PHARMACOLOGICAL TREATMENTS

Estimating the smallest worthwhile difference of antidepressants: a cross-sectional survey

Ethan Sahker,[1,2] Toshi A Furukawa [ORCID],[1] Yan Luo,[1,2] Manuela L Ferreira,[3] Kaori Okazaki,[4] Astrid Chevance [ORCID],[5,6] Sarah Markham,[7,8] Roger Ede,[7] Stefan Leucht [ORCID],[9] Andrea Cipriani [ORCID],[10,11,12] Georgia Salanti [ORCID] [13]

ABSTRACT

Background Approximately 30% of patients experience substantial improvement in depression after 2 months without treatment, and 45% with antidepressants. The smallest worthwhile difference (SWD) refers to an intervention's smallest beneficial effect over a comparison patients deem worthwhile given treatment burdens (harms, expenses and inconveniences), but is undetermined for antidepressants.

Objective Estimating the SWD of commonly prescribed antidepressants for depression compared to no treatment.

Methods The SWD was estimated as a patient-required difference in response rates between antidepressants and no treatment after 2 months. An online cross-sectional survey using Prolific, MQ Mental Health and Amazon Mechanical Turk crowdsourcing services in the UK and USA between October 2022 and January 2023 garnered participants (N=935) that were a mean age of 44.1 (SD=13.9) and 66% women (n=617).

Findings Of 935 participants, 124 reported moderate-to-severe depressive symptoms but were not in treatment, 390 were in treatment and 495 reported absent-to-mild symptoms with or without treatment experiences. The median SWD was a 20% (IQR=10–30%) difference in response rates for people with moderate-to-severe depressive symptoms, not in treatment, and willing to consider antidepressants, and 25% (IQR=10–35%) for the full sample.

Conclusions Our observed SWDs mean that the current 15% antidepressant benefit over no treatment was sufficient for one in three people to accept antidepressants given the burdens, but two in three expected greater treatment benefits.

Implications While a minority may be satisfied with the best currently available antidepressants, more effective and/or less burdensome medications are needed, with more attention given to patient perspectives.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Acute treatment of depression with antidepressant medications can benefit patients with an increased response over placebo, with ORs ranging between 1.5 and 1.9.

⇒ Antidepressant treatments are associated with burdens including side effects, expenses and may involve other inconveniences.

⇒ Patient preferences and trade-offs between the benefits and harms of pharmacological interventions for treating depression have never been estimated.

WHAT THIS STUDY ADDS

⇒ People with moderate-to-severe depressive symptoms who are not in treatment considered a 20 percentage point (IQR=10–30%) difference in treatment response rates, for antidepressant treatments over no treatment, to be worthwhile for initiating antidepressant treatment.

⇒ This study showed that one in three people would consider the efficacy of the antidepressant treatment worthwhile despite its burdens. However, the other two out of three would need to see greater efficacy or fewer burdens to consider the treatment.

⇒ There was large variability in expectations of worthwhile differences, but demographic and clinical variables were not materially associated with these estimates.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The smallest worthwhile difference (SWD), not the minimum important change, should inform the between-group difference to be aimed for in new antidepressant developments and to be detected in antidepressant trials.

⇒ We should consider the SWD in balancing the benefits and harms of antidepressant treatments both in individual decision-making and in guideline or policy developments.

⇒ We need more research estimating SWD for various treatments for depression (eg, psychotherapies, electroconvulsive therapies), for other mental disorders and for other diseases.

BACKGROUND

Depression is the second leading cause of disability worldwide[1] with millions seeking treatment through antidepressant medications.[2] Antidepressants are proven efficacious, backed by hundreds of randomised controlled trials and rigorous meta-analyses.[3 4] For example, a network meta-analysis including 21 commonly prescribed antidepressants

demonstrated response odds ratios (ORs) ranging between 1.49 and 1.85 favouring antidepressants over placebo.[3] However, effect size measures such as ORs miss the patient's perspective on the significance of intervention benefits.[5] The patient-deemed worthiness of antidepressant treatments, given the symptom improvement benefits on the one hand and the burdens (harms, expenses and inconveniences) on the other, remains controversial.[6 7]

The minimum important change (MIC) can help determine whether changes in health outcomes over time are important from the patient viewpoint. The MIC, also known as the minimal important difference or minimal important clinical difference, is the smallest change after treatment in a health outcome perceived as important.[8] Defining the MIC is a useful way to interpret patient-reported outcome measures.[9] By definition, the MIC is specific to a particular instrument or outcome measure,[10 11] generally lacks association with an intervention[12] and does not explicitly account for burdens and benefits relative to an alternative.[10 12] The MIC for depression scales is estimated to be a 6-point reduction for the Beck Depression Inventory-II[13] and a 7-point to 8-point reduction for the Hamilton Depression Rating Scale 17-item.[14] Estimates of the MIC concern intraindividual change and the mean change seen in participants in antidepressant trials (both on placebo and antidepressants) is usually larger than the MIC.[13 14]

