

Review

# Contribution of Synthetic Data Generation towards an Improved Patient Stratification in Palliative Care

Waldemar Hahn <sup>1,†</sup>, Katharina Schütte <sup>2,†</sup>, Kristian Schultz <sup>3</sup>, Olaf Wolkenhauer <sup>3,4,5</sup> , Martin Sedlmayr <sup>1</sup>, Ulrich Schuler <sup>2</sup> , Martin Eichler <sup>6,7,8,9</sup> , Saptarshi Bej <sup>3,4,†</sup> and Markus Wolfien <sup>1,\*</sup> 

- <sup>1</sup> Institute for Medical Informatics and Biometry, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Fetscherstraße 74, 01307 Dresden, Germany
- <sup>2</sup> University Palliative Center, University Hospital Carl Gustav Carus, Technische Universität Dresden, Fetscherstraße 74, 01307 Dresden, Germany
- <sup>3</sup> Department of Systems Biology and Bioinformatics, University of Rostock, Universitätsplatz 1, 18051 Rostock, Germany
- <sup>4</sup> Leibniz-Institute for Food Systems Biology, Technical University Munich, 85354 Freising, Germany
- <sup>5</sup> Stellenbosch Institute of Advanced Study, Wallenberg Research Centre, Stellenbosch University, Stellenbosch 7602, South Africa
- <sup>6</sup> National Center for Tumor Diseases Dresden (NCT/UCC), Fetscherstraße 74, 01307 Dresden, Germany
- <sup>7</sup> German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
- <sup>8</sup> Faculty of Medicine, University Hospital Carl Gustav Carus, Technische Universität Dresden, Fetscherstraße 74, 01307 Dresden, Germany
- <sup>9</sup> Helmholtz-Zentrum Dresden-Rossendorf (HZDR), Bautzner Landstraße 400, 01328 Dresden, Germany
- \* Correspondence: markus.wolfien@tu-dresden.de
- † These authors contributed equally to this work.



**Citation:** Hahn, W.; Schütte, K.; Schultz, K.; Wolkenhauer, O.; Sedlmayr, M.; Schuler, U.; Eichler, M.; Bej, S.; Wolfien, M. Contribution of Synthetic Data Generation towards an Improved Patient Stratification in Palliative Care. *J. Pers. Med.* **2022**, *12*, 1278. <https://doi.org/10.3390/jpm12081278>

Academic Editor: William Duddy

Received: 4 July 2022

Accepted: 1 August 2022

Published: 4 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** AI model development for synthetic data generation to improve Machine Learning (ML) methodologies is an integral part of research in Computer Science and is currently being transferred to related medical fields, such as Systems Medicine and Medical Informatics. In general, the idea of personalized decision-making support based on patient data has driven the motivation of researchers in the medical domain for more than a decade, but the overall sparsity and scarcity of data are still major limitations. This is in contrast to currently applied technology that allows us to generate and analyze patient data in diverse forms, such as tabular data on health records, medical images, genomics data, or even audio and video. One solution arising to overcome these data limitations in relation to medical records is the synthetic generation of tabular data based on real world data. Consequently, ML-assisted decision-support can be interpreted more conveniently, using more relevant patient data at hand. At a methodological level, several state-of-the-art ML algorithms generate and derive decisions from such data. However, there remain key issues that hinder a broad practical implementation in real-life clinical settings. In this review, we will give for the first time insights towards current perspectives and potential impacts of using synthetic data generation in palliative care screening because it is a challenging prime example of highly individualized, sparsely available patient information. Taken together, the reader will obtain initial starting points and suitable solutions relevant for generating and using synthetic data for ML-based screenings in palliative care and beyond.

**Keywords:** palliative care; screening; personalized medicine; synthetic data generation; GANs

## 1. Introduction and Definition of Palliative Care

Patients with advanced, incurable cancer suffer from changing psychological and physical symptoms in terms of type and severity. In addition, there are social burdens for both the patient and for the informal caregivers. As per the definition of the World Health Organization (WHO) (<https://www.who.int/news-room/fact-sheets/detail/palliative-care> (accessed on 27 July 2022)) and extended via Radbruch et al. [1], Palliative care (PC)

uses a team-oriented approach to improve the quality of life of patients and their families who are facing problems associated with a life-threatening illness. It prevents and relieves suffering through the early identification, correct assessment, and treatment of pain and other problems, whether physical, psychosocial, or spiritual. Thus, it offers a support system to help patients live as actively as possible until death [1]. Furthermore, PC values patients' needs to receive adequate, personally, and culturally sensitive information on their health status to make independent decisions about a treatment [2]. Palliative care is applicable throughout all health care settings (place of residence and institutions) and in all levels (primary to tertiary care) [3]. Primary care is performed by general practitioners, oncologists, and in outpatient structures, as well as in hospitals [4,5]. Secondary palliative care involves palliative-care specialists acting as consultants and is offered to all patients with a symptomatic advanced, progressive life-threatening disease and limited therapeutic options [6]. Furthermore, most guidelines refer to this collective [3]. Over the past five decades, PC has evolved from serving patients at the end of life into a highly specialized discipline focused on delivering supportive care to patients with life-limiting illnesses throughout the disease trajectory [4]. Still, there are different perceptions about the timing of palliative care in the course of disease, including the difficulty of a reliable and timely screening [7].

To the best of our knowledge, the herein presented review combining the ideas of synthetic data generation and its potential utilization towards the screening of PC needs does not exist in the current literature. Thus, we here give an introduction into both fields for an initial conjunction and motivation for the use of this quickly evolving computational field within an important medical domain, which will raise the overall awareness and open up the discussion for such novel technologies in PC or related disciplines in personalized medicine.

### *1.1. Literature Screening Methodology*

We conducted our literature research in publicly available databases (PubMed, Scopus, Web of Science, Google Scholar) for the search terms of "Palliative Care" AND "Screening" AND "GANs" OR "VAE" OR "Generative Adversarial Net" OR "Variational Autoencoder" and found no matching results that could be attributed to the specific scope of this domain. As a result, we felt highly motivated to connect these two important topics.

### *1.2. Current Screening for Patients in Need for Palliative Care*

Commonly, there are two screening approaches to trigger a palliative-care referral: one is based on the patient's prognosis and the other focuses primarily on PC needs. The rationale for focusing on prognosis is that for most patients with advanced cancer symptoms, as well as others, the palliative care needs an increase within the last two months of life. The main indicators of this final phase are a poor general condition, weight loss, clinical symptoms (e.g., anorexia, breathlessness, or confusion), and abnormalities on laboratory parameters (e.g., high white cell count, lymphopenia, hyopalbuminemia, elevated lactate dehydrogenase, or C-reactive protein and Vitamin B12) [8]. The prognosis can also be derived from scores assessing physical disabilities and patient mortality based on comorbidities or the prevalence of symptoms, as well as other individual parameters [9]. A systematic review of studies using prognostic tools for identification showed that mainly five tools were evaluated for accuracy over eight studies [10]. Both sensitivity and specificity diverged widely (sensitivity 3% to 94%, specificity 26% to 99%). The authors conclude that the ability of current screening tools to identify patients with advanced progressive diseases who are likely to have palliative care needs is limited.

The current gold standard to screen for patients' needs is the Patient Reported Outcome Measurement (PROM) [11]. To date, several instruments are recommended for symptom assessment, e.g., MIDOS [12], ESAS [13], and IPOS [14], as well as the Distress Thermometer (DT) [15], which are described in the following. The Minimal Documentation System for Patients in Palliative Care (MIDOS) includes ten questions about distressing physical

symptoms and also anxiety and depression [12]. The Edmonton Symptom Assessment System (ESAS) queries eight distressing physical symptoms and includes mood and well-being [13]. On the distress thermometer, patients can indicate their psychological distress on a scale of zero to ten [15]. The Integrated Palliative Care Outcome Scale (IPOS) represents a combination of physical symptoms with those from the psychosocial domain [14]. An integration of the outcome from such PROMs into AI-based Clinical Decision Support Systems (CDSS) may provide a significant contribution towards the identification of PC needs, as was recently indicated by Sandham et al. [16]. To date, numerous studies have been conducted to test screening tools that combine prognostic criteria (e.g., diagnosis, functional status, complications, comorbidities) with symptoms and needs (symptom management, distress, and support of family) [17–20]. Study results show limitations in the feasibility of the tools due to time-consuming questionnaires [20]. Currently, the German SCREBEL trial compares a simple screening with a symptom assessment tool (IPOS and DT) to a more detailed assessment [17]. The era of electronic health records may facilitate referrals by providing electronic alerts, pre-populated note templates, and order sets [21]. A study performed in Würzburg currently applies the nursing history of the digital health record (nutritional status, weight loss, functional status, and unplanned hospital admissions) and combines it with PROM (ESAS, MiDOS, DT, IPOS) to assess symptoms [20]. Taken together, in routine clinical practice, screening should be reliable and require as few human resources as possible.

