DOI: 10.1561/9781638284772.ch9

Chapter 9

Differential Privacy and Medical Data Analysis

By Vinith M. Suriyakumar, Nicolas Papernot and Anna Goldenberg

9.1 Introduction

Medical data analysis is crucial to advancing our understanding of human health and biology. Analyzing electronic health records (EHR) has helped provide insights into disease trajectories [Raj+18], lab and test efficacy [Gha+17], racial disparities [CSG19], and hospital operations [Wan+19]. Medical image analysis for x-rays, magnetic resonance images (MRIs), computed tomography (CTs), positron emission tomography (PET) have provided improved understanding of disease [SLMG20] and potential improvements to screening protocols to help detect disease earlier [TBL17]. Finally, analyzing *omics* data such as genomics, proteomics, and metabolomics have given us a much deeper of understanding of the biological mechanisms that underly disease progression [Mob+18], disease inheritance, and drug efficacy [Hon+18].

Medical data is incredibly personal and sensitive thus it requires strong privacy protections. Regulations around the world help govern the mechanisms used to protect the privacy of medical data. These regulations set the different levels of privacy required depending on who is accessing the data and the purpose of the data access. The current standard practice amongst most regulations is anonymization. However, this mechanism for protecting data privacy is not robust to a number of

attacks such as reconstruction attacks, differential attacks, and linkage attacks. The reader is referred to Chapters 1 and 5 for more details on privacy issues arising in statistical data release and machine learning systems, respectively.

In this chapter, we discuss regulations that govern the privacy of medical data for a variety of countries around the world and examples of medical data privacy breaches. We discuss why differential privacy is a promising framework for the next gold standard in medical data privacy. Finally, we discuss conducting both statistical analyses and machine learning on medical data with guarantees of differential privacy.

We survey a number of different case studies for different important statistical tasks in medicine. These include survival analysis, cohort identification, variant lookup, and genome-wide association studies. We present custom differentially private algorithms developed by researchers for these tasks, some of which provide optimal privacy-utility tradeoffs.

For machine learning, we present case studies of prediction and data synthesis across across different data modalities including electronic health records, medical images, and genomics data. We discuss some studies that show extreme loss of utility when incorporating differential privacy and others that have found minimal loss in utility. We contrast these studies to show some of the fundamental technical challenges that need work to help move differentially private machine learning to deployment in medicine. Given the global scale of medical research and the siloed nature of medical data we discuss applications of differentially private federated learning to medical data. It is clear that as research progresses, DP distributed learning will help learn higher utility models by allowing hospitals around the world to privately collaborate on model training.

Finally, we discuss the current challenges of applying differential privacy to medical data analysis and future opportunities for advancing applications of differential privacy in medical data analysis. This discussion will focus on methodological, ethical and interdisciplinary directions that could be explored to help differential privacy become the next standard practice for privacy in medical data analysis.

9.2 Data Privacy in Medicine

Medical data captures a historical view of chronic disease status, past procedures, lab values, imaging, genetics, and much more. Breaches of privacy in medicine can negatively impact an individual's dignity and cause them harm. An example is if a patient seeking mental health care or with positive HIV status experiences a privacy breach. They may face judgement and stigma from friends and family members. To prevent these harms, there are laws and regulations around the world that govern

how medical data privacy is protected. In contrast, there is a burgeoning need for medical data to be made publicly available to improve health research and general clinical knowledge. In many settings, the lack of such publicly available data for legitimate academic use can slow progress. We offer DP as a technical framework with many strong use cases in medical data and discuss when it is most appropriate given these contrasting needs.

9.2.1 Medical Data Privacy Laws Around the World

In the U.S., the Health Insurance Portability and Accountability Act (HIPAA) [Act96] governs individual rights for medical data privacy. The medical data that is protected under this law includes: information in electronic health records (EHRs), conversations between doctors and nurses often held in clinical notes, information about you that your health insurer stores, and billing information from interactions with the healthcare system [Offa]. There are different levels of access to medical data each with their own levels of privacy protections. Patients, healthcare providers, and any authorized family members or friends have full access to individual medical data. When this data is used for research or public health purposes it typically goes through a process called *anonymization*. This process removes certain pieces of data and adds noise to others to *de-identify* the data. We will discuss these two terms and the pitfalls of this process later on.

In Canada, since the healthcare system is managed provincially many provinces have their own variant of the Personal Information Protection and Electronic Documents Act (PIPEDA) which governs the privacy laws for medical data [Offb]. Broadly, the medical data protected under these laws includes: all information collected during interactions with the healthcare system (i.e. information inputted by doctors and nurses) and information collected by health insurance companies (e.g. prescriptions for medications). Similar to the U.S. there are varying levels of access and privacy protections.

In Europe, the General Data Protection Regulation (GDPR) governs the medical data privacy laws [Com18]. The medical data that is protected under the law is more broad than the previous regulations discussed: it covers all data generated from interactions with the healthcare system, data collected from wearable devices, data collected by health insurers, and health data that might be inferred from app usage. The levels of access and necessary permissions also differ from the previous regulations. Oftentimes, *explicit consent* is required for the processing of medical data.

In Asia, the regulations protecting medical data privacy differ between countries. China recently passed a set of regulations called the Personal Information Protection Law (PIPL) [Hor21] which took effect November 1, 2021. Similar to GDPR, these regulations protect medical data such as: data collected during interactions with

the healthcare system, data collected from wearables, and data collected from apps that can be used to infer health. Additionally, explicit consent is required to process health data. India does not currently have any explicit regulations which protect the privacy of medical data but is covered under the regulations for sensitive data from the Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules 2011 [GKN21]. The data protected by these laws is similar to GDPR. Explicit consent is not required for the analysis of health data according to these regulations.

Similar to Asia, the regulations protecting medical data privacy vary widely between countries in Africa. This ranges from no explicit data privacy laws to having similar laws to GDPR [Uni]. For example, in South Africa, the National Health Act governs the medical data privacy laws. These regulations protect medical data including: data collected from your doctors and nurses during your interactions and conversations between your health care providers. Processing of this data is governed by the Protection of Personal Information Act (POPIA) which states that only healthcare institutions, social services, insurance companies, schools, and anyone authorized by the individual whose healthcare is being processed.

As we can see, the regulations governing medical data privacy are similar in many ways around the world but there are differences. The kind of data covered and the levels of access allowed for secondary analysis of medical data (e.g. research purposes) are not the same across the regulations discussed. One common protection across many of these regulations is the use of anonymization to protect individual privacy for secondary analyses of medical data. The regulations above and anonymization are often considered strong methods for protecting patient privacy especially for public data releases. Next, we will discuss instances of medical data privacy breaches that have occurred despite the use of anonymization.

9.2.2 Medical Data Privacy Breaches

Despite the protections in place from the regulations discussed, medical data privacy breaches have and continue to happen. First, we discuss how often breaches occur and common reasons for these breaches. We end with examples of medical data privacy breaches in the research literature but also real-world data breaches. These breaches often correspond to a failure of anonymization. These breaches span public releases of medical data, medical data stored by hospitals, and more recently medical data collected by wearables and smart devices. In the U.S. from 2009-2020 as documented in the HIPAA Journals, there have been over 2705 privacy breaches of 500 records or more [Ald21]. This has resulted in the exposure of 268 189 693 medical records in the U.S. which equates to 81.72% of the population over 11 years. The common sources of the largest breaches during this time period included: hacking, loss, theft, and unauthorized access / disclosure. A combination of security

issues and privacy issues resulted in these breaches. They are evidence that we can help prevent some of them with stronger privacy protections but breaches due to security failing cannot be prevented with improved privacy protections.

A seminal example of such a breach in the research is that of Latanya Sweeney. Sweeney's research on re-identifying anonymized health data demonstrated that using only sex, zip code, and birth date that a large percentage of individuals are uniquely identifiable [Swe]. Sweeney demonstrated this using auxiliary data in what is known as a *linkage attack*. The study was able to uniquely link individuals in the Illinois Health Care Cost Containment Council data and the 1990 US Census using sex, zip code, and birth date with 87% accuracy. This study is a seminal example of the pitfalls of using anonymization for protecting medical data. See also Section 1.2 in Chapter 1 for additional discussion on why data anonymization fails to protect privacy.

In 2010, researchers showed that even genomic data could be re-identified using a combination of diagnostic codes, ethnicity, year of birth, and gender. The researchers in this study started from an anonymized sample of 1 174 793 patients from the Vanderbilt University Medical Center. In addition to this sample, they had a subset of 2762 of these patients who had been chosen for a genome wide association study (GWAS) on heart health. Their linkage attack leverages the fact that many of the people in the EHR are most likely to be candidates for the GWAS since most healthy individuals would not be in the EHR data.

In 2019, The Australian Department of Health released anonymized health data from 10% of the population (approx. 2.9 million people) which was re-identified six weeks later [CRT17]. To anonymize the data before release, an encrypted ID was used to represent each patient and all dates including date of birth were randomized to a day in a two week window. The researchers show that the anonymization and randomization procedures were not as protective of data privacy as originally thought. First, this was demonstrated using publicly available one-off information about individuals such as birth date, gender, and childbirth histories for women. Second, they demonstrate the efficacy of these attacks if there is access to a much larger dataset similar to Sweeney.