A conceptually different approach to facilitate interpretation of patient importance in the context of an intervention is to estimate the smallest worthwhile difference (SWD). The SWD is 'the smallest beneficial effect of an intervention that justifies the costs, risks and inconveniences of that intervention' over a treatment alternative.[10] The SWD represents a between-treatments assessment reflecting a trade-off of the benefits and burdens of two treatment options. It is patient-derived, intervention-specific, control-specific and expressed as an absolute difference between treatment options.[10 12] Two methods have been proposed to estimate the SWD. The discrete choice experiment asks individuals their preferences for hypothetical scenarios where benefits and costs vary. Regression models determine the preference threshold for one treatment over another.[15] The benefit-harm trade-off method (BHTM) asks individuals to state their preferences for scenarios, but benefits vary, while burdens remain constant. The BHTM ascertains how many benefits people are willing to trade-off for the expected burdens of one intervention over another.[16] The BHTM is easy to understand and has been used to estimate the SWD in treatments for respiratory disease,[12] fall prevention[15] and pain reduction therapies.[17] Despite the debate about whether the effect of antidepressants is large enough to justify their burdens, the SWD of antidepressants has never been estimated.[3 18]

OBJECTIVE

The present study aims to estimate the SWD of commonly prescribed antidepressants for depressive disorders compared with no treatment/natural course using the BHTM. We hypothesise that the SWD estimates may vary by participants' clinical needs and current/past-treatment experiences. Our primary focus was on people who were currently depressed but are not receiving treatment. We also surveyed people in treatment and those without depressive symptoms. This study aimed to help doctors understand patient expectations and establish evidence-based benchmarks for clinical trials testing new antidepressant therapies.

METHODS

Online supplemental 1 presents the prespecified protocol approved by the Kyoto University Graduate School of Medicine Ethics Committee on 9 September 2022 (R3574-1, changes in online supplemental 2), and all participants provided e-consent.

Study design

We conducted a cross-sectional survey using three research participant crowdsourcing services (RPCSs): Prolific, and MQ Mental Health (MQ) and Amazon Mechanical Turk (MTurk) between October 2022 and January 2023. We invited participation through RPCSs and linked participants fitting the inclusion criteria to an online survey. While Prolific and MTurk represent a general internet population, the MQ participant pool includes mainly people with lived experiences in mental health and healthcare professionals who volunteer to improve research representation. RPCS participants generally demonstrate high test–retest reliability and high convergent and concurrent validity in psychological tests,[19] although there remain concerns of careless or fraudulent responses.[20] Following previous methods to increase data quality,[19 20] we restricted MTurk participants to those in the USA rated by researchers as a 95%+approval and graded at the 'masters' level; the highest vetting. Prolific participants were limited to the UK or USA. MQ participants were limited to the UK only. These RPCSs provide monetary compensation for participation from their established participant pools. The compensation was £1.20 for Prolific, US$1 for MTurk and voluntary for MQ. Remuneration was commensurate with similar length RPCS studies, based on time-to-completion, and above the average amount as evaluated by Prolific.

Primary outcome and its measurement

Our primary outcome was the SWD, representing the patient-preferred efficacy of depression treatment with antidepressants that would be deemed worthwhile compared with no treatment, given the treatment burdens (harms, expenses and inconveniences).

We estimated the individual participant preference through the BHTM. We presented a major depressive episode symptoms summary and explained the benefits and burdens of antidepressant treatment and the no treatment/natural course alternative based on The Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) and US Federal Drug Administration (FDA) descriptors (online supplemental 3). We used selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs) in the antidepressants description because they are most frequently prescribed[21] and have similar tolerability/efficacy profiles.[3] We estimated that response, defined as 50% or greater reduction in depression severity, occurs in approximately 30 out of 100 people after 2 months without any treatment.[22–24]

Next, we asked participants if they believed antidepressant treatment was worthwhile given variable hypothetical response rates for antidepressants compared with the response rate for no treatment (ie, 30%) after 2 months. We asked them to weigh the benefits and burdens, then decide if they would accept the drug. Based on their yes/no response, we next asked the participants about lower or higher response rates, respectively. To eschew a default to heuristics and laziness, we randomly assigned participants to two different response algorithms, requiring different attention and concentration levels (online supplemental 4). As we observed no difference in SWD by algorithm, in the following we report the combined results. Then, the difference between

30% and the minimum antidepressant response the participant would consider taking defines the individual's SWD.