To summarize, PC is an interprofessional specialty to improve quality of life for patients and their families. Existing evidence supports that timely involvement of specialist PC teams can enhance the care delivered, but identification of patients in need of PC is insufficient. International guidelines of leading medical societies recommend performing screenings as well [22–25]. However, to date, no screening tools have been developed that identify reliably those patients with individual PC needs without requiring too many medical resources. For optimal screening, heterogeneous data from different domains should be used, including both disease phase and symptoms.

### *1.3. Data-Related Challenges That Limit A General Use of AI in Palliative Care*

Medical data are highly sensitive. They need proper protection and regulation. In general, data sharing is regulated under data privacy by the European General Data Protection Regulation (GDPR). With respect to the quickly evolving technology and all involved stakeholders, data sharing needs to be adequately and continuously improved by periodic adaptations of the implementations [2,26]. In terms of ethics, with the rise of novel technologies, such as Artificial Intelligence (AI), the problem also of re-identification from data, such as images and genomic information, becomes an essential aspect [27,28]. Thus, anonymization is one possibility to keep the data private. This is usually achieved by changing patient-specific identifiers through removal, substitution, distortion, generalization, or aggregation [29,30]. In contrast, data pseudonymization as another solution is a data management and de-identification procedure by which personally identifiable information fields within a data record are replaced by one or more artificial identifiers or pseudonyms [31]. Although sharing anonymized data meets the requirements of the GDPR, there have been incidents in the past where people of anonymized datasets were identified through linkage attacks [32,33]. To overcome the paucity of annotated medical data in real-world settings and (fully) save the patients' anonymity, synthetic data generation is being used more frequently in medicine and healthcare to increase the diversity in datasets and to enhance the robustness and adaptability of AI models [34]. To conform with ethical regulations in a research context, medical data should remain only available in a highly controlled manner and according to strict procedures (e.g., "systematic oversight" [35] or "embedded ethics" [36]). These points can be summarized as key challenges that need to be addressed:

1. *Clinical data are often few* [37,38]. The main contributing factors refer to the sparsity and scarcity of cases incident to a certain clinical problem, such as the need of palliative care that represents the very specific patient history
2. *Palliative care is a transient process and highly case specific*. There is an ongoing controversial debate on the most important parameters that are used to define and effectively screen the need for palliative care
3. *Patient data are subject to privacy issues* [2,27,28]. This hinders clinicians from sharing data with modelers, data scientists, and external clinical colleagues freely, even in an anonymized manner to improve patient classification

## 2. Existing and Prospective Applications of AI for Palliative Care

So far, research in artificial intelligence (AI) and machine learning (ML) dealing with PC have focused on survival prediction and mortality rates. To obtain an overview about these current developments, we briefly highlight and discuss the most prominent studies in the field. Random forests, feature selection, and logistic regression were applied to general patient electronic health records (EHR) [21]. In addition, a long short-term memory (LSTM) model was able to effectively predict mortality by using a combination of EHR data and administrative claims data [39]. A rapid review showed that ML approaches are powerful in predicting mortality in older and/or hospitalized adults [40]. Patients' outcome is dependent on the right timing of specialized PC referral. Palliative patients go through different phases of their disease (stable, unstable, deteriorating, terminal/dying, deceased) [41]. Data-driven ML and network analysis were expected to identify these phases through symptoms reported on IPOS [42]. ML was moderately successful to predict cases within phases. Precision-recall curves (PRCs) were calculated in addition to ROC area under curve (AUC). PRC figures decreased from stable to terminal, leading to reduced relevance of the model for the later stages due to greater proportions of patients being in earlier palliative stages [16]. Deep learning (DL), an area of ML that uses mathematical and statistical models, has also tried to predict mortality and beneficence from PC by using a combination of clinical features including disease diagnosis and patient demographics. A Deep Neural Network model was trained on the EHR data of patients from previous years, to predict the mortality of patients within the next 3–12 month period [43]. Another study used the information on symptom burden of free-text notes in the EHR [44]. Here, natural language processing (NLP) was able to identify hospitalized cancer patients with uncontrolled symptoms (pain, dyspnea, or nausea/vomiting) in the EHR. The accuracy was between 61% and 80% with low sensitivity for nausea/vomiting (21%) and dyspnea (22%). For this reason, this model also has to be further developed before it can be used to trigger early access to PC [44]. However, despite these existing success stories, specific screening tools or CDSS of patients in need for palliative care in early, intermediate, and late stages are missing because time-specific screening parameters and a reasonable amount of underlying data are not yet available to build such tools.

A starting point for important screening features can be obtained from the National Comprehensive Cancer Network (NCCN), which has proposed consensus criteria for screening of patients care needs and subsequent referral to specialized PC: (i) uncontrolled symptoms, (ii) moderate to severe distress related to cancer diagnosis and therapy, serious comorbid physical, psychiatric, and psychosocial conditions, (iii) life expectancy of six months or less, (iv) patient or family concerns about the disease course and decision-making, and/or (v) a specific request for palliative care by the patient or family [9]. Such a systematic screening can be carried out by using checklists [45–47]. These included different unspecified criteria like frequent hospital admission or hospital stays due to difficult-to-control symptoms, complex nursing care, or vast deterioration. In addition, there were more specific criteria like admission from a long-term care facility or medical foster home, chronic home oxygen use, current or past hospice program enrollee, limited social support, and a lack of an advance care planning document. Others used a checklist in patients with advanced cancer stage IV, including re-hospitalization in less than 30 days, hospitalization

longer than seven days, active symptoms of pain, nausea, vomiting, dyspnea, delirium, psychological distress [48]. Glare et al. [9] examined the use of six NCCN screening and further criteria (metastatic or locally advanced cancer, a limited prognosis, active source of suffering) and later included prolonged length hospital stay as an extra item [49]. Potential parameters for the screening of PC needs can thus be derived from the literature; however, the limited amount of available data across all facets is still missing.

As a supportive addition to sparse real-world data, novel synthetically generated data may serve PC in two different ways: (i) the model is trained using real-world clinical data and once trained, will not require any data in the future (fixed model approach), (ii) the model is constantly fed with data to generate synthetic data (continuous model approaches). There are three different categories of algorithms used in the generation of synthetic data: probabilistic models, machine learning, and deep learning methods. Currently, an implementation towards the field of PC screening is still missing.

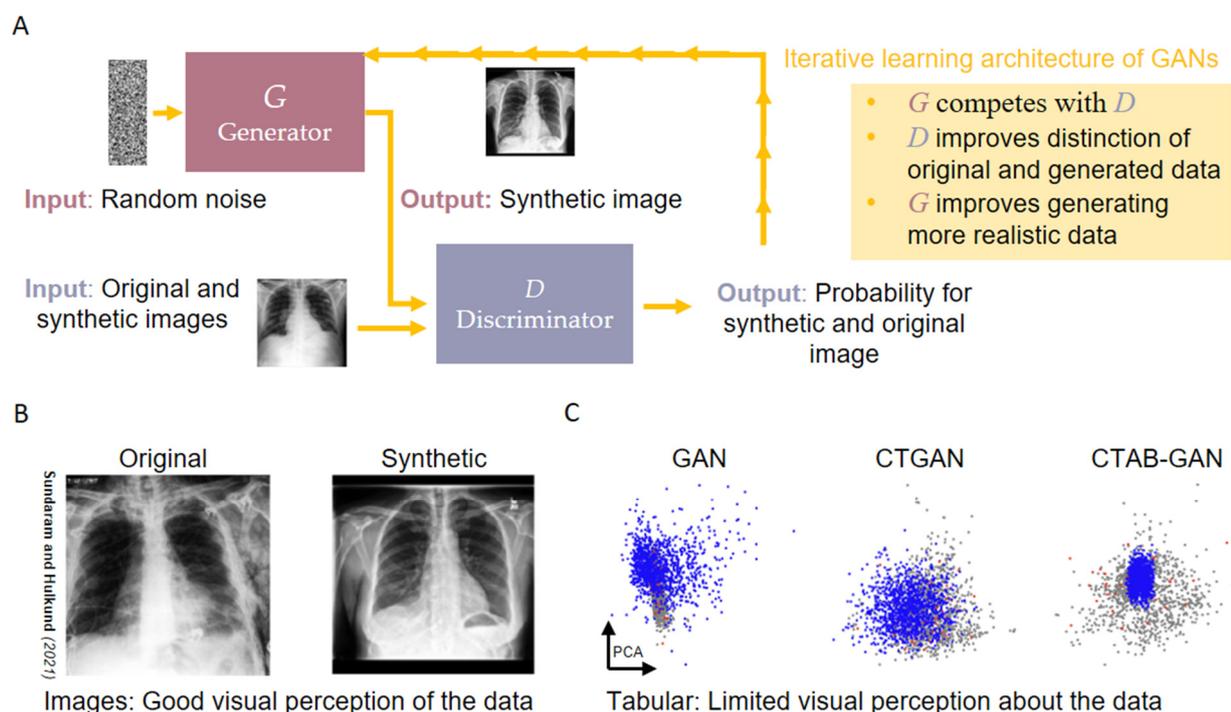
### 3. Potential Impact of Synthetic Data Generation Towards an Improved Identification of Patients in Need of Palliative Care

