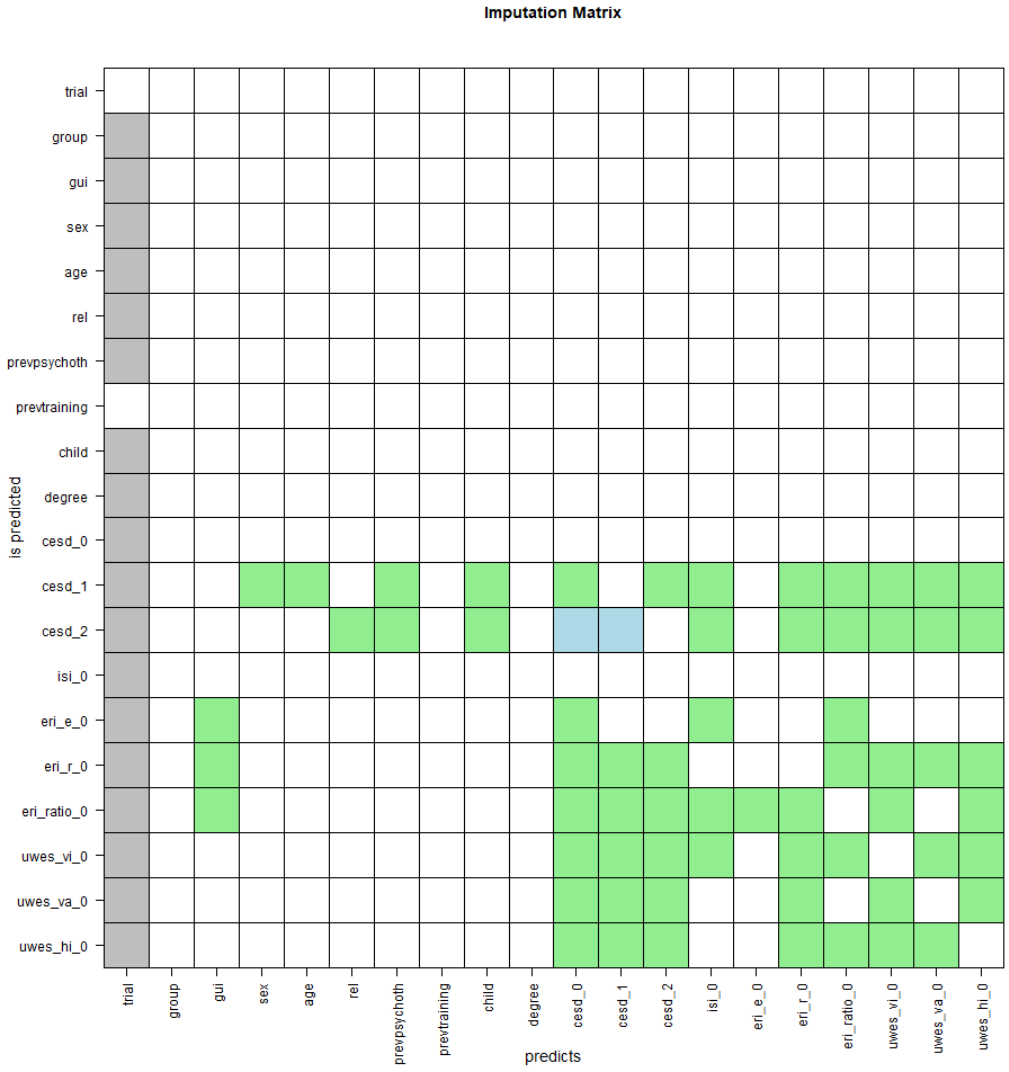
*Predictor matrix of the imputation model for post-treatment data*




*Note.* Variables were not used as predictors if their correlation with the imputed variables was lower than $r = 0.05$.

**Figure 2**

*Predictor matrix of the imputation model for follow-up data*

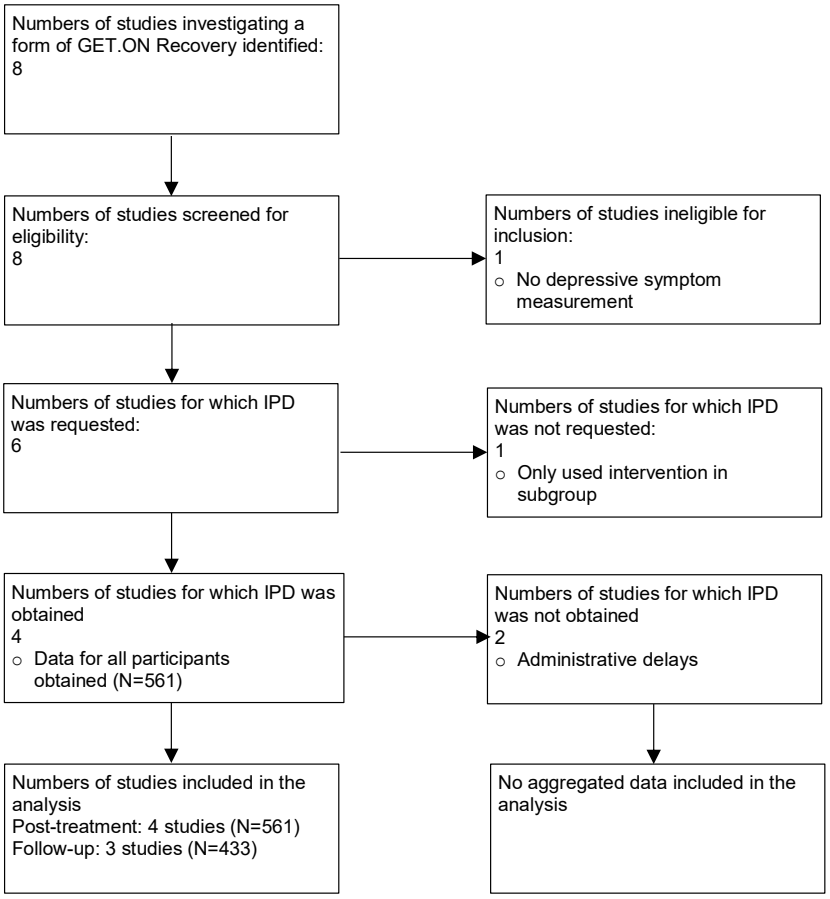



*Note.* Variables were not used as predictors if their correlation with the imputed variables was lower than $r = 0.05$.

**Supplement 3: Flow-Chart study inclusion**

**Figure 3**

*Flowchart of study selection and inclusion*



**Supplement 4: Participants' characteristics**

**Table 2**

*Baseline descriptive of sociodemographic, clinical and work-related characteristics*

| variable | total (N=561) | | intervention (N=280) | | control (N=281) | |
|---|---|---|---|---|---|---|
| | n | % | n | % | n | % |
| **sociodemographic variables** | | | | | | |
| age (M, SD) | 47.15 | 9.73 | 47.82 | 9.52 | 46.47 | 9.91 |
| sex | | | | | | |
| *female* | 381 | 67.91 | 186 | 66.43 | 195 | 69.4 |
| *male* | 179 | 31.91 | 94 | 33.57 | 85 | 30.25 |
| relationship status | | | | | | |
| *single* | 122 | 21.75 | 57 | 20.36 | 65 | 23.13 |
| *married/in a relationship* | 395 | 70.41 | 201 | 71.79 | 194 | 69.04 |
| *divorced/separated* | 42 | 7.49 | 21 | 7.5 | 21 | 7.47 |
| *widowed* | 2 | 0.36 | 1 | 0.36 | 1 | 0.36 |
| employment | 558 | 99.47 | 280 | 100 | 278 | 98.93 |
| children | 345 | 61.5 | 172 | 61.43 | 173 | 61.57 |
| education | | | | | | |
| *education up to high school only (7-9 years)* | 24 | 4.28 | 15 | 5.36 | 9 | 3.2 |
| *high school education (12-13 years)* | 77 | 13.73 | 40 | 14.29 | 37 | 13.17 |
| *education after high school* | 429 | 76.47 | 209 | 74.64 | 220 | 78.29 |
| *post graduate education (>17 years)* | 31 | 5.53 | 16 | 5.71 | 15 | 5.34 |
| **clinical variables** | | | | | | |
| CES-D (M, SD) | 22.13 | 7.99 | 22.24 | 8.05 | 22.02 | 7.94 |
| clinically relevant depression (CESD >=16) | 440 | 78.43 | 224 | 80 | 216 | 76.87 |
| ISI (M,SD) | 16.81 | 3.65 | 16.87 | 3.65 | 16.75 | 3.66 |
| clinically relevant insomnia (ISI >=15) | 425 | 75.76 | 211 | 75.36 | 214 | 76.16 |
| close-to-symptom-free (CES-D <16) | 121 | 21.57 | 56 | 20 | 65 | 23.13 |
| previous experience | | | | | | |
| *psychotherapy* | 216 | 38.5 | 98 | 35 | 118 | 41.99 |
| *health training* | 76 | 13.55 | 34 | 12.14 | 42 | 14.95 |
| **work related variables** | | | | | | |
| Effort-Reward-Ratio (ERI-S) | | | | | | |
| *effort (M, SD)* | 9.93 | 1.74 | 9.95 | 1.74 | 9.91 | 1.74 |
| *reward (M, SD)* | 17.09 | 3.27 | 17.1 | 3.32 | 17.09 | 3.23 |
| *effort-reward-Ratio (M, SD)* | 1.41 | 0.40 | 1.41 | 0.41 | 1.41 | 0.39 |
| Work engagement (UWES) | | | | | | |
| *vigor (M, SD)* | 3.02 | 1.26 | 3.02 | 1.29 | 3.02 | 1.23 |
| *absorption (M, SD)* | 3.07 | 1.49 | 3.04 | 1.54 | 3.09 | 1.44 |
| *dedication (M, SD)* | 3.38 | 1.36 | 3.37 | 1.37 | 3.4 | 1.34 |

*Abbreviations.* CES-D: Center for Epidemiological Studies Depression Scale (total score range 0-60), ISI: Insomnia Severity Index (total score range 0-28), ERI-S: Effort-Reward Imbalance Scale – Short form (range effort: 3-15, range reward: 7-35), UWES: Utrecht Work Engagement Scale (scale range 0-6)

**Table 3**

*Overview over all depressive outcomes at post-treatment (8 weeks post-randomization) and follow-up (24 weeks post-randomization) based on complete cases*

| variable | group | $n_{T0}$ | M | SD | ß [95%-CI] | $p_{adjusted}$ | D [95%-CI] | $\tau_{intercept}$ | $\tau_{group}$ | marginal $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| depression symptom severity (CES-D) | | | | | | | | | | |
| *baseline* | IG | 216 | 22.03 | 8.00 | | | | | | |
| | CG | 239 | 21.75 | 7.95 | | | | | | |
| *post* | IG | 216 | 13.75 | 8.02 | -6.11 [ -7.79, -4.43] | <0.001 | -0.70 [-0.51, -0.89] | 0.67 | 1.26 | 0.41 |
| | CG | 239 | 19.65 | 8.82 | | | | | | |
| *follow-up* | IG | 149 | 14.08 | 7.98 | -6.97 [ -9.21, -4.73] | <0.001 | -0.80 [-0.58, -1.03] | 0.67 | 1.26 | 0.34 |
| | CG | 178 | 21.05 | 9.26 | | | | | | |
| **variable** | **group** | **$n_{T0}$** | **$n_{event}$** | **%** | **OR [95%-CI]** | **$p_{adjusted}$** | **NNT [95%-CI]** | **$\tau_{intercept}$** | **$\tau_{group}$** | **marginal $R^2$** |
| clinical relevant depression (MDD onset)[a] | | | | | | | | | | |
| *post* | IG | 46 | 7 | 15.2 | 1.97 [0.70, 5.58] | 1.0 | 9.8 [-3.9, 20.0] | 0.04 | 0.04 | 0.11 |
| | CG | 59 | 15 | 25.4 | | | | | | |
| *follow-up* | IG | 33 | 8 | 24.2 | 1.76 [0.62, 5.03] | 1.0 | 8.5 [-3.2, 12.5] | 0.03 | 0.04 | 0.13 |
| | CG | 50 | 18 | 36.0 | | | | | | |
| Close-to-symptom-free status[b] | | | | | | | | | | |
| *post* | IG | 170 | 106 | 62.4 | 0.14 [0.08, 0.24] | <0.001 | 2.5 [3.3, 2.0] | 0.02 | 0.04 | 0.34 |
| | CG | 180 | 40 | 22.2 | | | | | | |
| *follow-up* | IG | 116 | 71 | 61.2 | 0.12 [0.06, 0.23] | <0.001 | 2.2 [3.0, 1.8] | 0.03 | 0.06 | 0.31 |
| | CG | 128 | 21 | 16.4 | | | | | | |
| RCI improvement | | | | | | | | | | |
| *post* | IG | 280 | 123 | 43.9 | 0.17 [0.11, 0.27] | <0.001 | 3.9 [5.4, 3.0] | 0.02 | 0.03 | 0.30 |
| | CG | 281 | 51 | 18.1 | | | | | | |
| *follow-up* | IG | 280 | 80 | 28.6 | 0.1 [0.05, 0.19] | <0.001 | 5.2 [7.7, 3.9] | 0.05 | 0.05 | 0.45 |
| | CG | 281 | 26 | 9.3 | | | | | | |

**Table 3 (continued)**

*Overview over all depressive outcomes at post-treatment (8 weeks post-randomization) and follow-up (24 weeks post-randomization) based on complete cases*

| variable | group | $n_{T0}$ | $n_{event}$ | % | OR [95%-CI] | $p_{adjusted}$ | NNT [95%-CI] | $\tau_{intercept}$ | $\tau_{group}$ | marginal $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| RCI deterioration | | | | | | | | | | |
| *post* | IG | 280 | 7 | 2.5 | 2.93 [1.18, 7.30] | 0.25 | 20.1 [-11.7, -71.6] | 0.04 | 0.05 | 0.12 |
| | CG | 281 | 21 | 7.5 | | | | | | |
| *follow-up* | IG | 280 | 7 | 2.5 | 2.87 [1.14, 7.23] | 0.30 | 18.8 [-11.2, -59.0] | 0.05 | 0.05 | 0.12 |
| | CG | 281 | 22 | 7.8 | | | | | | |
| anchor-based clinically relevant change | | | | | | | | | | |
| *post* | IG | 280 | 135 | 48.2 | 0.15 [0.10, 0.24] | <0.001 | 3.2 [4.3, 2.6] | 0.02 | 0.03 | 0.22 |
| | CG | 281 | 49 | 17.4 | | | | | | |
| *follow-up* | IG | 280 | 90 | 32.1 | 0.11 [0.06, 0.20] | <0.001 | 4.4 [6.1, 3.4] | 0.03 | 0.05 | 0.28 |
| | CG | 281 | 26 | 9.3 | | | | | | |

*Note.* Numbers are estimated based on complete cases analysis. [a]analyzed in subgroup considered close-to-symptom-free at baseline (CES-D <16), [b]analyzed in subgroup exceeding the cut-offs for clinically relevant cases at baseline (CES-D≥16). *Abbreviations.* CES-D: Center for Epidemiological Studies Depression Scale, MDD: Major depressive disorder, RCI: Reliable Change Index, IG: intervention group, CG: control group, NNT: Numbers Needed to treat, $n_{T0}$: case number at baseline assessment, $n_{T0}$: cases number with outcome at post-treatment or follow-up, $\tau_{intercept}$: intercept variance, $\tau_{group}$: slope variance for the treatment effect.

**Table 4**

*Overview over all depressive outcomes at post-treatment (8 weeks post-randomization) and follow-up (24 weeks post-randomization) based on CES-D scores excluding the sleep item*

| variable | group | $n_{T0}$ | M | SD | d [95%-CI] | $p_{adjusted}$ | NNT [95%-CI] | $\tau_{intercept}$ | $\tau_{group}$ | marginal $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Depression symptom severity (CES-D) | | | | | | | | | | |
| *baseline* | IG | 280 | 19.97 | 7.89 | | | | | | |
| | CG | 281 | 19.81 | 7.82 | | | | | | |
| *post* | IG | 280 | 12.68 | 6.87 | -5.28 [-6.96, -3.59] | < .0001 | -0.65 [-0.44, -0.86] | 0.69 | 1.07 | 0.35 |
| | G | 281 | 17.81 | 7.89 | | | | | | |
| *follow-up* | IG | 280 | 12.98 | 6.48 | -6.46 [-8.59, -4.33] | < .0001 | -0.78 [-0.52 to -1.03] | 1.89 | 1.27 | 0.32 |
| | CG | 281 | 19.08 | 8.28 | | | | | | |
| **variable** | **group** | **$n_{T0}$** | **$n_{event}$** | **%** | **OR [95%-CI]** | **$p_{adjusted}$** | **NNT [95%-CI]** | **$\tau_{intercept}$** | **$\tau_{group}$** | **marginal $R^2$** |
| clinical relevant depression (MDD onset)[a] | | | | | | | | | | |
| *post* | IG | 84 | 5 | 6 | 5.34 [16.15, 1.76] | 1.00 | -4.3 [-3.0, -8.0] | 0.04 | 0.07 | 0.17 |
| | CG | 93 | 27 | 29 | | | | | | |
| *follow-up* | IG | 64 | 10 | 15.6 | 3.06 [8.19, 1.14] | 1.00 | -3.3 [-2.2, -6.2] | 0.03 | 0.05 | 0.13 |
| | CG | 76 | 35 | 46.1 | | | | | | |
| Close-to-symptom-free status[b] | | | | | | | | | | |
| *post* | IG | 196 | 140 | 71.4 | 0.17 [0.31, 0.10] | <0.001 | 2.1 [1.8, 2.5 ] | 0.03 | 0.03 | 0.24 |
| | CG | 188 | 44 | 23.4 | | | | | | |
| *follow-up* | IG | 152 | 110 | 72.4 | 0.12 [0.25, 0.06] | <0.001 | 1.8 [1.6, 2.2] | 0.05 | 0.03 | 0.28 |
| | CG | 141 | 25 | 17.7 | | | | | | |
| RCI improvement | | | | | | | | | | |
| *post* | IG | 280 | 142 | 50.7 | 0.21 [0.33, 0.13] | <0.001 | 3.2 [2.6, 4.2] | 0.01 | 0.01 | 0.22 |
| | CG | 281 | 54 | 19.2 | | | | | | |
| *follow-up* | IG | 280 | 113 | 40.4 | 0.13 [0.27, 0.07] | <0.001 | 3.4 [2.8, 4.4] | 0.05 | 0.04 | 0.3 |
| | CG | 281 | 31 | 11 | | | | | | |

**Table 4 (continued)**

*Overview over all depressive outcomes at post-treatment (8 weeks post-randomization) and follow-up (24 weeks post-randomization) based on CES-D scores excluding the sleep item*

| variable | group | $n_{T0}$ | $n_{event}$ | % | OR [95%-CI] | $p_{adjusted}$ | NNT [95%-CI] | $\tau_{intercept}$ | $\tau_{group}$ | marginal $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| RCI deterioration | | | | | | | | | | |
| *post* | IG | 280 | 8 | 2.9 | 2.31 [5.64, 0.95] | 0.28 | -23.5 [-12.7, -147.9] | 0.03 | 0.05 | 0.03 |
| | CG | 281 | 20 | 7.1 | | | | | | |
| *follow-up* | IG | 280 | 7 | 2.5 | 2.61 [6.75, 1.01] | 0.3 | -20.1 [-11.7, -71.6] | 0.06 | 0.04 | 0.04 |
| | CG | 281 | 21 | 7.5 | | | | | | |
| anchor-based clinical change | | | | | | | | | | |
| *post* | IG | 280 | 177 | 63.2 | 0.21 [0.33, 0.13] | <0.001 | 2.4 [2.0, 2.9] | 0.04 | 0.04 | 0.15 |
| | CG | 281 | 60 | 21.4 | | | | | | |
| *follow-up* | IG | 280 | 138 | 49.3 | 0.12 [0.23, 0.07] | <0.001 | 2.6 [2.2, 3.1] | 0.03 | 0.07 | 0.24 |
| | CG | 281 | 29 | 10.3 | | | | | | |

*Note.* Numbers are estimated based on imputed data. [a]analyzed in subgroup considered close-to-symptom-free at baseline (CES-D <16), [b]analyzed in subgroup exceeding the cut-offs for clinically relevant cases at baseline (CES-D≥16). *Abbreviations.* CES-D: Center for Epidemiological Studies Depression Scale (excluding the sleep item), MDD: Major depressive disorder, RCI: Reliable Change Index, IG: intervention group, CG: control group, NNT: Numbers Needed to treat, $n_{T0}$: case number at baseline assessment, $n_{T0}$: cases number with outcome at post-treatment or follow-up, $\tau_{intercept}$: intercept variance, $\tau_{group}$: slope variance for the treatment effect.