Demographic and clinical variables

We collected demographic information including gender, race, education, employment, country of residence and insurance status. Insurance was categorised as Affordable Care Act (USA only), Medicare/Medicaid (USA only), national healthcare insurance, private health insurance, other and uninsured.

Clinical variables included lifetime depression prevalence, family history of depression, lifetime antidepressant treatment, lifetime psychotherapy, current antidepressant treatment, current psychotherapy, treatment preference (antidepressants vs psychotherapy) and current depression symptoms. Current depression symptom severity was assessed with the Patient Health Questionnaire-9 (PHQ-9).

Participants

We included participants aged 18 or older, residing in the UK or USA, who were fluent in English. We were interested in a general population, but we were most concerned with people experiencing depressive symptoms of at least moderate severity (PHQ-9≥10) but not currently engaged in any treatment. This group most closely resembles potential treatment seekers, and we can expect them to provide more accurate estimates of the SWD for a major depressive episode as depicted in the provided clinical scenario because of their current experiences and potential treatment needs. This group would represent the best estimate of a clinical sample taken from a general internet population. Thus, to explore how different experiences with depression and treatment could be associated with treatment-seeking judgements and SWD estimates, we included participants with four differing profiles, (1) *moderate-to-severe depressive symptoms but not in treatment*: PHQ-9≥10 and not receiving any treatment, the primary interest group for SWD estimation, (2) *currently in treatment*: ongoing antidepressant therapy or psychotherapy, (3) *absent-to-mild depressive symptoms with treatment experiences*: zero-to-mild depression symptoms (PHQ-9<10), no current treatment, but previous antidepressant treatment and (4) *absent-to-mild depressive symptoms without treatment experiences*: zero-to-mild depression symptoms and no current or previous antidepressant treatment.

Sample size

We set the sample size to achieve the expected precision in the estimate of the SWD. We assumed that the SD of the SWD would resemble that estimated in a pain study (SD=22 for SWDs of 20%)[17] because it is an SWD study investigating a subjective outcome, and no depression SWD study exists from which to base a power analysis. To obtain a 95% CI within 10 percentage points, we needed approximately 80 participants with at least moderate symptoms but not in treatment. We also anticipated similar precision for the three other groups as well. Estimates of depression incidence in RPCSs vary but are demonstrably higher than in the general population.[19] We therefore assumed approximately 20% of participants would present with moderate-to-severe depressive symptoms on the PHQ-9. Depending on these subgroup populations, approximately 800 participants may be necessary to reach n=80 in the groups with the smallest populations (online supplemental 5). Recruitment was stopped after all four groups included 80 participants. We accepted responses with no missing outcome variables.

Data analysis

We first presented the SWD with its distribution and percentile ranking for people with *moderate-to-severe depressive symptoms but not in treatment* and estimated its median and IQR. We then examined the SWD by participant groups for comparison using box-violin plots. Finally, we analysed the entire sample's SWD using demographic and clinical independent variables in 11 univariable regressions and a single multivariable regression investigating SWD predictors using the least absolute shrinkage and selection operator (LASSO) method.[25] Participants who reported they would not accept antidepressant treatment, even if the response were 100%, were removed from the primary analysis because they would not be real-world candidates for antidepressant treatment.[12] They may be philosophically opposed to taking mood-altering drugs, or any drugs. To examine the effect of this decision, we conducted sensitivity analyses by assigning them an SWD=71, which is an impossible value representing an antidepressant response rate over 100%. We used SAS V.9.4 (Cary, North Carolina, USA, SAS Institute) and R V.4.2.2 (R Core Team, 2022) for all statistical analyses.

Patient and public involvement

We piloted the BHTM script with two members of the patient and public involvement (PPI) group at the Oxford Precision Psychiatry Lab of the University of Oxford. PPI members (SM and RE) reviewed descriptions of a DSM-5 major depressive episode, and the benefits and burdens of antidepressant treatments. These experts by experience provided feedback on clarity, inclusivity and accuracy of patient experiences, and we modified the scripts accordingly. They further collaborated in the interpretation and write-up of the manuscript.

FINDINGS

Participants

The total sample included 935 participants from three RPCSs: MTurk (n=255), Prolific (n=395) and MQ (n=285). Participants had a mean age of 44.1 (SD=13.9), were mostly women (66%) and Caucasian (84%). Table 1 lists the demographic and clinical characteristics for those with *moderate-to-severe symptoms but not in treatment* and for the full sample.

Ninety-five participants (10.2%; n=20 with *moderate-to-severe symptoms but not in treatment*, n=20 *currently in treatment*, n=14 with *absent-to-mild symptoms with treatment experiences* and n=41 with *absent-to-mild symptoms without treatment experiences*) reported that they would not consider taking antidepressants even if these drugs achieved 100% response.