#### 3.1. Synthetic Data Generation via Generative Adversarial Networks

If only a small amount of data can be made available to the AI model, that oftentimes is not enough for optimizing, training, and testing a precise and robust decision support model at a clinical scale. Synthetic data generation would be a sensible approach to tackle this problem. Here, relevant medical data (pseudonymized, anonymized or actual) is used as an input for an ML-model to learn the underlying data structure, which is utilized in a subsequent step to generate new artificial data that is close to the original. Thus, instead of providing the AI model only with a small amount of data, a larger amount of synthetic data can be provided for the purpose to improve the training of ML-based decision support models, e.g., for patient stratification. Deep generative models, such as Variational Autoencoders (VAE) [50] and Generative Adversarial Networks (GAN) [51,52], play a key role in this. Although VAEs are also widely applied for generative modeling studies, especially with respect to sparse and scarce data in the medical/health domain for images [53,54] and data integration [55], relatively few examples for tabular data exist [56–58]. GANs are currently seen as most promising according to the findings of Xu et al. [57]. They see GANs as better suited for privacy preserving data generation in comparison to VAEs, since these are easier to integrate with respect to differential privacy. Several of such models have been developed over the past few years and a current technical review of Hernandez et al. [58] presents the different synthetic data generation methods for tabular healthcare datasets. A comparable work of Georges-Filteau and Cirillo investigates the possibility of synthetic data generation via GANs to ultimately obtain digital twins [59]. However, deep generative models are more popular for synthetic data generation from image datasets and there are only relatively few models relying on tabular patient data as yet [60].

Traditionally, a generative network for adversarial learning consists of a Generator  $G$  and a Discriminator  $D$  (Figure 1). The Generator is realized as  $G: N \rightarrow X$ , meaning that, ideally, the generative model  $G$  maps random noise to the data space,  $X$ . The Discriminator  $D: G(N) \rightarrow [0,1]$  ensures that the synthetic samples generated by the Generator  $G$  are realistic enough. The two neural networks  $G$  and  $D$  compete throughout the training process with  $G$  generating synthetic samples from random noise and  $D$  ensuring that with each iteration,  $G$  learns to generate more realistic synthetic samples. However, like every neural network model, GANs require a lot of data to be trained. Thus, for smaller tabular datasets, they are often not the best option for synthetic data generation. These might be addressed by specific linear interpolation-based algorithms that take explicitly rare cases into account. Interpolation-based methods applied to small data neighborhoods are commonly used in the context of imbalance tabular datasets of smaller size. Although these methods are developed in the context of synthetic data generation to tackle class imbalance, the underlying philosophy of linear interpolation can be applied to generate synthetic data from tabular datasets, in general. Imbalanced datasets are characterized by unequal distribution of samples over classes. Since some classes have fewer examples,

synthetic samples are generated for such classes to create balanced classifiers over such datasets. Our recently proposed algorithms LoRAS [61] and ProWRAS [62], as well as generalizations of the SMOTE algorithms [63], propose a way to control the variance of the synthetic samples by generating them as convex combinations of multiple shadowsamples (Gaussian noise added to original samples) from data neighborhoods. Tabular datasets typically have well-defined features following a distribution in every data neighborhood, ensuring a synthetic sample generated as a convex combination/weighted average, which is an unbiased estimator of the local mean for every feature distribution. Thus, challenges for synthetic data generation from small datasets still remain, but it is essential that these challenges are directly addressed by using real-world medical datasets to likewise identify further specific hurdles and finally ensure a versatile use on a clinical scale.



**Figure 1.** Main principle of the GAN architecture. (A) Refers to the iterative learning principle of Generative Adversarial Networks (GANs) using two Neural Networks (Generator  $G$ , and Discriminator  $D$ ). (B) Shows examples of synthetically generated images, adapted from Sundaram and Hulkund [64]. (C) Shows a current example of a current benchmark for synthetic tabular data, adapted from Schultz et al. [65].

### 3.2. Domain Level Challenges Concerning the Use of GANs for Clinical Problems

For an improved, realistic representation of the current limitations, we present data and domain-related challenges in PC to motivate the importance of the conducted research in this area: Firstly, the diverse data types that are usually present in clinical tabular data, i.e., continuous and categorical data can pose a challenge in model building. In particular, categorical data highly increase the complexity because they can be further divided into nominal and ordinal data-types. This requires the ML-model to handle potentially complex continuous and discrete distributions at the same time. Additionally, continuous features can follow different distributions and have multiple modes. Secondly, considering that synthetic data generation is a feasible solution to support data privacy, the development and comparison of algorithms, metrics, and protocols that can quantify how reliably the synthetic data represent the original data would be crucial for a practical realization. Finally, the usability and technical acceptance of clinicians using the developed models are often not adequately addressed right from the beginning.

(i) Since 2017, there exist multiple deep generative models focusing on synthetic data generation on tabular datasets. MedGAN, the first of such architectures, can handle either Boolean or count data [66]. After the initial release, there were several adaptations of this architecture to enable the generation of categorical values and to boost performance (e.g., changing the loss from vanilla GAN loss to Wasserstein loss) [52,67,68]. Another model, TableGAN, proposed shortly after, is based on deep convolutional GAN (DCGAN), uses an additional third neural network called classifier that predicts labels, and can generate numerical and categorical values [67]. TGAN or tabular GAN is yet another contemporary model that handles multiple modes in continuous variables through Gaussian Mixture Models (GMM) and can create categorical values with the help of gumbel softmax as activation function [69]. It also uses a LSTM as a generator. The authors also published an improved model in 2019, called CTGAN, which is based on the conditional GAN architecture, in which conditional vectors for categorical values are introduced [57]. In comparison to TGAN, CTGANs no longer use LSTMs as the generator network. A more complete list of different GAN architectures for tabular data that were published until the end of 2020 can be found in the work of Coutinho-Almeida et al. [70]. Since then, other GAN-based architectures were proposed [71–75]. Among these, one interesting recent work refers to CTAB-GAN from 2021, which combines the ideas of TableGAN and CTGAN [74]. It uses convolutions and a DCGAN architecture in addition to GMM and conditional vector construction. Additionally, it adds to the sampling mechanism a random selection of the mode of multi-modal continuous variables and can handle more data types. Besides GANs, other generative approaches, like Variational Autoencoders (VAE), Classification And Regression Trees (CART), Bayesian Networks (BN), and Copulas, etc., exist. However, the flexibility of GANs to handle complex distributions, and their success in the generation of other types of data (especially images) make it one of the most promising approaches for the generation of tabular data as well.

(ii) For tabular data, there is yet no consensus in science on how to evaluate synthetic data. Therefore, it is still an open research field. Loosely, existing evaluation metrics can be divided into four categories as follows:

Firstly, the statistical similarity of generated data can be compared to real data. Since the features are consistent and well defined across all data points, statistical hypothesis tests can be used to compare feature-wise distributions among the synthetic and original data. To measure the relationships between multiple features, pairwise correlation, k-way-marginals, or results of clustering approaches can be compared. Additionally, it is also possible to compare the similarity of the joined probability distributions of all features through metrics like Wasserstein distance, Kullback–Leibler Divergence (KL divergence), or Jason–Shannon Divergence (JSD).

Secondly, the generated data can be compared to real data with a specific task in mind. Usually, the task is to predict a specific feature given all the other features (ML efficiency). Therefore, the generation model is trained on a partition of the original data and afterwards, used to generate synthetic data. A predictive model is then trained on these data. Additionally, another model is trained on the same partition of the original data, which was used to train the generative model. Both predictive models are then compared on the test set of the original data.