While the breaches we've presented so far have primarily been on statistical databases, similar attacks have been successful on machine learning models trained using medical data. Researchers have used membership inference attacks on large language models such as BERT [DCLT18] finetuned with clinical notes to identify whether a specific patient was in the training data [JRY21]. They show that these large clinical language models were susceptible to leaking up to 7% of the original training data. While this percentage is low, training data extraction is still a nascent field of research. So we can expect that these attacks will likely improve and more data will be leaked without protections.

All of these breaches are concerned with anonymized datasets that have been publicly released thus the intention of the users is not understood. For research purposes, the intent of researchers are well understood and defined. Still the appropriate security measures (such as trusted computing environments) are required to prevent any adversarial users from gaining unauthorized access to these systems. In these specific settings, the risk of breaches with anonymized data and access defined by the law is often lower. This is because it is assumed that researchers are not adversarial users who will attempt to re-identify individuals and the computing systems are secured. Once the data is meant to be made publicly available then DP is crucial to protecting patient privacy.

9.2.3 Differential Privacy as the New Gold Standard

The breaches and attacks presented are clear examples for why stronger privacy protections are needed for medical data analysis. Differential privacy is a prime candidate as the new gold standard because it is a framework for measuring and bounding the privacy leakage of running an algorithm on data without making assumptions about how an adversarial user tries to gain access and what the access already knows about the users in the data. In the rest of the chapter, we will present a series of case studies on analyzing medical data with differential privacy and training models with differential privacy for medical machine learning tasks.

9.3 Statistical Analysis

Statistical analysis of medical data is the cornerstone of the secondary use of medical data. These analyses have far reaching impacts including: improving hospital operations, treatment recommendations, understanding of chronic diseases, policy recommendations, and understanding of global pandemics. The most commonly analyzed sources of medical data include: electronic health records, randomized control trials, and genomics. There are different types of statistical analyses performed on these types of data as well. We will survey different types of differentially private statistical analyses on these types of data. Some of these algorithms will be extensions of those presented in previous chapters.

9.3.1 Electronic Health Records

Cohort Identification

Identifying cohorts of patients for clinical trials is an important use of electronic health record data. Most studies require a minimum number of patients to be enrolled. This starts with researchers identifying the total number of individuals

eligible for the study. This question is a counting query. As discussed previously in Chapter 4, there have been many methods developed over time for privately answering counting queries. Methods that have been developed specifically for cohort identification typically allow analysts to leverage domain knowledge to improve the utility of the algorithms. This was first done using the exponential mechanism with a utility function defined by a set of user-defined parameters such as the expected upper and lower bounds for the counts [VSB12].

Their method was improved upon using a truncated geometric mechanism and post-processing to achieve optimal privacy-utility tradeoffs [CSKB20]. The mechanism differs from the original exponential mechanism in the noise distribution that is used and that it is optimal for asymmetric utility functions. First, the analyst defines the count query $q : \mathcal{R}^d \rightarrow \{0, 1\}$ (e.g. "How many individuals have Type 1 Diabetes?") and the associated privacy budget ε. Additionally, the analyst has a prior belief $\pi(c)$ (typically based on knowledge of disease prevalence) over the true count. Both the count query and the privacy budget are sent to a trusted data curator. Next, the data curator computes the true count c for the query posed by the analyst. After this, noise z is sampled from a truncated geometric distribution, defined as

$$\begin{aligned} Pr[Y = y] &= \frac{1-\alpha}{1+\alpha}\alpha^{|y|} \\ z &= \min\{\max(0, y), n\}, \end{aligned} \tag{9.1}$$

where $\alpha = \frac{1}{\varepsilon}$), and is added to the true count. This produces the noised count $\hat{c}$ (i.e. $\hat{c} = c + z$). The accuracy of the mechanism is evaluated using a loss function $\ell(c, \hat{c})$ that evaluates the error between the true count and the noisy count, A post-processing step is applied to $\hat{c}$ to maximize the utility of the final noised count. This was inspired by the user-specific post-processing step in [GRS12]. This step consists of minimizing the expected loss under the posterior belief of the true count c conditioned on the noisy count $\hat{c}$. The conditional output distribution is defined as $q_{TGM}(z|x; \alpha, n)$ n is the size of the database. The posterior belief over the true count c is then expressed as:

$$q(c|\hat{c}; \pi, \alpha, n) \propto \pi(c) q_{TGM}(\hat{c}|c; \alpha, n). \tag{9.2}$$

Finally, a post-processing map of $\hat{c}$ is defined by $T : [n] \rightarrow [n]$:

$$T(\hat{c}|\pi, \ell, \alpha, n) = \arg\min_{y} \sum_{x} q(c|\hat{c}; \pi, \alpha, n)\ell(c, \hat{c}). \tag{9.3}$$

This optimization problem is solved using the fast Fourier transform based convolution algorithms where the loss function is the difference between the true and

noised count expressed as $\ell(c, \hat{c}) = (\hat{c} - c)$. This algorithm allows analysts to incorporate their beliefs about the true counts using knowledge of prevalence in general populations which is important for maximizing the utility in cohort identification. For example, if the query is about individuals with a rare disease then the analyst can reasonably believe the counts to be quite low. Knowing that the counts are low improves the utility by choosing a prior which places high probability on low counts. Thus ensuring that the noised counts do not diverge too far away from lower counts after the post-processing step.

This algorithm was evaluated on identifying a cohort for a study conducted on the effectiveness of genetic risk score in improving health outcomes associated with coronary heart disease (CHD) [Kul+16]. The study was looking to recruit participants in the Mayo Clinic Biobank with the following criteria: "age 45-65 years, non-Hispanic White ethnicity, no itory of atherosclerotic cardiovascular disease, not on statins, at intermediate risk for CHD (10 year CHD risk 5%-20%), and residents of Olmsted County Minnesota." 2026 subjects in the Biobank were eligible for the study, from which 216 subjects were enrolled. Focusing on queries of similar size, the authors of the above counting query algorithm aimed to estimate the number of eligible subjects in the Biobank. They demonstrate that for large counts ($c = 2000$) such as the query in [Kul+16] their algorithm returns 1998 at $\varepsilon = 0.05$ which is highly accurate. For smaller counts ($c = 100$) which require more noise to maintain the same level of privacy ($\varepsilon = 0.05$) the counts are still quite accurate at 88. These results show the efficacy of custom private counting query algorithms for cohort identification.

Survival Analysis

Survival analysis is concerned with predicting the time until death or an event of interest in a variety of situations in medicine (e.g. time to death after an organ transplant). Researchers developed a differentially private version of the nonparametric Kaplan-Meier survival model [BJO20] that provides minimal drops in utility. Survival analyses use event data up to time t to estimate the survival probability at time t. The algorithm they developed was inspired by work on continual release in differential privacy. They consider an event stream $S = (e_1, e_2, ..., e_t)$ where each event is defined as $e_i = (c_i, u_i, t_i)$ where c_i defines whether censoring occurred, u_i defines whether an event is uncensored, r_i denotes the patients remaining at time t and t_i is the time. To start we describe the K-M survival model without differential privacy and describe the changes made by the researchers to guarantee differential privacy. The Kaplan-Meier model uses the estimator defined as

$$S(t) = \prod_{i:t_i \leq t} (1 - \frac{u_i}{r_i}). \tag{9.4}$$

This estimator determines the probability of survival up to time t. Typically, the notion of adjacent data in DP is defined as two databases differing in one example. This notion does not apply directly to survival analyses. Thus, the study defines a new notion of neighboring streams as follows for DP survival analyses:

Definition 9.1. *Two streams S_t and S'_t are neighboring if at most one time $t_i \in \{1...t\}$ we find $|c_i - c'_i| + |u_i - u'_i| \leq 1$.*

The original definition of differential privacy is then used based on this definition of neighboring streams. The differentially private algorithm is composed of three parts: data partitioning, the survival curve computation, and post-processing. For data partitioning, the original stream S is separated into different groups that are made up of multiple events. This is necessary to organize the events by each of the individuals instead of treating each event as a separate piece of data. Each event e_i in S is processed individually into a group until $\Theta - 1$ events are processed. Both the number of events and the threshold θ are perturbed with noise to protect the privacy of the counts and size of each partition. A privacy budget of ε_1 is allocated to this procedure. These steps are required to ensure the proceeding steps in survival analysis are differentially private.

As described in Equation 9.4 the survival probability at each time is computed using the number of uncensored u_i and censored c_i events in a partition. The algorithm computes both of these quantities in a differentially private manner to compute $S(t)$. This is done using the binary tree mechanism for aggregating counts in an online fashion [CSS11; DNPR10]. This ensures that the noise added to the inputs only grows logarithmically instead of linearly. The new estimator is defined in Equation 9.5:

$$S(t) = S(t-1) \cdot \frac{N - \hat{u}_i - \hat{c}_{i-1}}{N - \hat{u}_{i-1} - \hat{c}_{i-1}}, \tag{9.5}$$

where $\hat{u}_i$ and $\hat{c}_i$ are the total number of uncensored and censored events up to partition i.