**Supplement 7: Results from random-forest-analysis (variable importance)**

**Figure 4**

*Variable importance of potential moderators for the treatment effect at post-treatment in random-forest models with 300 bootstrap samples (N= 561)*

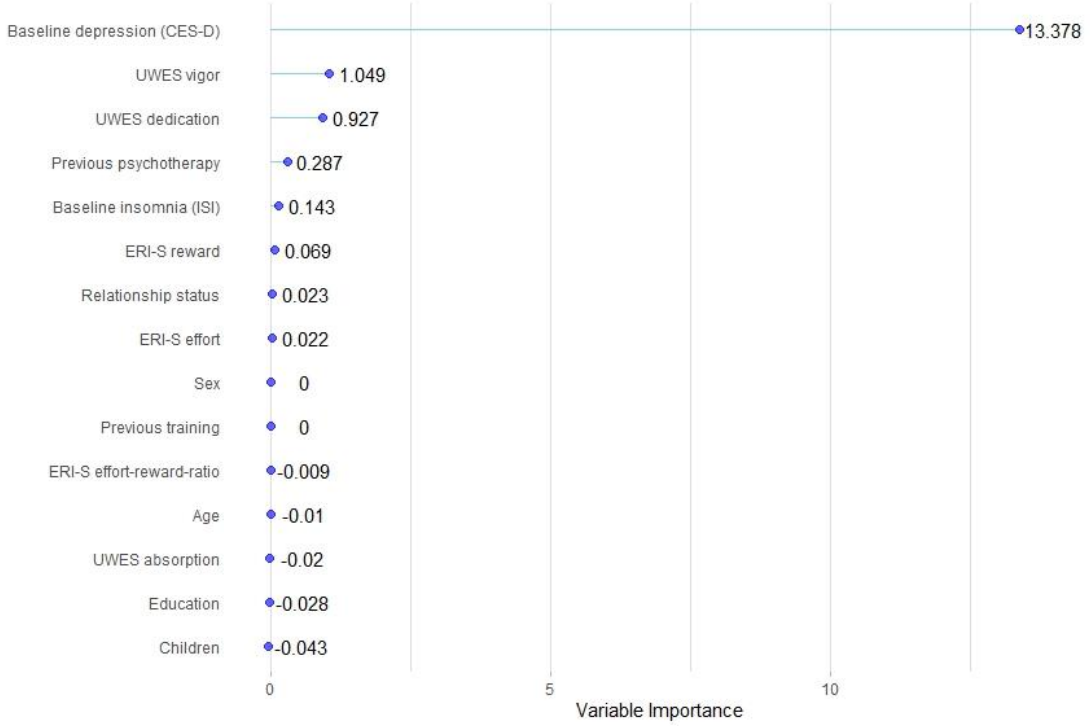
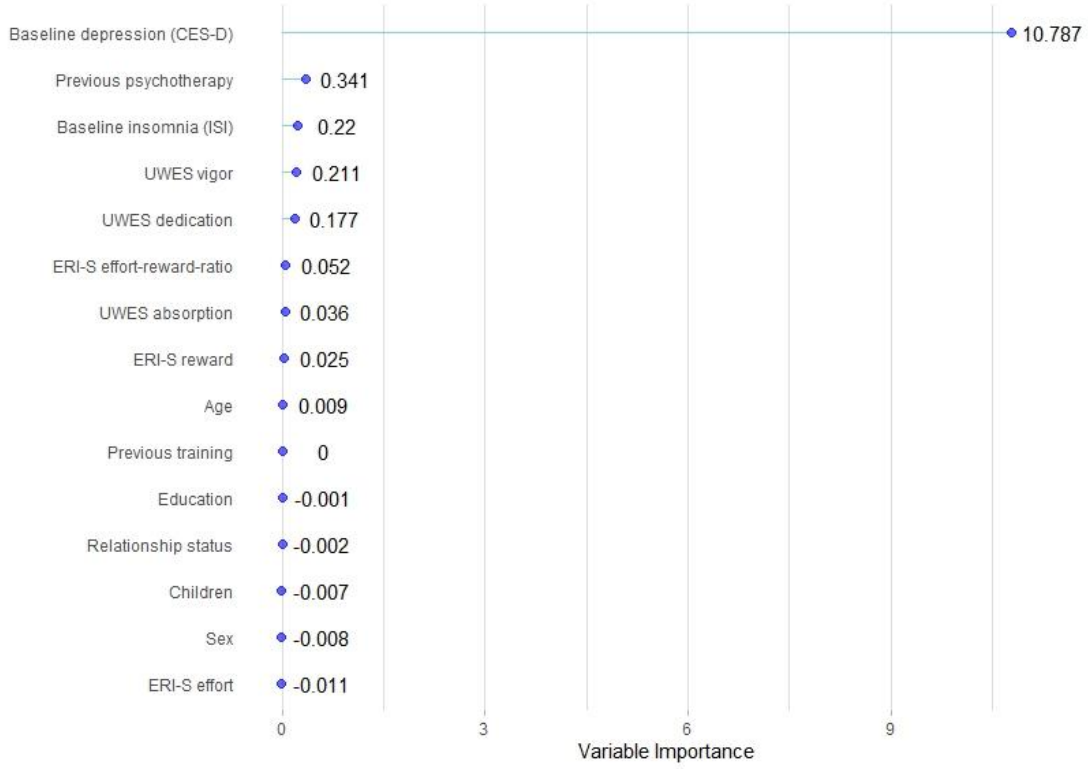


**Figure 5**

*Variable importance of potential moderators for the treatment effect at follow-up in random-forest models with 300 bootstrap samples (N= 433)*



# Part III

# GENERAL DISCUSSION

# 9

# DISCUSSION

## 9.1 CONTRIBUTION OF THIS WORK TO THE FIELD OF DIGITAL INTERVENTIONS IN THE PREVENTION OF DEPRESSION

In this work I set out to, on the one hand, to increase the evidence for the effectiveness of DMHIs and especially examine the chances of the telephone to deliver preventive interventions to at risk individuals. On the other hand, I wanted to address some of the challenges currently faced by DMHIs, namely slow implementation and low uptake, unknown predictors of treatment outcome and differential treatment efficacy.

In Articles 1 – 3, the telephone coaching RCT, I could show that a personalized telephone coaching for farmers could reduce depressive symptoms when compared to enhanced TAU (single psychoeducational e-mail) for up to 18-months after randomization. The coaching also improved several other psychological outcomes including perceived stress, emotional exhaustion and somatization. However, no group differences in depression onset were observed, likely due to limited power and the personalized but less depression-specific intervention.

In addition to the RCT, in Article 4, a qualitative interview study was performed which aimed to identify factors that supporting the implementation of the personalized telephone coaching into routine care and its uptake by the target group. The article described participant's experience with the telephone coaching with a focus on acceptance and satisfaction. According to participants' reports, they accepted the telephone coaching since they expected help with their psychosocial problems from it and appreciated the perceived low effort through the telephone conduct. Satisfaction was accelerated by the good relationship with the coach as well as the high degree of personalization. The results are limited to very satisfied participants, as the intervention was generally well received, and the convenience sampling method.

In order to examine predictors of the treatment outcome in DMHIs, a secondary data analysis of two randomized trials evaluating a web-based intervention for individuals with subthreshold depression was done in Article 5. Results on the effect outcome expectancy can have on depression outcomes after using a DMHI were inconclusive

but indicated a potential influence in the long run. A research agenda of what needs to be considered in order to determine the potential of outcome expectancy as a predictor of treatment outcome and a possible treatment mechanism was outlined. More dedicated studies are needed, especially those that clearly define expectancy, describe those in a preventive setting (both qualitatively and quantitatively) and examine their influence on the process and result of DMHIs using validated instruments and appropriate control conditions.

In Article 6, an IPD approach was used to investigate the differential effects of web-based interventions for insomnia in the reduction of depression symptom severity. I managed to replicate findings suggesting the efficacy of sleep intervention as an indirect approach on depressive symptom severity but not on potential MDD onset. Additionally, the study showed that this approach works for all levels of baseline depression severity and across a wide range of sociodemographic, clinical and work-related variables. Due to the available IPD the results are restricted to a working population with high work stress and insomnia.

Summarizing, all studies supported the idea of using DMHIs for the treatment of subthreshold depressive symptoms in different population (i.e. general population, farmers, employees) and were generally in line with prior research. The results from the RCT add evidence for a new digital technology (i.e. stand alone telephone interventions) in the field of depression prevention. Following the general trend, the importance of finding new ways to personalize preventive offers and target them to the individual situation and needs of participants was noticeble across all studies. Effects on potential depressive onset in the included studies remain unclear, mostly given the limited number of events in relatively small samples, limiting the statistical power. Several such aspects that need to be considered when studying or implementing DMHIs, are discussed in the next sections to create a roadmap for using DMHIs to enhance depression prevention in routine care.

## 9.2 LIMITATIONS OF THE CURRENT RESEARCH

The studies included in the current work are not without their limitations. While the limitations of each study are discussed in the associated section of the individual articles, overarching concerns are addressed in this section.

### 9.2.1 *Disentangling prevention and (early) treatment*

Baseline assessment of depressive symptoms varied across the TEC-A trials and the trials used in the secondary analyses (Articles 5 & 6). Besides the insomnia trials, all original trials were indicated

prevention based on elevated depression scores and the insomnia trials targeted individuals with high work stress, which arguably, could be an indication for depression prevention as well.

In the PREV-DEP I and II trials [1, 2, 3] and one of the studies included in Article 6 [4], absence and onset of MDD were established using diagnostic interviews which should be considered the gold standard. In the PREV-DEP trials, individuals fulfilling the criteria for MDD in the interview were excluded, clearly defining all participants as subthreshold depression cases. In all other trials cut-offs on validated self-report scales were used to assess at least subthreshold depression symptoms but none applied an upper cut-off on depression. This procedure mimics routine care prevention where depression scores or diagnostic interviews are not commonly used to exclude interested individuals from parttaking in an intervention. However, that does not allow researchers to control for inclusion of individuals with clinically relevant levels of depression or already diagnosable MDD. Cuijpers et al. [5] reported that when openly recruiting for indicated prevention in the general population 39% fulfilled DSM-IV criteria in a diagnostic interview. Similarly, in TEC-A 28% of participants reported clinically relevant levels of depressive symptoms on the QIDS-SR16 at baseline. This has to be considered when comparing effects with studies more closely defining "subthreshold depression" as a range of depression scores or where MDD is ruled out via interview [6], given that higher baseline depressive symptom severity offers more room for improvement and often lines up with greater symptom reduction [7, 8].

Similarly, the handling of individuals with suicidal thoughts differed across studies included in this work. All individual studies used the suicidality item of the Beck Depression Inventory II (BDI-II) [9] to address suicidal thoughts because it distinguishes between different levels of risk for suicide attempts [10]. While in TEC-A all levels of suicidality were included (provided that participants could distance themselves from suicidal ideas at the moment), the other studies excluded participants with an item score > 1. Different approaches in handling suicidality and exclusion of individuals with suicidal thoughts might diminish representativeness of included participants, given that suicidality is associated with higher depression severity [11, 12]. Neither severe baseline depressive symptoms nor suicidal thoughts should be a reason to exclude individuals from internet interventions for depression per se [13, 7]. However, experts highlighted the need to consider human support and the consultation of local ethics committees, and advocate for including elements addressing suicidal thoughts in internet interventions for depression [13]. Special attention should also be paid in less established treatment approaches like indirect prevention and treatment, considered that e.g. insomnia on its own is associated with suicidality [14, 15].

Lastly, all studies reported on prior use of psychotherapy and/or health training and some on prior MDD episodes, but prior experiences were not considered further in the web-based intervention and handling in the telephone coaching is unknown. Therefore, no clear distinction between prevention of first onset and relapse/reoccurrence prevention could be made. This might however be crucial to the design of intervention given that Bücker et al. [16] observed that participation in another online intervention after initial participation in one did not result in any further effects compared to a treatment as usual control group.

Enrolling a mix of individuals with subthreshold and clinically relevant levels of depression in one sample might represent a more realistically care scenario, but it raises concerns both for research and practice. From a research perspective it enhances the problem of comparability as mentioned before, but it also leads to a power problem when trying to investigate subgroups of differing baseline severity. Thus if not combined in an IPD, trials with mixed-severity samples are unlikely to be large enough and have sufficient power to inform about the effectiveness for different baseline severity (for further discussion see section 9.2.3.1). Researchers should therefore carefully consider if they are interested in a more naturalistic care scenario or if the research question is in need of a specific symptom severity.

In practice that simultaneously raises the question whether and how different levels of symptoms need to be addressed differently by interventions. Guidelines for practitioners generally recommend DMHIs primarily for prevention and treatment of mild conditions [17, 18, 19]. However, evidence is good that more severe cases do profit at least to the same degree as milder cases from internet interventions [7, 8] and results from standardized telephone CBT for MDD seem promising as well [20]. More research on direct comparisons and concepts for stepped care approaches are warranted and are discussed under the practical implications in section 9.4.1.3.

Concerning depressive symptom reduction, a general critique point of indicated prevention is that if symptoms are present, it is not prevention but early treatment in a prodromal phase of depression, which is even more valid when not controlling for current MDD at enrollment. However, this would narrow the term prevention down to universal and selective prevention in only those individuals receiving it that have absolutely no symptoms to start with [21]. Given that more people reached by preventive interventions have or will develop subthreshold depressive symptoms than they will ever fulfill the diagnostic criteria for MDD, intervention efforts to reduce subthreshold symptoms are highly warranted given the associated burden [22]. While researchers might be especially interested in the distal events of non-occurring MDD onset, participants in prevention

interventions might be more interested in more proximal improvement of current problems, given that this is one of the main factors assumed for motivation to join a prevention intervention [23].

### 9.2.2 *Measuring depression*

When planning to assess depression in clinical trials – either by self-report or by clinician report – a vast amount of available instruments are available to do so, adding further comparability problems to the one descried in the last section. For example, in TEC-A, three different depression measures were used: PHQ-9 for short initial screening and comparability with a parallel running study on web-based interventions recruiting from the same population [24], the suicide item from the BDI-II to access degree of suicidality and the QIDS-SR16 as primary outcome. All instruments have their own validated severity categories, which nevertheless do not necessarily imply the same severity [25] and cut-offs for clinically relevant cases. The CES-D on the other hand, used in the other articles, has no validated categorical description of severity but a multitude of potential cut-off scores to indicate clinically relevant cases of MDD [26, 27] with a cut-off of 16 [28] or 22 [29, 17] most commonly used in Germany.

One proposed solution for the heterogeneity in depression measures is the use of common metrics [25], a standardized metric which is informed by an item base of 11 commonly used depression questionnaires. The common metrics cover the full range of depression severity which thresholds for mild, moderate and severe depression [25]. However not all instruments are included in the common metrics (e.g. the QIDS-SR16 is missing) and cut-offs vary slightly across established thresholds. For example, based on common metrics, the TEC-A inclusion criterion of PHQ $\geq$ 5 still would be considered slightly below the cut-off of mild depression and in the realm of the general population mean while the criterion CES-D $\geq$ 16 used in all other studies is just above the cut-off of mild depression, even though both studies aimed to include participants with elevated depressive symptoms. Reaching a consensus on which instruments and cut-off to use for prevention research would facilitate comparability across studies and future evidence synthesis (see also section 9.4.4.2 for a discussion of core outcome sets (COS) as a possible solution).

The multitude of instruments and the (arbitrary) cut-offs for severity and diagnosis also point to larger underlying problems in the way depression is conceptualized and measured. First, it highlights a persisting effort to categorize MDD dichotomously, even though considering depressive symptoms on a continuous level or as different phenotypes might be more appropriate and helpful, when trying to develop more personalized and effective care [30, 31]. Second, it raises the question how and why these measurements differ. In a

critical appraisal of depression measurements, Fried et al. [32] report that the overlap of which symptoms assessed in the different self- or clinician reported measurements has been shown to be low. Similarly, overlap of symptoms assessed and symptoms defined in the DSM-5 was moderate and some aspects identified as relevant to patients, caregivers and healthcare professionals were not adequately covered at all. However, most of these study discuss the meassurment in context of MDD patients and research rarely considers which instruments might be best to assess subthreshold depressive symptoms.

The consequences of different operationalizations became especially obvious in the discussion on literature on insomnia interventions and depression in Article 6. Given that most depression measurements include some kind of assessment of sleep problems, different measurements might result in different depression scores depending on how they include sleep. For example, when one sleep item (range 0 – 3) is included in the PHQ-9, which has total score between 0 – 27, and one sleep item (range 0 – 3) on the CES-D, which has a total score between 0 – 60, changes in insomnia will affect the CES-D sum score less than the PHQ-9 one, which has to be considered in interpretation and comparisons. Practical help to reflect on measurements for research purpose can be found by Flake and Fried [33], who provide six questions which should be considered when choosing and reporting on measurements, including the question how the construct of interest relates to the measurement used. While these questions might help to make more informed and conscious decisions about measurements in RCTs in the short term, the question of how depression is conceptualized remains absolutely crucial to consider in prevention research in a long-term perspective.

### 9.2.3 *Methodological considerations*

#### 9.2.3.1 *Statistical power*

Only one study included in this dissertation, the TEC-A trial, used an a-priori power calculation to determine sample size. The power calculation was based on an estimated effect sizes of interest of around d = 0.35 previously found in a meta-analysis of psychological interventions in subthreshold depression [34]. The secondary data analysis was based on the availability of data rather than a-priori power calculations. At time of study planning for Article 6, less than half of the identified studies had published aggregated data on relevant outcomes and covariates which could have been used to conduct a power and recommendations for power calculations were still limited [35].

Power calculations for prediction models or moderation analyses based on IPD are challenging because many factors such as different sample sizes, variability in treatment effects, and the distribution of

covariates need to be considered [35, 36]. Further power calculations would need to be conducted for continuous and binary outcomes separately as well as for each individual predictor, which hinders exploratory analysis considering multiple moderators. However, this point limits the reliability of the results of the analysis. Given the overall exploratory nature of the secondary analyses this might be justifiable but, as concluded in both Articles, the results should be replicated in dedicated study for which the effort of a power calculation should be considered.

The topic of statistical power might be especially important when investigating potential MDD onset. The incidence in indicated prevention is probably highest compared to universal or selective prevention [37] and has the best chance to evaluate effects on MDD onset in well-sized and feasible RCTs, as demonstrated by the PREV-DEP I trial [1, 2]. This is however not automatically given in studies powered to detect group differences in depressive symptom severity as these (depending on the assumed effect sizes) need fewer participants than survival analysis. The concerns about achieving a sufficient sample size is especially present in mixed-severity samples. Given that for the evaluation of onset of potential depression, individuals already fulfilling the criteria for depression at baseline should be excluded, the sample size can drastically be reduced by 30-40% [5, Article 2] in studies without upper cut-off on depressive symptom severity at baseline and thus lead to reduced power and inconclusive results. Consequently, trials enrolling mixed severity samples would need to be substantially increased in sample size in order to allow for statistically sound subgroup analyses, or should refrain from analyzing MDD onset. This might be even more true for studies focusing on indirect effects in individual with comorbid problematics, where baseline symptoms might be high in most of included individuals such as in Article 6.

Given that it has historically been debated *if* MDD onset can actually be prevented [38], researchers should be especially careful to present results on MDD onset as underpowered secondary outcomes or in very short time frames in trials not specifically designed for that research question. Rather, they should adequately reason and report depressive symptom severity change and support synthesizing projects such as IPDs [39] – which at least in theory have more potential to deliver more robust results.

#### 9.2.3.2 *Testing of multiple outcomes*

Besides their primary research question targeting depressive symptom severity, which determined the design of the study, all of the studies involved multiple secondary outcomes, either including scales for other mental health problems or using different metrics for depression (e.g. MDD onset, RCI). Investigating multiple research questions

within the same larger trial is reasonable with regard to available resources to organize such trials [40], but it raises the question if and how researchers should handle testing multiple outcomes in these circumstances in order to avoid an increase in the likelihood of obtaining at least statistically significant result by chance alone (i.e. "Type I error accumulation").

A review and survey on publicly funded RCTs in United Kingdom highlights the lack of consensus regarding this topic in clinical trialists [41]. In that survey of 27 (of 47 invited) registered clinical trial units, just over half of statisticians 16/27 (59%) would consider adjusting when two primary outcomes are tested and the trial hypotheses required that at least one null hypothesis would need to be rejected. Only 3/27 (11%) would adjust for multiple testing in secondary outcomes, however that was influenced by the type of analysis (e.g. efficacy, cost-effectiveness). Most often in the discussion about whether or not to adjust for multiple testing in clinical trials is the distinction if the analysis is confirmatory or exploratory in nature [40, 42, 41].

In confirmatory analysis, a limited number of pre-defined hypotheses based on a priori knowledge or theoretical predictions are tested and results are usually used to guide decision making. Therefore, Type I error rate (false positives) should be avoided. If all multiple outcomes are considered confirmatory, adjustment is recommended. In exploratory analysis, however, research questions are more open-ended and usually used to generate new research ideas and hypotheses and should inform future studies testing potentially generated hypothesis in a confirmatory setting. Therefore, adjusting for multiple testing is considered less important [40, 42]. An overview of possible scenarios (including multiple treatments and multiple assessment time points) and advice for on handling multiple testing can be found in Li et al. [42].

Adjusting for multiple testing was handled independently and differently in the articles included in this dissertation. In TEC-A, no adjustment of alpha level was done when testing group differences in depression presented in different ways (RCI, close-to-symptom-free status) or in other mental health outcomes (questionnaire battery including stress, anxiety, sleep, quality of life and more). The decision to not adjust for multiple testing was based on the assumption that the additional mental health outcomes in TEC-A were exploratory in nature, given that the whole trial design was based on the aim to evaluate the telephone coaching as an intervention to reduce depressive symptom severity. In Article 5, where all models evaluated the predictive value of outcome expectancy on different depressive outcomes (i.e. symptom severity, close-to-symptom-free status and MDD onset), Bonferroni-Holm correction [43] was applied. Similarly, in Article 6 where all outcomes were based on the CES-D, the Bonferroni [44]

correction was used. The more conservative Bonferroni correction was used in this Article to align the group comparison to the correction in the model-based decision-tree, where it is a crucial part of the of the method to determine meaningful sub-groups (see section 2.3.1.1).

After revisiting the literature, adjusting for multiple outcomes in the secondary data analyses, seems overly conservative and cautious. The reasons for adjusting were based on the fact that one variable (outcome expectancy or CES-D scores in different presentations) was repeatedly used in the analysis. However, as Sjölander et al. [45] point out, this logic is limited because it would be neither reasonable nor feasible to adjust for all tests ever conducted with this outcome and analysis of the same outcome in different analyses should rather be treated as sensitivity analyses and not necessarily include adjustment for multiplicity [42]. The way of reporting both adjusted and unadjusted p-values, together with the CIs might have therefore be a good solution to give the readers the best information basis to interpret the results.