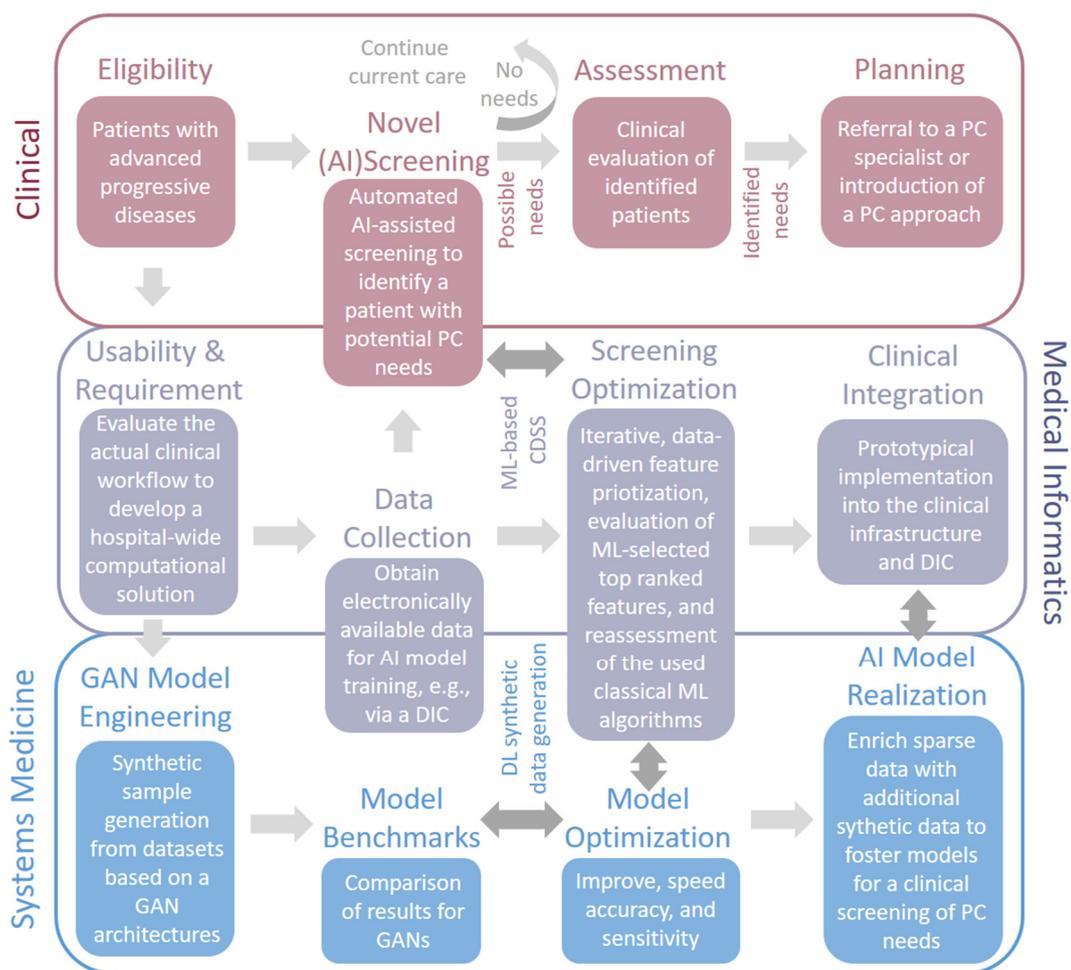
Thirdly, unsupervised learning approaches can also be used to assess the similarity of synthetic data in relation to the original data. In particular, tabular data contain diverse feature types, such as continuous (e.g., BMI, height—variables that can take theoretically any real value), ordinal (e.g., patients having hypertension or diabetes—categorical variables have a sense of order associated with them), and nominal (e.g., sex of a patient—categorical variables do not have a sense of order associated with them). Recent studies indicate that a conventional application of state-of-the-art dimension reduction algorithms, like UMAP, on such heterogeneous data lead to a biased embedding generation dimension reduction, in a sense, that the similarity among data with respect to the continuous features have a higher influence in the low dimensional embedding generation. A novel empirical

feature-distributed approach has been proposed by Bej et al. that accounts for this bias [76]. In brief, the method uses separate distance measures for available feature embeddings and finally, combines these into a single embedding, which is used to detect and visualize clusters in a more robust manner. This method could also be adopted to extract embeddings from the original data, which in combination with a supervised Neural Network, can be applied to assess the similarity between synthetic and original samples.

Finally, synthetic data can be evaluated regarding privacy. One possible way for this is to simulate membership inference attacks, where an attacker tries to predict which record was used for the training of the generative model. This can be done by calculating the distance to the closest record (DCR) in the real data for each synthetic record. Another way to evaluate the privacy of a given model is to perform attribute disclosure attacks, where the attacker uses a set of non-sensitive attributes to predict a sensitive attribute. To mitigate the risks of privacy leakage, several techniques like differential privacy were proposed.

(iii) A successful software implementation into the clinical routine requires in addition to the expert-in-the-loop and the knowledge of the technical infrastructure (e.g., clinical information system and stored data types) also the practical aspects of usability, feasibility, and technology acceptance. This can be achieved through user-centered design (UCD) processes, which involve clinicians as the later users in the development at an early stage because this highly facilitates the acceptance and user-friendliness [77–79]. A UCD is described, among other things, by DIN EN ISO 9241-210, “Processes for the design of usable systems” [80]. In general, at the beginning of the UCD, the application context and the exact user requirements must be specified to be understood. These activities are carried out by means of a user-centered requirement survey (stakeholder analysis). For example, guideline-supported interviews are used for stakeholder analysis and expert workshops for the development of prototypical user interfaces.

As shown above, open questions remain in the field of synthetic data generation and its application. In Figure 2, we demonstrate the potential synergistic activities and current developments in Systems Medicine and Medical Informatics to improve the clinical outcome of PC screening. The image highlights that all domains share specific, essential processes, such as Data Collection, Screening, and Optimization, because these processes need a more interdisciplinary approach. The overall aim should be an approach that can be offered to all identified patients with a symptomatic advanced, progressive life-threatening disease, and limited therapeutic options. The highest level of evidence relates to cancer patients but it would be not limited to these outcomes. Importantly, two thirds of advanced cancer patients have unmet palliative care needs [81]. Specialized algorithms for the generation of such heterogeneous tabular patient data would thus highly facilitate the early identification of actual patients leading to an actual clinical impact.



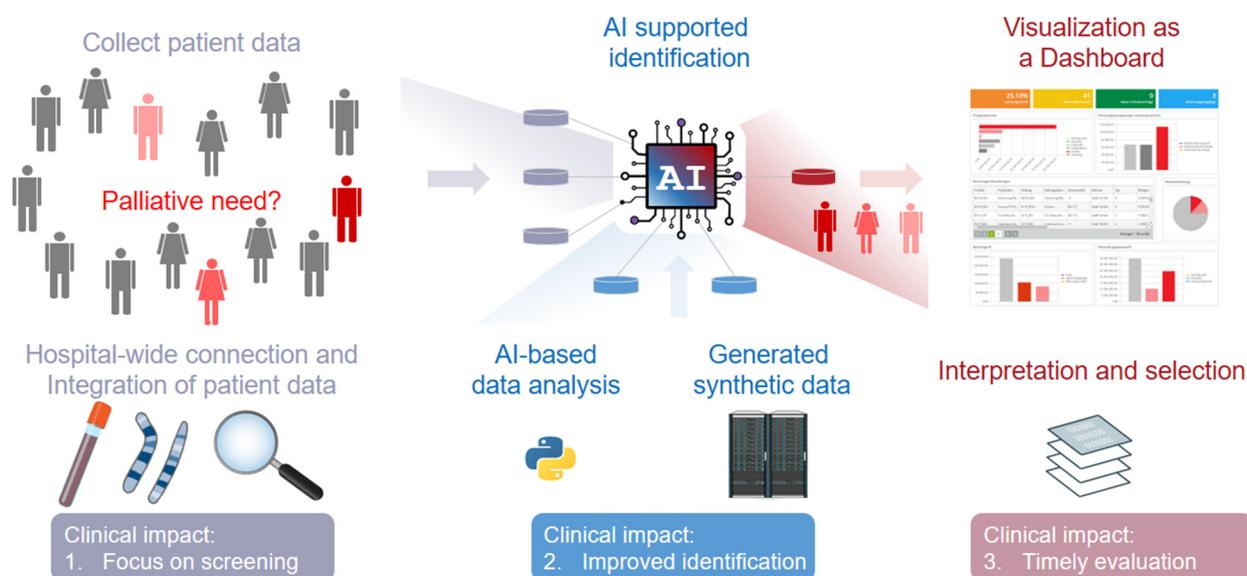
**Figure 2.** Scheme to illustrate the mutual cooperativity between the fields of Systems Medicine, Medical Informatics, and the clinical domain of Palliative Care (PC). The schematic representation is a content-wise extension of ElMokhallalati et al. [82], who only considered the clinical aspect. The specific introduction of Medical Informatics results in an advanced access of digitized medical data, e.g., through a Clinical IT Center like a Data Integration Center (DIC). Thus, underlying Artificial Intelligence (AI) approaches, i.e., Machine Learning (ML) and Deep Learning (DL), as well as Generative Adversarial Networks (GANs) for data generation, are able to foster the overall screening process in PC.