Finally, the post-processing step occurs as the addition of noise might violate some of the required properties of survival curves. The values of the curve must be monotonically decreasing as $t \to \infty$ and $S(t) \in [0, 1]$. First, another survival curve which respects these constraints $s^*(t)$ is computed that best matches the original differentially private curve $s(t)$. This is formulated as an optimization problem similar to previous work [HRMS09; BB72] and solved as an isotonic regression problem. A privacy budget of ε_2 is allocated for this computation. This method was tested on time to death of breast cancer patients in the Surveillance Epidemiology and End Results (SEER) dataset [HRE99]. The mean absolute error for the algorithm with a privacy budget of $\varepsilon = 1$ is no greater than 0.1 when the size of the

dataset is 10000. They also demonstrate using the Komolgorov-Smirnov test that the differentially private curves are not statistically different. Again, by developing a custom DP algorithm based on the needs of survival analysis of medical data there isn't too much loss in utility.

9.3.2 Clinical Trials

Randomized control trials are the gold standard for testing new medications, vaccines, treatments and confirming findings in previous studies. After these trials are conducted, meta-analyses are conducted to determine actionable changes for practice and policy. Thus, there has been an effort to make the data from clinical trials more transparent. Currently, ClinicalTrials.gov [TWZ09; HC15] is the largest registry in the world containing data from more than 393 097 studies. At the core of analyzing this data is hypothesis testing. Scientists are interested in understanding for example whether a new cancer treatment A performs better than the existing cancer treatment B. One question that is often answered using clinical trials is an association between a disease and a drug [DDB11]. Many of these studies use contingency tables to represent the results of the trial, lending them to use the performing Pearson's χ^2 test to test independence. A concern with conducting these tests is the power of the test, partly determined by the sample size. Thus an important task for clinical trials is determining the necessary sample size for a specific amount of statistical power and confidence. An algorithm for determining the sample size with differential privacy was developed [VS09]. This algorithm is one of the only ones inspired by statistical tasks for analyzing clinical trial data. This is an important area of research in differentially private statistics and medical data.

Determining Sample Size

We start with the problem of proving that a drug is effective for treating a disease. Typically a certain confidence is required (e.g. $\alpha = 0.05$) for FDA approval. We setup a simple example from the study for when the outcome is binary (e.g. success or failure of a drug dosage in lowering blood pressure) [VS09]. We define our null hypothesis and alternative hypothesis below. For notation, the true sample size calculated without privacy is N while the sample size calculated with differential privacy is $\hat{N}$.

Consider a sample $x_1, x_2, ..., x_N \sim$ Bernoulli(p) and the test statistic of interest be the mean $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i$. Our hypothesis test is formulated as follows:

$$H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_a : \mu = \mu_0 + \delta,$$

where in our example $\hat{\mu}$ represents the proportion of people who responded to the drug treatment. To guarantee differential privacy, the study uses the the addition

of Laplace noise $Z \sim L(\frac{\sqrt{2}}{\varepsilon N})$. The sampling distribution of $\hat{\mu}$ with this noise addition under the null hypothesis is approximated by $\mathcal{N}(\mu_0, \frac{\sigma^2}{N} + Z)$. Under the the alternate hypothesis it is approximated by $\mathcal{N}(\mu_0 + \delta, \frac{\sigma^2}{N} + Z)$. $\sigma^2 = \bar{\mu}(1 - \bar{\mu})$ where $\bar{\mu} = \mu_0 + \frac{\delta}{2}$. To calculate the true sample size N for confidence $1 - \alpha$ and power $1 - \beta$ the following equation must be solved:

$$\mu_0 + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{N} + \frac{2}{\varepsilon^2 N^2}} = \mu_0 + z_{1-\beta}\sqrt{\frac{\sigma^2}{N} + \frac{2}{\varepsilon^2 N^2}}, \tag{9.6}$$

which results in the following expression for the non-private sample size:

$$N = \frac{(z_{1-\frac{\alpha}{2}} + z_{1-\beta})^2 \sigma^2}{\delta^2}. \tag{9.7}$$

Using these two results, they derive the the expression for the private sample size N' is:

$$N' = N \cdot (\frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{8\delta^2}{\varepsilon^2 (z_{1-\frac{alpha}{2}} + z_{1-\beta})^2 \sigma^4}}). \tag{9.8}$$

Finally, to compute the exact sample size correction factor (i.e. the factor to multiply the true sample size to get the private sample size) numerical methods are used. Instead of noise from a Laplace distribution, the Normal Laplace is used such that: $X_1 \sim NL(\mu_0, \frac{\sigma^2}{N'}, \varepsilon N', \varepsilon N', 1)$ and $X_2 \sim NL(\mu_0 + \delta, \frac{\sigma^2}{N'}, \varepsilon N', \varepsilon N', 1)$. The exact private sample size N' is calculated by using a unique root-finder for the equation $F_{X_1}^{-1}(1-\frac{\alpha}{2}) = F_{X_2}^{-1}(1-\beta)$. The exact sample correction factor is found by $K = \frac{N'}{N}$. The authors evaluate the algorithm on synthetic tasks and for the Pearson χ^2 test of independence, demonstrating promising utility with the DP estimator. Further research should explore the utility of this method on real clinical trial data and the implications of using such a differentially private sample size calculation on the efficacy of trials.

9.3.3 Genomics

As biotechnology for genomic analysis has advanced, the utility of data analysis on genetic data has become incredibly apparent. Analysis of genomic data is broadly concerned with understanding the information contained in DNA and making predictions based on this data. Genomic data is important in advancing understanding of human disease and the underlying biological processes of these diseases. Given that genomic data is so unique to each individual any analysis of such data must provide strong privacy protections. Recent studies have shown the susceptibility

of anonymized genomics data to reidentification [BHO20; AAC21; Goo09], suggesting a need for strong privacy protections. As discussed previously, the broad protections DP provides against any type of adversarial entity and auxiliary information make it well-suited for genomic data privacy. We present two examples of differentially private algorithms for the analysis of genomic data in two settings: looking up variants and genome-wide association studies.

Variant Lookup

In biomedicine, *beacons* are web services that scientists query for information about specific alleles. The development of beacons has been championed by The Beacon Project, an initiative from the Global Alliance for Genomics and health. Their goal is to enable better genomic and clinical data sharing. The queries posed by users of Beacons can be formulated as membership queries (discussed in Chapter 5). In this setting, scientists are interested in understanding if a variant of interest is included in a database. Both the exponential and the Laplace mechanism would suffice for these queries, but a mechanism with optimal privacy-utility tradeoffs based on the truncated geometric mechanism was developed [CSKB20]. We revisit the algorithm and problem setup described in Section 9.3.1. In this setting $c \in \{0, 1\}$ represents the true membership answer and $\hat{c}$ represents the membership answer returned from the truncated geometric mechanism. The user's prior belief over c is defined as $\pi(c)$.

The loss function in this setting is defined in Equation 9.9:

$$\ell(c, \hat{c}) = \begin{cases} \ell(c, 0) & \text{if } \hat{c} = 0, \\ \ell(c, 1) & \text{if } \hat{c} > 0 \end{cases}. \tag{9.9}$$

The post-processing step is formulated as the exact same optimization problem defined in Equation 9.3 which we restate below:

$$T(\hat{c}|\pi, \ell, \alpha, n) = \arg\min_{y} \sum_{x} q(c|\hat{c}; \pi, \alpha, n)\ell(c, \hat{c}). \tag{9.10}$$

This optimization problem is also solved using fast Fourier transform based convolution algorithms since c and $\hat{c}$ are discrete and the loss function for cohort identification is typically $\ell(c, \hat{c}) = (\hat{c} - c)$. The authors evaluated this algorithm on identifying the top variants of autism spectrum disorder (ASD) studied by [Vel+19] in the ClinVar database [Lan+20]. The mechanism was able to correctly answer 11 out of 17 membership queries at $\varepsilon = 0.2$. This is compared to the Laplace mechanism which produces comparable results and the exponential mechanism which produces much worse results. As expected, when the number of occurrences of the

variant in the clinical database is larger the mechanism has a higher probability of answering correctly.

Genome-Wide Association Studies

Genome-wide association studies (GWAS) are used by scientists to capture associations between specific genetic markers and particular diseases [Uff+21]. These studies involve scanning the genomes of many different people to find genetic markers that are predictive of the presence of a disease. These findings are crucial to inform new prevention and treatment strategies. The cornerstone of these studies is performing statistical tests to measure the importance of a genetic marker for the presence of a disease. One of the challenges in performing differentially private GWAS is that the genome is extremely high dimensional. Thus the loss in utility incurred is high [USF13]. A promising algorithm called the neighbor method [JS13] was improved giving the current state of the art algorithm in identifying the top-k most significant genetic markers [SB16].

In GWAS, an allelic test statistic is used to test for associations between a genetic marker and disease status. Going forward we will refer to the genetic marker as a single nucleotide polymorphism (SNP). SNPs represent a variation at a single position in a DNA sequence for an individual. First, access to a case control cohort is assumed. For a given SNP, s_0, s_1 and s_2 are defined to be the number of individuals with 0, 1, or 2 copies of the minor allele in the control. We define r_0, r_1 and r_2 to be the same counts in the case cohort. Finally, n_0, n_1 and n_2 are the same quantities over the entire study population. R, N, S are the total number of case, study, and control participants. The allele test statistic used [SB16] is defined as:

$$Y(x, y) = \frac{2N(xS - yR)^2}{RS(x + y)(2N - x - y)}, \tag{9.11}$$

where $x = 2r_0 + r_1$ and $y = 2s_0 + s_1$.