Considering the pragmatic trial TEC-A is more difficult. While we maintained the preregistered argument, that all other outcomes reported besides the primary outcome of depressive symptom severity are non-confirmative and exploratory, this approach does not take into account the nature of this pragmatic RCT and the broader context. The trial is embedded in a larger, parallel implementation program [46] and was meant to inform decisions within the implementation in a timely manner. Therefore, it was naïve to believe that any real effort will be made in the foreseeable future to further evaluate or replicate the reported results and that these are the results guiding implementation and dissemination in routine care. This makes it highly likely that the range of secondary outcomes might also be interpreted as confirmatory in nature, especially by stakeholders involved that are not formally trained in statistics. Under these circumstances it might have been more responsible to adjust for multiple testing and should highlight the need to consider the broader circumstances in study planning.

#### 9.2.3.3 *Control conditions*

In order to determine effect sizes for power calculations or compare effects across study, the control condition in RCTs is an important factor to discuss. A multitude on possible control condition exist including active comparators in form of different interventions, minimal interventions, psychological placebos and no treatment. However, most commonly waiting lists, in which the control group receives the studied intervention after post-treatment assessment or TAU, which allows participants to use every treatment available in routine care, are used in research on depression prevention. TAU, however, can encompass a range of different things from highly specialized mental

health care to no intervention at all [47, 48]. Further, low intensity TAU and waiting list controls were shown to yield similar treatment effects in a meta-analysis on depression psychotherapy [49]. Waiting list controls should be regarded with special caution, because they have been suggested in one network meta-analysis to potentially be harmful to participants because they knew they were not receiving treatment at the time ("nocebo effect") [50].

Unlike in pharmacology and psychotherapy research, pill placebos are not commonly used in preventive trials, given that recommendations for subthreshold depression do not include medication [19, 17]. Psychological placebos, which equal number and frequency of sessions as in the intervention but without therapeutically active methods are hard to realize and ethically questionable because of their deceptive nature [51, 52, 53]. For prevention purposes, education on broader health topics, as for example used by Christensen et al. [54] could pose an alternative. For confirmatory trials informing decision about routine care use, effectiveness compared to active control(s) should be considered standard in therapy research [52] but might be difficult to determine in preventive context where low-intensity TAU might be the reality.

Concerning effect sizes, waiting-list controls and TAU consistently yield the largest treatment effects in meta-analyses studying control conditions of psychotherapy trials [52, 55]. (Online) psychoeducation as a possible minimally active comparator has been shown to be efficacious [56, 57, 58] with small effect sizes around d = 0.20 [95%–CI: 0.01, 0.40] [56] and is therefore likely to reveal smaller treatment effects compared to TAU or waiting list if used as a control group [50].

In the TEC-A trial and the trials used in the secondary analyses either waiting list control (waiting time 8 weeks to 3 months) or psychoeducation were used, both with no restriction to primary care use. Psychoeducation was used in TEC-A and PREV-DEP I where it was deemed unethical to let individuals with subthreshold depression wait for at least 6 months without any support. Psychoeducation in TEC-A encompassed a one-time e-mailed pdf by the study team that contained information about stress, risk factors for psychological disorders, information about depression (symptoms) and definitions of prevention, treatment and relapse prevention and how to find a therapy place or self-help offers. More concretely, this meant a link to an online available audiobook on stress management at work [59], a link to the search website for (in-person) prevention courses funded by general healthcare (ZPP, see section 9.3.1) and some tips to regulate stress in everyday life (e.g. learning relaxation techniques, taking breaks). In PREV-DEP I, psychoeducation was delivered on the same platform as the intervention *GET.ON Mood Enhancer*, and informed participants about depression symptoms, evidence-based treatments and potential help services. Not restricting TAU and additionally pro-

viding information on how to receive further help might be especially important for trials such as TEC-A, where no limit on depressive severity at baseline was posed. If the premise of studies using DMHIs for depression prevention is at least partly to overcome barriers to care and underrecognition and undertreatment in routine care (see section 9.3.1 for a discussion of current options and limitations in routine care prevention), comparison against TAU alone seems almost cynical and should be discussed in the interpretation as comparing to no treatment.

To the best of my knowledge, no study exists that investigated the effect of providing additional information about routine care offers to TAU which might increase uptake of these interventions. Similarly, not much is known on how giving information on (mostly in-person) interventions is perceived by individuals signing up for DMHIs and if these might have demotivating effects. This however underlines the need to define and monitor TAU more closely in clinical trials [48, 52, 55, 49]. While I reported on TAU use in Articles 2 & 3, given that potentially both groups could receive some motivation to use (further) routine care services, the effect of providing this information as a control condition could not clearly be established.

When investigating mechanisms of treatment effects, different considerations should be made when choosing a control condition. Where common factors such as outcome expectancy or working alliance are of interest, a psychological placebo would be highly warranted and might be more ethically appropriate in prevention intervention than in treatment due to an overall lower risk [52]. If specific components (e.g. behavioral activation, worrying or sleep hygiene) should be tested, especially highly modular web-based intervention allow to compare different components or sequences thereof to each other in a factorial design without an inactive control group [60, 61, 62]. Alternatively, a stepped-up intervention with varying "doses" of the intervention could be an alternative.

### 9.2.4 *Considering digital health literacy*

In hindsight, a shortcoming in all studies is the lack of any measurement of health or technology related literacy relevant to the preventive context. Van Kessel et al. [63] summarize *health literacy* as the skills to gain, understand, and use health-related information to make informed health decisions and *digital literacy* as the cognitive and technological skills to gain, evaluate, and communicate information from digital sources. The authors enhance these to the concept of *digital health literacy* to further encompass the ability to appraise the trustworthiness of remotely received and shared health information, and to transfer digital information into different settings.

Fostering (digital) health literacy as part of universal prevention might help increase mental health literacy and decrease stigma around mental health problems which might have a trickle-down effect on further prevention and treatment. Raising awareness for risk factors associated with mental health problems in at-risk groups could increase uptake in selective prevention offers and enable individuals to act earlier on first signs of the disorder could bringing them to use indicated prevention offers or seek treatment earlier [37]. This claim is however difficult to test under real world conditions.

As Koh et al. [64] discuss, it is not enough to offer a digital health intervention for health promotion or in that fact prevention, but it is equally important to close the digital divide in technology access, enable individuals to use technologies and create a supportive and trustworthy environment for their use. At the moment, both digital and health literacy vary highly across countries and cultural groups with low literacy levels correlating with older age, lower education, lower socioeconomic status and poorer health [65, 66, 67]. Additionally, males might have slightly lower scores for health literacy [68]. Different levels of health and/or digital literacy might affect an individual's ability to make an informed choice when using a DMHI in routine care or taking part in treatment studies and should be considered in future study planning and conduct.

Given their potential influence on health related behavior, consistently assessing (digital) health literacy in the studies discussed in this dissertation could have shed some more light on uptake of DMHIs or would have helped to critically discuss the consequences of including mostly female, highly educated individuals from high-income countries in RCTs, especially in the studies on web-based interventions in employees (Article 6).

Especially with regard to farmer's mental health, researchers and practitioners alike should be careful in blindly adapting the stance that farmers have especially low health literacy based on international studies from other continents [69, 70, 71], which usually is not compared to health literacy in other occupations or the general population. This just might add to a stereotypical view of a very heterogeneous occupational group that already suffers from perceived bad image and lower status in society (Article 3). Instead, measuring these variables would enable more informed choices in intervention design and communication. For the ongoing implementation it would be interesting to know whether health or digital literacy influences the uptake of the different interventions offered. For example technical difficulties were barely mentioned in the interview study with telephone coaching participants (Article 4), but were perceived as a relevant barrier in the interviews on online interventions in the same occupational group [72, 73] which might affect uptake in routine care, where both interventions are offered in parallel [46].

In the study evaluating outcome expectancy (Article 5) as a predictor of depression treatment outcomes, (digital) health literacy would have been an interesting construct in addition to expectancy. Having higher digital health literacy might lead to more realistic or positive outcome expectancy in studies of web-based interventions. For example, in a study by Wilhelm & Euteneuer [74], health literacy moderated the reported outcome expectancy in patients with occasional headaches when evaluating one of two different treatment options (conventional vs. homeopathic) for headaches in vignettes [74]. It might also have allowed for an discussion on how informed participants might have been when making their choice to participate in the study and how they evaluate the different groups (web-based intervention vs. web psychoeducation) in PREV-DEP I, which could have brought interesting insights for future study designs. In Article 6 presumably high health literacy might have played a role in picking up an web-based sleep intervention with comorbid stress and depressive problems.

## 9.3 IMPLICATIONS FOR PRACTICE

### 9.3.1 *Implementation never ends*

Slow implementation in routine care and low uptake of DMHI was one of the challenges addressed in the introduction and it remains a challenge to date. The interventions examined in the different studies in this work — namely the personalized telephone coaching by IVPNetworks and the two GET.ON web-based interventions *Recovery* and *Mood Enhancer* — are all available in routine care at the moment, however in three different ways.

The telephone coaching by IVPNetworks is routinely available for individuals insured with the SVLFG, who finances the service for their clients and similar agreements seem to be in place with other health care insurers according to the company's website [75]. Similarly, *GET.ON Mood Enhancer* is available as a preventive self-pay service which is covered by some cooperating health insurance providers [76]. *GET.ON Recovery* on the other hand has become a certified *DiGA* ("Digitale Gesundheitsanwendung", Engl. Digital Health Application) [77]. DiGas must adhere to strict federal regulations (e.g. concerning data security, efficacy proof when compared to a control group) controlled by the German Federal Agency for Drugs and Medical Devices (BfArM) [78, 79]. They are considered a medical product and can be prescribed by all doctors and psychotherapists based on a current diagnosis [78, 79]. In this case they are covered by all statutory health

care insurances in Germany but they are explicitly *not* meant to be preventive interventions [79].[1]

None of the aforementioned interventions are registered with the centralized prevention course registry (German: Zentrale Prüfstelle Prävention, ZPP) [80], which inspects all courses or intervention concepts for preventive interventions which are then subsidized or fully covered by all health care insurances in Germany [18]. However, the course registry has to be considered more universal than indicated prevention and explicitly focuses on promoting movement, healthy diet, stress- and resource management and drug use and not directly on specific psychological disorders.

How far general "availability in routine care" can be used to describe "successful implementation" is difficult. According to the RE-AIM model [81], which postulates the five stages for implementation research of reach, efficacy, adaptation, implementation and maintenance, the "maintenance" dimension suggests that new health promoting interventions need to be actively cared for even after initial implementation. While an initial 6+ months period after study-funding is suggested for evaluating maintenance, there is not really one defined criterion or time point when an intervention can be considered "implemented" once and for all (and it is questionable if that should be the goal for all implementation efforts). However, availability in a prevention context still largely depends on a person's specific health care insurance and their individual contracts with providers of DMHIs or the ability and motivation to use self-pay services. This limits generalizability of information available on DMHIs and makes it very complex for service providers or general practitioners to recommend or refer help seeking individuals to DMHIs. Access to these preventive DMHIs is not yet as low-threshold as envisioned by experts in the field or observed in other countries [82, 83, 84, 62, 85]. Thus, it would still be far-fetched to consider DMHIs as part of routine care in Germany.

Participants in the individual studies included in this dissertation were largely recruited via active outreach by health care insurances or employers and retention was monitored. For example, in the telephone coaching RCT, over 90% of participants in the initial screening reported that an individual invitation letter drew them to seek more information about the study and the telephone coaching (prior to implementation) [86] and only a handful reported counselling from insurer or the members magazine as information source.

In routine care on the other hand, individuals usually need to actively seek out information on preventive interventions or DiGas [87] on their own initiative and much less is known about uptake of the interventions. The results expected from the "With us in balance" implementation study [46] will allow some comparison between the

1 This is however only a momentary snapshot at time of writing this dissertation and might change quickly given the rapidly developing market for DMHIs.

actively recruited effectiveness study and routine care distribution in terms of which individuals are reached for a study or a routine care preventive offer and which ways of informing individuals can effectively be used in routine care. More information like retention rate and cost of recruiting in both settings should be made available to researchers conducting evaluation and implementation research.

For the development of further implementation of DMHIs in routine care it can be considered a good signal that both the new guideline for prevention [18] as well as the newest treatment guidelines for depression in Germany [17] promote and recommend digital and low-intensity psychological interventions. It remains, however, unclear who is supposed to delivered these interventions or guide individuals towards using them, given that health care resources are already limited [88]. The requirements for a ZPP certification in the field of stress- and resource management (at least pre-registered, one group open-label study including three measurements of perceived stress or resilience and the presumed mechanism of change; availability of highly qualified staff for user contact within 48h [18]) are beyond the scope of what individual trainers or coaches can develop and maintain. These regulations seem to be easier to fulfill by already commercialized interventions, but companies might be more interested in being registered as a medical product in Germany than in prevention because it could be more economical due to greater uptake in treatment and better reimbursement options. Streamlining processes for DiGa and ZPP certification by for example sharing criteria for mild symptom levels or accepting the same efficacy proofs if analyzed across different symptom levels [89, Article 6] might make the certification in prevention more attractive.

Additionally, important stakeholders, such as health care insurers, practitioners and policy makers are still skeptical of the potential of DMHIs but are generally more open for an application in preventive use, for mild cases or in blended formats [90, 62]. More initial training and on-job training for healthcare workers across disciplines, easier accessibility and reimbursement of activities related to DMHIs, have been identified as important next steps for a wide-spread use of DMHIs [91, 62, 92].

Further, a central and trusted information source about DMHIs is necessary to easily inform potential users as well as healthcare workers about the options of different DMHIs and support informed decision making. This should include information about the effectiveness of DMHIs but also on quality criteria to help make an informed decision on their uses for example in form of quality badges or rating scales [93, 94, 95]. This could also help to distinguish evidence-based and scientifically evaluated DMHIs as the ones discussed in this dissertation from more readily available programs available over the app-stores which might be less evidence-based but available to users on the spot [94, 96].

Research should pay attention to how different offers are advertised and perceived by help-seeking individuals. Potential dangers may arise from rivaling non-evidence-based and unregulated interventions when these are the more easily accessible ones but might not fulfill the individual needs and reduce trust, readiness or outcome expectancy for DMHIs in general.

### 9.3.2 *The role of coaching interventions*

The situation of digital interventions in the German health care system described in the last section is especially critical for telephone coaching interventions. Currently, no form of in-person or remote one on one coaching is, generally speaking, covered by health care insurances in Germany (besides exclusive contracts) and it is not applicable for consideration as a DMHIs under regulations of the ZPP or as a DiGa.

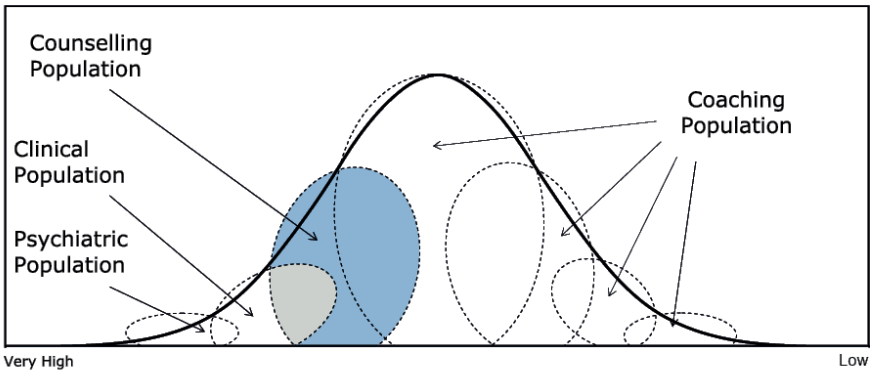
However, the results from the RCT and the qualitative interviews indicate that telephone coaching can be a valuable contribution to the preventive landscape in reducing depression burden in a population with mostly subthreshold depression. The telephone coaching in this dissertation combined a high degree of personalization and synchronous human contact with most of the convenience and perceived anonymity of web-based interventions. This raises the question where telephone coaching (not only the one intervention evaluated in this dissertation but in more general terms or even using different media for example video calls) could be envisioned in the German health care system as a preventive intervention for depression and what its unique contribution could be.

In an approach to distinguish coaching activities from psychological counselling and therapy, Grant and Green [97] tried to operationalize the populations for coaching vs. clinical practice based on the psychopathology in the general population with coaching covering the large part of non-clinical cases (Figure 1)[2].

Using this graph as an example, a (telephone) coaching aiming at the upper end of the non-clinical spectrum (highlighted in blue) is likely to represent an indicated prevention approach for individuals with elevated depressive symptoms. One could argue, even without more information about the cost of such interventions, that a one-on-one coaching in the lower (right) part of the graph, which would represent universal prevention, is unlikely to be (cost-)effective. However, in the realm of indicated prevention coaching might be in a

2 In some contexts coaching and counselling might be used interchangeably while counselling (psychologist) can also imply specific post-graduate training in English speaking countries [98, 99]. "Counselling" or in German "psychologische Beratung", is similarly to "Coaching", not a regulated occupation in Germany and is thus not considered to deliver treatment yet. Therefore, in this hypothetical use case I will consider them interchangeable. For a detailed discussion of differences in qualifications and implications for practice see Bachkirova & Baker [100].

Figure 1: Theoretical distribution of psychopathology in the general population as presented by Grant and Green (2018) [97]




*Note*. Reprint kindly permitted by John Wiley & Sons – Books. (see Appendix A). Color-highlights added. Scale label at lower end adapted to better represent original meaning.

unique position to not only reduce depression burden in this population but also pay special attention to individuals who might be on the threshold for need of more clinical care (highlighted in grey in Figure 1). This is of extra importance given that indicated prevention as mentioned before [5], but also coaching as an intervention in itself [101], attracts individuals with a clinical level of mental health symptoms. The latter might be because "coaching" in comparison to "therapy" might be more socially acceptable [101], which in itself could be discussed as an indirect approach as well.

Following the idea of a stepped-care-approach, a coach in a one-on-one setting could continuously adapt the intensity of intervention according to assessed needs and could facilitate transfer into treatment (e.g. help participants find therapy place and use the waiting time to work on behavioral activation), thus addressing the fact that a significant part of individuals attracted to preventive offers/coaching might already present clinically relevant depression symptoms. In addition, this kind of intervention might be easily adaptable to act as after-care and relapse prevention or improve mental health other than depression.

In reality, even though the possibility to referral (e.g. to socio-economic counselling or change to on-site services) was already given as part of the telephone coaching program evaluated in the dissertation and has since then been under constant development within the model project "With us in balance" and its implementation, to date not enough data is available to clearly evaluate the feasibility nor effectiveness of such an approach. Building such professional networks and integrating different parts of the healthcare system and psychosocial services remains challenging [102].

Further, if such a coaching explicitly aims to operate in the vague border area of indicated prevention and early treatment, coaches would need to identify and support individuals with care needs beyond a preventive scope, which needs to be reflected in coach qualification and treatment guidelines [100]. In that case coaches need to be qualified to assess these needs correctly (i.e. have diagnostic knowledge or access to a diagnostic service) and guidelines need to define what happens in case of enhanced needs and how coaches could refer to specialized clinical care. Therefore, such an intervention would be in need of new health care-policies as well as re-imbursement regulations [103].

The dilemma with (telephone) coaching at this point is the overall lack and diffusion of definition, qualification and evidence- base. Coaching, since its beginning rise in popularity with the turn of the last century, has had a problem with defining and distinguishing itself from counselling/therapy and (as a consequence) to synthesize evidence for it [104]. As Smoktunowicz et al. [105] point out, "CBT" — even though it is also an umbrella term for a wide range of techniques — has become a widely accepted term to describe a certain kind of therapeutically process, theoretic foundation and assumed evidence. While this kind of simplification might create its own problems, nothing similar to such an agreed on approach exists in the realm of "coaching", underlining the inconsistencies within the field and the lack of professional organization [104, 106]. While research on coaching increased and developed [107], especially in the field of executive coaching with a recent meta-analysis summarizing 37 RCTs [108], less development can be observed in the field of mental health [109]. As coaching is most often focused around patient-centered goals [109, 97], one possible beginning could be to closer analyse what kind of goals are routinely set in preventive coaching, with which psychological methods these are addressed and which might be directly (e.g. improving self-worth) or indirectly (e.g. resolving work conflicts) be related to depression.

These definition problems make it hard to argue for a dedicated space for telephone coaching in the German health care system at the moment – regardless of all the potential benefits. It is nevertheless almost astonishing that the question has to be raised which role telephone coaching can play in a health care system, when one of the largest systematically implemented services for improving mental health care worldwide has telephone interventions as one of its core elements: the Improving Access to Psychological Therapy (IAPT) program in the United Kingdom.

The IAPT program was introduced in the United Kingdom in 2008 with the aim of increasing access to evidence-based psychological therapies for individuals with common mental health problems, such as depression and anxiety. The main premises of the IATP are a

stepped care approach, highly standardized evidence-based treatments and routine outcome monitoring [110]. The steps include primary care by the general practitioner (step 1), referral to guided CBT self-help for depression and anxiety (step 2) and high-intensity psychological therapies including formal CBT, interpersonal therapy and eye movement desensitization and reprocessing (step 3).

Telephone support is primarily used in assessment and in step-2 intervention and consists of short-term, low-cost and low-intensity CBT with a focus on psychoeducation. Alternatives in this stage include computerized CBT, group CBT, face to face delivery or combinations thereof [110]. No differences between face-to-face and telephone delivery could be observed in terms of effectiveness in this step [111]. Research explicitly focusing on the effectiveness of the telephone support in reducing depression is missing. If these step-2 interventions do not yield improvement, patients are stepped-up to high-intensity therapy. The health care professionals delivering step-2 services do not necessarily need a university degree, but they need to obtain specialized training as a *Psychological Wellbeing Practitioner* (PWP), which is a certified and regulated occupation [112].

After practicing this model for over 10 years in routine care, the IAPT service is estimated to receive around 1.25 million referrals per year and the concept has been adapted by other countries, including France, Australia and Canada [110, 113]. In a recent meta-analysis synthesizing evidence of 10 years' worth of research the effectiveness of the IAPT service as a whole is estimated at $d_{within} = 0.87$ [95%–CI: 0.78, 0.96] for depression measured with the PHQ-9 [110].