#### 4. Clinical Impact of AI and GAN-Based Screening Solutions in Palliative Care

To assign the presented AI-based methods towards a more specific clinical outcome for palliative care screening, we summarized the computational tasks and their attribution to potentially arising clinical impacts (Figure 3). Here, the interplay and synergistic effects of the involved research areas, namely, Medical Informatics, Systems Medicine, and Clinics, for the domain can be conceived on a broader scale.

##### 4.1. Clinical Impact 1: Set A Focus on Screening Rather Than Prognosis

ElMokhallalati et al. [82] conclude that existing screening tools are not adequate to represent palliative care needs, particularly because the focus is on prognosis. The rationale for focusing on prognosis is that for most patients with advanced cancer, symptoms, and thus PC, need to increase within the last two months of life [10]. Here, it would be of pivotal interest to explicitly identify those patients with an accurate screening rather than predict the remaining life expectancy or individual prognosis.



**Figure 3.** Scheme to illustrate the arising clinical impacts of synthetic data generation and Artificial Intelligence (AI) in general for palliative care. Such a concept would focus on patient screening and would use AI-based methods for an improved identification of patients, including a timelier screening (represented as different shades of red for the patients).

#### 4.2. Clinical Impact 2: Identification of Patients with Palliative Care Needs and Its Barriers

As already pointed out, two thirds of advanced cancer patients have unmet palliative care needs [81]. A study of inpatients with and without cancer revealed that 6.9% of them had palliative care needs, but only 2% of these received specialized palliative consultation. Especially, older patients without relatives who suffered from metastatic cancer and/or liver cirrhosis had the highest risk of developing PC needs. Often, those patients only request PC themselves if they have high symptom burden. However, patients are more likely to pursue specialized PC if recommended by their oncologist [83]. Of note, oncologists can also have over optimistic estimates of survival [10], a mistaken concern about a shortening of survival [49], a misconception of PC as synonymous with end-of-life care, as well as insecurities in the communication about PC, which often results in late referrals [23,84–87]. Therefore, a physician independent screening with a recommendation to re-evaluate the individual needs can significantly support physicians to improve the treatment of patients with unmet palliative care needs.

#### 4.3. Clinical Impact 3: Evaluation of the Correct Timing to Specialized Palliative Care

Early (within 2 to 3 months of diagnosis of advanced diseases) [88] provision of palliative care concomitant to life-prolonging treatment is associated with better quality of life, fewer depressive symptoms, less aggressive care at the end of life [88], and improved quality of life, symptom burden, and patient satisfaction compared to standard oncological care [22,24,25,89,90]. Contrarily, these patients are often in a good performance status [7,91–94]. A recent subgroup analysis of the early-integration Zimmermann trial showed that only patients with higher symptom burden at baseline derived a benefit from the palliative care intervention [95]. Although a too-late PC intervention may shorten survival and worsen quality of life [88], it is not possible to provide early PC for all patients with advanced disease due to the scarcity of resources [96]. Timely integration of PC is included in the European Society of Medical Oncology (ESMO) [24], as well as in the German-language palliative care guideline [97], the recommendations of the German Comprehensive Cancer Centers (CCC), and the American Society of Clinical Oncology (ASCO) recommendations for best oncology practice [7]. In summary, the importance of novel solutions is clearly given and needed.

## 5. Conclusions

Palliative care has evolved from serving patients only at the end of life into a highly specialized discipline focused on delivering supportive care to patients with life-limiting illnesses throughout their patient journey. This very individual track needs specific attention and awareness for a proper and timely screening, which is a time-intensive and domain-expertise-driven process that is difficult to achieve in clinical routine at all times. Therefore, a physician-independent automatic screening, supporting the physician's assessment, would be essential to improve the referral of patients with unmet palliative care needs. Current AI solutions already provide a well-suited tool set, but are still limited in terms of data availability and, thus, a versatile clinical applicability. A highly promising approach to filling this gap can be attributed to GAN-based synthetic data generation to provide AI classification models with an enriched set of anonymous, heterogeneous patient data to achieve likewise a high degree of data security and an accurate model performance. As was initially shown within this review article, synthetic data generation and PC have both so far only a limited number of common grounds. However, as other medical domains already show promising results and GANs are used more and more for data sharing in data sensitive domains, this review might contribute towards examples in the near future. In general, the high amount of methods and restricted consensus of evaluation metrics for synthetic data remain the main limitations that have to be solved from a computational perspective. In contrast, for PC, the main limitation is the availability of enough individual patient data, for which synthetic data could be one possible, existing solution. This novel combination can therefore lead to more precise AI-based models and finally, to improved clinical screening tools in palliative care.

**Author Contributions:** Conceptualization: W.H., K.S. (Katharina Schütte), M.E., S.B. and M.W.; Supervision: O.W., M.S. and U.S.; Visualization: W.H., K.S. (Katharina Schütte), O.W., S.B. and M.W.; Roles/Writing—original draft: W.H., K.S. (Katharina Schütte), M.E., S.B. and M.W.; Writing—review & editing: K.S. (Kristian Schultz), O.W., M.S. and U.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Radbruch, L.; De Lima, L.; Knaul, F.; Wenk, R.; Ali, Z.; Bhatnagar, S.; Blanchard, C.; Bruera, E.; Buitrago, R.; Burla, C.; et al. Redefining Palliative Care—A New Consensus-Based Definition. *J. Pain Symptom Manag.* **2020**, *60*, 754–764. [[CrossRef](#)] [[PubMed](#)]
2. Lopes, I.M.; Guarda, T.; Oliveira, P. General Data Protection Regulation in Health Clinics. *J. Med. Syst.* **2020**, *44*, 1–9. [[CrossRef](#)] [[PubMed](#)]
3. Kamal, A.H.; Bausewein, C.; Casarett, D.J.; Currow, D.C.; Dudgeon, D.J.; Higginson, I.J. Standards, Guidelines, and Quality Measures for Successful Specialty Palliative Care Integration into Oncology: Current Approaches and Future Directions. *J. Clin. Oncol.* **2020**, *38*, 987–994. [[CrossRef](#)] [[PubMed](#)]
4. Hui, D.; Bruera, E. Integrating palliative care into the trajectory of cancer care. *Nat. Rev. Clin. Oncol.* **2016**, *13*, 159. [[CrossRef](#)] [[PubMed](#)]
5. Rangachari, D.; Smith, T.J.; Kimmel, S. Integrating Palliative Care in Oncology: The Oncologist as a Primary Palliative Care Provider. *Cancer J.* **2013**, *19*, 373. [[CrossRef](#)]
6. Schenker, Y.; Arnold, R. The Next Era of Palliative Care. *JAMA* **2015**, *314*, 1565. [[CrossRef](#)]
7. Schenker, Y.; Crowley-Matoka, M.; Dohan, D.; Rabow, M.W.; Smith, C.B.; White, D.B.; Chu, E.; Tiver, G.A.; Einhorn, S.; Arnold, R.M. Oncologist Factors That Influence Referrals to Subspecialty Palliative Care Clinics. *J. Oncol. Pract.* **2014**, *10*, e37. [[CrossRef](#)]
8. Buurman, B.M.; Van Munster, B.C.; Korevaar, J.C.; Abu-Hanna, A.; Levi, M.; De Rooij, S.E. Prognostication in acutely admitted older patients by nurses and physicians. *J. Gen. Intern. Med.* **2008**, *23*, 1883–1889. [[CrossRef](#)]