The original neighbor method starts with a user defined threshold γ and selects all the SNPs where the allelic test statistic is higher than the threshold. To operationalize this, a notion of a neighbor distance is needed. For picking top k SNPs, the distance is defined as minimum number of individuals whose genotypes have to have be different for the SNP to be determined as significant. The algorithm uses this distance as presented in Algorithm 2. This distance works well for SNPs because it closely resembles the allele test statistic defined above. The major drawback of the original neighbor method is that the distance chosen sometimes gives different orderings than if we were to use the allelic test statistic.

To improve upon this issue, Next, we describe the improved neighbor method for picking the top k SNPs. The method starts with a user defined threshold γ. Significant SNPs are those with an allelic test statistic greater than this threshold

(i.e. $Y(D)_i > \gamma$ where $Y(D)_i$ represents the allelic test statistic for the ith SNP in the study cohort D). Next a neighbor distance is defined as the minimum number of individuals whose genomes need to change in the database for SNP i to become significant. This translates to minimum Hamming distance. The neighbor method leverages the intuition that the test statistic and the neighbor distance are closely related. If the SNP has a strong association with the disease than it will take many more changes to make the allele not insignificant. We present the modified neighbor algorithm in Algorithm 1.

Algorithm 1 Neighbor Method for Picking Top k SNPs [SB16]

Require: Study D, number of SNPs to return k, privacy budgets ε_1 and ε_2
Ensure: List of k SNPs that is $\varepsilon_1 + \varepsilon_2$ differentially private
Let γ be the mean score of the kth and $k+1$th highest scoring SNP.
Let $\gamma_{private}$ be a private estimate of γ using the Laplacian mechanism and privacy budget ε_1
Return list of SNPs from subroutine Algorithm 2 with privacy budget ε_2 and threshold value $\gamma_{private}$

Algorithm 1 runs in constant time the details of which can be found in [SB16]. The last step of the algorithm is to return the allelic test estimates for the chosen top k SNPs. Instead of using output perturbation, the authors use input perturbation before computing the test statistic which guarantees differential privacy via post-processing. The only difference in Equation 9.11 is that noise sampled from $Lap(\frac{2}{\varepsilon})$ is added to both x and y. This algorithm is evaluated on a rheumatoid arthritis dataset, NARAC-1 [Ple+07]. The dataset contains 893 cases and 1244 controls, with a total of 62 441 SNPs. The method performs significantly better than the traditional Laplacian mechanism in terms of accuracy (i.e. percentage of SNPs correctly identified). When $k = 3, \varepsilon = 0.5$ the accuracy of the method is 80%. For a larger number of SNPs (i.e. $k = 15$), the method achieves 80% accuracy at $\varepsilon = 5.0$.

In this section, we presented custom differentially private algorithms for common types of statistical analyses in medicine. This included: cohort identification, survival analyses, determining sample size for clinical trials, variant lookup, and performing genome-wide association studies. We discuss how the utility of standard differentially private algorithms can be improved significantly by incorporating medical domain knowledge into the algorithm design. This is a common theme throughout applying differential privacy to medical data analysis. An important and upcoming area of research is developing differentially private algorithms to analyze data generated from wearable devices. Analyzing this type of data is becoming an

important part of understanding human health with lots of opportunities for interesting DP algorithm development beyond what has initially been done [KJY18; Lin+16; Wu+20; Ren+16; Sal+16; UJM19].

Algorithm 2 Subroutine for Picking Top k SNPs [JS13]

Require: Study D, number of SNPs to return k, privacy budgets ε, threshold γ
Ensure: List of k SNPs that is ε differentially private
  for i = 0, ..., m **do**
    if $Y(D)_i > \gamma$ **then**
      $d_i = \min_{D'}(\{|D - D'| : Y(D')_i < \gamma, |D'| = |D|\})$
    else
      $d_i = 1 - \min_{D'}(\{|D - D'| : Y(D')_i > \gamma, |D'| = |D|\})$
    end if
  end for
  $\gamma_i = \exp(\frac{\varepsilon}{2k} d)_i$ for all i
  Without replacement, choose k SNPs where $Pr(\text{SNP}_i) \propto \gamma_i$
  Return the list of chosen SNPs

Many of the problems we discussed in this section are focused on membership in a database or counting the number of events in a database. Recently, there has been surging interest in developing diagnostic and prognostic models for different diseases or adverse events. This is motivated by the broad use of electronic health records and rapid success in developing high performing machine learning models for different modalities (i.e. time series, images, and natural language). The algorithms presented above are not designed for making predictions about disease presence or risk of developing a disease from labs, vitals, images, and clinical notes. Thus, we need to use more complex algorithms from machine learning to support creating models for diagnosis and risk prediction. In the next section, we discuss the application of DP to machine learning for medical data.

9.4 Machine Learning

Machine learning has demonstrated great potential to learn clinically relevant patterns from medical data across a wide variety of tasks. These tasks include disease / acuity prediction (e.g. mortality, breast cancer, kidney failure) [Tom+19; Gul+16; Wu+19; Raj+18; Gha+15], improvements to hospital operations (length of stay) [Wan+19], drug response prediction [Ram+19; Kue+20], and tumor segmentations [Hav+17; Koh+18]. However, as discussed in Chapter 5, machine learning models are susceptible to privacy attacks such as membership inference

and attribute inference [JRY21]. The potential for these attacks to occur is especially concerning when analyzing medical data. Without the use of differential privacy, sensitive health information such as HIV status of patients included in training data may be leaked. In this section, we will review different applications of differentially private machine learning (discussed in Chapters 7 and 8) for analyzing medical data. We will focus on: prediction and generating synthetic data.

9.4.1 Prediction

There have been various applications of DP machine learning and deep learning to different types of medical data such as: electronic health records, medical images, and genomics data. This subsection will focus on differentially private prediction in the central setting.

Electronic Health Records

A follow up study [SPGG21], examined the impact of applying DP machine learning to these prediction tasks. To train these models with differential privacy, the study uses differentially private stochastic gradient descent (DP-SGD) . The main changes to standard SGD are the addition of clipping individual gradients and adding Gaussian noise to these gradients. Further details are presented in Chapter 7. They demonstrate for tasks such as mortality where only 7% of patients passed away for both linear models and deep learning the drop in area under the curve (AUC) is quite high (22% and 26% respectively) for privacy budget $\varepsilon = 3.54, \delta = 10^{-5}$. These drops in utility are too high for the models to be useful in practice. This work demonstrates the need for more research on DP machine learning for high-dimensional multivariate time series medical data.

Another important prediction task from EHR data is 30-day readmission. Differentially private deep learning was applied to this prediction task [Cho+18] showing more promise for differentially private deep learning than the previous study. The authors use a dataset [Str+14] from the UCI Machine Learning Resposiitory that contains 101 000 medical records over 10 years with information such as demographics, potential risk factors such as diabetes, and the readmission label. The major difference between this dataset and the one in the previous study is that it is larger and lower dimensional. Both of these factors are important for reducing the utility loss from private training. The model trained is a small neural network with one hidden layer of size 32 and one output layer. They fix the privacy parameters as $\varepsilon = 1.0$ and $\delta = 10^{-5}$. The non-private model gives an AUC of 0.67 while the private one gives an AUC of 0.63. This is a much more reasonable drop in utility. Thus, algorithms such as DP-SGD are showing promise on predicting important tasks from electronic health record data when the dataset is larger and

lower dimensional. This is expected given the known tradeoffs between utility and data dimension for DP-SGD.

These are the two main studies that have applied differentially private machine learning and deep learning to prediction from EHR data. Both studies show the settings that DP-SGD is currently performing well and ones where more work is necessary. This is an active area of research that we hope readers will take an interest in.

Imaging

Next, we focus on applications of differentially private deep learning for disease prediction from medical images such as chest x-rays, magnetic resonance images (MRIs) and computed tomography (CTs) images. Similar to EHR data there are a limited number of studies on the application of current methods. Chest x-rays are often used by radiologists to diagnose issues such as pneumonia, collapsed lungs, pleural effusion, and the presence of tumors in the chest cavity and lungs. Pretrained DenseNet-121 models finetuned on chest x-ray datasets have shown promise in multi-label disease prediction achieving an AUC of 0.86 [SLMG20]. Following a similar procedure using differentially private fine-tuning, extreme drops in utility were observed (AUC = 0.50) at all privacy levels [SPGG21]. The authors cite the large dimensionality of the model and the long tailedness of the label distribution as reasons for such a large drop in utility. Work towards making the error DP learning algorithms independent of dimensionality will surely help improve the utility of DP deep learning models for medical imaging. This seems quite difficult in full generality but there have been several works [MMZ22; ZWB20; KDRT21] that have proven dimension independence or near dimension independence. This is essential for medical data broadly due to its high dimensionality across different modalities such as multivariate time series, images, and genomics.