The IAPT service is not without criticism. According to Scott [114, 113] a major critique point is that independent research on the IAPT is missing. Further critique points entail, that the IAPT's own research fails to prove the IAPT's added value in comparison to treatment before the service was introduced, that there is a lack in formal diagnoses, that the evaluation of extremely heterogeneous samples hinders comparison and that there is a general shortcoming to address service drop-outs and treatment fidelity.

Obviously, the IAPT telephone support program in step-2 differs from the telephone coaching under investigation in Articles 1 – 3 in several aspects, including a dedicated focus on mild-moderate clinical cases of depression and anxiety, so that it is not a preventive intervention. Furthermore, IAPT is focused on standardized CBT with a relatively short timeframe of six times 30 minutes to begin with and has been criticized for being too rigorous in its protocol [115]. However, this example shows that telephone interventions can be implemented on a large scale and fulfill a dedicated function within a health care system.

Thus the "problem" with finding a place for telephone coaching in the German within the health care system has less to do with the

telephone as the digital medium by which it is delivered, but with the practice of coaching itself. Evidence from psychotherapy research has promising results from meta-analysis [116] of showing telephone delivered psychotherapy being non-inferior to face-to-face delivery and single studies (partly derived from the IAPT service) suggesting similar working alliance compared to face-to-face [117, 118] and potentially higher adherence [116, 119, 120]. Similar comparisons could be drawn from psychotherapy research on videoconferencing [121, 122]. Thus, if coaching, in a well-defined, organized and widely available form, could prove its (cost-)effectiveness in depression prevention, the choice of medium for delivery is unlikely to pose a big threat to its success in health care but rather shows potential to enhance it due to its accessibility. To realize such an integration into the health care system, it would need serious collaboration efforts from coaching practitioners, their federations, researchers, policy makers and health care system representatives, but then it has the potential to improve access to mental health care for some of the individuals most in need of professional help.

### 9.3.3 *Considering different approaches to depression prevention*

With increasing availability in routine care, the question of which intervention suits which individual best, becomes more and more urgent. How preventive options are presented to potential users and which formats would fit best for the given use can be important aspects when considering improving uptake and use of DMHIs. With the explored intervention in this dissertation we can try to filling this statement with recommendations.

#### 9.3.3.1 *Addressing mental health*

As seen in the intervention discussed in this dissertation, the degree to which depression is directly addressed varied from the PREV-DEP trials, which were openly framed as depression prevention interventions to the sleep interventions which were not promoted with relation to depression at all. When recruiting study participants for TEC-A, more invited individuals initially responded (i.e. clicked on a link for more information) when the invitation letter was framed with a health promotion focus using terms like "enhancing quality of life and well-being" instead of preventive focus with terms such as "preventing depression" [86]. This could suggest that wording and framing can play an important role when trying to enhance help-seeking and uptake of mental health intervention and would be inline with an indirect preventive approach.

On the other hand, Kruzan et al. [123] interviewed young adults with mild-moderate symptoms of depression who used an online available screening tool for depression and describe that a common

reason to look for information online was that they perceived elevated depression symptoms that deviated from the personal norm and that they sought validation from using the screening. The authors identify the moment of looking for information and taking a screening as an important step in an individual's mental health journey, where support might be especially beneficial for uptake of further treatment. In the case that individuals already suspect a specific mental health problem, not adequately naming or labeling it might be perceived negatively or demotivating and hinder them from seeking further help.

For clinical practice that raises the question when to use an indirect approach and when to openly talk about depression from the beginning. When looking for help online, individuals are most likely to already use wording related to the problem that they are most concerned about at that moment in their search terms. Therefore, it might be helpful to highlight different problems (e.g. sleep, worrying, stress, mood) to increase the fit of perceived needs to information and intervention offered to facilitate uptake. It would nevertheless be important to clearly communicate indications and contraindications for any suggested intervention in detail, taking into account different health literacy levels.

Another approach could be screening for common mental health problems, either as a (dynamic) questionnaire as mentioned before, in form of a conversational agents online or at the beginning of a consultation at a general practitioners office. Such a screening could potentially also consider the problem that motivates current help seeking (e.g. worrying or sleep vs. assumed depression diagnosis), treatment preferences (self-help vs. human support) and preferred modality (in-person vs. remote) next to measuring symptoms. Recommendations could then target the most relevant problem with a short explanation on which answers the recommendation is based.

Even though treatment recommendations in early stages of help-seeking might profit from focusing on the problem reported to be most urgent by an individual, individuals should not be left in the dark about symptoms of mental health problems or downplaying those. Withholding health information or information about intervention option in favor of a more socially acceptable framing or consideration of financial aspects could be seen as deceptive and unethical [124], violate the principles of informed consent and shared decision making in health care, and might influence the further process of help-seeking and intervention uptake [125, 115]. In fact, qualitative studies investigating whether individuals would like to know their personal risk for developing a depression suggest that individuals would like to be informed about their risk factors which might have motivating qualities to take actions to improve mental health [126, 127]. Receiving information about viable next steps instead of just being informed

about the status quo was perceived as crucial in both screening and risk-assessment [123, 126].

These ethical aspects might be especially important in personal conversations between a help-seeking individual and a health care professional (e.g. coach or general practitioner) or counselling agent (e.g. health care insurance service center). For example, the employees working in the telephone service at the SVLFG are not health care professionals nor employ a screening instrument and have to work with the information disclosed by the caller to inform them about available treatment options in their health insurance plan and motivate uptake of a service. Sometimes these calls can be the first time individuals disclose their problems to another person and this can be crucial for the further process. Obviously, at that point no diagnosis takes place and recommendation for interventions have to be made on the problems described and possibly asking about the assumed reasons for the problems or preferences. For example, an individual describing problems with worrying and falling at sleep at night, could be recommended to start an online training for sleep problems, but should be informed that this can also help with low mood and depressive symptoms without the need to discuss these any further. On the other hand, a caller already suggesting that they experience signs of depression could be recommended an online training for depression or a telephone coaching based on their preference of professional guidance and medium. Considering the results from Article 4, this recommendation could easily be personalized by asking about additional sleep problems which could be treated combined with a web-based intervention for insomnia. In addition, in a situation where only a limited number of preventive services can be offered, a web-based sleep intervention can be a good choice given the potential to reduce multiple complaints[3].

Professionals in the position of recommending and facilitating uptake should be aware of the importance of this step and their role. In a one-on one-consultation on different treatment options, experiences from the IAPT service show that the PWP responsible for assessment and recommendation of further treatment had a substantial effect on further treatment uptake [129]. Additionally, as pointed out by Irvine et al. [125], the way different options are presented (i.e. single choice vs. choice from a selection either parallel or after each other after one option has been rejected) can influence how offers are perceived and reacted to. However, professionals could also quickly find themselves

3 This actually is a current development in the realm of the SVLFG preventive offers. Persons employed at a farm (and which are not the entrepreneur, their spouse or a family member and considered under TEC-A or ImplementIt) do not have access to all interventions in the model project "With us in balance" due to underlying insurance regulations. They only have access to a limited offer of preventive on-site interventions and the web-based intervention *GET.ON Recovery* in a version adapted to the farming population [128].

struggling to balance the available resources, individual preferences and expected effectiveness [125, 115]. This highlights the need for further research in how to best allocate available resources (see further discussion in section 9.3.3.2).

#### 9.3.3.2 *Offering different media*

Besides addressing different problems that are more or less obviously related to depression, the modality of the different interventions that could address these problems could be crucial for actual use and adherence. The interviews in Article 4 highlighted the importance of the telephone conduct for actually use of a mental health service in the farming community. Therefore, choosing and recommending a delivery medium is already an important step in overcoming structural and attitudinal barriers and in personalizing a preventive offer and enhancing potential uptake and use.

In case where mostly structural barriers (e.g. long distance to next offer, reduced mobility) are the main barrier hinder conventional help-seeking but there is a general wish for personal contact or even peer contact, synchronous offers such a telephone or video coaching or virtual group offers might be most appropriate. This might also be the case when problems are not yet well defined and the synchronous dialogue can facilitate problem definition. Increasing convenience of participation by offering remote but very personal intervention might be especially important for preventive services (and studies thereof[4]), where the level of distress or intrinsic motivation for intervention uptake might be lower than for treatment.

When attitudinal barriers are more relevant, highlighting the anonymity, flexibility and self-help character of web-based intervention could be more promising. These aspects are most likely to be true when the intervention is directly accessible from home without consultation of any kind, however, human support might still be crucial in order to support uptake and completion of these interventions.

At the moment, little is known about comparative effects of different technological solutions for the same mental health problems or preventive goals to make an informed recommendation. Two aspects that might be important to consider are the amount of guidance to support adherence and the knowledge on available DMHI formats. Human support has been shown to increase adherence to DMHIs [130, 131] as well as efficacy [132, 131]. For web-based interventions, a study based on a large routine care data set from an Australian

4 Anecdotal side note: Even though rarely used in everyday university contexts, offering the option to send back the informed consent for study participation via fax for the study enrolling farmers was a very well used and extremely important way to facilitate enrollment procedure. This highlights the need to consider different media options — even antiquated ones — to facilitate uptake and use for different populations.

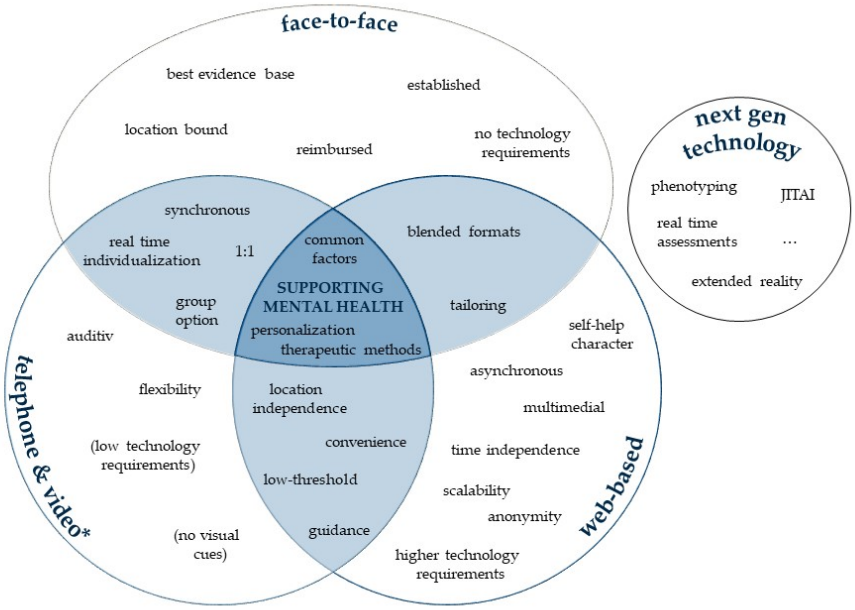
online clinic [133] showed that higher baseline severity, younger age, less education, and being single were negatively associated with uptake, completion and effectiveness of a transdiagnostic intervention targeting depression and anxiety after a first assessment. In a German sample with mild-moderate depressive symptoms [134], age was similarly predictive of adherence to a web-based intervention. The effect of baseline severity on adherence was mostly resolved by severer cases receiving guidance. More research like this is needed in order to inform allocation of resources in prevention for example for closer monitoring, more guidance or stepping-up approaches.

If participants report having tried self-help methods in the past but have struggled sticking to their intended goals, an intervention with more human support (either as a synchronous intervention or with different level of asynchronous guidance) should be recommended. The problem of not wanting to ask for help but rather be able to solve ones' problems independently, might not only be relevant for initial uptake of an intervention but also for adjusting the amount of support within an intervention and adhering to it. At the moment, the result on whether regularly scheduled support or on-demand-support in web-based intervention influence adherence and/or efficacy of interventions differently is still inconclusive [60, 135, 136] and should be discussed with regard to underlying mechanisms [130].

When relying on individuals' preferences for a delivery medium, this might also be influenced by the available knowledge about different DMHIs and thus their digital health literacy. Acceptability of DMHIs is often associated with previous use and experience of such a service, indicating a lack of knowledge upfront [137, 138, 139, 140]. This can hinder outcome expectancy and therefore diminish intervention uptake or its efficacy as discussed in Article 5. While a telephone coaching might be easy to envision (especially when already in a telephone call with a health insurance worker), a web-based intervention might not directly be known or imaginable. Delivering information about DMHIs or offering them should take into account digital health literacy and could profit from more explanations, comparison to better known technology (e.g. chat-functions, online learning platforms) or brief demonstrations and test-phases. Overviews of features of different delivery methods, as exampled in Figure 2, might help guide recommendations in practice.

Offering different media is also a way to support health equity – especially if different low- and high-tech offers are made [141]. Even in generally high-income countries, not every individual has the same technological prerequisites such as the (highspeed) bandwidth to video call or follow psychoeducational videos but rather has to rely on toll-free calls or more static psychoeducational web-pages. The recent COVID-19 pandemic has dramatically shown that technology inequity is often faced by groups already facing other sorts of inequity,

Figure 2: Overview of features associated with different media to deliver psychological interventions for depression prevention




*Note.* *aspects in brackets not applicable to video-conferencing

for example in health or education [142, 143, 144]. Researchers, practitioners and health care providers alike need to be aware of the danger for further inequity when promoting further digitalization of mental health care.

#### 9.3.3.3 *Planning ahead*

The interventions discussed in this dissertation can be all considered "conventional, but digital" following the taxonomy by de Witte et al. [145], meaning that the principal methods of the interventions (e.g. CBT, rules for sleep hygiene, goal-setting) are well-stablished in in-person care but delivered via a different synchronous medium or are presented in a stand-alone web-based program. None of the discussed interventions employed any kind of uniquely therapeutic component, e.g. with using information from smartphone sensors or digital prompts in everyday life (Figure 2).

While the delivery of the interventions like the DMHIs discussed in this dissertation can overcome structural and attitudinal barriers as discussed above and might create valuable opportunities to study mechanisms in those "conventional" DMHIs, so-called "emerging technologies" like sensors [146, 147], digital phenotyping [148, 149] and artificial intelligence [150, 151] might open the door for completely new diagnostic and therapeutic approaches. This is a very promising notion, given that not everyone engaging with conventional DMHIs

(or with psychotherapy for that matter) is likely to profit from it [152], which makes fundamentally different approaches employing unique technological options highly warranted.

The current challenges faced by DMHIs should serve as a cautionary tale to the next generation of DMHIs. It is highly important to foster a positive attitude and build trust towards DMHIs in routine care now, to learn about the individuals using and profiting from them and share what can be learned from implementing these more conventional interventions to inform and support the implementation of next generation technologies.

## 9.4 FUTURE RESEARCH

In this last chapter, I want to address some questions that are not yet answered completely or that occurred during my research for this dissertation. These questions could guide the future research in the field of digital depression prevention.

### 9.4.1 *How much personalization do we need in DMHIs?*

In the course of this dissertation, the idea of adapting interventions to best fit an individual's situation has been a core aspect and was often named "personalization". While traditionally the idea of *personalized medicine* (see section 1.2) entails the choice of treatment being based on an individual's risk profile, Article 6 showed that identifying such risk profiles is not that easy. Articles 2 & 3 highlighted how results from so called personalized telephone coaching are difficult to generalize given their high adaptation to the individual's situation in terms of coaching aim and integration into everyday life. Finally, in Article 4 participants of the telephone coaching described exactly that adaptation to their private and professional situation was a very personal approach.

In this regard, summarizing all kinds of digital interventions under one term of DMHIs really blurs the differences in the options provided. A one-on-one telephone coaching is usually a synchronous interaction with possibility for immediate feedback. In such an approach one might distinguish more between standardized or manualized coaching procedures vs. a free flowing coaching conversation [97, 109]. In web-based programs, personalization could mean that specific modules are pre-selected or that individuals can choose between different pre-programmed content. It could also relate to the human support given and if the support contains additional therapeutic action and has the aim to help adapt the intervention content to the individual's situation or is again more standardized feedback focusing on adherence of the individual and motivating further interaction with the web-based-intervention.

Thus, that term already included a wide range of meanings and different goals. For psychotherapy research (face-to-face and online), a recent meta-analysis summarized that most personalization happens upfront, before intervention start [153] and is more effective both in cases of treatment matching (i.e. interventions matched to subgroups based on hypothesized interactions) and individual tailoring (i.e. based on clinical profile and idiosyncratic case conceptualisations) compared to standardized care. The influence of involvig participants' views in the tailoring seems less studied but might be important to consider in the realm of DMHIs. The following section aims to distinguish the different adaptation option in preventive DMHIs and to create new ideas how to further investigate their effectiveness.

Therfore, I will borrow the terminology used in information and marketing science [154] in the next part. Therefore, I will use *tailoring* as the broadest term to describe the general effort to optimize the fit of an intervention to an individual's personal situation, needs and preferences, including but not limited to personal risk factors. The term *personalization* will be used to describe adaptations made on a systems level (e.g. adaptations made by clinicians or the intervention) which usually refers to the approach of personalized medicine and the adaptation to individual risk factors. *Customization* is then used to describe adaptations made by the individual user or client for example by (de-)selection of an optional model or feedback on the content. Lastly, I will discuss the term *individualization* as a way to describe synchronous interventions entailing parts of both, personalization and customization.

In all cases, the goal of these tailoring approaches should be critically considered in light of the resources needed for the tailoring. Effectiveness might most likely be improved by better targeting individuals according to their risk-profiles and symptoms (personalization). On the other hand, aspects such as satisfaction and adherence might profit from enhancing the individuals' influence on treatment decisions and intervention design (customization).

#### 9.4.1.1 *Personalization in DMHIs*

In the realm of psychiatry and DMHIs, the question of which set of characteristics best predicts ideal treatment outcome is still inconclusive [155, 156, 157] but might be especially relevant for preventive interventions in order to enhance their impact on depression burden [158].

With Article 6 I tried to support the effort to determine which participants might profit most from an indirect prevention and treatment approach across a wide range of potential characteristics. However, only depressive symptom severity at baseline predicted outcomes which were overall similar to what could be expected in the different symptom range of depression. While that proved the concept of using

web-based interventions for insomnia as an indirect approach for both prevention and treatment for depression, no further targeting options could be identified.

So far, only baseline symptom severity and age have been repeatedly identified as potential moderators in (online) psychological interventions for depression [8, 6]. To further enhance precision in intervention recommendation and therefore personalization, different approaches need to be considered, for example including biological risk factors [159] or individual symptoms or identified subtypes that might have higher potential of profiting from specific treatment [157]. For example, Blanken et al. [160] found in a 6-year prospective network analysis that from all sleep measurements only the item assessing difficulties in initiating sleep was predictive of depression onset. Likewise, Leerssen et al. [161] showed that CBT-I interventions targeting insomnia subtypes with heightened risk for depression (including but not limited to difficulties in initiating sleep) prevent worsening of depressive symptoms thorough a year.

Another approach would be to improve the prediction of depression onset based on characteristics beyond the usual baseline questionnaires. The WARN-D project by Fried et al. [162] for example uses technology-based assessments to collect a wide range of phenotype data and real time experience in addition to baseline characteristics to build an early warning system for depression. However, this project will initially reveal potential new predictors or combinations thereof for potential MDD onset which then can help inform new preventive interventions to best address the identified risk factors. Improving prediction of the course of depression as well as improving the targeted addressing of the risk factors and complaints remains the main challenge for depression prevention in the next decade and is the basis for further personalization.

#### 9.4.1.2 *Customization and individualization in DMHIs*

The web-based interventions examined in this dissertation did not involve any further adaptation to individuals' needs after being selected based on their sleep, stress or depression test scores. This is however generally possible in web-based interventions, and a recent study from Andersson et al. [60] showed that if treatment components at treatment start were selected by participants (customization) instead of clinicians (personalization), small but significant effects on treatment outcomes were observed in favor of the customized treatment. This is a hint that participant involvement in the intervention can be an important factor to perceive interventions as fitting and to enhance the effectiveness.

Interventions that rely on a high degree of interaction between user and service provider such as the telephone coaching might present an mixture of both – personalization and customization. The consideration of the individual's background and the one-on-one

setting is likely to be perceived as tailoring or personalization, and coaches are likely to have a general procedure on how they address different topics or identified risk factors. The central coaching practice of mutual goal setting [97] on the other hand already suggests a high degree of customization at the outset of the intervention. For example, if a participant's report included symptoms of anxiety in addition to depressive symptoms, coach and participant could decide to focus on these symptoms first or instead of depressive symptoms. Such an adaptation was not possible in the web-based interventions focusing on depression or sleep as the main target.

Further, given the synchronous interaction, participants can directly give feedback throughout the intervention (e.g. if certain exercises or parts of psychoeducation are already known or are not perceived as fitting for an individual's situation), ask for changes or even adapt the goal throughout the intervention, which can be considered another way of customization. The delivery of such an interactive intervention could best be described as *individualized*[5], implying that it is nearly impossible that two individuals will receive the same intervention even if they work with the same coach. This interactive adjustment might be the reason why the telephone coaching in Article 4 was perceived as very much tailored to the individual's need in all addressed aspects, while a lack of perceived fit was a complaint with the web-based interventions for farmers [72]. This kind of adaptation could also be crucial in cases where individuals already have a lot of prior treatment experience or recent experience with a similar intervention and are less likely to profit from a repetition [16].

While the question of upfront personalization (or customization) can be integrated in research designs relatively easily [153], for example by specifying subgroup analyses or factorial designs, individualized interventions are more difficult to evaluate. Wolever et al. [109], after reviewing 284 health and wellness coaching-related studies, conclude that evidence synthesis in the realm of coaching remains difficult, due to unstandardized definitions and conduct in coaching interventions. They also point out that a highly individualized process does not necessarily hinder standardized evaluation. They name motivational interviewing as an example, which, besides its individual process, has a high degree of standardization, allowing for comparability and for the development of scales for outcomes, coach skills, and coaching fidelity [109]. Similarly, in the field of depression prevention, Bellon et al. [163, 164] established methods to monitor assessment of risk factors and individualized prevention intervention in primary

5 In hindsight, the telephone coaching discussed in this dissertation would better be named "individualized" given that the tailoring is more based on the participants' living circumstances and acute stressors than on underlying risk factors. However, I retained the wording of "personalized telephone coaching" as used in the original articles to not add unnecessarily confusion in wording.

care in Spain by using a pre-specified set of interventions available and analyze general practitioners documentation about their use.