9. Glare, P.; Plakovic, K.; Schloms, A.; Egan, B.; Epstein, A.S.; Kelsen, D.; Saltz, L. Study using the NCCN guidelines for palliative care to screen patients for palliative care needs and referral to palliative care specialists. *J. Natl. Compr. Cancer Netw.* **2013**, *11*, 1087–1096. [[CrossRef](#)]
10. Weissman, D.E.; Meier, D.E. Identifying patients in need of a palliative care assessment in the hospital setting: A consensus report from the Center to Advance Palliative Care. *J. Palliat. Med.* **2011**, *14*, 17–23. [[CrossRef](#)]
11. Trout, A.; Kirsh, K.L.; Peppin, J.F. Development and implementation of a palliative care consultation tool. *Palliat. Support. Care* **2012**, *10*, 171–175. [[CrossRef](#)] [[PubMed](#)]
12. Stiel, S.; Matthes, M.E.; Bertram, L.; Ostgathe, C.; Elsner, F.; Radbruch, L. Validierung der neuen Fassung des Minimalen Dokumentationssystems (MIDOS2) für Patienten in der Palliativmedizin: Deutsche Version der Edmonton Symptom Assessment Scale (ESAS). *Schmerz* **2010**, *24*, 596–604. [[CrossRef](#)] [[PubMed](#)]
13. Bruera, E.; Kuehn, N.; Miller, M.J.; Selmser, P.; Macmillan, K. The Edmonton Symptom Assessment System (ESAS): A simple method for the assessment of palliative care patients. *J. Palliat. Care* **1991**, *7*, 6–9. [[CrossRef](#)]
14. Murtagh, F.E.M.; Ramsenthaler, C.; Firth, A.; Groeneveld, E.I.; Lovell, N.; Simon, S.T.; Denzel, J.; Guo, P.; Bernhardt, F.; Schildmann, E.; et al. A brief, patient- and proxy-reported outcome measure in advanced illness: Validity, reliability and responsiveness of the Integrated Palliative care Outcome Scale (IPOS). *Palliat. Med.* **2019**, *33*, 1045–1057. [[CrossRef](#)] [[PubMed](#)]
15. Mehnert, A.; Lehmann, C.; Cao, P.; Koch, U. Die erfassung psychosozialer belastungen und ressourcen in der onkologie—Ein literaturüberblick zu screeningmethoden und entwicklungstrends. *PPmP Psychother. Psychosom. Med. Psychol.* **2006**, *56*, 462–479. [[CrossRef](#)] [[PubMed](#)]
16. Sandham, M.H.; Hedgecock, E.A.; Siegert, R.J.; Narayanan, A.; Hocaoglu, M.B.; Higginson, I.J. Intelligent Palliative Care Based on Patient-Reported Outcome Measures. *J. Pain Symptom Manag.* **2022**, *63*, 747–757. [[CrossRef](#)] [[PubMed](#)]
17. Alt-Epping, B.; Solar, S.; Lordick, F.; Mehnert-Theuerkauf, A.; Oechsle, K.; van Oorschot, B.; Thomas, M.; Asendorf, T. Niederschwelliges Screening versus multidimensionales Assessment von Symptomen und psychosozialen Belastungen bei Krebspatienten ab dem Zeitpunkt der Inkurabilität (SCREBEL). *Forum Fam. Plan. West. Hemisph.* **2020**, *35*, 143–144. [[CrossRef](#)]
18. Moss, A.H.; Lunney, J.R.; Culp, S.; Auber, M.; Kurian, S.; Rogers, J.; Dower, J.; Abraham, J. Prognostic significance of the “surprise” question in cancer patients. *J. Palliat. Med.* **2010**, *13*, 837–840. [[CrossRef](#)]
19. Bausewein, C.; Fegg, M.; Radbruch, L.; Nauck, F.; Von Mackensen, S.; Borasio, G.D.; Higginson, I.J. Validation and clinical application of the german version of the palliative care outcome scale. *J. Pain Symptom Manag.* **2005**, *30*, 51–62. [[CrossRef](#)]
20. Roch, C.; Heckel, M.; van Oorschot, B.; Alt-Epping, B.; Tewes, M. Screening for Palliative Care Needs: Pilot Data From German Comprehensive Cancer Centers. *JCO Oncol. Pract.* **2021**, *17*, e1584–e1591. [[CrossRef](#)]
21. Cava, W.L.; Bauer, C.; Moore, J.H.; Pendergrass, S.A. Interpretation of machine learning predictions for patient outcomes in electronic health records. *AMIA Annu. Symp. Proc.* **2019**, *2019*, 572. [[PubMed](#)]
22. Simon, S.T.; Pralong, A.; Radbruch, L.; Bausewein, C.; Voltz, R. The Palliative Care of Patients with Incurable Cancer. *Dtsch. Arztebl. Int.* **2020**, *117*, 108. [[CrossRef](#)] [[PubMed](#)]
23. Levy, M.H.; Adolph, M.D.; Back, A.; Block, S.; Codada, S.N.; Dalal, S.; Deshields, T.L.; Dexter, E.; Dy, S.M.; Knight, S.J.; et al. Palliative care. *J. Natl. Compr. Cancer Netw.* **2012**, *10*, 1284–1309. [[CrossRef](#)] [[PubMed](#)]
24. Hui, D.; Cherny, N.I.; Wu, J.; Liu, D.; Latino, N.J.; Strasser, F. Indicators of integration at ESMO Designated Centres of Integrated Oncology and Palliative Care. *ESMO Open* **2018**, *3*, e000372. [[CrossRef](#)] [[PubMed](#)]
25. Smith, T.J.; Temin, S.; Alesi, E.R.; Abernethy, A.P.; Balboni, T.A.; Basch, E.M.; Ferrell, B.R.; Loscalzo, M.; Meier, D.E.; Paice, J.A.; et al. American Society of Clinical Oncology provisional clinical opinion: The integration of palliative care into standard oncology care. *J. Clin. Oncol.* **2012**, *30*, 880–887. [[CrossRef](#)]
26. Copen, R.; Van Veen, E.B.; Groenewegen, P.P.; Hazes, J.M.W.; De Jong, J.D.; Kievit, J.; De Neeling, J.N.D.; Reijneveld, S.A.; Verheij, R.A.; Vroom, E. Will the trilogue on the EU Data Protection Regulation recognise the importance of health research? *Eur. J. Public Health* **2015**, *25*, 757. [[CrossRef](#)]
27. Murdoch, B. Privacy and artificial intelligence: Challenges for protecting health information in a new era. *BMC Med. Ethics* **2021**, *22*, 1–5. [[CrossRef](#)]
28. Liaw, S.T.; Liyanage, H.; Kuziemy, C.; Terry, A.L.; Schreiber, R.; Jonnagaddala, J.; de Lusignan, S. Ethical Use of Electronic Health Record Data and Artificial Intelligence: Recommendations of the Primary Care Informatics Working Group of the International Medical Informatics Association. *Yearb. Med. Inform.* **2020**, *29*, 51. [[CrossRef](#)]
29. Olatunji, I.E.; Rauch, J.; Katzensteiner, M.; Khosla, M. A Review of Anonymization for Healthcare Data. *Big Data* **2022**; online ahead of print. [[CrossRef](#)]
30. Csányi, G.M.; Nagy, D.; Vági, R.; Vadász, J.P.; Orosz, T. Challenges and Open Problems of Legal Document Anonymization. *Symmetry* **2021**, *13*, 1490. [[CrossRef](#)]
31. Zuo, Z.; Watson, M.; Budgen, D.; Hall, R.; Kennelly, C.; Al Moubayed, N. Data Anonymization for Pervasive Health Care: Systematic Literature Mapping Study. *JMIR Med. Inf.* **2021**, *9*, e29871. [[CrossRef](#)] [[PubMed](#)]
32. Narayanan, A.; Shmatikov, V. Robust de-anonymization of large sparse datasets. *Proc. IEEE Symp. Secur. Priv.* **2008**, 111–125. [[CrossRef](#)]
33. Douriez, M.; Doraiswamy, H.; Freire, J.; Silva, C.T. Anonymizing NYC taxi data: Does it matter? In Proceedings of the 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Montreal, QC, Canada, 17–19 October 2016; pp. 140–148. [[CrossRef](#)]