SWD for people with moderate-to-severe depressive symptoms but not in treatment

Figure 1 shows the distribution and percentile rank of the SWD as reported by participants with *moderate-to-severe symptoms but not in treatment*. The median was 20% (IQR=10–30%, n=104).

SWD for participant groups according to depression symptoms and treatment experience

Table 2 shows the average and dispersion data of the SWD for the four groups and the full sample. Distributions of individual estimates overlapped to a great degree across the groups, but dispersion varied considerably (figure 2). The participants *currently in treatment* comprised the largest proportion of the sample (44%) and showed the largest dispersion.

Table 1 Demographic and clinical characteristics of participants

	Moderate-to-severe depressive symptoms but not in treatment		Full sample	
	n=124		n=935	
	n	% or mean	n	% or mean
Demographic variables				
Age mean (SD)	124	40.2 (SD:14.9) range: 19–91	935	44.1 (SD: 13.9) range: 18–91
Female	74	59.7	617	66.0
Race				
White/Caucasian	98	79.0	785	84.0
Black/African American	8	6.5	48	5.1
Asian	3	2.4	35	3.7
Latino/Hispanic	6	4.8	31	3.3
Multiracial	7	5.7	24	2.6
Other	2	1.6	12	1.3
Education				
Less than high school	3	2.4	27	2.9
High school graduate/equivalent	20	16.1	123	13.2
Some college	34	27.4	250	26.7
2-year degree	9	7.3	111	11.9
4-year degree	50	40.3	314	33.6
Master's degree	8	6.5	90	9.6
Doctorate	0	0.0	20	2.1
Employment				
Disabled	6	4.8	65	7.0
Homemaker	5	4.0	38	4.1
Retired	8	6.5	86	9.2
Student	5	4.0	37	4.0
Unemployed and looking	16	12.9	48	5.1
Working full-time	58	46.8	478	51.1
Working part-time	26	21.0	183	19.6
Country				
UK	71	57.3	569	60.9
USA	53	42.7	366	39.1
Insurance				
Affordable Care Act/Obamacare	5	4.0	21	2.3
Medicare/Medicaid	14	11.3	81	8.7
National healthcare insurance	43	34.7	364	38.9
Private health insurance	25	20.2	237	25.4
Other	1	0.8	17	1.8
Uninsured	36	29.0	215	23.0
Clinical variables				
PHQ-9 score	124	15.2 (SD: 4.6) range: 10–27	935	8.4 (SD: 7.0) range: 0–27
Lifetime depression	67	54.0	521	55.7
Family history of depression	74	59.7	552	59.0
Lifetime antidepressant treatment	64	51.6	537	57.4
Lifetime psychotherapy	80	64.5	595	63.6
Current antidepressant treatment	0	0.0	339	36.3
Current psychotherapy	0	0.0	155	16.6
Treatment preference				
Antidepressants	48	38.7	339	36.3
Psychotherapy	76	61.3	596	63.7
Unwilling to take antidepressants even if the response rate were 100%	20	16.1	95	10.2

PHQ-9, Patient Health Questionnaire-9.

To further investigate potential correlates of the SWD, we conducted univariable regressions and a single multivariable linear regression between all baseline covariates and the SWD (online supplemental 6). LASSO indicated that only treatment preference could be an important predictor of the SWD. Participants who preferred antidepressants at outset reported a lower SWD (median=20%, IQR=10–35%) than those who preferred psychotherapy (median=25%, IQR=15–35%).