Future studies should compare coaching interventions with a set of pre-selected goals and different degree of individualization. A highly individualized intervention, such as the one in Articles 1 – 3, could then be compared with a customized, modular coaching where topics for each session are pre-planned together in a personal coaching plan and a more standardized version with the same predefined session topic for each participant targeting the same condition (similar to the IAPT). Such an approach would however need to consider if such a standardization and pre-defined goal would still be an option accepted by coaches and seen as in line with their practice [97]. With an adequate description of the methods used in all arms and documentation of time needed from the coach to prepare an deliver the intervention, such a study could help to describe the content of a highly individualized coaching, the resources needed for individualization and show the impact of it on potential outcomes. This would inform the necessary qualifications for coaches and would help to create the much needed evidence-based for coaching in a preventive/mental health context (see section 9.3.2).

#### 9.4.1.3 *Stepped care as a special form of tailoring*

Stepped care approaches can be seen as a special form of personalization, given that clinicians can monitor and react to symptom change during the treatment. Stepped care approaches try to optimize the use of available resources (i.e. clinician time) by matching symptom severity to intervention intensity [165]. Stepped care approaches for the treatment of mild to moderate depression have been shown favorable when compared to TAU in two meta-analyses [166, 165]. However, heterogeneity in both studies was high so no general recommendations can be made. Most notably, the included studies usually had switches between delivery modes (e.g. unguided self-help to group face-to-face intention).

One large-scale study by Mohr et al. [167] compared manualized CBT via telephone with a digital stepped care model (1. step: web-based CBT with intense guidance, 2. step manualized CBT via telephone) in the treatment of MDD and found the two interventions were non-inferior in their efficacy, both before first potential stepping-up after 5 weeks [168] and after the 20 weeks of either telephone CBT or stepped care [167]. While clinician time was halved in the stepped care conditions and thus associated with lower cost and increased scalability, satisfaction was lower in the stepped care arm without affecting treatment completion[167].

In the preventive context, studies on stepped care interventions are much more scarce and to the best of my knowledge do not include DMHIs. One review on stepped care interventions in the prevention

of depression revealed that preventive stepped care was mostly targeted at the elderly and involved watchful-waiting, (bibliography) self-help, CBT, and primary care. The review concluded that there is no general evidence for the effectiveness in depression prevention for stepped care interventions when compared to TAU [169]. However, three of the four studies reviewed were placed in nursing homes and clinics, suggesting that TAU might include relatively active treatment. Therefore, there is an urgent need to investigate stepped care models in indicated preventive outside of nursing homes or clinics and within digital interventions. Such a stepped care model could range from self-help, over adherence focused guidance to therapeutic guidance/blended care to a one-on-one psychological intervention over telephone or video-conferencing (with the option to facilitate primary care or special mental health care as needed as discussed in section 9.3.2). Additionally, opt-in option for different delivery or contact modalities might be a way to further customize such an approach and keep the balance between the possibility of anonymous use and ease of getting more support if needed.

At the moment, a similar range of options is available as part of the implementation project "With us in balance" for German farmers where web-based interventions, telephone coaching and on-site services can be offered for choice in a telephone counselling by the SVLFG (see section 9.3.1). The process of how decisions are made under these circumstances is, however, neither standardized nor studied. Therefore, not enough is known about what individuals with elevated symptoms perceive as the best fit to their symptoms, what preferences for interventions exist, and how the decision is managed and supported by the health care insurance workers. Studying this process with an approach similar to the one used by Irvine et al. in IAPT [125] could inform how tailoring can take place in practice and could inform resources management. Just starting with the least intense and cost-effective intervention without taking into account the individual's situation would still follow the idea of "one size fits all" instead of a tailored preventive approach.

### 9.4.2 *What motivates individuals to start preventive DMHIs and what do they expect?*

Combining results from the interviews in Article 4 and the investigation of outcome expectancy in Article 5 both suggest that not every individual starting an indicated prevention DMHIs has a very clear idea of what kind of outcome they can expect from taking part. The lack of concrete expectations with regard to improvement of psychological symptoms was also observed in a parallel study within the larger project "With us in balance", interviewing farmers using a web-based intervention for the prevention of depression [73]. Motivation

and expectation remain two very important aspects for acceptance and uptake of DMHIs [170] which are worth investigating further in order to find ways to improve uptake.

As mentioned before, identifying personal goals and motivations for taking part in a preventive DMHIs can inform personalization efforts as well as identify potential ways to address depression via indirect approaches. In an attempt to standardize motivation assessment, Allert-van Dam et al. [23] created the Nijmegen Motivation List for Prevention (NML-P) which has been shown to predict adherence in a web-based depression intervention [171] but has mostly been used in addiction research. Motivation in this assessment include the factors "readiness for participation", "doubt concerning participation", "perceived social support" and "burden" indicating a relation to constructs such as acceptance and outcome expectancy. Further development and testing of the questionnaire (e.g. in terms of construct validity) would allow for a relatively easy way to assess this information at the start of the intervention. For further tailoring it would also important to be able to distinguish expectations in prevention of first onset and relapse prevention for example by assessing prior experience.

Efforts to counter doubt and enhance positive expectations that the available DMHIs can help achieving the personal goals can then enhance uptake. Coming back to the unclear effects of outcome expectancy in Article 5, a clarification of expectation in the beginning of an intervention might be important in order to have a positive effect on intervention outcomes (see Shedden-Mora et al. [172] for a recent conceptual approach and measurement). An expectation that cannot be defined is less likely able to positively influence the outcome in question. While Constantino et al. [173] give examples on how therapists can explicitly address expectations in patients during face-to-face psychotherapy in order to foster expectancies and enhance them, this is more difficult in asynchronous DMHIs. In guided interventions, expectations and their fulfillment could be addressed by the coach after the initial 1 – 3 sessions and could potentially be used as a stepping-up indicator to try other approaches or media, given that treatment preference (which could be related to expectations) has been shown to predict the need of stepping up in an web-based intervention [174].

Depression prevention research might also learn from other preventive efforts and marketing initiatives. For example, van Voorhees et al. [175] used a marketing approach to identify strategies to recruit for their depression prevention study for adolescents which identified needs of the target group (e.g. focusing on the possibility of missing out on something rather than depression itself) and estimated costs necessary for recruitment. Commercial meditation apps advertise with special offers and the option to send gift cards or vouchers to friends. Personal recommendation is usually not promoted in evidence-based DMHIs even though the role of personal testimonies or recommenda-

tion of intervention by the word of mouth could be interesting aspects to consider in further research. This might be especially possible in indirect treatment approaches, where directly addressing mental health problems is not necessary.

In other fields aiming to improve health behavior, for example physical activity, fitness trackers make progress of training visible and add a gamification component [176]. Even though the avoidance of MDD onset or "living a healthy life" might be the ultimate goal, due to its distal and hard to measure nature it might not be the most useful goal for building behavior change motivation. Inclusion of tracking components which make progress or lifestyle changes more visible might have positive effects on some mental health related behavior. However, further research is needed to identify for whom and for which condition behavior tracking can safely be recommended [177].

### 9.4.3 *What happens to individuals who do not profit from preventive DMHIs?*

Ideally, in indicated prevention, the health service chosen by the individual is the first contact with a health service for an emerging problem. It is a noble quest to target the individuals who might profit most from an intervention and optimize interventions accordingly, but knowing about differential effects of these interventions also raises the question of what happens to individuals who do not or only to a low degree profit from an intervention or even experience a deterioration in symptoms while using a DMHI?

Perceived ineffectiveness was the most common reason for dropout from mental health interventions identified in a large-scale WHO survey [178]. Due to the nature of dropout, not much is known about what individuals do after dropout for their health – if anything. In cases, where a single preventive intervention did not "do the trick" (which might mean multiple things, e.g., reducing depressive symptoms down to close-to-symptom free status, preventing or delaying the onset of an MDE or at least lead to a change perceived as positive and meaningful by the individual), this intervention might be either the starting point of a mental health journey, or – in the worst case – a dead end if the individual generalizes the ineffectiveness of the one intervention to for example all psychological interventions and stops using any health service. Therefore, even in DMHIs with low human support, symptom change should be closely monitored and treatment stepped-up where necessary. The stepping-up process could either be "more of the same" (e.g., more guidance, higher intensity) or a different approach (e.g. other therapeutic approach) (see section 9.4.1.3).

The stepped care idea could be supported by knowing more about how individuals choose their health services and how prior

experience guides those decisions, given that positive health care experience has been associated with various long-term outcomes (incl. further treatment adherence) across various somatic and psychological disorders [179]. For preventive DMHIs, this could be studied for example in routine data from health insurance companies. Insurers have information of their own and subsidized preventive interventions and they may also have information about succeeding preventive offers, diagnoses, and treatments (within the realm of insurance coverage). However, the sole information on whether a service was used or not does not inform about the motivation behind it and its perceived usefulness. Therefore, they would need to include a helpfulness rating of the preventive offer for example in the reimbursement process at the ZPP or request feedback. Additionally, electronic health records could be designed to include preferences and experiences with prior treatment. While DMHIs provides opportunities for prevention and early intervention for subthreshold depressed individuals, it also comes with a responsibility to help find adequate care for everyone whose needs might go beyond the initial preventive intervention. Enabling individuals to take care of their needs and mental health in the long run as well as advocating for themselves in health care services should be included in all preventive interventions.

### 9.4.4 *Thinking big (data)*

Not only our understanding of mental disorders and our technological possibilities keep changing, but so does —- and must — our research practice. Therefore, in this last section of this dissertation, I want to discuss two aspects crucial for the development of depression prevention research.

Prevention of mental health problems still remains peripheral in mental health care with an estimated 1 – 3% of publicly funded health care in Germany is going into prevention and health promotion projects across all indications [180, 181, 182]. Similar representation can be expected in related research and development where only 4.4% of DMHIs regulated by central agencies worldwide focus on prevention [183]. Therefore rigorous research integrity and good evidence synthesis might help elevate the impact of prevention research [184, 185].

#### 9.4.4.1 *Accumulating real world evidence*

The evidence for the efficacy of DMHIs is constantly growing. However, with the ongoing implementation into routine care, the efforts cannot stop there. More effectiveness studies in real world settings are needed to validate the results from the trials [186]. Especially questions concerning uptake and adherence are important to investigate outside of a RCT which will almost always suffer from self-selection of moti-

vated individuals willing to endure the extra effort of research studies. Uptake and completion in routine care showed to be highly variable and lower than in RCTs for freely available (unguided) DMHIs [187] and guided online CBT [188]. However, comparability suffers from different definitions of drop-out and completers.

Regulation bodies such as the Food and Drug Administration (FDA) in the United States of America or the BfArM in Germany could play an important role in promoting the use of real-world data (RWD) and real-world evidence (RWE) with the integration of such in their regulatory processes. The FDA has valued RWE for a longer time now and for example consider RWD as a basis for enhancing indication of already registered interventions or would consider the comparison to an existing registration as a valid control condition [189]. While the regulations for DiGas in Germany in principle allow for RWE in the application process, the permanently listed DiGas at the moment still relied on classic RCTs as their evidence basis and there is to date no real example of an intervention receiving formal listing via RWE [190]. Regulating bodies and funders might also be in the position to enforce at least some standardized assessments and reporting of use cases in routine care.

Using RWD nevertheless provides great potential for highly efficient and agile development and evaluation of DMHIs [190] which could revolutionize research on DMHIs (for example see the framework proposed by Kim et al. [191]). This kind of evidence will also show who is not reached by classic RCTs and might therefore improve recruiting and evaluation in RCTs as well. While these regulations are not intended for preventive interventions, RWE generated for treatment of depression might also inform the further development of preventive intervention and set example for new methods to create evidence.

Additional to intervention data from routine care, with the rise of electronic health records and improved statistical and computational abilities (e.g. machine learning), new possibilities derive from using routine care health data or "big data" [192, 193]. Even though not without their limitations themselves, they will add a new perspective on healthcare and can counteract some of the potential bias introduced by self-report (see sections 2.3.3 and 9.2.3.3). For the German social insurance system, Effertz [194] drew up an agenda on how health insurance information could be used to enhance preventive efforts in a highly automated and routine manner. While they already highlight the potential of using "mhealth" for the communication of preventive measures, this could even go further and directly recommend or give access to DMHIs based on a personal risk-assessment. Ant any rate, this highlights once more the need to identify individuals most likely to profit from preventive DMHIs.

#### 9.4.4.2 *Synthesizing Evidence*

Some of the addressed research perspectives and gaps are beyond the scope of a single RCT or a single qualitative study. As shown in the last years [8] and this dissertation (Articles 5 & 6) the combination of IPD in meta-analyses can be a valuable tool to overcome some limitations of small scale studies and facilitate research on treatment effect predictors and moderators. They are at the moment nevertheless often still long-lasting endeavors due to the effort needed to acquire and harmonize data.

Throughout this dissertation, it has become obvious that the sheer multitude in different assessments of arguably the same construct hinders comparability, both in clinical trials as well as in routine care. Core outcome sets (COS) are consensus-driven collections of outcomes and instruments for specific health conditions, which – if agreed on and realized in clinical studies and practice – would drastically facilitate future evidence synthesis [195].

For depression treatment, a stakeholder panel involving clinicians, patients, government and industry representatives in the United States of America already suggested a set of 10 outcome measures [196]. A similar effort was made by the International Consortium for Health Outcomes Measurement (ICHOM) including mostly researchers and clinicians besides patient advocates [197]. The two proposals differ quite extensively, especially in that the ICHOM set includes more demographic and functioning variables and – interestingly – outcome expectancy. Unfortunately, outcome expectancy is reduced to one credibility item from the CEQ raising question about its validity and the definition of the underlying construct (see Article 5). It further recommends the use of common metrics in order to ensure comparability and consider item-based differences (see section 9.2.2 for details on the importance of that).

While these COS partly could be used in a preventive context, the main important outcome of how MDD onset should be determined would be missing and as pointed out in section 9.4.2, not enough is known about what kind of outcomes are expected or important to participants in preventive interventions. None of the COS are specifically developed for DMHIs, and similar agreements should be reached in how to report on the use of DMHIs e.g. concerning dropout or minimal use [188] in both research and routine care. The standards for developing new COS could guide this process in fields without such consensus and would allow including prior effort in related fields (i.e. depression treatment / psychotherapy) [198].

Improved and standardized reporting is at the same time only one part of the solution. In order to include (standardized) data into IPD meta-analyses, they also must be *FAIR* (findable, accessibel, interoperable, and reusable) [199] to researchers. As seen in Article 6 and as generally acknowledged [200, 201, 202, 203] actually obtaining

IPD from other researchers is still highly administration-heavy, time-consuming and not always successful, which threatens the potential of this method [203, 204]. This could be drastically improved with open data becoming more common in everyday research practice [205, 206] and more mandated by funders and publishers [207].

Besides facilitating IPD meta-analyses, which is a quite specific use case, open science, including but not limited to open access publishing, sharing of material data and analysis code has the potential to improve the quality of the research conducted [206] as well as freeing up resources, facilitating and accelerating evidence generation [208, 209]. Speeding up the research process while improving the quality of its output are especially critically aspects for the field of DMHIs, which might quickly become desperate to keep up with the technological developments they are trying to use to improve mental health.

## 9.5 FINAL THOUGHTS

This work highlights the opportunities for — and contributed to — potential solutions to overcome some of the challenges of DMHIs for the indicated prevention of depression. The big, distal goal, however, remains the reduction of overall depression burden. Working towards this goal has to be a global initiative to reducing early risk factors, promoting general mental health, detecting mental health problems early and offering diverse, available, effective and targeted interventions to everyone. This needs the commitment from mental health professionals, researchers, policy makers, society as well as individuals alike in a world that becomes more and more interconnected through digitalization.

## REFERENCES


1. Buntrock C, Ebert DD, Lehr D, Cuijpers P, Riper H, Smit F, and Berking M. Evaluating the Efficacy and Cost-Effectiveness of Web-Based Indicated Prevention of Major Depression: Design of a Randomised Controlled Trial. BMC Psychiatry. 2014; 14:25. DOI: 10.1186/1471-244X-14-25
2. Buntrock C, Ebert DD, Lehr D, Smit F, Riper H, Berking M, and Cuijpers P. Effect of a Web-Based Guided Self-Help Intervention for Prevention of Major Depression in Adults with Subthreshold Depression a Randomized Clinical Trial. JAMA - Journal of the American Medical Association. 2016; 315:1854–63. DOI: 10.1001/jama.2016.4326
3. Ebert DD, Buntrock C, Lehr D, Smit F, Riper H, Baumeister H, Cuijpers P, and Berking M. Effectiveness of Web- and Mobile-Based Treatment of Subthreshold Depression With Adherence-Focused Guidance: A Single-Blind Randomized Controlled Trial. Behavior Therapy. 2018; 49:71–83. DOI: 10.1016/j.beth.2017.05.004
4. Thiart H, Lehr D, Ebert DD, Berking M, and Riper H. Log in and Breathe out: Internet-based Recovery Training for Sleepless Employees with Work-Related Strain – Results of a Randomized Controlled Trial. Scandinavian Journal of Work, Environment and Health. 2015; 41:164–74. DOI: 10.5271/sjweh.3478
5. Cuijpers P, van Straten A, Warmerdam L, and van Rooy MJ. Recruiting Participants for Interventions to Prevent the Onset of Depressive Disorders: Possibile Ways to Increase Participation Rates. BMC Health Services Research. 2010; 10:181. DOI: 10.1186/1472-6963-10-181
6. Reins JA, Buntrock C, Zimmermann J, Grund S, Harrer M, Lehr D, Baumeister H, Weisel K, Domhardt M, Imamura K, Kawakami N, Spek V, Nobis S, Snoek F, Cuijpers P, Klein JP, Moritz S, and Ebert DD. Efficacy and Moderators of Internet-Based Interventions in Adults with Subthreshold Depression: An Individual Participant Data Meta-Analysis of Randomized Controlled Trials. Psychotherapy and Psychosomatics. 2021; 90:94–106. DOI: 10.1159/000507819
7. Bower P, Kontopantelis E, Sutton A, Kendrick T, Richards DA, Gilbody S, Knowles S, Cuijpers P, Andersson G, Christensen H, Meyer B, Huibers M, Smit F, van Straten A, Warmerdam L, Barkham M, Bilich L, Lovell K, and Liu ETH. Influence of Initial Severity of Depression on Effectiveness of Low Intensity Interventions: Meta-Analysis of Individual Patient Data. BMJ. 2013; 346:f540–f540. DOI: 10.1136/bmj.f540
8. Cuijpers P, Ciharova M, Quero S, Miguel C, Driessen E, Harrer M, Purgato M, Ebert D, and Karyotaki E. The Contribution of "Individual Participant Data" Meta-Analyses of Psychotherapies for Depression to the Development of Personalized Treatments: A Systematic Review. Journal of Personalized Medicine. 2022; 12:93. DOI: 10.3390/jpm12010093
9. Hautzinger M, Keller F, and Kühner C. Das Beck Depressionsinventar II. Deutsche Bearbeitung Und Handbuch Zum BDI-II. 2nd ed. Germany, Göttingen: Hogrefe, 2009

10. Green KL, Brown GK, Jager-Hyman S, Cha J, Steer RA, and Beck AT. The Predictive Validity of the Beck Depression Inventory Suicide Item. The Journal of Clinical Psychiatry. 2015; 76:1683–6. DOI: 10.4088/JCP.14m09391

11. Hawton K, Casañas i Comabella C, Haw C, and Saunders K. Risk Factors for Suicide in Individuals with Depression: A Systematic Review. Journal of Affective Disorders. 2013; 147:17–28. DOI: 10.1016/j.jad.2013.01.004

12. Wang X, Cheng S, and Xu H. Systematic Review and Meta-Analysis of the Relationship between Sleep Disorders and Suicidal Behaviour in Patients with Depression. BMC Psychiatry. 2019; 19:303. DOI: 10.1186/s12888-019-2302-5

13. Sander L, Gerhardinger K, Bailey E, Robinson J, Lin J, Cuijpers P, and Mühlmann C. Suicide Risk Management in Research on Internet-Based Interventions for Depression: A Synthesis of the Current State and Recommendations for Future Research. Journal of Affective Disorders. 2020; 263:676–83. DOI: 10.1016/j.jad.2019.11.045

14. McCall WV, Blocker JN, D'Agostino R, Kimball J, Boggs N, Lasater B, and Rosenquist PB. Insomnia Severity Is an Indicator of Suicidal Ideation during a Depression Clinical Trial. Sleep Medicine. 2010; 11:822–7. DOI: 10.1016/j.sleep.2010.04.004

15. Simmons Z, Erickson LD, Hedges D, and Kay DB. Insomnia Is Associated With Frequency of Suicidal Ideation Independent of Depression: A Replication and Extension of Findings From the National Health and Nutrition Examination Survey. Frontiers in Psychiatry. 2020; 11:561564. DOI: 10.3389/fpsyt.2020.561564

16. Bücker L, Schnakenberg P, Karyotaki E, Moritz S, and Westermann S. Diminishing Effects After Recurrent Use of Self-Guided Internet-Based Interventions in Depression: Randomized Controlled Trial. Journal of Medical Internet Research. 2019; 21:e14240. DOI: 10.2196/14240

17. Bundesärztekammer BÄK, Kassenärztliche Bundesvereinigung KBV, and Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften AWMF. Nationale VersorgungsLeitlinie Unipolare Depression - Langfassung. Version 3.0. 2022. DOI: 10.6101/AZQ/000493

18. GKV-Spitzenverband. Leitfaden Prävention – Handlungsfelder und Kriterien nach § 20 Abs. 2 SGB V zur Umsetzung der §§ 20, 20a und 20b SGB V vom 21. Juni 2000 in der Fassung vom 21. Dezember 2022. 2022. Accessed on: 09-03-2023. Available from: https://www.gkv-spitzenverband.de/krankenversicherung/praevention_selbsthilfe_beratung/praevention_und_bgf/leitfaden_praevention/leitfaden_praevention.jsp