34. Chen, R.J.; Lu, M.Y.; Chen, T.Y.; Williamson, D.F.K.; Mahmood, F. Synthetic data in machine learning for medicine and healthcare. *Nat. Biomed. Eng.* **2021**, *5*, 493–497. [[CrossRef](#)] [[PubMed](#)]
35. Vayena, E.; Blasimme, A. Health Research with Big Data: Time for Systemic Oversight. *J. Law. Med. Ethics* **2018**, *46*, 119. [[CrossRef](#)] [[PubMed](#)]
36. McLennan, S.; Fiske, A.; Tigard, D.; Müller, R.; Haddadin, S.; Buyx, A. Embedded ethics: A proposal for integrating ethics into the development of medical AI. *BMC Med. Ethics* **2022**, *23*, 1–10. [[CrossRef](#)]
37. Bohr, A.; Memarzadeh, K. The rise of artificial intelligence in healthcare applications. *Artif. Intell. Healthc.* **2020**, *25*, 25–60. [[CrossRef](#)]
38. Meskó, B.; Görög, M. A short guide for medical professionals in the era of artificial intelligence. *NPJ Digit. Med.* **2020**, *3*, 1–8. [[CrossRef](#)]
39. Maragatham, G.; Devi, S. LSTM Model for Prediction of Heart Failure in Big Data. *J. Med. Syst.* **2019**, *43*, 1–13. [[CrossRef](#)]
40. Storick, V.; O’Herlihy, A.; Abdelhafeez, S.; Ahmed, R.; May, P. Improving palliative care with machine learning and routine data: A rapid review. *HRB Open Res.* **2019**, *2*, 13. [[CrossRef](#)] [[PubMed](#)]
41. Mather, H.; Guo, P.; Firth, A.; Davies, J.M.; Sykes, N.; Landon, A.; Murtagh, F.E.M. Phase of Illness in palliative care: Cross-sectional analysis of clinical data from community, hospital and hospice patients. *Palliat. Med.* **2018**, *32*, 404. [[CrossRef](#)] [[PubMed](#)]
42. Lind, S.; Wallin, L.; Fürst, C.J.; Beck, I. The integrated palliative care outcome scale for patients with palliative care needs: Factors related to and experiences of the use in acute care settings. *Palliat. Support. Care* **2019**, *17*, 561–568. [[CrossRef](#)]
43. Avati, A.; Jung, K.; Harman, S.; Downing, L.; Ng, A.; Shah, N.H. Improving palliative care with deep learning. *BMC Med. Inform. Decis. Mak.* **2018**, *18*, 55–64. [[CrossRef](#)] [[PubMed](#)]
44. Mashima, Y.; Tamura, T.; Kunikata, J.; Tada, S.; Yamada, A.; Tanigawa, M.; Hayakawa, A.; Tanabe, H.; Yokoi, H. Using Natural Language Processing Techniques to Detect Adverse Events From Progress Notes Due to Chemotherapy. *Cancer Inform.* **2022**, *21*, 11769351221085064. [[CrossRef](#)] [[PubMed](#)]
45. Swan, A.; Azhar, A.; Anderson, A.E.; Williams, J.L.; Liu, D.; Bruera, E. Empowering the Health and Well-Being of the Palliative Care Workforce: Evaluation of a Weekly Self-Care Checklist. *J. Pain Symptom Manag.* **2021**, *61*, 817–823. [[CrossRef](#)] [[PubMed](#)]
46. Kuosmanen, L.; Hupli, M.; Ahtiluoto, S.; Haavisto, E. Patient participation in shared decision-making in palliative care—An integrative review. *J. Clin. Nurs.* **2021**, *30*, 3415–3428. [[CrossRef](#)]
47. Forbat, L.; Chapman, M.; Lovell, C.; Liu, W.M.; Johnston, N. Improving specialist palliative care in residential care for older people: A checklist to guide practice. *BMJ Support. Palliat. Care* **2018**, *8*, 347–353. [[CrossRef](#)] [[PubMed](#)]
48. Tai, S.Y.; Lee, C.Y.; Wu, C.Y.; Hsieh, H.Y.; Huang, J.J.; Huang, C.T.; Chien, C.Y. Symptom severity of patients with advanced cancer in palliative care unit: Longitudinal assessments of symptoms improvement. *BMC Palliat. Care* **2016**, *15*, 1–7. [[CrossRef](#)]
49. Glare, P.A.; Chow, K. Validation of a Simple Screening Tool for Identifying Unmet Palliative Care Needs in Patients with Cancer. *J. Oncol. Pract.* **2015**, *11*, e81–e86. [[CrossRef](#)]
50. Kingma, D.P.; Welling, M. Auto-Encoding Variational Bayes. *arXiv* **2013**. [[CrossRef](#)]
51. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *Commun. ACM* **2014**, *63*, 139–144. [[CrossRef](#)]
52. Baowaly, M.K.; Lin, C.C.; Liu, C.L.; Chen, K.T. Synthesizing electronic health records using improved generative adversarial networks. *J. Am. Med. Inform. Assoc.* **2019**, *26*, 228–241. [[CrossRef](#)]
53. Elbattah, M.; Loughnane, C.; Guérin, J.L.; Carette, R.; Cilia, F.; Dequen, G. Variational Autoencoder for Image-Based Augmentation of Eye-Tracking Data. *J. Imaging* **2021**, *7*, 83. [[CrossRef](#)] [[PubMed](#)]
54. García-Ordás, M.T.; Benítez-Andrades, J.A.; García-Rodríguez, I.; Benavides, C.; Alaiz-Moretón, H. Detecting Respiratory Pathologies Using Convolutional Neural Networks and Variational Autoencoders for Unbalancing Data. *Sensors* **2020**, *20*, 1214. [[CrossRef](#)] [[PubMed](#)]
55. Simidjievski, N.; Bodnar, C.; Tariq, I.; Scherer, P.; Andres Terre, H.; Shams, Z.; Jamnik, M.; Liò, P. Variational Autoencoders for Cancer Data Integration: Design Principles and Computational Practice. *Front. Genet.* **2019**, *10*, 1205. [[CrossRef](#)] [[PubMed](#)]
56. Akrami, H.; Aydore, S.; Leahy, R.M.; Joshi, A.A. Robust Variational Autoencoder for Tabular Data with Beta Divergence. *arXiv* **2020**. [[CrossRef](#)]
57. Xu, L.; Skoularidou, M.; Cuesta-Infante, A.; Veeramachaneni, K. Modeling Tabular data using Conditional GAN. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 7335–7345. [[CrossRef](#)]
58. Hernandez, M.; Epelde, G.; Alberdi, A.; Cilla, R.; Rankin, D. Synthetic data generation for tabular health records: A systematic review. *Neurocomputing* **2022**, *493*, 28–45. [[CrossRef](#)]
59. Georges-Filteau, J.; Cirillo, E. Synthetic Observational Health Data with GANs: From slow adoption to a boom in medical research and ultimately digital twins? *arXiv* **2020**. [[CrossRef](#)]
60. Goncalves, A.; Ray, P.; Soper, B.; Stevens, J.; Coyle, L.; Sales, A.P. Generation and evaluation of synthetic patient data. *BMC Med. Res. Methodol.* **2020**, *20*, 1–40. [[CrossRef](#)]
61. Bej, S.; Davtyan, N.; Wolfien, M.; Nassar, M.; Wolkenhauer, O. LoRAS: An oversampling approach for imbalanced datasets. *Mach. Learn.* **2021**, *110*, 279–301. [[CrossRef](#)]
62. Bej, S.; Schulz, K.; Srivastava, P.; Wolfien, M.; Wolkenhauer, O. A Multi-Schematic Classifier-Independent Oversampling Approach for Imbalanced Datasets. *IEEE Access* **2021**, *9*, 123358–123374. [[CrossRef](#)]