Another study showed great success in predicting pneumonia from chest x-rays and semantic segmentation of liver CT [Zil+21]. The chest x-ray dataset used is from the Pediatric Pneumonia dataset [Ker+18] which contains 5232 images. 3883 of these images depicted pneumonia while the other 1349 were normal. The entire dataset was split into 85% for training and 15% for testing. Given the class imbalance the loss was reweighted as $w_i = \frac{c_i}{n_i}$ where i represents the class and n_i represents the total dataset size and c_i represents the number of samples of class i. The authors trained a VGG-11 [SZ14] model on images of size 224×224. The Gaussian Differential Privacy (GDP) accountant [BDLS20] was used for privacy accounting which differs from the Renyi DP Accountant [Aba+16; MTZ19] that is typically used. Their results show a modest drop in AUC of 11.2% at a privacy budget of $\varepsilon = 0.52$.

For semantic segmentation, they use the Medical Segmentation Decathlon Liver Segmentation Dataset [Sim+19]. The prediction task of importance is tumor segmentation which is incredibly important for treatment planning in oncology. The dataset consisted of 5184 training samples. A U-Net using VGG-11 was trained for this semantic segmentation task. This task was evaluated using the dice score which measures the similarity between the predicted segmentation map and the ground-truth segmentation (performed by radiologists). For $\varepsilon = 0.35$ they find that the drop in the dice score between the private and non-private model is minimal at 0.007. Both of these results show a lot of promise in applying differentially private deep learning to medical image segmentation.

Finally, we discuss an application of DP-SGD to a well-studied medical imaging deep learning problem which is prediction of diabetic retinopathy from fundus photography [SSKT20]. The authors study the privacy-utility tradeoffs of training residual networks with DP-SGD on this task. The dataset used is the APTOS 2019 Blindness Detection Dataset [KMD19] which contains 3600 training images which are labeled on a scale from 0 to 4 where 0 is no diabetic retinopathy and 4 is proliferative diabetic retinopathy (i.e severe disease). The models that were trained are pretrained 18 layer residual networks that were finetuned on the fundus images. This study notes similar drops in performance seen in [SPGG21] where the accuracy of models are around 50% or below. Thus differential privacy results in the models no longer being useful. The differences in these results are a direct product of issues such as dimensionality which are fundamental technical challenges that the differential privacy research community is addressing currently.

Genomics

Personalization of drug recommendations is important because individuals often have very different reactions to the same drug. Gene expression data has become an important source of data for predicting personalized drug sensitivity. The application of differentially private machine learning to this task has been studied in two related studies on drug sensitivity prediction [NHHK19; Hon+18]. This is the main example of DP machine learning applied to genomic data in the current literature. The authors focus on DREAM-NCI drug sensitivity prediction challenge [Cos+14]. This challenge focuses on identifying the best treatments based on genomic data for breast cancer cell lines. After pre-processing the gene expression data the dimensionality of the data is reduced to $d = 64$. The non-private algorithm for predicting drug sensitivity on this task is Bayesian linear regression. This study produces a differentially private version of Bayesian linear regression which experiences only modest drops of about 2–4% in accuracy from the non-private model.

9.4.2 Data Synthesis and Sharing

Making medical data publicly available is vital to advancing medical discoveries, helping improve transparency in medical research, and to making the research community more inclusive. Publicly releasing medical data while maintaining patient privacy is incredibly difficult. In Section 9.2, we discussed failures of publicly releasing medical data. An upcoming area of research for being able to publicly release medical data is differentially private data synthesis. This area of research has become especially promising due to the large amount of positive results from deep generative models such as diffusion models, GANs, and VAEs. In this section, we will discuss some of the recent successes in synthesizing medical data with differential privacy, some of the failures in doing so, and what the ongoing challenges are.

One recent success has been in simulating patients from the SPRINT (Systolic Blood Pressure Trial) Data Analysis Change held by the *New England Journal of Medicine* [BM17]. The SPRINT clinical trial examined the effect of intensive lowering of systolic blood pressure (< 120 mmHg) instead of aiming for a standard systolic blood pressure (< 140 mmHg). Researchers used this tabular data of 6000 patients to train an auxiliary classifier GAN (AC-GAN) with differential privacy to generate synthetic data similar to the original SPRINT data [Bea+19]. They compare the synthetic data generated with and without privacy guarantees quantitatively and qualitatively. The quantitative evaluation examined variable correlation structure and whether models trained on the synthetic achieved similar performance. The qualitative evaluation consisted of visualizing blood pressure trajectories and having clinicians attempt to differentiate between real and synthetic data. The visual inspection of the blood pressure trajectories showed very minimal differences for both intensive and standard patients between the private and non-private synthetic data. The accuracy of the four different models (logistic regression, random forest, nearest neighbors, and SVMs) dropped slightly more than 10% in the worst scenario. This is only a modest drop performance for differential privacy signalling that the private synthetic data still maintains a large amount of utility. It should be noted that the success of this study is promising but limited since the dataset was much simpler and lower dimensional than what we see in other domains such as medical imaging and genomics. In these higher dimensional settings, DP synthetic data generation is much more difficult. One key contribution of this study is a set of guidelines for evaluating the utility of private synthetic medical data. We summarize the guidelines below:

Quantitative

- Measure how similar the variable correlations are between the real data and the the private synthetic data

- Measure the performance of a classifier trained on the private synthetic data vs trained on real data and tested on real data
- Measure whether features have the same importance level between the classifiers trained on real and private synthetic data

Qualitative

- Visualize differences between real and private synthetic data (method will depend on data type)
- Have clinical experts attempt to differentiate between real and private synthetic data samples

Another study focused on generating synthetic EHR data (i.e. high dimensional multivariate time series) using dual adversarial autoencoders [Lee+20]. This study used the MIMIC-III [Joh+16] and the UT Physicians Clinical databases. The dual adversarial autoencoder (DAAE) contains three components: a sequence-to-sequence autoencoder (seq2seq AE), an inner GAN, and an outer GAN. The seq2seq AE is responsible for learning the latent space the encodes semantic features of target sequences and the temporal dynamics of these sequences. The inner GAN aims to match the learned space and the outer GAN aims to distinguish the real samples from the ones generated using the seq2seq AE. All of these different model components are trained using DP-SGD to guarantee differential privacy. The two evaluation components similar to the ones described above are evaluating the performance of models trained on the real vs synthetic data and having clinicians evaluate the plausibility of the synthetic sequences. The prediction task of interest is predicting top-k diagnosis codes. On this task the DAAE method is only marginally worse (0.1 on average) than the model trained on the real data across recall and precision at 10,20,30 on the MIMIC-III dataset. This study shows promise for training accurate prediction models on DP synthetic data for this specific task. It is unclear from the study though whether correlations in the original data are maintained and if feature importance remained consistent. Both of these are important evaluations for understanding if this method can be of use more broadly for other datasets and tasks.

Finally, we discuss an area where more research is needed to realize the potential of differentially private synthetic medical data which is medical images. Generating high quality synthetic medical images such as chest x-rays or dermatological images (e.g. pictures of skin lesions) has proven to be incredibly difficult [Che+21]. In this study, deep convolutional GANs trained with DP-SGD generate images of both quantitative and qualitative quality. Additionally, the classifiers trained on these images face significant drops in utility. This study demonstrates the additional work

needed from the differential privacy community to demonstrate the promise of private synthetic medical images.

9.5 Federated Learning

The majority of medical data remains in private silos around the world. This inhibits the advancement of health research and the potential for developing high utility predictive models. Privacy is the paramount concern when sharing data across hospitals in an attempt to learn a predictive model that is better than each hospital using their individual data. Federated learning, is a technique that learns a central model and series of local models that keeps data in each of their respective silos. The technique alone is not private, requiring differentially private versions of FL algorithms (further details in Chapter 8) to protect privacy of the data even though it does not leave the silos. Here we discuss a couple of applications of DP federated learning to common medical machine learning tasks.

Similar to the work in the centralized setting [SPGG21], another study analyzed the efficacy of predicting prolonged length of stay and in hospital mortality from electronic records using DP federated learning [PDH19]. This study used the eICU database which contains EHR data from 52 different hospitals across the U.S. For both prolonged length of stay and mortality prediction DP federated averaging performs quite poorly unfortunately, especially in the class imbalanced mortality tasks. Many of the AUC values for each local hospital are below 0.5 at reasonable privacy budgets such as $\varepsilon < 10$. While these results are discouraging especially given the high number of hospitals in the dataset, other studies have shown more promise. It is unclear from this study what the different sources of utility loss were. For example, heterogeneity between the different hospitals may be a significant issue that leads to large utility loss for applying DP FL to larger cohorts of hospitals.

Another study developed a variant of DP-SGD that incorporates cyclical weight transfer to train neural networks with differential privacy in a distributed setting [BYFW18]. The algorithm used in this study differed from the traditional DP FL algorithm used in the prior study by using cyclical weight transfer. This means that the central weights are transferred to each institution after the aggregation of local steps. This study also focused on in hospital mortality for the eICU dataset and on classifying two different subtypes of breast cancer from gene expression data found in The Cancer Genome Atlas (TCGA) [Can+13]. At a privacy budget of $\varepsilon = 3.84$ and $\delta = 10^{-5}$ the drop in AUC between the non-private and private distributed model was 0.9% when the number of collaborating institutions was 5 and was a

drop of 6% when only two institutions were collaborating. This is promising as it incentivizes more institutions to collaborate since it will help improve the utility at a specific privacy budget for all. For the subtype prediction task, for a privacy budget of $\varepsilon = 6.11$ and $\delta = 10^{-5}$ the drop in AUC went from 0.7% to 1.7% when going from one to three sites collaborating. The success suggests that cyclical weight transfer may be promising for reducing utility loss in DP FL. Although, it is hard to compare the two studies because of the differences in scale. The first study we discussed was across 52 different hospitals whereas this second study was across 5 at most. As mentioned previously, as the heterogeneity increases due to larger cohorts it might be that it is much more difficult to prevent drops in utility even for algorithms that use cyclical weight transfer.