19. National Institute for Health and Care Excellence NICE. Depression in Adults: Treatment and Management. 2022. Available from: www.nice.org.uk/guidance/ng222

20. Coughtrey AE and Pistrang N. The Effectiveness of Telephone-Delivered Psychological Therapies for Depression and Anxiety: A Systematic Review. 2018; 24:65–74. DOI: 10.1177/1357633X16686547

21. Cuijpers P. Why Primary Prevention Often Is No Prevention at All. European Neuropsychopharmacology. 2022; 58:1–3. DOI: 10.1016/j.euroneuro.2022.01.004

22. Volz HP, Stirnweiß J, Kasper S, Möller HJ, and Seifritz E. Subthreshold Depression – Concept, Operationalisation and Epidemiological Data. A Scoping Review. International Journal of Psychiatry in Clinical Practice. 2022; online first:1–15. DOI: 10.1080/13651501.2022.2087530

23. Allart-van Dam E, Hosman CM, and Keijsers GP. A New Instrument to Assess Participant Motivation for Involvement in Preventive Interventions. Journal of Clinical Psychology. 2004; 60:555–65. DOI: 10.1002/jclp.10236

24. Braun L, Titzler I, Ebert DD, Buntrock C, Terhorst Y, Freund J, Thielecke J, and Baumeister H. Clinical and Cost-Effectiveness of Guided Internet-Based Interventions in the Indicated Prevention of Depression in Green Professions (PROD-A): Study Protocol of a 36-Month Follow-up Pragmatic Randomized Controlled Trial. BMC Psychiatry. 2019; 19:278. DOI: 10.1186/s12888-019-2244-y

25. Wahl I, Löwe B, Bjorner JB, Fischer F, Langs G, Voderholzer U, Aita SA, Bergemann N, Brähler E, and Rose M. Standardization of Depression Measurement: A Common Metric Was Developed for 11 Self-Report Depression Measures. Journal of Clinical Epidemiology. 2014; 67:73–86. DOI: 10.1016/j.jclinepi.2013.04.019

26. Park SH and Yu HY. How Useful Is the Center for Epidemiologic Studies Depression Scale in Screening for Depression in Adults? An Updated Systematic Review and Meta-Analysis. Psychiatry Research. 2021; 302:114037. DOI: 10.1016/j.psychres.2021.114037

27. Vilagut G, Forero CG, Barbaglia G, and Alonso J. Screening for Depression in the General Population with the Center for Epidemiologic Studies Depression (CES-D): A Systematic Review with Meta-Analysis. PLoS ONE. 2016; 11:e0155431. DOI: 10.1371/journal.pone.0155431

28. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Applied Psychological Measurement. 1977; 1:385–401. DOI: 10.1177/014662167700100306

29. Hautzinger M, Bailer M, Hofmeister D, and Keller F. Allgemeine Depressionsskala. 2nd ed. Germany, Göttingen: Hogrefe, 2012

30. Fried E. Moving Forward: How Depression Heterogeneity Hinders Progress in Treatment and Research. Expert Review of Neurotherapeutics. 2017; 17:423–5. DOI: 10.1080/14737175.2017.1307737

31. Maj M, Stein DJ, Parker G, Zimmerman M, Fava GA, De Hert M, Demyttenaere K, McIntyre RS, Widiger T, and Wittchen HU. The Clinical Characterization of the Adult Patient with Depression Aimed at Personalization of Management. World Psychiatry. 2020; 19:269–93. DOI: 10.1002/wps.20771

32. Fried EI, Flake JK, and Robinaugh DJ. Revisiting the Theoretical and Methodological Foundations of Depression Measurement. Nature Reviews Psychology. 2022; 1:358–68. DOI: 10.1038/s44159-022-00050-2

33. Flake JK and Fried EI. Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them. Advances in Methods and Practices in Psychological Science. 2020; 3:456–65. DOI: 10.1177/2515245920952393

34. Cuijpers P, Koole SL, Van Dijke A, Roca M, Li J, and Reynolds CF. Psychotherapy for Subclinical Depression: Meta-analysis. British Journal of Psychiatry. 2014; 205:268–74. DOI: 10.1192/bjp.bp.113.138784

35. Riley RD and Ensor J. Power Calculations for Planning an IPD Meta-Analysis. *Individual Participant Data Meta-Analysis*. Ed. by Riley RD, Tierney JF, and Stewart LA. 1st ed. 291-309. UK, Chichester: Wiley, 2021

36. Riley RD, Hattle M, Collins GS, Whittle R, and Ensor J. Calculating the Power to Examine Treatment-covariate Interactions When Planning an Individual Participant Data Meta-analysis of Randomized Trials with a Binary Outcome. Statistics in Medicine. 2022; 41:4822–37. DOI: 10.1002/sim.9538

37. Cuijpers P. Examining the Effects of Prevention Programs on the Incidence of New Cases of Mental Disorders: The Lack of Statistical Power. American Journal of Psychiatry. 2003; 160:1385–91. DOI: 10.1176/appi.ajp.160.8.1385

38. Muñoz RF, Beardslee WR, and Leykin Y. Major Depression Can Be Prevented. American Psychologist. 2012; 67:285–95. DOI: 10.1037/a0027666

39. Schuster R, Kaiser T, Terhorst Y, Messner EM, Strohmeier LM, and Laireiter AR. Sample Size, Sample Size Planning, and the Impact of Study Context: Systematic Review and Recommendations by the Example of Psychological Depression Treatment. Psychological Medicine. 2021; 51:902–8. DOI: 10.1017/S003329172100129X

40. Althouse AD. Adjust for Multiple Comparisons? It's Not That Simple. The Annals of Thoracic Surgery. 2016; 101:1644–5. DOI: 10.1016/j.athoracsur.2015.11.024

41. Pike K, Reeves BC, and Rogers CA. Approaches to Multiplicity in Publicly Funded Pragmatic Randomised Controlled Trials: A Survey of Clinical Trials Units and a Rapid Review of Published Trials. BMC Medical Research Methodology. 2022; 22:39. DOI: 10.1186/s12874-022-01525-9

42. Li G, Taljaard M, van den Heuvel ER, Levine MA, Cook DJ, Wells GA, Devereaux PJ, and Thabane L. An Introduction to Multiplicity Issues in Clinical Trials: The What, Why, When and How. International Journal of Epidemiology. 2016; 46:746–55. DOI: 10.1093/ije/dyw320

43. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics. 1979; 6:65–70. Available from: http://www.jstor.org/stable/4615733

44. Emerson RW. Bonferroni Correction and Type I Error. Journal of Visual Impairment & Blindness. 2020; 114:77–8. DOI: 10.1177/0145482X20901378

45. Sjölander A and Vansteelandt S. Frequentist versus Bayesian Approaches to Multiple Testing. European Journal of Epidemiology. 2019; 34:809–21. DOI: 10.1007/s10654-019-00517-2

46. Freund J, Titzler I, Thielecke J, Braun L, Baumeister H, Berking M, and Ebert DD. Implementing Internet- and Tele-Based Interventions to Prevent Mental Health Disorders in Farmers, Foresters and Gardeners (ImplementIT): Study Protocol for the Multi-Level Evaluation of a Nationwide Project. BMC Psychiatry. 2020; 20:424. DOI: 10.1186/s12888-020-02800-z

47. Cuijpers P, Quero S, Papola D, Cristea IA, and Karyotaki E. Care-as-Usual Control Groups across Different Settings in Randomized Trials on Psychotherapy for Adult Depression: A Meta-Analysis. Psychological Medicine. 2021; 51:634–44. DOI: 10.1017/S0033291719003581

48. Freedland KE, Mohr DC, Davidson KW, and Schwartz JE. Usual and Unusual Care: Existing Practice Control Groups in Randomized Controlled Trials of Behavioral Interventions. Psychosomatic Medicine. 2011; 73:323–35. DOI: 10.1097/PSY.0b013e318218e1fb

49. Munder T, Geisshüsler A, Krieger T, Zimmermann J, Wolf M, Berger T, and Watzke B. Intensity of Treatment as Usual and Its Impact on the Effects of Face-to-Face and Internet-Based Psychotherapy for Depression: A Preregistered Meta-Analysis of Randomized Controlled Trials. Psychotherapy and Psychosomatics. 2022; 91:200–9. DOI: 10.1159/000521951

50. Furukawa TA, Noma H, Caldwell DM, Honyashiki M, Shinohara K, Imai H, Chen P, Hunot V, and Churchill R. Waiting List May Be a Nocebo Condition in Psychotherapy Trials: A Contribution from Network Meta-Analysis. Acta Psychiatrica Scandinavica. 2014; 130:181–92. DOI: 10.1111/acps.12275

51. Berufsverband Deutscher Psychologen BDP and Deutsche Gesellschaft für Psychologie DGP s. Berufsethische Richtlinien. 2016. Accessed on: 15-04-2023. Available from: https://zwpd.transmit.de/zwpd-dienstleistungen/zwpd-ethikkommission

52. Gold SM, Enck P, Hasselmann H, Friede T, Hegerl U, Mohr DC, and Otte C. Control Conditions for Randomised Trials of Behavioural Interventions in Psychiatry: A Decision Framework. The Lancet Psychiatry. 2017; 4:725–32. DOI: 10.1016/S2215-0366(17)30153-0

53. Stang A, Hense HW, Jöckel KH, Turner EH, and Tramèr MR. Is It Always Unethical to Use a Placebo in a Clinical Trial? PLoS Medicine. 2005; 2:e72. DOI: 10.1371/journal.pmed.0020072

54. Christensen H, Batterham PJ, Gosling JA, Ritterband LM, Griffiths KM, Thorndike FP, Glozier N, O'Dea B, Hickie IB, and Mackinnon AJ. Effectiveness of an Online Insomnia Program (SHUTi) for Prevention of Depressive Episodes (the GoodNight Study): A Randomised Controlled Trial. The Lancet Psychiatry. 2016; 3:333–41. DOI: 10.1016/S2215-0366(15)00536-2

55. Michopoulos I, Furukawa TA, Noma H, Kishimoto S, Onishi A, Ostinelli EG, Ciharova M, Miguel C, Karyotaki E, and Cuijpers P. Different Control Conditions Can Produce Different Effect Estimates in Psychotherapy Trials for Depression. Journal of Clinical Epidemiology. 2021; 132:59–70. DOI: 10.1016/j.jclinepi.2020.12.012

56. Donker T, Griffiths KM, Cuijpers P, and Christensen H. Psychoeducation for Depression, Anxiety and Psychological Distress: A Meta-Analysis. BMC Medicine. 2009; 7:79. DOI: 10.1186/1741-7015-7-79

57. Reins JA, Boß L, Lehr D, Berking M, and Ebert DD. The More I Got, the Less I Need? Efficacy of Internet-based Guided Self-Help Compared to Online Psychoeducation for Major Depressive Disorder. Journal of Affective Disorders. 2019; 246:695–705. DOI: 10.1016/j.jad.2018.12.065

58. De Souza Tursi MF, von Werne Baes C, Camacho FR, de CarvalhoTofoli SM, and Juruena MF. Effectiveness of Psychoeducation for Depression: A Systematic Review. Australian & New Zealand Journal of Psychiatry. 2013; 47:1019–31. DOI: 10.1177/0004867413491154

59. Gesundheit am Arbeitsplatz psyGA P für psychische. Kein Stress Mit Dem Stress Hörbuch: Handeln, Bevor Der Job Krank Macht. Accessed on: 07-03-2023. Available from: https://www.psyga.info/unsere-angebote#c764

60. Andersson G, Käll A, Juhlin S, Wahlström C, de Fine Licht E, Färdeman S, Franck A, Tholcke A, Nachtweij K, Fransson E, Vernmark K, Ludvigsson M, and Berg M. Free Choice of Treatment Content, Support on Demand and Supervision in Internet-Delivered CBT for Adults with Depression: A Randomized Factorial Design Trial. Behaviour Research and Therapy. 2023; 162:104265. DOI: 10.1016/j.brat.2023.104265

61. Domhardt M, Cuijpers P, Ebert DD, and Baumeister H. More Light? Opportunities and Pitfalls in Digitalized Psychotherapy Process Research. Frontiers in Psychology. 2021; 12:544129. DOI: 10.3389/fpsyg.2021.544129

62. Andersson G, Titov N, Dear BF, Rozental A, and Carlbring P. Internet-delivered Psychological Treatments: From Innovation to Implementation. World Psychiatry. 2019; 18:20–8. DOI: 10.1002/wps.20610

63. Van Kessel R, Wong BLH, Clemens T, and Brand H. Digital Health Literacy as a Super Determinant of Health: More than Simply the Sum of Its Parts. Internet Interventions. 2022; 27:100500. DOI: 10.1016/j.invent.2022.100500

64. Koh A, Swanepoel DW, Ling A, Ho BL, Tan SY, and Lim J. Digital Health Promotion: Promise and Peril. Health Promotion International. 2021; 36:i70–i80. DOI: 10.1093/heapro/daab134

65. Sørensen K, Pelikan JM, Röthlin F, Ganahl K, Slonska Z, Doyle G, Fullam J, Kondilis B, Agrafiotis D, Uiters E, Falcon M, Mensing M, Tchamov K, Broucke S van den, and Brand H. Health Literacy in Europe: Comparative Results of the European Health Literacy Survey (HLS-EU). The European Journal of Public Health. 2015; 25:1053–8. DOI: 10.1093/eurpub/ckv043

66. Levin-Zamir D and Bertschi I. Media Health Literacy, eHealth Literacy, and the Role of the Social Environment in Context. International Journal of Environmental Research and Public Health. 2018; 15:1643. DOI: 10.3390/ijerph15081643

67. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, and Crotty K. Low Health Literacy and Health Outcomes: An Updated Systematic Review. Annals of Internal Medicine. 2011; 155:97–107. DOI: 10.7326/0003-4819-155-2-201107190-00005

68. Schaeffer D, Berens EM, Gille S, Griese L, Klinger J, de Sombre S, Vogt D, and Hurrelmann K. Gesundheitskompetenz Der Bevölkerung in Deutschland Vor Und Während Der Corona Pandemie: Ergebnisse Des HLS-GER 2. Germany, Bielefeld: Universität Bielefeld, Interdisziplinäres Zentrum für Gesundheitskompetenzforschung (IZGK), 2021. DOI: 10.4119/UNIBI/2950305

69. Hagen BNM, Harper SL, O'Sullivan TL, and Jones-Bitton A. Tailored Mental Health Literacy Training Improves Mental Health Knowledge and Confidence among Canadian Farmers. International Journal of Environmental Research and Public Health. 2020; 17:3807. DOI: 10.3390/ijerph17113807

70. Yazd D, Wheeler SA, and Zuo A. Key Risk Factors Affecting Farmers' Mental Health: A Systematic Review. International Journal of Environmental Research and Public Health. 2019; 16:4849. DOI: 10.3390/ijerph16234849

71. Younker T and Radunovich HL. Farmer Mental Health Interventions: A Systematic Review. International Journal of Environmental Research and Public Health. 2021; 19:244. DOI: 10.3390/ijerph19010244

72. Braun L, Freund J, Thielecke J, Baumeister H, Ebert DD, and Titzler I. Barriers and Facilitators to Engaging and Adhering to Guided Internet-based Interventions for Depression Prevention and Reduction of Pain-related Disability in Green Professions: A Mixed Methods Study. JMIR Mental Health. 2022; 9:e39122. DOI: 10.2196/39122

73. Freund J, Buntrock C, Braun L, Thielecke J, Baumeister H, Berking M, Ebert DD, and Titzler I. Digital Prevention of Depression for Farmers? A Qualitative Study on Participants' Experiences Regarding Determinants of Acceptance and Satisfaction with a Tailored Guided Internet Intervention Program. Internet Interventions. 2022; 29:100566. DOI: 10.1016/j.invent.2022.100566

74. Wilhelm M and Euteneuer F. Does Health Literacy Make a Difference? Comparing the Effect of Conventional Medicine Versus Homeopathic Prescribing on Treatment Credibility and Expectancy. Frontiers in Psychology. 2021; 12:581255. DOI: 10.3389/fpsyg.2021.581255

75. IVPNetworks GmbH. Website IVPNetworks GmbH - Unsere Versorgungsprogramme. Accessed on: 09-03-2023. Available from: https://www.ivpnetworks.de/ivp-versorgungsprogramme-zur-unterstuetzung/

76. Online Gesundheitstrainings GmbH GI für. Website GET.ON Institute - Depression Prävention. Accessed on: 09-03-2023. Available from: https://hellobetter.de/online-kurse/depression-stimmung/

77. Bundesinstitut für Arzneimittel und Medizinprodukte BfArM. DiGA Verzeichnis. HelloBetter Schlafen. Accessed on: 09-03-2023. Available from: https://diga.bfarm.de/de/verzeichnis/01772/

78. Sauermann S, Herzberg J, Burkert S, and Habetha S. DiGA – A Chance for the German Healthcare System. Journal of European CME. 2022; 11:2014047. DOI: 10.1080/21614083.2021.2014047

79. Bundesinstitut für Arzneimittel und Medizinprodukte BfArM. DiGa Leitfaden Version 3.1. 2022. Accessed on: 29-03-2023. Available from: https://www.bfarm.de/SharedDocs/Downloads/DE/Medizinprodukte/diga_leitfaden.html

80. Zentrale Prüfstelle Prävention ZPP. Zentrale Prüfstelle Prävention - Finden Sie Ihren Gesundheitskurs! Accessed on: 10-03-2023. Available from: https://portal.zentrale-pruefstelle-praevention.de/portfolio/gkv-sv/suche

81. Glasgow RE, Vogt TM, and Boles SM. Evaluating the Public Health Impact of Health Promotion Interventions: The RE-AIM Framework. American Journal of Public Health. 1999; 89:1322–7. DOI: 10.2105/AJPH.89.9.1322

82. Clark DM. Implementing NICE Guidelines for the Psychological Treatment of Depression and Anxiety Disorders: The IAPT Experience. International Review of Psychiatry. 2011; 23:318–27. DOI: 10.3109/09540261.2011.606803

83. Muñoz RF, Pineda BS, Barrera AZ, Bunge E, and Leykin Y. Digital Tools for Prevention and Treatment of Depression: Lessons from the Institute for International Internet Interventions for Health. Clinica y Salud. 2021; 32:37–40. DOI: 10.5093/CLYSA2021A2

84. Muñoz RF, Bunge EL, Chen K, Schueller SM, Bravin JI, Shaughnessy EA, and Pérez-Stable EJ. Massive Open Online Interventions: A Novel Model for Delivering Behavioral-Health Services Worldwide. Clinical Psychological Science. 2016; 4:194–205. DOI: 10.1177/2167702615583840

85. Ebert DD, Van Daele T, Nordgreen T, Karekla M, Compare A, Zarbo C, Brugnera A, Øverland S, Trebbi G, Jensen KL, Kaehlke F, and Baumeister H. Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health. European Psychologist. 2018; 23:167–87. DOI: 10.1027/1016-9040/a000318

86. Thielecke J, Freund J, Buntrock C, Braun L, Baumeister H, Berking M, Ebert D, and Titzler I. Wie erreicht man Landwirt*innen für eine Teilnahme an digitalen Präventionsangeboten? ‚Lessons learned' aus einem nationalen Projekt zur Prävention von Depression. *Gesundheitswesen*. Conference presentation at the Gemeinsame Jahrestagung der DGSMP und der DGMS. 850. Magdeburg: Thieme, 2022. DOI: 10.1055/s-0042-1753945

87. Dahlhausen F, Zinner M, Bieske L, Ehlers JP, Boehme P, and Fehring L. Physicians' Attitudes Toward Prescribable mHealth Apps and Implications for Adoption in Germany: Mixed Methods Study. JMIR mHealth and uHealth. 2021; 9:e33012. DOI: 10.2196/33012

88. Bundespsychotherapeutenkammer. Ein Jahr Nach Der Reform Der Psychotherapie-Richtlinie. Wartezeiten 2018. 2019. Accessed on: 21.10.2022. Available from: https://www.bptk.de/publikationen/bptk-studie/

89. Weisel KK, Lehr D, Heber E, Zarski AC, Berking M, Riper H, and Ebert DD. Severely Burdened Individuals Do Not Need to Be Excluded from Internet-Based and Mobile-Based Stress Management: Effect Modifiers of Treatment Outcomes from Three Randomized Controlled Trials. Journal of Medical Internet Research. 2018; 20:e211. DOI: 10.2196/jmir.9387

90. Topooco N, Riper H, Araya R, Berking M, Brunn M, Chevreul K, Cieslak R, Ebert DD, Etchmendy E, Herrero R, Kleiboer A, Krieger T, García-Palacios A, Cerga-Pashoja A, Smoktunowicz E, Urech A, Vis C, and Andersson G. Attitudes towards Digital Treatment for Depression: A European Stakeholder Survey. Internet Interventions. 2017; 8:1–9. DOI: /10.1016/j.invent.2017.01.001

91. Thomas EE, Haydon HM, Mehrotra A, Caffery LJ, Snoswell CL, Banbury A, and Smith AC. Building on the Momentum: Sustaining Telehealth beyond COVID-19. Journal of Telemedicine and Telecare. 2020. DOI: 10.1177/1357633X20960638

92. Dahlhausen F, Zinner M, Bieske L, Ehlers JP, Boehme P, and Fehring L. There's an App for That, but Nobody's Using It: Insights on Improving Patient Access and Adherence to Digital Therapeutics in Germany. DIGITAL HEALTH. 2022; 8:1–12. DOI: 10.1177/20552076221104672

93. Messner EM, Terhorst Y, Barke A, Baumeister H, Stoyanov S, Hides L, Kavanagh D, Pryss R, Sander L, and Probst T. The German Version of the Mobile App Rating Scale (MARS-G): Development and Validation Study. JMIR mHealth and uHealth. 2020; 8:e14479. DOI: 10.2196/14479