As a sensitivity analysis, we assigned an SWD of 71% to participants unwilling to ever accept antidepressants, which increased the median SWD for the people with *moderate-to-severe symptoms but not in treatment* (median=25%, IQR=15–42.5%,
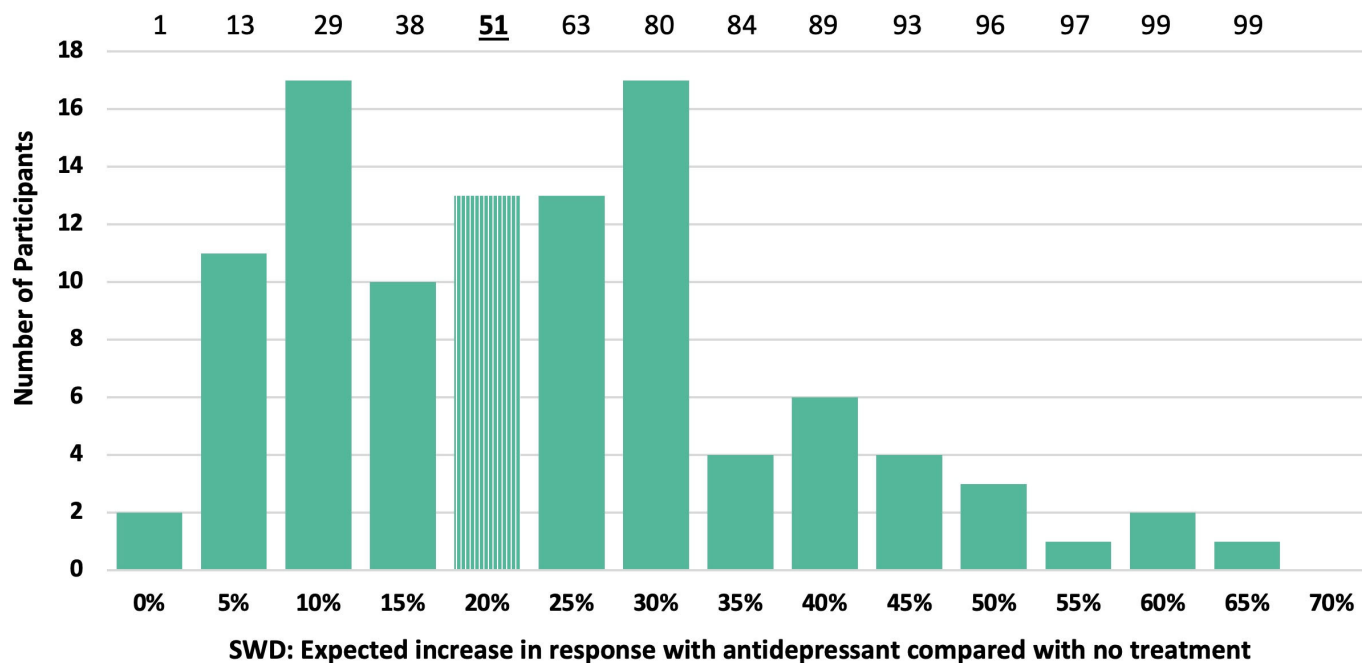
Figure 1 Frequency distribution of the smallest worthwhile difference (SWD) for people with moderate-to-severe depressive symptoms but not in treatment (n=104). The top row shows the cumulative percentiles of the distribution. The stripped bar represents the median SWD.

n=124) and for the entire sample (median=30%, IQR=15–40%, n=935).

DISCUSSION

We used the BHTM to estimate the SWD for initiation of antidepressant treatment for depression in an online cross-sectional survey. The median SWD among participants with *moderate-to-severe depressive symptoms but not in treatment* was an additional 20 percentage points (IQR: 10–30%) over the assumed natural response rate of 30% for no treatment. Other groups showed a median of 25% with considerable dispersion: People *currently in treatment* showed the largest variability (IQR=10–40%), and people with *absent-to-mild symptoms without treatment experiences* showed the smallest variability (IQR=20–30%). About 10% of the participants reported they would not take antidepressants even if these drugs achieved 100% response. Treatment preference, either for drug therapy or for psychotherapy, was the only important predictor of the SWD (median of 20% vs 25%, respectively). These findings, and their implications for the relevance with existing antidepressants, should be interpreted considering the natural response rate of no treatment (assumed to be 30% in our study) and the average greater response rate of antidepressants (assumed to be about 45% according to the literature, see below).

We queried the expected difference in antidepressants response rates versus the no treatment natural course because placebo is not a real-world treatment alternative. However, there are a very limited number of studies examining natural response rates among depressed seeking-treatment patients. The survival curve of 393 incident depressive episodes observed in a naturalistic community cohort study suggests a response rate between 10% and 40% after 2 months.[23] Compared with placebo, a systematic review comprising 252 randomized controlled trials (RCTs) showed the average response rate was 37%.[22] Some evidence suggests that placebo produces greater response than no treatment, due to the placebo effect. For example, the OR for pill placebo response over no treatment was 1.38 (0.75–2.55) in a network meta-analysis in depression psychotherapy trials.[26] One large pragmatic trial compared watchful waiting versus antidepressants among primary care patients and found watchful waiting response rate to be 29%.[24] Taken together, we conservatively set the no treatment response rate to be 30%.

A network meta-analysis of 522 randomised trials of antidepressants showed that the response ORs for SSRIs and SNRIs over placebo ranged between 1.5 and 1.9.[3] Assuming placebo response rates of 30 or 40%, these ORs would translate into antidepressants response rates of 39–45% or 50–56%, constituting risk differences between 9% and 16%. However, it is generally believed that more responsive patients are selectively enrolled in RCTs, and antidepressant real-world effectiveness may be smaller than these 9–16% risk differences. Alternatively, a greater probability of receiving placebo in RCTs is associated with lesser response rates for the same antidepressant, with an observed response rate difference up to 10%.[27] In practice we have no placebo condition, and the antidepressant response rate may be somewhat higher than the above estimations. A recent individual participant data meta-analysis of 73 388 patients from 232 placebo controlled RCTs reported an overall 15% risk difference for antidepressants over placebo when individual