Finally, we summarize an extensive study demonstrating the effectiveness of differentially private federated averaging at the individual patient level and at the hospital level [Sad+21]. They choose both logistic regression and neural networks across a variety of tasks including: survival prediction from heart failure, diabetes prediction, mortality prediction, SARS-CoV-2 positivity prediction, and adverse event prediction in individuals treated with Azithromycin. Across these tasks they demonstrate minimal drops in performance when using DP-FedAvg but that this also outperforms its centralized counterpart. Applications of differentially private federated learning to medical data are a new and active area of research that has the opportunity to help change the way hospitals collaborate around the world. There are still open problems to be addressed before wide-scale use such as large heterogeneity amongst hospitals.

In the previous two sections, we provided case studies on the successes and failures of differentially private machine learning and federated learning applied to medical data analysis. These studies provide guidance on fruitful technical directions for researchers to help make these technologies more widely applicable. Differentially private machine learning has the potential to help advance health research by making public sharing of data and models safer and easier. As discussed at the beginning of this chapter, anonymization for data release are susceptible to attacks. Currently, model sharing / publishing is rare due to privacy concerns. We end this chapter by outlining some of the challenges not yet mentioned and opportunities for helping realize differentially private statistics and machine learning into practice for medical data analysis.

9.6 Considerations for Deployment in Medicine

All of the case studies that we presented have been done in research settings. There has yet to be a large-scale deployment of differential privacy in a medical setting.

Here we discuss some of the hesitations that exist and areas of research that may be fruitful to help support the first large-scale deployment of DP for medical data analysis.

9.6.1 Health Inequities

There are many inequities in the healthcare system that have resulted in marginalized communities such as women and Black individuals receiving poor quality of care. Examples of this include: women suffering higher mortality rates from heart attacks [GCH18], Black patients receiving worse care for kidney disease [VEJ20], and Black women being 3× more likely to die during pregnancy than white women [Cre+14]. While medical machine learning is showing promise it is important for it to not exacerbate existing inequities. It could even potentially help be part of existing efforts to remove these inequities [CSG19]. Currently, differential privacy has a negative impact on fairness in medical machine learning [SPGG21]. There are potential solutions that may help alleviate these tradeoffs such as using publicly available data during differentially private training.

9.6.2 Robustness to Distribution Shift

Another important consideration especially for differentially private machine learning is the non-stationarity of medical data. Often times the underlying data distribution is changing due to different populations over time, policy changes, and new treatment recommendations. This means that the data distribution a model was trained on will differ from the test distribution. If models are not robust to distributional shifts then they will silently fail when deployed in practice [Nes+19; SSS18]. While it is well understood that differential privacy provides out-of-sample generalization, the connection to distributional robustness is not well understood. An initial study [Kul+] has proven that differential privacy guarantees distributional robustness but empirical understanding is still limited. Research which empirically explores the current impact of differential privacy on distributional robustness and develops algorithms to maintain both will be important for deploying differentially private machine learning in medical settings.

9.6.3 Education, Audits, and Policy

An incredibly important and underresearched topic in helping deploy differential privacy for medical data analysis is education and policy. Many individuals such as IT professionals, legal experts, data analysts, bioethicists, doctors, nurses, and

patients do not know what differential privacy really means and what they should expect from it. For many individuals it is unclear what are the exact privacy protections that differential privacy does and does not provide for their medical data. This is incredibly important as misinterpretations of these protections can undermine the need for differential privacy over the current standard of anonymization. Data analysts are unsure of how implementing differential privacy would impact their current workflow for performing research and interacting with medical data. Legal and policy experts are unsure of what an acceptable level for the privacy budget ε is and what this means for the probability of an attack being successful. This leads to broader questions in medical data analysis which are also important in other contexts of how should one tradeoff between utility and a smaller privacy budget? Studies to understand the gap between expectations of differential privacy and the realities in medicine are needed similar to this [CKR21]. Research auditing differentially private statistics and machine learning [JUO20; Kif+20] is vital to providing legal and policy experts in medicine better understanding of the technology.

9.7 Concluding Remarks

In this chapter, we presented a broad overview of differentially private statistics and machine learning applied to medical data analysis. We discussed the need for differential privacy as the new gold standard for privacy in medicine. This was highlighted by the overwhelming amount of privacy breaches and re-identification of anonymized data that has occurred. We presented a series of case studies on differentially private statistical analyses applied to medical data. These case studies demonstrated that often times custom algorithms are needed for medical data analysis to achieve reasonable privacy-utility tradeoffs. These case studies covered important statistical tasks including cohort identification, survival analysis, variant lookup, and GWAS. We then examined applications of differentially private machine learning and deep learning to prediction and data synthesis. Given the global scale of medical research we discussed the promise of differentially private distributed learning for learning higher utility models and allowing hospitals to share data without comprising privacy. Finally, we discuss important research directions and issues to address for supporting large scale deployments of differential privacy for medical data analysis. As research progresses and more interdisciplinary work with the healthcare system takes place, deployments will start to occur, helping to provide better privacy protections for patient data.

References

[AAC21] K. Ayoz, E. Ayday, and A. E. Cicek. "Genome reconstruction attacks against genomic data-sharing beacons". In: Proceedings on Privacy Enhancing Technologies. Privacy Enhancing Technologies Symposium. Vol. 2021. 3. NIH Public Access. 2021, p. 28 (cit. on p. 328).

[Aba+16] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang. "Deep Learning with Differential Privacy". In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. CCS '16. Vienna, Austria: ACM, 2016, pp. 308–318 (cit. on p. 333).

[Act96] A. Act. "Health insurance portability and accountability act of 1996". In: Public law 104 (1996), p. 191 (cit. on p. 319).

[Ald21] S. Alder. Healthcare Data Breach Statistics. en. https://www.hipaajournal.com/healthcare-data-breach-statistics/. Accessed: 2021-10-23. Jan. 2021 (cit. on p. 320).

[BB72] R. E. Barlow and H. D. Brunk. "The isotonic regression problem and its dual". In: Journal of the American Statistical Association 67.337 (1972), pp. 140–147 (cit. on p. 325).

[BDLS20] Z. Bu, J. Dong, Q. Long, and W. J. Su. "Deep learning with Gaussian differential privacy". In: Harvard data science review 2020.23 (2020) (cit. on p. 333).

[Bea+19] B. K. Beaulieu-Jones, Z. S. Wu, C. Williams, R. Lee, S. P. Bhavnani, J. B. Byrd, and C. S. Greene. "Privacy-preserving generative deep neural networks support clinical data sharing". In: Circulation: Cardiovascular Quality and Outcomes 12.7 (2019), e005122 (cit. on p. 335).

[BHO20] L. Bonomi, Y. Huang, and L. Ohno-Machado. "Privacy challenges and research opportunities for genomic data sharing". In: Nature genetics 52.7 (2020), pp. 646–654 (cit. on p. 328).

[BJO20] L. Bonomi, X. Jiang, and L. Ohno-Machado. "Protecting patient privacy in survival analyses". In: Journal of the American Medical Informatics Association 27.3 (2020), pp. 366–375 (cit. on p. 324).

[BM17] N. S. Burns and P. W. Miller. "Learning what we didn't know—the SPRINT data analysis challenge". In: New England Journal of Medicine 376.23 (2017), pp. 2205–2207 (cit. on p. 335).

[BYFW18] B. K. Beaulieu-Jones, W. Yuan, S. G. Finlayson, and Z. S. Wu. "Privacy-preserving distributed deep learning for clinical data". In: arXiv preprint arXiv:1812.01484 (2018) (cit. on p. 337).

[Can+13] Cancer Genome Atlas Research Network, J. N. Weinstein, E. A. Collisson, G. B. Mills, K. R. M. Shaw, B. A. Ozenberger, K. Ellrott, I. Shmulevich, C. Sander, and J. M. Stuart. "The Cancer Genome Atlas Pan-Cancer analysis project". en. In: Nat. Genet. 45.10 (Oct. 2013), pp. 1113–1120 (cit. on p. 337).

[Che+21] V. Cheng, V. M. Suriyakumar, N. Dullerud, S. Joshi, and M. Ghassemi. "Can You Fake It Until You Make It? Impacts of Differentially Private Synthetic Data on Downstream Classification Fairness". In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. 2021, pp. 149–160 (cit. on p. 336).

[Cho+18] E. Chou, T. Nguyen, J. Beal, A. Haque, and L. Fei-Fei. "A fully private pipeline for deep learning on electronic health records". In: arXiv preprint arXiv:1811.09951 (2018) (cit. on p. 332).