94. Terhorst Y, Rathner EM, Baumeister H, and Sander L. «Hilfe Aus Dem App-Store?»: Eine Systematische Übersichtsarbeit Und Evaluation von Apps Zur Anwendung Bei Depressionen. Verhaltenstherapie. 2018; 28:101–12. DOI: 10.1159/000481692

95. Klein JP, Gerlinger G, Knaevelsrud C, Bohus M, Meisenzahl E, Kersting A, Röhr S, Riedel-Heller SG, Sprick U, Dirmaier J, Härter M, Hegerl U, Hohagen F, and Hauth I. Internetbasierte Interventionen in der Behandlung psychischer Störungen: Überblick, Qualitätskriterien, Perspektiven. Der Nervenarzt. 2016; 87:1185–93. DOI: 10.1007/s00115-016-0217-7

96. Ruwaard J and Kok RN. Wild West eHealth: Time to Hold Our Horses? The European Health Psychologist. 2015; 17:45–9. Available from: http://ehps.net/ehp/index.php/contents/article/view/765/pdf_44

97. Grant AM and Green RM. Developing Clarity on the Coaching - Counselling Conundrum: Implications for Counsellors and Psychotherapists. Counselling and Psychotherapy Research. 2018; 18:347–55. DOI: 10.1002/capr.12188

98. Institute for Apprenticeships & Technical Education IATE. Counselling Psychologist. Accessed on: 16-03-2023. Available from: https://www.healthcareers.nhs.uk/explore-roles/psychological-therapies/roles/counselling-psychologist

99. Norcross JC and Sayette MA. Insider's Guide to Graduate Programs in Clinical and Counseling Psychology. 1st ed. USA, New York, NY: The Guilford Press, 2022

100. Bachkirova T and Baker S. Revisiting the Issue of Boundaries between Coaching and Counselling. *Handbook of Coaching Psychology - a Guide for Practitioners*. Ed. by Palmer S and Whybrow A. 487-499. USA, New York, NY: Routledge, 2019

101. Spence GB and Grant AM. Professional and Peer Life Coaching and the Enhancement of Goal Striving and Well-Being: An Exploratory Study. The Journal of Positive Psychology. 2007; 2:185–94. DOI: 10.1080/17439760701228896

102. Van Rinsum CE, Gerards SMPL, Rutten GM, van de Goor IAM, Kremers SPJ, and Mercken L. Lifestyle Coaches as a Central Professional in the Health Care Network? Dynamic Changes over Time Using a Network Analysis. BMC Health Services Research. 2021; 21:247. DOI: 10.1186/s12913-021-06252-3

103. Kassenärztliche Bundesvereinigung KVB. Praxisinformation Videosprechstunde. 20231223742023. Accessed on: 07-04-2023. Available from: https://www.kbv.de/html/videosprechstunde.php

104. Bennett JL. An Agenda for Coaching-Related Research: A Challenge for Researchers. Consulting Psychology Journal: Practice and Research. 2006; 58:240–9. DOI: 10.1037/1065-9293.58.4.240

105. Smoktunowicz E, Barak A, Andersson G, Banos RM, Berger T, Botella C, Dear BF, Donker T, Ebert DD, Hadjistavropoulos H, Hodgins DC, Kaldo V, Mohr DC, Nordgreen T, Powers MB, Riper H, Ritterband LM, Rozental A, Schueller SM, Titov N, Weise C, and Carlbring P. Consensus Statement on the Problem of Terminology in Psychological Interventions Using the Internet or Digital Components. Internet Interventions. 2020; 21:100331. DOI: 10.1016/j.invent.2020.100331

106. Grant AM and Cavanagh MJ. Evidence-Based Coaching: Flourishing or Languishing? Australian Psychologist. 2007; 42:239–54. DOI: 10.1080/00050060701648175

107. Passmore J and Theeboom T. Coaching Psychology Research: A Journey of Development in Research. *Coaching Psychology: Meta-theoretical Perspectives and Applications in Multicultural Contexts*. Ed. by van Zyl LE, Stander MW, and Odendaal A. 27-46. USA, New York, NY: Springer, 2016. DOI: 10.1007/978-3-319-31012-1_2

108. De Haan E and Nilsson VO. What Can We Know about the Effectiveness of Coaching? A Meta-Analysis Based Only on Randomized Controlled Trials. Academy of Management Learning & Education. 2023; online first. DOI: 10.5465/amle.2022.0107

109. Wolever RQ, Simmons LA, Sforzo GA, Dill D, Kaye M, Bechard EM, Southard ME, Kennedy M, Vosloo J, and Yang N. A Systematic Review of the Literature on Health and Wellness Coaching: Defining a Key Behavioral Intervention in Healthcare. Global Advances in Health and Medicine. 2013; 2:38–57. DOI: 10.7453/gahmj.2013.042

110. Wakefield S, Kellett S, Simmonds-Buckley M, Stockton D, Bradbury A, and Delgadillo J. Improving Access to Psychological Therapies (IAPT) in the United Kingdom: A Systematic Review and Meta-analysis of 10-years of Practice-based Evidence. British Journal of Clinical Psychology. 2021; 60:1–37. DOI: 10.1111/bjc.12259

111. Hammond GC, Croudace TJ, Radhakrishnan M, Lafortune L, Watson A, McMillan-Shields F, and Jones PB. Comparative Effectiveness of Cognitive Therapies Delivered Face-to-Face or over the Telephone: An Observational Study Using Propensity Methods. PLoS ONE. 2012; 7:e42916. DOI: 10.1371/journal.pone.0042916

112. Institute for Apprenticeships & Technical Education IATE. Psychological Wellbeing Practitioner. Accessed on: 16-03-2023. Available from: https://www.instituteforapprenticeships.org/apprenticeship-standards/psychological-wellbeing-practitioner-v1-2

113. Scott MJ. Ensuring That the Improving Access to Psychological Therapies (IAPT) Programme Does What It Says on the Tin. British Journal of Clinical Psychology. 2021; 60:38–41. DOI: 10.1111/bjc.12264

114. Scott MJ. Improving Access to Psychological Therapies (IAPT) - The Need for Radical Reform. Journal of Health Psychology. 2018; 23:1136–47. DOI: 10.1177/1359105318755264

115. Drew P, Irvine A, Barkham M, Faija C, Gellatly J, Ardern K, Armitage J, Brooks H, Rushton K, Welsh C, Bower P, and Bee P. Telephone Delivery of Psychological Interventions: Balancing Protocol with Patient-Centred Care. Social Science & Medicine. 2021; 277:113818. DOI: 10.1016/j.socscimed.2021.113818

116. Castro A, Gili M, Ricci-Cabello I, Roca M, Roca M, Gilbody S, Perez-Ara MÁ, Maria Ángeles Perez-Ara, Pérez-Ara M, Seguí A, and McMillan D. Effectiveness and Adherence of Telephone-Administered Psychotherapy for Depression: A Systematic Review and Meta-Analysis. Journal of Affective Disorders. 2020; 260:514–26. DOI: 10.1016/j.jad.2019.09.023

117. Irvine A, Drew P, Bower P, Brooks H, Gellatly J, Armitage CJ, Barkham M, McMillan D, and Bee P. Are There Interactional Differences between Telephone and Face-to-Face Psychological Therapy? A Systematic Review of Comparative Studies. Journal of Affective Disorders. 2020; 265:120–31. DOI: 10.1016/j.jad.2020.01.057

118. Stiles-Shields C, Kwasny MJ, Cai X, and Mohr DC. Therapeutic Alliance in Face-to-Face and Telephone-Administered Cognitive Behavioral Therapy. Journal of Consulting and Clinical Psychology. 2014; 82:349–54. DOI: 10.1037/a0035554

119. Mohr DC, Ho J, Duffecy J, Reifler D, Sokol L, Burns MN, Jin L, and Siddique J. Effect of Telephone-Administered vs Face-to-Face Cognitive Behavioral Therapy on Adherence to Therapy and Depression Outcomes among Primary Care Patients: A Randomized Trial. JAMA - Journal of the American Medical Association. 2012; 307:2278–85. DOI: 10.1001/jama.2012.5588

120. Mohr DC, Vella L, Vella L, Hart SL, Heckman TG, and Simon GE. The Effect of Telephone-Administered Psychotherapy on Symptoms of Depression and Attrition: A Meta-Analysis. Clinical Psychology-science and Practice. 2008; 15:243–53. DOI: 10.1111/j.1468-2850.2008.00134.x

121. Markowitz JC, Milrod B, Heckman TG, Bergman M, Amsalem D, Amsalem D, Zalman H, Ballas T, and Neria Y. Psychotherapy at a Distance. American Journal of Psychiatry. 2020; 178:240–6. DOI: 10.1176/appi.ajp.2020.20050557

122. Scott AM, Clark J, Greenwood H, Krzyzaniak N, Cardona M, Peiris R, Sims R, and Glasziou P. Telehealth v. Face-to-Face Provision of Care to Patients with Depression: A Systematic Review and Meta-Analysis. Psychological Medicine. 2022; 52:2852–60. DOI: 10.1017/S0033291722002331

123. Kruzan KP, Meyerhoff J, Nguyen T, Reddy M, Mohr DC, and Kornfield R. "I Wanted to See How Bad It Was": Online Self-screening as a Critical Transition Point Among Young Adults with Common Mental Health Conditions. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22: CHI Conference on Human Factors in Computing Systems. 328:1-16. USA, New Orleans, LA: ACM, 2022. DOI: 10.1145/3491102.3501976

124. Gaab J, Blease C, Locher C, and Gerger H. Go Open: A Plea for Transparency in Psychotherapy. Psychology of Consciousness: Theory, Research, and Practice. 2016; 3:175–98. DOI: 10.1037/cns0000063

125. Irvine A, Drew P, Bower P, Ardern K, Armitage CJ, Barkham M, Brooks H, Connell J, Faija CL, Gellatly J, Rushton K, Welsh C, and Bee P. 'So Just to Go through the Options…': Patient Choice in the Telephone Delivery of the NHS *Improving Access to Psychological Therapies* Services. Sociology of Health & Illness. 2021; 43:3–19. DOI: 10.1111/1467-9566.13182

126. Bellón JÁ, Moreno-Peral P, Moreno-Küstner B, Motrico E, Aiarzagüena JM, Fernández A, Fernández-Alonso C, Montón-Franco C, Rodríguez-Bayón A, Ballesta-Rodríguez MI, Rüntel-Geidel A, Payo-Gordón J, Serrano-Blanco A, Oliván-Blázquez B, Araujo L, Muñoz-García MdM, King M, Nazareth I, and Amezcua M. Patients' Opinions about Knowing Their Risk for Depression and What to Do about It. The PredictD-Qualitative Study. PLoS ONE. 2014; 9:e92008. DOI: 10.1371/journal.pone.0092008

127. Wang J, Eccles H, Nannarone M, Schmitz N, Patten S, and Lashewicz B. Does Providing Personalized Depression Risk Information Lead to Increased Psychological Distress and Functional Impairment? Results from a Mixed-Methods Randomized Controlled Trial. Psychological Medicine. 2022; 52:2071–9. DOI: 10.1017/S0033291720003955

128. Sozialversicherung für Landwirtschaft Forsten und Gartenbau SVLFG. Online-Gesundheitstraining Für Arbeitnehmer. Accessed on: 07-04-2023. Available from: https://www.svlfg.de/online-gesundheitstraining-schlaf

129. Saxon D, Barkham M, Bee P, Gellatly J, Faija C, and Bower P. Telephone Treatments in Improving Access to Psychological Therapies Services: An Analysis of Use and Impact on Treatment Uptake. BMC Psychiatry. 2023; 23:95. DOI: 10.1186/s12888-022-04404-1

130. Mohr DC, Cuijpers P, and Lehman K. Supportive Accountability: A Model for Providing Human Support to Enhance Adherence to eHealth Interventions. Journal of Medical Internet Research. 2011; 13:e30. DOI: 10.2196/jmir.1602

131. Koelen J, Vonk A, Klein A, de Koning L, Vonk P, de Vet S, and Wiers R. Man vs. Machine: A Meta-Analysis on the Added Value of Human Support in Text-Based Internet Treatments ("e-Therapy") for Mental Disorders. Clinical Psychology Review. 2022; 96:102179. DOI: 10.1016/j.cpr.2022.102179

132. Cuijpers P, Noma H, Karyotaki E, Cipriani A, and Furukawa TA. Effectiveness and Acceptability of Cognitive Behavior Therapy Delivery Formats in Adults With Depression: A Network Meta-analysis. JAMA Psychiatry. 2019; 76:700–7. DOI: 10.1001/jamapsychiatry.2019.0268

133. Cross SP, Karin E, Staples LG, Bisby MA, Ryan K, Duke G, Nielssen O, Kayrouz R, Fisher A, Dear BF, and Titov N. Factors Associated with Treatment Uptake, Completion, and Subsequent Symptom Improvement in a National Digital Mental Health Service. Internet Interventions. 2022; 27:100506. DOI: 10.1016/j.invent.2022.100506

134. Fuhr K, Schröder J, Berger T, Moritz S, Meyer B, Lutz W, Hohagen F, Hautzinger M, and Klein JP. The Association between Adherence and Outcome in an Internet Intervention for Depression. Journal of Affective Disorders. 2018; 229:443–9. DOI: 10.1016/j.jad.2017.12.028

135. Kleiboer A, Donker T, Seekles W, van Straten A, Riper H, and Cuijpers P. A Randomized Controlled Trial on the Role of Support in Internet-based Problem Solving Therapy for Depression and Anxiety. Behaviour Research and Therapy. 2015; 72:63–71. DOI: 10.1016/j.brat.2015.06.013

136. Zagorscak P, Heinrich M, Sommer D, Wagner B, and Knaevelsrud C. Benefits of Individualized Feedback in Internet-Based Interventions for Depression: A Randomized Controlled Trial. Psychotherapy and Psychosomatics. 2018; 87:32–45. DOI: 10.1159/000481515

137. Apolinário-Hagen J, Harrer M, Kählke F, Fritsche L, Salewski C, and Ebert DD. Public Attitudes Toward Guided Internet-Based Therapies: Web-Based Survey Study. JMIR Mental Health. 2018; 5:e10735. DOI: 10.2196/10735

138. Rushton K, Ardern K, Hopkin E, Welsh C, Gellatly J, Faija C, Armitage CJ, Lidbetter N, Lovell K, Bower P, and Bee P. 'I Didn't Know What to Expect': Exploring Patient Perspectives to Identify Targets for Change to Improve Telephone-Delivered Psychological Interventions. BMC Psychiatry. 2020; 20:1–13. DOI: 10.1186/s12888-020-02564-6

139. Perski O and Short CE. Acceptability of Digital Health Interventions: Embracing the Complexity. Translational Behavioral Medicine. 2021; 11:1473–80. DOI: 10.1093/tbm/ibab048

140. Gun SY, Titov N, and Andrews G. Acceptability of Internet Treatment of Anxiety and Depression. Australasian Psychiatry. 2011; 19:259–64. DOI: 10.3109/10398562.2011.562295

141. Richardson S, Lawrence K, Schoenthaler AM, and Mann D. A Framework for Digital Health Equity. npj Digital Medicine. 2022; 5:119. DOI: 10.1038/s41746-022-00663-0

142. Khilnani A, Schulz J, and Robinson L. The COVID-19 Pandemic: New Concerns and Connections between eHealth and Digital Inequalities. Journal of Information, Communication and Ethics in Society. 2020; 18:393–403. DOI: 10.1108/JICES-04-2020-0052

143. Zheng Y and Walsham G. Inequality of What? An Intersectional Approach to Digital Inequality under Covid-19. Information and Organization. 2021; 31:100341. DOI: 10.1016/j.infoandorg.2021.100341

144. Farrer LM, Batterham PJ, Gulliver A, Morse A, Calear AL, McCallum S, Banfield M, Shou Y, Newman E, and Dawel A. Factors Associated with Telehealth Use and Avoidance during the COVID-19 Pandemic. Journal of Medical Internet Research. 2022; 25:e43798. DOI: 10.2196/43798

145. De Witte NAJ, Joris S, Van Assche E, and Van Daele T. Technological and Digital Interventions for Mental Health and Wellbeing: An Overview of Systematic Reviews. Frontiers in Digital Health. 2021; 3:754337. DOI: 10.3389/fdgth.2021.754337

146. Koinis L, Mobbs RJ, Fonseka RD, and Natarajan P. A Commentary on the Potential of Smartphones and Other Wearable Devices to Be Used in the Identification and Monitoring of Mental Illness. Annals of Translational Medicine. 2022; 10:24. DOI: 10.21037/atm-21-6016

147. Schick A, Rauschenberg C, Ader L, Daemen M, Wieland LM, Paetzold I, Postma MR, Schulte-Strathaus JCC, and Reininghaus U. Novel Digital Methods for Gathering Intensive Time Series Data in Mental Health Research: Scoping Review of a Rapidly Evolving Field. Psychological Medicine. 2022; 53:55–65. DOI: 10.1017/S0033291722003336

148. Birk RH and Samuel G. Digital Phenotyping for Mental Health: Reviewing the Challenges of Using Data to Monitor and Predict Mental Health Problems. Current Psychiatry Reports. 2022; 24:523–8. DOI: 10.1007/s11920-022-01358-9

149. Ettore E, Müller P, Hinze J, Benoit M, Giordana B, Postin D, Lecomte A, Lindsay H, Robert P, and König A. Digital Phenotyping for Differential Diagnosis of Major Depressive Episode: Narrative Review. JMIR Mental Health. 2023; 10:e37225. DOI: 10.2196/37225

150. Lennon MJ and Harmer C. Machine Learning Prediction Will Be Part of Future Treatment of Depression. Australian & New Zealand Journal of Psychiatry. 2023; online first. DOI: 10.1177/00048674231158267

151. Hamdoun S, Monteleone R, Bookman T, and Michael K. AI-Based and Digital Mental Health Apps: Balancing Need and Risk. IEEE Technology and Society Magazine. 2023; 42:25–36. DOI: 10.1109/MTS.2023.3241309

152. Gazzola N and Iwakabe S. Psychotherapy Failures: To Err Is Human. Counselling Psychology Quarterly. 2022; 35:719–23. DOI: 10.1080/09515070.2022.2142383

153. Nye A, Delgadillo J, and Barkham M. Efficacy of Personalized Psychological Interventions: A Systematic Review and Meta-Analysis. Journal of Consulting and Clinical Psychology. 2023; preprint. DOI: 10.1037/ccp0000820

154. Kim KJ, Shin DH, and Yoon H. Information Tailoring and Framing in Wearable Health Communication. Information Processing & Management. 2017; 53:351–8. DOI: 10.1016/j.ipm.2016.11.005

155. Cuijpers P, Ebert DD, Acarturk C, Andersson G, and Cristea IA. Personalized Psychotherapy for Adult Depression: A Meta-Analytic Review. Behavior Therapy. 2016; 47:966–80. DOI: 10.1016/j.beth.2016.04.007

156. Berrouiguet S, Perez-Rodriguez MM, Larsen M, Baca-García E, Courtet P, and Oquendo M. From eHealth to iHealth: Transition to Participatory and Personalized Medicine in Mental Health. Journal of Medical Internet Research. 2018; 20:e2. DOI: 10.2196/jmir.7412

157. Kessler RC, van Loo HM, Wardenaar KJ, Bossarte RM, Brenner LA, Ebert DD, de Jonge P, Nierenberg AA, Rosellini AJ, Sampson NA, Schoevers RA, Wilcox MA, and Zaslavsky AM. Using Patient Self-Reports to Study Heterogeneity of Treatment Effects in Major Depressive Disorder. Epidemiology and Psychiatric Sciences. 2017; 26:22–36. DOI: 10.1017/S2045796016000020

158. Cuijpers P. Prevention: An Achievable Goal in Personalized Medicine. Dialogues in Clinical Neuroscience. 2009; 11:447–54. DOI: 10.31887/DCNS.2009.11.4/pcuijpers

159. Kraus C, Kadriu B, Lanzenberger R, Zarate CA, and Kasper S. Prognosis and Improved Outcomes in Major Depression: A Review. Translational Psychiatry. 2019; 9:127. DOI: 10.1038/s41398-019-0460-3

160. Blanken TF, Borsboom D, Penninx BW, and Van Someren EJ. Network Outcome Analysis Identifies Difficulty Initiating Sleep as a Primary Target for Prevention of Depression: A 6-Year Prospective Study. Sleep. 2020; 43:1–6. DOI: 10.1093/sleep/zsz288

161. Leerssen J, Lakbila-Kamal O, Dekkers LM, Ikelaar SL, Albers AC, Blanken TF, Lancee J, Van Der Lande GJ, Maksimovic T, Mastenbroek SE, Reesen JE, Van De Ven S, Van Der Zweerde T, Foster-Dingley JC, and Van Someren EJ. Treating Insomnia with High Risk of Depression Using Therapist-Guided Digital Cognitive, Behavioral, and Circadian Rhythm Support Interventions to Prevent Worsening of Depressive Symptoms: A Randomized Controlled Trial. Psychotherapy and Psychosomatics. 2021; 91:168–79. DOI: 10.1159/000520282

162. Fried EI, Rieble C, and Proppert RKK. Building an Early Warning System for Depression: Rationale, Objectives, and Methods of the WARN-D Study. PsyArXiv. 2022; preprint. DOI: DOI:10.31234/osf.io/9qcvs

163. Bellón JÁ, Conejo-Cerón S, Moreno-Peral P, King M, Nazareth I, Martín-Pérez C, Fernández-Alonso C, Ballesta-Rodríguez MI, Fernández A, Aiarzaguena JM, Montón-Franco C, Ibanez-Casas I, Rodríguez-Sánchez E, Rodríguez-Bayón A, Serrano-Blanco A, Gómez MC, Lafuente P, del Mar Muñoz-García M, Mínguez-Gonzalo P, Araujo L, Palao D, Espinosa-Cifuentes M, Zubiaga F, Navas-Campaña D, Mendive J, Aranda-Regules JM, Rodriguez-Morejón A, Salvador-Carulla L, and de Dios Luna J. Preventing the Onset of Major Depression Based on the Level and Profile of Risk of Primary Care Attendees: Protocol of a Cluster Randomised Trial (the predictD-CCRT Study). BMC Psychiatry. 2013; 13:171. DOI: 10.1186/1471-244X-13-171