Table 2 The SWD for participant groups according to depression symptoms and treatment experience

	Sample size	Median	IQR
Groups			
Moderate-to-severe depressive symptoms but not in treatment	104	20%	10–30%
Currently in treatment	370	25%	10–40%
Absent-to-mild depressive symptoms with treatment experiences	91	25%	15–35%
Absent-to-mild depressive symptoms without treatment experiences	275	25%	20–30%
Total	840	25%	10–35%

Figure 2 Box-violin plots of the SWD distributions for four participant groups. Boxes represent the second and third quartiles with the median line in the middle. Whiskers represent the range. Violins represent the frequency distribution, enhanced with raw data points. The sample used to estimate the figure does not include 95 participants who reported they would not take antidepressants for depression even if response were 100%. SWD, smallest worthwhile difference.

response patterns were considered.[28] Everything considered, we assume 15% as the realistic response rate difference between the currently most efficacious antidepressants and no treatment/natural course in the following discussion.

The estimated SWD of 20% (IQR: 10–30%) for people with *moderate-to-severe symptoms but not in treatment,* or 25% (IQR: 10–35%) for the entire sample, are greater than the 15% greater response rate to be expected on currently available best antidepressant drugs. However, there was wide variability in individual SWDs, both for the participants with *moderate-to-severe symptoms but not in treatment* and for the entire sample. Approximately one in three (40/104, or 276/840) would be willing to take antidepressants for a depressive episode at the currently expected response rate (ie, 15 percentage points greater than no treatment), in exchange for the potential burdens. Another one-third would need double the current antidepressant effect (ie, 30%) before they initiate antidepressant treatment. The remaining third would need to see greater response rates or fewer burdens.

To explore potential sources of variability in the SWD values, we examined all demographic and clinical variables. Only treatment preference demonstrated a robust association. Preconceived notions about antidepressants can affect confidence in therapies and motivation to seek psychiatric treatment. Depressed but not yet treated patients would likely be seeking treatment in the real world. Our findings suggest that these people may show the smallest average SWD, but the group SWD distributions largely overlapped. The fact that the SWD estimates did not substantively differ among those with or without clinical needs, and with or without lived experiences, corroborate the appropriateness of our method co-produced by people with lived experiences. The observed substantial variability in the participants' requirements to accept antidepressants, coupled with the increasing prescription rates,[22] highlights the need for an explanation of the high SWD observed in the current study. We can only speculate, but perhaps there is a lack of understanding from patients and a lack of communication from doctors (understating burdens

or overstating efficacy) that factor into hasty decision-making in prescription acceptance. Much more research is needed in this area about actual transactions in the real world and their appropriateness.

The SWD may provide a benchmark for efficacy to be expected for future antidepressant drugs over placebo. The median SWD of a 20–25 percentage points greater response rate than the 30% for no treatment would correspond with an OR between 2.3 and 2.9 and an SMD between 0.47 and 0.58[29] (online supplemental 7 for calculations). Unfortunately, a systematic review of currently approved general medicine and psychiatric drugs shows that only a minority of drugs have evidence for this magnitude of efficacy.[30] It should also be noted that this SMD of 0.5–0.6 would translate into a mean difference of 3–4 points on the Hamilton Rating Scale for Depression, because the average SD of this scale is around 7 points.[3] The MIC of the same rating scale is reported to be 7–8,[14] and is therefore much larger than the SWD. Thus, the MIC is both conceptually and pragmatically distinct from the SWD. We must use the SWD, not the MIC, as a benchmark to evaluate the observed group differences in RCTs and to estimate the sample size of future trials.

Limitations

There were some noteworthy limitations. Foremost is our use of an RPCS sample, which may not represent real-world populations.[19] While the SWD from a clinical sample is important and may potentially produce a different SWD, evaluation of a general population serves as foundation for further research in the perception of people with lived experiences. Our study focused on people observed in an RPCS sample with moderate-to-severe depression symptoms and represents a group who may or may not access psychiatric services. Participants recruited through RPCSs tend to be younger, more educated and report greater psychopathology, when compared with the general population.[20] However, RPCSs produce greater generalisability than convenience samples.[20] We balanced our population by including

two RPCSs based in two different countries and we also sampled from MQ Mental Health, a UK-based charity specifically aimed at improving mental health outcomes research. By sampling the RPCSs with performance restrictions, we were able to focus on those with more pressing clinical needs and evaluate other groups with potentially different needs.[20] The fact that the two different sequences of questions in the BHTM method yielded the same estimates attests to the attentiveness and trustworthiness of our internet collaborators.