63. Han, H.; Wang, W.Y.; Mao, B.H. Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning. In Proceedings of the International Conference on Intelligent Computing, Hefei, China, 23–26 August 2005; pp. 878–887. [\[CrossRef\]](#)
64. KammounAmina; SlamaRim; TabiaHedi; OuniTarek; AbidMohmed Generative Adversarial Networks for face generation: A survey. *ACM Comput. Surv.* **2021**. [\[CrossRef\]](#)
65. Schultz, K.; Bej, S.; Hahn, W.; Wolfien, M.; Srivastava, P.; Wolkenhauer, O. Convex space learning improves deep-generative oversampling for tabular imbalanced classification on smaller datasets. *arXiv* **2022**. [\[CrossRef\]](#)
66. Choi, E.; Biswal, S.; Malin, B.; Duke, J.; Stewart, W.F.; Org, S.; Sun, J. Generating Multi-label Discrete Patient Records using Generative Adversarial Networks. In Proceedings of the 2nd Machine Learning for Healthcare Conference, Boston, MA, USA, 18–19 August 2017. [\[CrossRef\]](#)
67. Park, N.; Mohammadi, M.; Gorde, K.; Jajodia, S.; Park, H.; Kim, Y. Data Synthesis based on Generative Adversarial Networks. *Proc. VLDB Endow.* **2018**, *11*, 1071–1083. [\[CrossRef\]](#)
68. Camino, R.D.; Hammerschmidt, C.A.; State, R. Generating Multi-Categorical Samples with Generative Adversarial Networks. *arXiv* **2018**. [\[CrossRef\]](#)
69. Xu, L.; Veeramachaneni, K. Synthesizing Tabular Data using Generative Adversarial Networks. *arXiv* **2018**. [\[CrossRef\]](#)
70. Coutinho-Almeida, J.; Rodrigues, P.P.; Cruz-Correia, R.J. GANs for Tabular Healthcare Data Generation: A Review on Utility and Privacy. *Lect. Notes Comput. Sci.* **2021**, *12986*, 282–291. [\[CrossRef\]](#)
71. Wen, B.; Colon, L.O.; Subbalakshmi, K.P.; Chandramouli, R. Causal-TGAN: Generating Tabular Data Using Causal Generative Adversarial Networks. *arXiv* **2021**. [\[CrossRef\]](#)
72. Kim, J.; Jeon, J.; Lee, J.; Hyeong, J.; Park, N. OCT-GAN: Neural ODE-based Conditional Tabular GANs. *arXiv* **2021**, *10*. [\[CrossRef\]](#)
73. Kunar, A.; Birke, R.; Zhao, Z.; Chen, L. DTGAN: Differential Private Training for Tabular GANs. *arXiv* **2021**. [\[CrossRef\]](#)
74. Zhao, Z.; Kunar, A.; Van der Scheer, H.; Birke, R.; Chen, L.Y. CTAB-GAN: Effective Table Data Synthesizing. *arXiv* **2021**. [\[CrossRef\]](#)
75. Tantipongpipat, U.T.; Waites, C.; Boob, D.; Siva, A.A.; Cummings, R. Differentially Private Synthetic Mixed-Type Data Generation For Unsupervised Learning. *Intell. Decis. Technol.* **2019**, *15*, 779–807. [\[CrossRef\]](#)
76. Bej, S.; Sarkar, J.; Biswas, S.; Mitra, P.; Chakrabarti, P.; Wolkenhauer, O. Identification and epidemiological characterization of Type-2 diabetes sub-population using an unsupervised machine learning approach. *Nutr. Diabetes* **2022**, *12*, 1–11. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Marcy, T.W.; Kaplan, B.; Connolly, S.W.; Michel, G.; Shiffman, R.N.; Flynn, B.S. Developing a Decision Support System for Tobacco Use Counseling Using Primary Care Physicians. *Inform. Prim. Care* **2008**, *16*, 101. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Fraccaro, P.; Casteleiro, M.A.; Ainsworth, J.; Buchan, I. Adoption of clinical decision support in multimorbidity: A systematic review. *JMIR Med. Inform.* **2015**, *3*, e3503. [\[CrossRef\]](#) [\[PubMed\]](#)
79. Brunner, J.; Chuang, E.; Goldzweig, C.; Cain, C.L.; Sugar, C.; Yano, E.M. User-centered design to improve clinical decision support in primary care. *Int. J. Med. Inform.* **2017**, *104*, 56–64. [\[CrossRef\]](#) [\[PubMed\]](#)
80. Jokela, T. Assessments of Usability Engineering Processes: Experiences from Experiments. In Proceedings of the 36th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 6–9 January 2003.
81. Hui, D.; Meng, Y.-C.; Bruera, S.; Geng, Y.; Hutchins, R.; Mori, M.; Strasser, F.; Bruera, E. Referral Criteria for Outpatient Palliative Cancer Care: A Systematic Review. *Oncologist* **2016**, *21*, 895–901. [\[CrossRef\]](#)
82. ElMokhallalati, Y.; Bradley, S.H.; Chapman, E.; Ziegler, L.; Murtagh, F.E.M.; Johnson, M.J.; Bennett, M.I. Identification of patients with potential palliative care needs: A systematic review of screening tools in primary care. *Palliat. Med.* **2020**, *34*, 989–1005. [\[CrossRef\]](#)
83. Gaertner, J.; Wolf, J.; Hallek, M.; Glossmann, J.P.; Voltz, R. Standardizing integration of palliative care into comprehensive cancer therapy—a disease specific approach. *Support. Care Cancer* **2011**, *19*, 1037–1043. [\[CrossRef\]](#)
84. *Erweiterte S3-Leitlinie Palliativmedizin für Patienten mit Einer Nicht-Heilbaren Krebserkrankung*; Kohlhammer: Stuttgart, Germany, 2020; ISBN 978-3-17-038390-6.
85. Adelson, K.; Paris, J.; Horton, J.R.; Hernandez-Tellez, L.; Ricks, D.; Morrison, R.S.; Smith, C.B. Standardized Criteria for Palliative Care Consultation on a Solid Tumor Oncology Service Reduces Downstream Health Care Use. *J. Oncol. Pract.* **2017**, *13*, e431–e438. [\[CrossRef\]](#)
86. Ostgathe, C.; Wendt, K.N.; Heckel, M.; Kurkowski, S.; Klein, C.; Krause, S.W.; Fuchs, F.S.; Bayer, C.M.; Stiel, S. Identifying the need for specialized palliative care in adult cancer patients—development and validation of a screening procedure based on proxy assessment by physicians and filter questions. *BMC Cancer* **2019**, *19*, 646. [\[CrossRef\]](#) [\[PubMed\]](#)
87. Bentler, S.E.; Liu, L.; Obrizan, M.; Cook, E.A.; Wright, K.B.; Geweke, J.F.; Chrischilles, E.A.; Pavlik, C.E.; Wallace, R.B.; Ohsfeldt, R.L.; et al. The aftermath of hip fracture: Discharge placement, functional status change, and mortality. *Am. J. Epidemiol.* **2009**, *170*, 1290–1299. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Davis, M.P.; Temel, J.S.; Balboni, T.; Glare, P. A review of the trials which examine early integration of outpatient and home palliative care for patients with serious illnesses. *Ann. Palliat. Med.* **2015**, *4*, 99–121. [\[CrossRef\]](#) [\[PubMed\]](#)
89. Haun, M.W.; Estel, S.; Rücker, G.; Friederich, H.C.; Villalobos, M.; Thomas, M.; Hartmann, M. Early palliative care for adults with advanced cancer. *Cochrane Database Syst. Rev.* **2017**, *6*, 6. [\[CrossRef\]](#) [\[PubMed\]](#)
90. Berendt, J.; Stiel, S.; Simon, S.T.; Schmitz, A.; van Oorschot, B.; Stachura, P.; Ostgathe, C. Integrating Palliative Care into Comprehensive Cancer Centers: Consensus-Based Development of Best Practice Recommendations. *Oncologist* **2016**, *21*, 1241–1249. [\[CrossRef\]](#)

91. Tewes, M.; Rettler, T.; Wolf, N.; Hense, J.; Schuler, M.; Teufel, M.; Beckmann, M. Predictors of outpatients' request for palliative care service at a medical oncology clinic of a German comprehensive cancer center. *Support. Care Cancer* **2018**, *26*, 3641–3647. [[CrossRef](#)]
92. Zimmermann, C.; Swami, N.; Krzyzanowska, M.; Leighl, N.; Rydall, A.; Rodin, G.; Tannock, I.; Hannon, B. Perceptions of palliative care among patients with advanced cancer and their caregivers. *CMAJ* **2016**, *188*, E217–E227. [[CrossRef](#)]
93. Kavalieratos, D.; Corbelli, J.; Zhang, D.; Dionne-Odom, J.N.; Ernecoff, N.C.; Hanmer, J.; Hoydich, Z.P.; Ikejiani, D.Z.; Klein-Fedyshin, M.; Zimmermann, C.; et al. Association Between Palliative Care and Patient and Caregiver Outcomes: A Systematic Review and Meta-analysis. *JAMA* **2016**, *316*, 2104–2114. [[CrossRef](#)]
94. Bennardi, M.; Diviani, N.; Gamondi, C.; Stüssi, G.; Saletti, P.; Cinesi, I.; Rubinelli, S. Palliative care utilization in oncology and hemato-oncology: A systematic review of cognitive barriers and facilitators from the perspective of healthcare professionals, adult patients, and their families. *BMC Palliat. Care* **2020**, *19*, 1–17. [[CrossRef](#)] [[PubMed](#)]
95. Okuyama, T.; Akechi, T.; Yamashita, H.; Toyama, T.; Nakaguchi, T.; Uchida, M.; Furukawa, T.A. Oncologists' recognition of supportive care needs and symptoms of their patients in a breast cancer outpatient consultation. *Jpn. J. Clin. Oncol.* **2011**, *41*, 1251–1258. [[CrossRef](#)]
96. Gaertner, J.; Siemens, W.; Meerpohl, J.J.; Antes, G.; Meffert, C.; Xander, C.; Stock, S.; Mueller, D.; Schwarzer, G.; Becker, G. Effect of specialist palliative care services on quality of life in adults with advanced incurable illness in hospital, hospice, or community settings: Systematic review and meta-analysis. *BMJ* **2017**, *357*, 2925. [[CrossRef](#)]
97. Hui, D.; De La Cruz, M.; Mori, M.; Parsons, H.A.; Kwon, J.H.; Torres-Vigil, I.; Kim, S.H.; Dev, R.; Hutchins, R.; Liem, C.; et al. Concepts and definitions for “supportive care”, “best supportive care”, “palliative care”, and “hospice care” in the published literature, dictionaries, and textbooks. *Support. Care Cancer* **2013**, *21*, 659–685. [[CrossRef](#)] [[PubMed](#)]