164. Bellón JÁ, Conejo-Cerón S, Moreno-Peral P, King M, Nazareth I, Martín-Pérez C, Fernández-Alonso C, Rodríguez-Bayón A, Fernández A, Aiarzaguena JM, Montón-Franco C, Ibanez-Casas I, Rodríguez-Sánchez E, Ballesta-Rodríguez MI, Serrano-Blanco A, Gómez MC, La-Fuente P, Muñoz-García MdM, Mínguez-Gonzalo P, Araujo L, Palao D, Bully P, Zubiaga F, Navas-Campaña D, Mendive J, Aranda-Regules JM, Rodriguez-Morejón A, Salvador-Carulla L, and de Dios Luna J. Intervention to Prevent Major Depression in Primary Care: A Cluster Randomized Trial. Annals of Internal Medicine. 2016; 164:656–66. DOI: 10.7326/M14-2653

165. Van Straten A, Hill J, Richards DA, and Cuijpers P. Stepped Care Treatment Delivery for Depression: A Systematic Review and Meta-Analysis. Psychological Medicine. 2015; 45:231–46. DOI: 10.1017/S0033291714000701

166. Rivero-Santana A, Perestelo-Perez L, Alvarez-Perez Y, Ramos-Garcia V, Duarte-Diaz A, Linertova R, Garcia-Perez L, and Serrano-Aguilar P. Stepped Care for the Treatment of Depression: A Systematic Review and Meta-Analysis. Journal of Affective Disorders. 2021; 294:391–409. DOI: 10.1016/j.jad.2021.07.008

167. Mohr DC, Lattie EG, Tomasino KN, Kwasny MJ, Kaiser SM, Gray E, Alam N, Jordan N, and Schueller SM. A Randomized Noninferiority Trial Evaluating Remotely-Delivered Stepped Care for Depression Using Internet Cognitive Behavioral Therapy (CBT) and Telephone CBT. Behaviour Research and Therapy. 2019; 123. DOI: 10.1016/j.brat.2019.103485

168. Nicholas J, Knapp AA, Vergara JL, Graham AK, Gray EL, Lattie EG, Kwasny MJ, and Mohr DC. An Exploratory Brief Head-To-Head Non-Inferiority Comparison of an Internet-Based and a Telephone-Delivered CBT Intervention for Adults with Depression. Journal of Affective Disorders. 2021; 281:673–7. DOI: 10.1016/j.jad.2020.11.093

169. Ho FYY, Yeung WF, Ng THY, and Chan CS. The Efficacy and Cost-Effectiveness of Stepped Care Prevention and Treatment for Depressive and/or Anxiety Disorders: A Systematic Review and Meta-Analysis. Scientific Reports. 2016; 6:29281. DOI: 10.1038/srep29281

170. Philippi P, Baumeister H, Apolinário-Hagen J, Ebert DD, Hennemann S, Kott L, Lin J, Messner EM, and Terhorst Y. Acceptance towards Digital Health Interventions – Model Validation and Further Development of the Unified Theory of Acceptance and Use of Technology. Internet Interventions. 2021; 26:100459. DOI: 10.1016/j.invent.2021.100459

171. Farrer L, Griffiths KM, Christensen H, Mackinnon A, and Batterham PJ. Predictors of Adherence and Outcome in Internet-Based Cognitive Behavior Therapy Delivered in a Telephone Counseling Setting. Cognitive Therapy and Research. 2014; 38:358–67. DOI: 10.1007/s10608-013-9589-1

172. Shedden-Mora MC, Alberts J, Petrie KJ, Laferton JAC, von Blanckenburg P, Kohlmann S, Nestoriuc Y, and Löwe B. The Treatment Expectation Questionnaire (TEX-Q): Validation of a Generic Multidimensional Scale Measuring Patients' Treatment Expectations. PLoS ONE. 2023; 18:e0280472. DOI: 10.1371/journal.pone.0280472

173. Constantino MJ, Vîslă A, Coyne AE, and Boswell JF. A Meta-Analysis of the Association between Patients' Early Treatment Outcome Expectation and Their Posttreatment Outcomes. Psychotherapy. 2018; 55:473–85. DOI: 10.1037/pst0000169

174. Nicholas J, Ringland KE, Graham AK, Knapp AA, Lattie EG, Kwasny MJ, and Mohr DC. Stepping Up: Predictors of 'Stepping' within an iCBT Stepped-Care Intervention for Depression. International Journal of Environmental Research and Public Health. 2019; 16:4689. DOI: 10.3390/ijerph16234689

175. Van Voorhees BW, Watson N, Bridges JFP, Fogel J, Galas J, Kramer C, Connery M, McGill A, Marko M, Cardenas A, Landsback J, Dmochowska K, Kuwabara SA, Ellis J, Prochaska M, and Bell C. Development and Pilot Study of a Marketing Strategy for Primary Care/Internet–Based Depression Prevention Intervention for Adolescents (The CATCH-IT Intervention). The Primary Care Companion to The Journal of Clinical Psychiatry. 2010; 12:e1–e9. DOI: 10.4088/PCC.09m00791blu

176. Cho I, Kaplanidou K, and Sato S. Gamified Wearable Fitness Tracker for Physical Activity: A Comprehensive Literature Review. Sustainability. 2021; 13:7017. DOI: 10.3390/su13137017

177. Kelley C, Lee B, and Wilcox L. Self-Tracking for Mental Wellness: Understanding Expert Perspectives and Student Experiences. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17: CHI Conference on Human Factors in Computing Systems. 629-641. USA, Denver, CO: ACM, 2017. DOI: 10.1145/3025453.3025750

178. Andrade LH et al. Barriers to Mental Health Treatment: Results from the WHO World Mental Health Surveys. Psychological Medicine. 2014; 44:1303–17. DOI: 10.1017/S0033291713001943

179. Doyle C, Lennox L, and Bell D. A Systematic Review of Evidence on the Links between Patient Experience and Clinical Safety and Effectiveness. BMJ Open. 2013; 3:e001570. DOI: 10.1136/bmjopen-2012-001570

180. Altgeld T. Finanzierung der Gesundheitsförderung. Leitbegriffe der Gesundheitsförderung und Prävention: Glossar zu Konzepten. 2016. Accessed on: 20-03-2023. Ed. by Bundeszentrale für gesundheitliche Aufklärung BZgA. DOI: 10.17623/BZGA:224-I018-1.0

181. Martin S, Lomas J, and Claxton K. Is an Ounce of Prevention Worth a Pound of Cure? A Cross-Sectional Study of the Impact of English Public Health Grant on Mortality and Morbidity. BMJ Open. 2020; 10:e036411. DOI: 10.1136/bmjopen-2019-036411

182. Statistische Bundesamt DESTATIS. Gesundheitsausgaben in Deutschland in Mio. €. Gliederungsmerkmale: Jahre, Art Der Einrichtung, Art Der Leistung, Ausgabenträger. 2023. Accessed on: 23-3-2023. Available from: https://www.gbe-bund.de/gbe/pkg_isgbe5.prc_menu_olap?p_uid=gastd&p_aid=53287072&p_sprache=D&p_help=0&p_indnr=322&p_indsp=&p_ityp=H&p_fid=

183. Wang C, Lee C, and Shin H. Digital Therapeutics from Bench to Bedside. npj Digital Medicine. 2023; 6:38. DOI: 10.1038/s41746-023-00777-z

184. Valentine JC, Biglan A, Boruch RF, Castro FG, Collins LM, Flay BR, Kellam S, Mościcki EK, and Schinke SP. Replication in Prevention Science. Prevention Science. 2011; 12:103–17. DOI: 10.1007/s11121-011-0217-6

185. Hennessy EA, Acabchuk RL, Arnold PA, Dunn AG, Foo YZ, Johnson BT, Geange SR, Haddaway NR, Nakagawa S, Mapanga W, Mengersen K, Page MJ, Sánchez-Tójar A, Welch V, and McGuinness LA. Ensuring Prevention Science Research Is Synthesis-Ready for Immediate and Lasting Scientific Impact. Prevention Science. 2022; 23:809–20. DOI: 10.1007/s11121-021-01279-8

186. Porzsolt Franz F, Galito N, Toledo-Arruda A, Thomaz T, Moraes C, Guerra T, Leão M, Migowski A, da Silva ARA, and Weiß C. Efficacy and Effectiveness Trials Have Different Goals, Use Different Tools, and Generate Different Messages. Pragmatic and Observational Research. 2015; 6:47–54. DOI: 10.2147/POR.S89946

187. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, and Merry S. Beyond the Trial: Systematic Review of Real-World Uptake and Engagement With Digital Self-Help Interventions for Depression, Low Mood, or Anxiety. Journal of Medical Internet Research. 2018; 20:e199. DOI: 10.2196/jmir.9275

188. Etzelmueller A, Vis C, Karyotaki E, Baumeister H, Titov N, Berking M, Cuijpers P, Riper H, and Ebert DD. Effects of Internet-Based Cognitive Behavioral Therapy in Routine Care for Adults in Treatment for Depression and Anxiety: Systematic Review and Meta-Analysis. Journal of Medical Internet Research. 2020; 22:e18100. DOI: 10.2196/18100

189. Food and Drug Administration FDA. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices - Guidance for Industry and Food and Drug Administration Staff. 2017. Accessed on: 27-03-2023. Available from: https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

190. Stern AD, Brönneke J, Debatin JF, Hagen J, Matthies H, Patel S, Clay I, Eskofier B, Herr A, Hoeller K, Jaksa A, Kramer DB, Kyhlstedt M, Lofgren KT, Mahendraratnam N, Muehlan H, Reif S, Riedemann L, and Goldsack JC. Advancing Digital Health Applications: Priorities for Innovation in Real-World Evidence Generation. The Lancet Digital Health. 2022; 4:e200–e206. DOI: 10.1016/S2589-7500(21)00292-2

191. Kim M, Patrick K, Nebeker C, Godino J, Stein S, Klasnja P, Perski O, Viglione C, Coleman A, and Hekler EB. The Digital Therapeutics Real World Evidence Framework: An Approach for Guiding Evidence-Based DTx Design, Development, Testing, and Monitoring. 2022; preprint. DOI: 10.31219/osf.io/8t5xy

192. Ross MK, Wei W, and Ohno-Machado L. "Big Data" and the Electronic Health Record. Yearbook of Medical Informatics. 2014; 23:97–104. DOI: 10.15265/IY-2014-0003

193. Saeed S, Moodie EEM, Strumpf EC, and Klein MB. Evaluating the Impact of Health Policies: Using a Difference-in-Differences Approach. International Journal of Public Health. 2019; 64:637–42. DOI: 10.1007/s00038-018-1195-2

194. Effertz T. Prävention und Kostenkontrolle im Gesundheitswesen: Routinedatenanalyse und effektives Gesundheitsmarketing als Vorteile im Krankenkassenwettbewerb? Prävention und Gesundheitsförderung. 2023. DOI: 10.1007/s11553-023-01021-y

195. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, and Tugwell P. Developing Core Outcome Sets for Clinical Trials: Issues to Consider. Trials. 2012; 13:132. DOI: 10.1186/1745-6215-13-132

196. Gliklich RE, Leavy MB, Cosgrove L, Simon GE, Gaynes BN, Peterson LE, Olin B, Cole C, DePaulo JR, Wang P, Crowe CM, Cusin C, Nix M, Berliner E, and Trivedi MH. Harmonized Outcome Measures for Use in Depression Patient Registries and Clinical Practice. Annals of Internal Medicine. 2020; 172:803–9. DOI: 10.7326/M19-3818

197. Obbarius A, van Maasakkers L, Baer L, Clark DM, Crocker AG, de Beurs E, Emmelkamp PMG, Furukawa TA, Hedman-Lagerlöf E, Kangas M, Langford L, Lesage A, Mwesigire DM, Nolte S, Patel V, Pilkonis PA, Pincus HA, Reis RA, Rojas G, Sherbourne C, Smithson D, Stowell C, Woolaway-Bickel K, and Rose M. Standardization of Health Outcomes Assessment for Depression and Anxiety: Recommendations from the ICHOM Depression and Anxiety Working Group. Quality of Life Research. 2017; 26:3211–25. DOI: 10.1007/s11136-017-1659-5

198. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, and Williamson PR. Core Outcome Set-STAndards for Development: The COS-STAD Recommendations. PLOS Medicine. 2017; 14:e1002447. DOI: 10.1371/journal.pmed.1002447

199. Boeckhout M, Zielhuis GA, and Bredenoord AL. The FAIR Guiding Principles for Data Stewardship: Fair Enough? European Journal of Human Genetics. 2018; 26:931–6. DOI: 10.1038/s41431-018-0160-0

200. Nevitt SJ and Tudur Smith C. Practical Considerations and Challenges When Conducting an Individual Participant Data (IPD) Meta-Analysis. *Meta-Research*. Ed. by Evangelou E and Veroniki AA. 263-278. USA, New York, NY: Springer, 2022. DOI: 10.1007/978-1-0716-1566-9_16

201. Rogozińska E, Marlin N, Thangaratinam S, Khan KS, and Zamora J. Meta-Analysis Using Individual Participant Data from Randomised Trials: Opportunities and Limitations Created by Access to Raw Data. Evidence Based Medicine. 2017; 22:157–62. DOI: 10.1136/ebmed-2017-110775

202. Van Middelkoop M, Lohmander S, and Bierma-Zeinstra SMA. Sharing Data–Taming the Beast: Barriers to Meta-Analyses of Individual Patient Data (IPD) and Solutions. British Journal of Sports Medicine. 2020; 54:822–4. DOI: 10.1136/bjsports-2019-101892

203. Ventresca M, Schünemann HJ, Macbeth F, Clarke M, Thabane L, Griffiths G, Noble S, Garcia D, Marcucci M, Iorio A, Zhou Q, Crowther M, Akl EA, Lyman GH, Gloy V, DiNisio M, and Briel M. Obtaining and Managing Data Sets for Individual Participant Data Meta-Analysis: Scoping Review and Practical Guide. BMC Medical Research Methodology. 2020; 20:113. DOI: 10.1186/s12874-020-00964-6

204. Gabelica M, Bojčić R, and Puljak L. Many Researchers Were Not Compliant with Their Published Data Sharing Statement: A Mixed-Methods Study. Journal of Clinical Epidemiology. 2022; 150:33–41. DOI: 10.1016/j.jclinepi.2022.05.019

205. Forero DA, Curioso WH, and Patrinos GP. The Importance of Adherence to International Standards for Depositing Open Data in Public Repositories. BMC Research Notes. 2021; 14:405. DOI: 10.1186/s13104-021-05817-z

206. Pisani E, Aaby P, Breugelmans JG, Carr D, Groves T, Helinski M, Kamuya D, Kern S, Littler K, Marsh V, Mboup S, Merson L, Sankoh O, Serafini M, Schneider M, Schoenenberger V, and Guerin PJ. Beyond Open Data: Realising the Health Benefits of Sharing Data. BMJ. 2016; 355:i5295. DOI: 10.1136/bmj.i5295

207. Burgelman JC, Pascu C, Szkuta K, Von Schomberg R, Karalopoulos A, Repanas K, and Schouppe M. Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century. Frontiers in Big Data. 2019; 2:43. DOI: 10.3389/fdata.2019.00043

208. Lowndes JSS, Best BD, Scarborough C, Afflerbach JC, Frazier MR, O'Hara CC, Jiang N, and Halpern BS. Our Path to Better Science in Less Time Using Open Data Science Tools. Nature Ecology & Evolution. 2017; 1:0160. DOI: 10.1038/s41559-017-0160

209. Arza V and Fressoli M. Systematizing Benefits of Open Science Practices. Information Services & Use. 2018; 37:463–74. DOI: 10.3233/ISU-170861

Part IV

# APPENDIX

# A

## REPRINT PERMISSIONS

Permission to reprint the edited version of the following publication was obtained from the publisher:
Thielecke J, Buntrock C, Titzler I, Braun L, Freund J, Berking M, Baumeister H, and Ebert DD. Clinical and Cost-Effectiveness of Personalized Telephone Coaching for Farmers, Foresters and Gardeners to Prevent Depression: Study Protocol of an 18-Month Follow-Up Pragmatic Randomized Controlled Trial (TEC-A). Frontiers in Psychiatry. 2020; 11:125. doi: 10.3389/fpsyt.2020.00125

Special Requests > Special Request Details

**Journal of telemedicine and telecare**

**Article:** Telephone coaching for the prevention of depression in depression in farmers: Results from a ...

**GENERAL INFORMATION**

| Request ID | Request Date |
|---|---|
| 600085984 | 28 Jun 2022 |
| **Request Status** | **Deny Reason** |
| Denied | Other (Authors able to reuse own work in thesis/di... ) |

ALL DETAILS

**COMMENTS**

Add Comment / Attachment

28 Jun 2022 8:33:04 PM, by Craig Myles

Dear Janika Thielecke,

Thank you for your order. I am pleased to report we can grant your request without a fee as part of your thesis or dissertation.

Please accept this email as permission for your request as you've detailed below. Permission is granted for the life of the edition on a non-exclusive basis, in the English language, throughout the world in all formats provided full citation is made to the original SAGE publication. Permission does not include any third-party material found within the work. Please contact us for any further usage of the material.

As no fee is needed, we will cancel this order. If you have any questions or should you require formal written permission, please place an order via our Customer Permission Portal (https://jira.sagepub.com/servicedesk/customer/portal/9)

Best,
SAGE Permissions Team

View Less

28 Jun 2022 9:11:53 AM, by Janika Thielecke

I would like to include the edited version of the mauscript in my PhD thesis which will be made public according to my universities regulation.

No permission to reprint the edited version of the following publication had to be obtained, because it was published under the terms of the Creative Commons Attribution License (CC BY):
Thielecke J, Buntrock C, Titzler I, Braun L, Freund J, Berking M, Baumeister H, and Ebert DD. Telephone Coaching for the Prevention of Depression in Farmers: Results from a Pragmatic Randomized Controlled Trial. Journal of Telemedicine and Telecare. 2022; online first. doi: 10.1177/1357633X221106027

Permission to reprint the graphic from Grant & Green (2018) on page 296 was obtained from John Wiley & Sons - Books:

Marketplace

This is a License Agreement between Janika Thielecke ("User") and Copyright Clearance Center, Inc. ("CCC") on behalf of the Rightsholder identified in the order details below. The license consists of the order details, the Marketplace Permissions General Terms and Conditions below, and any Rightsholder Terms and Conditions which are included below.

All payments must be made in full to CCC in accordance with the Marketplace Permissions General Terms and Conditions below.

| | | | |
|---|---|---|---|
| Order Date | 17-Mar-2023 | Type of Use | Republish in a thesis/dissertation |
| Order License ID | 1335318-1 | | |
| ISSN | 1473-3145 | Publisher | ROUTLEDGE |
| | | Portion | Image/photo/illustration |

## LICENSED CONTENT

| | | | |
|---|---|---|---|
| Publication Title | Counselling and psychotherapy research | Publication Type | Journal |
| | | Start Page | 347 |
| Article Title | Developing clarity on the coaching-counselling conundrum: Implications for counsellors and psychotherapists | End Page | 355 |
| | | Issue | 4 |
| | | Volume | 18 |
| Author/Editor | British Association for Counselling and Psychotherapy. | URL | http://www.tandfonline.com/loi/rcpr20 |
| Date | 01/01/2001 | | |
| Language | English | | |
| Country | United Kingdom of Great Britain and Northern Ireland | | |
| Rightsholder | John Wiley & Sons - Books | | |

## REQUEST DETAILS

| | | | |
|---|---|---|---|
| Portion Type | Image/photo/illustration | Distribution | Worldwide |
| Number of Images / Photos / Illustrations | 1 | Translation | Original language of publication |
| Format (select all that apply) | Electronic | Copies for the Disabled? | No |
| | | Minor Editing Privileges? | Yes |
| Who Will Republish the Content? | Academic institution | Incidental Promotional Use? | No |
| Duration of Use | Life of current edition | Currency | EUR |
| Lifetime Unit Quantity | Up to 999 | | |
| Rights Requested | Main product and any product related to main product | | |

## NEW WORK DETAILS

| | | | |
|---|---|---|---|
| Title | Opportunities and challenges for digital interventions in the prevention of depression | Institution Name | Technical University Munich |
| | | Expected Presentation Date | 2023-11-15 |
| Instructor Name | Janika Thielecke | | |

# B

## ABOUT THE AUTHOR

Janika Thielecke is a researcher and psychologist interested in improving access to and effectiveness of digital interventions to improve mental health. Janika was born and raised near Bonn, Germany. She obtained her university entrance qualification (Abitur) from the Nicolaus-Cusanus Gymnasium in Bonn in the bilingual branch (German – English) in 2011. She holds a Bachelor and Master degree in psychology (finished with honors in 2016) from the RWTH Aachen University. She worked as a research associate in the field of occupational rehabilitation at the RWTH Aachen from 2016-2018 and investigated personality factors that influences return-to-work in different settings.

Janika joined the Chair of Clinical Psychology and Psychotherapy at the Friedrich-Alexander-Universität of Erlangen-Nürnberg as a PhD candidate in the E-Mental Health research group in 2018. In the project "With us in balance" she was responsible for the randomized controlled trial evaluating the effectiveness of a personalized telephone coaching for farmers.

In 2021, Janika also joined the research group for Psychology & Digital Mental Health Care at the Technical University Munich to finish her PhD there. There she mainly conducted secondary data analysis on trials evaluating online interventions and focused on aspect to enhance the effectiveness of digital mental health interventions. She was a lecturer for introductory psychology courses at FOM Nürnberg (University of Applied Science) from 2018-2021.

Janika is a trained course instructor for stress management and relaxation techniques with a passion for teaching.