Second, we excluded participants who replied that they would not take antidepressants even if antidepressant response rates were 100%, while the expected natural response rate was 30%. We followed established methods citing they would not be candidates for treatment in clinical practice.[12] A sensitivity analysis including them with an improbable SWD, reflecting greater than 100% response, raised the estimate by 5 percentage points. Third, participants with *moderate-to-severe symptoms but not in treatment* were not necessarily diagnosed with major depression. We asked them to estimate the SWD for people with diagnosed depression as depicted in the clinical scenario. We assumed that while not diagnosed, they might provide personally invested SWD estimates because of their current depressive symptoms and potential treatment needs. Finally, systematic differences in depression treatment burdens must be considered. Policies differ between national healthcare systems, and between individuals within one country, based on insurance status and access to mental health services. The indirect costs (eg, loss in productivity) due to depression may also differ between individuals, and between societies. To evaluate these differences, we conducted exploratory analyses and found no appreciable association between the SWD and individual variables such as age, sex, race, nationality or insurance status.

Clinical implications

One-in-three of those who might consider antidepressant treatment and experience at least moderate depression symptoms would find the currently available antidepressants worthwhile, in exchange for the expected burdens of treatment including side effects, expenses and other inconveniences. Two-in-three require greater response rates or fewer burdens. There was great variability in individual SWD estimates but this variability could not be explained by demographic or clinical variables. Thus, while a minority may be satisfied with the best currently available antidepressants, more effective and/or less burdensome medications are needed. This is the first evaluation of the effectiveness that patients require for initiating antidepressant treatment. Greater value must be placed on the patient perspective for antidepressants in the treatment of depression. This information can help clinicians and researchers to understand patient expectations of therapies, evaluate the worthiness of antidepressants and establish evidence-based benchmarks for all medical research.

Author affiliations

[1]Department of Health Promotion and Human Behavior, Graduate School of Medicine / School of Public Health, Kyoto University, Kyoto, Japan
[2]Population Health and Policy Research Unit, Medical Education Center, Graduate School of Medicine, Kyoto University, Kyoto, Japan
[3]The University of Sydney, Sydney Musculoskeletal Health and The Kolling Institute, School of Health Sciences, Faculty of Medicine and Health, Sydney, New South Wales, Australia
[4]School of Medicine, Kyoto University, Kyoto, Japan
[5]Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), Paris, France
[6]Centre d'Epidémiologie Clinique, AP-HP, Hôpital Hôtel Dieu, Paris, France
[7]Patient and Public Involvement (PPI) Group, Oxford Precision Psychiatry Lab, NIHR Oxford Health Biomedical Research Centre, University of Oxford, Oxford, UK
[8]Department of Biostatistics & Health Informatics, Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College, London, UK
[9]Department of Psychiatry and Psychotherapy, Technical University of Munich, School of Medicine, München, Germany
[10]Department of Psychiatry, University of Oxford, Oxford, UK
[11]Oxford Precision Psychiatry Lab, NIHR Oxford Health Biomedical Research Centre, Oxford, UK
[12]Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK
[13]Institute of Social and Preventive Medicine and Clinical Trials, University of Bern, Bern, Switzerland

Twitter Toshi A Furukawa @Toshi_FRKW, Astrid Chevance @ChevanceAstrid, Sarah Markham @DrSMarkham and Andrea Cipriani @And_Cipriani

Patient consent for publication Not applicable.

Ethics approval The present study was approved by the Kyoto University Graduate School of Medicine Ethics Committee (R3574), and all participants provided e-consent. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data collected for this study and the code used to analyse these data is available upon request from the corresponding author, for purposes of reproducing the results. The lead author affirms that the manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned have been explained.

ORCID iDs
Toshi A Furukawa http://orcid.org/0000-0003-2159-3776
Astrid Chevance http://orcid.org/0000-0002-0852-4322
Stefan Leucht http://orcid.org/0000-0002-4934-4352
Andrea Cipriani http://orcid.org/0000-0001-5179-8321
Georgia Salanti http://orcid.org/0000-0002-3830-8508

REFERENCES

1 GBD 2019 Mental Disorders Collaborators. Mental disorders collaborators. global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet Psychiatry* 2022;9:137–50.

2 Herrman H, Patel V, Kieling C, et al. Time for United action on depression: A lancet-world psychiatric Association Commission. *The Lancet* 2022;399:957–1022.

3 Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *The Lancet* 2018;391:1357–66.

4 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.

5 Furukawa TA, Scott I, Guyatt G, et al. Measuring patients' experience. In: Guyatt G, Rennie D, Meade M, eds. *Users' guides to the medical literature: a manual for evidence-based clinical practice*. New York: The McGraw-Hill Companies, Inc, 2014: 219–34.

6 Hengartner MP, Plöderl M. Statistically significant antidepressant-placebo differences on subjective symptom-rating scales do not prove that the drugs work: effect size and method bias matter *Front Psychiatry* 2018;9:517.