[CKR21] R. Cummings, G. Kaptchuk, and E. M. Redmiles. "" I need a better description": An Investigation Into User Expectations For Differential Privacy". In: arXiv preprint arXiv:2110.06452 (2021) (cit. on p. 340).

[Com18] E. Commission. 2018 reform of EU data protection rules. May 25, 2018. URL: https://ec.europa.eu/commission/sites/beta-political/files/data-protection-factsheet-changes_en.pdf (visited on 06/17/2019) (cit. on p. 319).

[Cos+14] J. C. Costello, L. M. Heiser, E. Georgii, M. Gönen, M. P. Menden, N. J. Wang, M. Bansal, P. Hintsanen, S. A. Khan, J.-P. Mpindi, et al. "A community effort to assess and improve drug sensitivity prediction algorithms". In: Nature biotechnology 32.12 (2014), pp. 1202–1212 (cit. on p. 334).

[Cre+14] A. A. Creanga, C. J. Berg, J. Y. Ko, S. L. Farr, V. T. Tong, F. C. Bruce, and W. M. Callaghan. "Maternal mortality and morbidity in the United States: where are we now?" In: Journal of women's health 23.1 (2014), pp. 3–9 (cit. on p. 339).

[CRT17] C. Culnane, B. I. P. Rubinstein, and V. Teague. "Health Data in an Open World". In: (Dec. 2017). arXiv: 1712.05627 [cs.CY] (cit. on p. 321).

[CSG19] I. Y. Chen, P. Szolovits, and M. Ghassemi. "Can AI Help Reduce Disparities in General Medical and Mental Health Care?" In: AMA Journal of Ethics 21.2 (2019), pp. 167–179 (cit. on pp. 317, 339).

[CSKB20] H. Cho, S. Simmons, R. Kim, and B. Berger. "Privacy-preserving biomedical database queries with optimal privacy-utility tradeoffs". In: Cell systems 10.5 (2020), pp. 408–416 (cit. on pp. 323, 328).

[CSS11] T.-H. H. Chan, E. Shi, and D. Song. "Private and continual release of statistics". In: ACM Transactions on Information and System Security (TISSEC) 14.3 (2011), pp. 1–24 (cit. on p. 325).

[DCLT18] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: arXiv:1810.04805 [cs] (2018). arXiv: 1810.04805. URL: http://arxiv.org/abs/1810.04805 (visited on 05/22/2019) (cit. on p. 321).

[DDB11] J. T. Dudley, T. Deshpande, and A. J. Butte. "Exploiting drug–disease relationships for computational drug repositioning". In: Briefings in bioinformatics 12.4 (2011), pp. 303–311 (cit. on p. 326).

[DNPR10] C. Dwork, M. Naor, T. Pitassi, and G. N. Rothblum. "Differential privacy under continual observation". In: Proceedings of the forty-second ACM symposium on Theory of computing. 2010, pp. 715–724 (cit. on p. 325).

[GCH18] B. N. Greenwood, S. Carnahan, and L. Huang. "Patient–physician gender concordance and increased mortality among female heart attack patients". In: Proceedings of the National Academy of Sciences 115.34 (2018), pp. 8569–8574 (cit. on p. 339).

[Gha+15] M. Ghassemi, M. A. F. Pimentel, T. Naumann, T. Brennan, D. A. Clifton, P. Szolovits, and M. Feng. "A Multivariate Timeseries Modeling Approach to Severity of Illness Assessment and Forecasting in ICU with Sparse, Heterogeneous Clinical Data". en. In: Twenty-Ninth AAAI Conference on Artificial Intelligence. Feb. 2015. URL: https://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9393 (visited on 05/21/2019) (cit. on p. 331).

[Gha+17] M. Ghassemi, M. Wu, M. C. Hughes, P. Szolovits, and F. Doshi-Velez. "Predicting intervention onset in the ICU with switching state space models". In: AMIA Summits on Translational Science Proceedings 2017 (2017), p. 82 (cit. on p. 317).

[GKN21] N. Gandhi, S. Kapoor, and S. Nailwal. At a glance: data protection and management of health data in India. https://www.lexology.com/library/detail.aspx?g=0fdcef36-61a8-4e00-9bed-8abcf6866c96. Accessed: 2021-10-23. Jan. 2021 (cit. on p. 320).

[Goo09] M. T. Goodrich. "The mastermind attack on genomic data". In: 2009 30th IEEE Symposium on Security and Privacy. IEEE. 2009, pp. 204–218 (cit. on p. 328).

[GRS12] A. Ghosh, T. Roughgarden, and M. Sundararajan. "Universally utility-maximizing privacy mechanisms". In: SIAM Journal on Computing 41.6 (2012), pp. 1673–1693 (cit. on p. 323).

[Gul+16] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega, and D. R. Webster. "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs". en. In: JAMA 316.22 (2016), pp. 2402–2410 (cit. on p. 331).

[Hav+17] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle. "Brain tumor segmentation with deep neural networks". In: Medical image analysis 35 (2017), pp. 18–31 (cit. on p. 331).

[HC15] K. L. Hudson and F. S. Collins. "Sharing and reporting the results of clinical trials". In: Jama 313.4 (2015), pp. 355–356 (cit. on p. 326).

[Hon+18] A. Honkela, M. Das, A. Nieminen, O. Dikmen, and S. Kaski. "Efficient differentially private learning improves drug sensitivity prediction". In: Biology direct 13.1 (2018), pp. 1–12 (cit. on pp. 317, 334).

[Hor21] J. Horwitz. "China passes new personal data privacy law, to take effect Nov. 1". In: Reuters (Aug. 2021) (cit. on p. 319).

[HRE99] B. F. Hankey, L. A. Ries, and B. K. Edwards. "The surveillance, epidemiology, and end results program: a national resource". In: Cancer Epidemiology and Prevention Biomarkers 8.12 (1999), pp. 1117–1121 (cit. on p. 325).

[HRMS09] M. Hay, V. Rastogi, G. Miklau, and D. Suciu. "Boosting the accuracy of differentially-private histograms through consistency". In: arXiv preprint arXiv:0904.0942 (2009) (cit. on p. 325).

[Joh+16] A. E. W. Johnson, T. J. Pollard, L. Shen, L.-w. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark. "MIMIC-III, a freely accessible critical care database". en. In: Scientific Data 3.1 (May 2016), pp. 1–9. ISSN: 2052-4463. URL: https://www.nature.com/articles/sdata201635 (visited on 11/04/2019) (cit. on p. 336).

[JRY21] A. Jagannatha, B. P. S. Rawat, and H. Yu. "Membership Inference Attack Susceptibility of Clinical Language Models". In: arXiv preprint arXiv:2104.08305 (2021) (cit. on pp. 321, 332).

[JS13] A. Johnson and V. Shmatikov. "Privacy-preserving data exploration in genome-wide association studies". In: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. 2013, pp. 1079–1087 (cit. on pp. 329, 331).

[JUO20] M. Jagielski, J. Ullman, and A. Oprea. "Auditing differentially private machine learning: How private is private sgd?" In: arXiv preprint arXiv:2006.07709 (2020) (cit. on p. 340).

[KDRT21] P. Kairouz, M. R. Diaz, K. Rush, and A. Thakurta. "(Nearly) Dimension Independent Private ERM with AdaGrad Rates via Publicly Estimated Subspaces". In: Conference on Learning Theory. PMLR. 2021, pp. 2717–2746 (cit. on p. 333).

[Ker+18] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, et al. "Identifying medical diagnoses and treatable diseases by image-based deep learning". In: Cell 172.5 (2018), pp. 1122–1131 (cit. on p. 333).

[Kif+20] D. Kifer, S. Messing, A. Roth, A. Thakurta, and D. Zhang. "Guidelines for implementing and auditing differentially private systems". In: arXiv preprint arXiv:2002.04049 (2020) (cit. on p. 340).

[KJY18] J. W. Kim, B. Jang, and H. Yoo. "Privacy-preserving aggregation of personal health data streams". In: PloS one 13.11 (2018), e0207639 (cit. on p. 331).

[KMD19] Karthik, Maggie, and S. Dane. APTOS 2019 Blindness Detection. https://kaggle.com/competitions/aptos2019-blindness-detection. Kaggle. 2019 (cit. on p. 334).

[Koh+18] S. A. Kohl, B. Romera-Paredes, C. Meyer, J. De Fauw, J. R. Ledsam, K. H. Maier-Hein, S. Eslami, D. J. Rezende, and O. Ronneberger. "A probabilistic u-net for segmentation of ambiguous images". In: arXiv preprint arXiv:1806.05034 (2018) (cit. on p. 331).

[Kue+20] B. M. Kuenzi, J. Park, S. H. Fong, K. S. Sanchez, J. Lee, J. F. Kreisberg, J. Ma, and T. Ideker. "Predicting drug response and synergy using a deep learning model of human cancer cells". In: Cancer cell 38.5 (2020), pp. 672–684 (cit. on p. 331).

[Kul+] B. Kulynych, Y.-Y. Yang, Y. Yu, J. Błasiok, and P. Nakkiran. "What you see is what you get: Distributional generalization for algorithm design in deep learning". In: () (cit. on p. 339).