7 Moncrieff J. What does the latest meta-analysis really tell us about antidepressants *Epidemiol Psychiatr Sci* 2018;27:430–2.

8 Jaeschke R, Singer J, Guyatt GH. Measurement of health status ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407–15.

9 Carrasco-Labra A, Devji T, Qasim A, et al. Minimal important difference estimates for patient-reported outcomes: A systematic survey. *J Clin Epidemiol* 2021;133:61–71.

10 Ferreira ML, Herbert RD, Ferreira PH, et al. A critical review of methods used to determine the smallest worthwhile effect of interventions for low back pain. *J Clin Epidemiol* 2012;65:253–61.

11 Devji T, Carrasco-Labra A, Qasim A, et al. Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: instrument development and reliability study. *BMJ* 2020;369:m1714.

12 McNamara RJ, Elkins MR, Ferreira ML, et al. Smallest worthwhile effect of land-based and water-based pulmonary rehabilitation for COPD. *ERJ Open Res* 2015;1:00007-2015.

13 Hiroe T, Kojima M, Yamamoto I, et al. Gradations of clinical severity and sensitivity to change assessed with the Beck depression inventory-II in Japanese patients with depression. *Psychiatry Res* 2005;135:229–35.

14 Leucht S, Fennema H, Engel R, et al. What does the HAMD mean? *J Affect Disord* 2013;148:243–8.

15 Franco MR, Howard K, Sherrington C, et al. Smallest worthwhile effect of exercise programs to prevent falls among older people: estimates from benefit-harm trade-off and discrete choice methods. *Age Ageing* 2016;45:806–12.

16 Barrett B, Brown D, Mundt M, et al. Sufficiently important difference: expanding the framework of clinical significance. *Med Decis Making* 2005;25:250–61.

17 Ferreira ML, Herbert RD, Ferreira PH, et al. The smallest worthwhile effect of nonsteroidal anti-inflammatory drugs and Physiotherapy for chronic low back pain: a benefit-harm trade-off study. *J Clin Epidemiol* 2013;66:1397–404.

18 Kendrick T, Collinson S. Antidepressants and the serotonin hypothesis of depression. *BMJ* 2022;378:1993. 10.1136/bmj.o1993 Available: https://doi.org/10.1136/bmj.o1993

19 Chandler J, Shapiro D. Conducting clinical research using Crowdsourced convenience samples. *Annu Rev Clin Psychol* 2016;12:53–81.

20 Mortensen K, Hughes TL. Comparing Amazon's mechanical Turk platform to conventional data collection methods in the health and medical research literature. *J Gen Intern Med* 2018;33:533–8.

21 Luo Y, Kataoka Y, Ostinelli EG, et al. National prescription patterns of antidepressants in the treatment of adults with major depression in the US between 1996 and 2015: A population representative survey based analysis. *Front Psychiatry* 2020;11:35.

22 Furukawa TA, Cipriani A, Atkinson LZ, et al. Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies. *Lancet Psychiatry* 2016;3:1059–66.

23 Ten Have M, Penninx BWJH, Tuithof M, et al. Duration of major and minor depressive episodes and associated risk indicators in a psychiatric Epidemiological cohort study of the general population. *Acta Psychiatr Scand* 2017;136:300–12.

24 Kendrick T, Chatwin J, Dowrick C, et al. n.d. Randomised controlled trial to determine the clinical effectiveness and cost-effectiveness of selective serotonin reuptake inhibitors plus supportive care, versus supportive care alone, for mild to moderate depression with somatic symptoms in primary care. *Health Technol Assess*;13.

25 Tibshirani R. Regression shrinkage and selection via the lasso. *J Royal Statist Soc: Series B (Methodological)* 1996;58:267–88. 10.1111/j.2517-6161.1996.tb02080.x Available: https://rss.onlinelibrary.wiley.com/toc/25176161/58/1

26 Michopoulos I, Furukawa TA, Noma H, et al. Different control conditions can produce different effect estimates in psychotherapy trials for depression. *J Clin Epidemiol* 2021;132:59–70.

27 Salanti G, Chaimani A, Furukawa TA, et al. Impact of placebo arms on outcomes in antidepressant trials: systematic review and meta-regression analysis. *Int J Epidemiol* 2018;47:1454–64.

28 Stone MB, Yaseen ZS, Miller BJ, et al. Response to acute monotherapy for major depressive disorder in randomized, placebo controlled trials submitted to the US food and Drug Administration: individual participant data analysis. *BMJ* 2022:e067606.

29 Furukawa TA. From effect size into number needed to treat. *Lancet* 1999;353:1680.

30 Leucht S, Hierl S, Kissling W, et al. Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. *Br J Psychiatry* 2012;200:97–106.