[Kul+16] I. J. Kullo, H. Jouni, E. E. Austin, S.-A. Brown, T. M. Kruisselbrink, I. N. Isseh, R. A. Haddad, T. S. Marroush, K. Shameer, J. E. Olson, et al. "Incorporating a genetic risk score into coronary heart disease risk estimates: effect on low-density lipoprotein cholesterol levels (the MI-GENES clinical trial)". In: Circulation 133.12 (2016), pp. 1181–1188 (cit. on p. 324).

[Lan+20] M. J. Landrum, S. Chitipiralla, G. R. Brown, C. Chen, B. Gu, J. Hart, D. Hoffman, W. Jang, K. Kaur, C. Liu, et al. "ClinVar: improvements to accessing data". In: Nucleic acids research 48.D1 (2020), pp. D835–D844 (cit. on p. 328).

[Lee+20] D. Lee, H. Yu, X. Jiang, D. Rogith, M. Gudala, M. Tejani, Q. Zhang, and L. Xiong. "Generating sequential electronic health records using dual adversarial autoencoder". In: Journal of the American Medical Informatics Association 27.9 (2020), pp. 1411–1419 (cit. on p. 336).

[Lin+16] C. Lin, Z. Song, H. Song, Y. Zhou, Y. Wang, and G. Wu. "Differential privacy preserving in big data analytics for connected health". In: Journal of medical systems 40.4 (2016), p. 97 (cit. on p. 331).

[MMZ22] Y.-A. Ma, T. V. Marinov, and T. Zhang. "Dimension independent generalization of dp-sgd for overparameterized smooth convex optimization". In: arXiv preprint arXiv:2206.01836 (2022) (cit. on p. 333).

[Mob+18] P. Mobadersany, S. Yousefi, M. Amgad, D. A. Gutman, J. S. Barnholtz-Sloan, J. E. V. Vega, D. J. Brat, and L. A. Cooper. "Predicting cancer outcomes from histology and genomics using convolutional networks". In: Proceedings of the National Academy of Sciences 115.13 (2018), E2970–E2979 (cit. on p. 317).

[MTZ19] I. Mironov, K. Talwar, and L. Zhang. "R\'enyi differential privacy of the sampled gaussian mechanism". In: arXiv preprint arXiv:1908.10530 (2019) (cit. on p. 333).

[Nes+19] B. Nestor, M. B. A. McDermott, W. Boag, G. Berner, T. Naumann, M. C. Hughes, A. Goldenberg, and M. Ghassemi. Feature Robustness in Non-stationary Health Records: Caveats to Deployable Model Performance in Common Clinical Machine Learning Tasks. 2019 (cit. on p. 339).

[NHHK19] T. Niinimäki, M. A. Heikkilä, A. Honkela, and S. Kaski. "Representation transfer for differentially private drug sensitivity prediction". In: Bioinformatics 35.14 (2019), pp. i218–i224 (cit. on p. 334).

[Offa] Office for Civil Rights (OCR). Your Rights Under HIPAA. https://www.hhs.gov/hipaa/for-individuals/guidance-materials-for-consumers/index.html. Accessed: 2021-10-23 (cit. on p. 319).

[Offb] Office of the Privacy Commissioner of Canada. Provincial laws that may apply instead of PIPEDA. https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/r_o_p/prov-pipeda/. Accessed: 2021-10-23 (cit. on p. 319).

[PDH19] S. R. Pfohl, A. M. Dai, and K. Heller. "Federated and Differentially Private Learning for Electronic Health Records". In: arXiv preprint arXiv:1911.05861 (2019) (cit. on p. 337).

[Ple+07] R. M. Plenge, M. Seielstad, L. Padyukov, A. T. Lee, E. F. Remmers, B. Ding, A. Liew, H. Khalili, A. Chandrasekaran, L. R. Davies, et al. "TRAF1–C5 as a risk locus for rheumatoid arthritis—a genomewide study". In: New England Journal of Medicine 357.12 (2007), pp. 1199–1209 (cit. on p. 330).

[Raj+18] A. Rajkomar, E. Oren, K. Chen, A. M. Dai, N. Hajaj, M. Hardt, P. J. Liu, X. Liu, J. Marcus, M. Sun, et al. "Scalable and accurate deep learning with electronic health records". In: NPJ Digital Medicine 1.1 (2018), p. 18 (cit. on pp. 317, 331).

[Ram+19] L. Rampášek, D. Hidru, P. Smirnov, B. Haibe-Kains, and A. Goldenberg. "Dr. VAE: improving drug response prediction via modeling of drug perturbation effects". In: Bioinformatics 35.19 (2019), pp. 3743–3751 (cit. on p. 331).

[Ren+16] H. Ren, H. Li, X. Liang, S. He, Y. Dai, and L. Zhao. "Privacy-enhanced and multifunctional health data aggregation under differential privacy guarantees". In: Sensors 16.9 (2016), p. 1463 (cit. on p. 331).

[Sad+21] A. Sadilek, L. Liu, D. Nguyen, M. Kamruzzaman, S. Serghiou, B. Rader, A. Ingerman, S. Mellem, P. Kairouz, E. O. Nsoesie, et al. "Privacy-first health research with federated learning". In: NPJ digital medicine 4.1 (2021), pp. 1–8 (cit. on p. 338).

[Sal+16] N. Saleheen, S. Chakraborty, N. Ali, M. M. Rahman, S. M. Hossain, R. Bari, E. Buder, M. Srivastava, and S. Kumar. "mSieve: differential behavioral privacy in time series of mobile sensor data". In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2016, pp. 706–717 (cit. on p. 331).

[SB16] S. Simmons and B. Berger. "Realizing privacy preserving genome-wide association studies". In: Bioinformatics 32.9 (2016), pp. 1293–1300 (cit. on pp. 329, 330).

[Sim+19] A. L. Simpson, M. Antonelli, S. Bakas, M. Bilello, K. Farahani, B. Van Ginneken, A. Kopp-Schneider, B. A. Landman, G. Litjens, B. Menze, et al. "A large annotated medical image dataset for the development and evaluation of segmentation algorithms". In: arXiv preprint arXiv:1902.09063 (2019) (cit. on p. 334).

[SLMG20] L. Seyyed-Kalantari, G. Liu, M. McDermott, and M. Ghassemi. "CheXclusion: Fairness gaps in deep chest X-ray classifiers". In: arXiv preprint arXiv:2003.00827 (2020) (cit. on pp. 317, 333).

[SPGG21] V. M. Suriyakumar, N. Papernot, A. Goldenberg, and M. Ghassemi. "Chasing Your Long Tails: Differentially Private Prediction in Health Care Settings". In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. 2021, pp. 723–734 (cit. on pp. 332–334, 337, 339).

[SSKT20] S. Singh, H. Sikka, S. Kotti, and A. Trask. "Benchmarking differentially private residual networks for medical imagery". In: arXiv preprint arXiv:2005.13099 (2020) (cit. on p. 334).

[SSS18] A. Subbaswamy, P. Schulam, and S. Saria. “Preventing failures due to dataset shift: Learning predictive models that transport”. In: arXiv preprint arXiv:1812.04597 (2018) (cit. on p. 339).

[Str+14] B. Strack, J. P. DeShazo, C. Gennings, J. L. Olmo, S. Ventura, K. J. Cios, and J. N. Clore. “Impact of HbA1c measurement on hospital readmission rates: analysis of 70,000 clinical database patient records”. In: BioMed research international 2014 (2014) (cit. on p. 332).

[Swe] L. Sweeney. Simple demographics often identify people uniquely. https://dataprivacylab.org/projects/identifiability/paper1.pdf. Accessed: 2021-10-23 (cit. on p. 321).

[SZ14] K. Simonyan and A. Zisserman. “Very deep convolutional networks for large-scale image recognition”. In: arXiv preprint arXiv:1409.1556 (2014) (cit. on p. 333).

[TBL17] A. D. Trister, D. S. Buist, and C. I. Lee. “Will machine learning tip the balance in breast cancer screening?” In: JAMA oncology 3.11 (2017), pp. 1463–1464 (cit. on p. 317).

[Tom+19] N. Tomašev, X. Glorot, J. W. Rae, M. Zielinski, H. Askham, A. Saraiva, A. Mottram, C. Meyer, S. Ravuri, I. Protsyuk, A. Connell, C. O. Hughes, A. Karthikesalingam, J. Cornebise, H. Montgomery, G. Rees, C. Laing, C. R. Baker, K. Peterson, R. Reeves, D. Hassabis, D. King, M. Suleyman, T. Back, C. Nielson, J. R. Ledsam, and S. Mohamed. “A clinically applicable approach to continuous prediction of future acute kidney injury”. en. In: Nature 572.7767 (Aug. 2019), pp. 116–119 (cit. on p. 331).

[TWZ09] T. Tse, R. J. Williams, and D. A. Zarin. “Reporting “basic results” in ClinicalTrials. gov”. In: Chest 136.1 (2009), pp. 295–303 (cit. on p. 326).

[Uff+21] E. Uffelmann, Q. Q. Huang, N. S. Munung, J. de Vries, Y. Okada, A. R. Martin, H. C. Martin, T. Lappalainen, and D. Posthuma. “Genome-wide association studies”. In: Nature Reviews Methods Primers 1.1 (2021), p. 59. URL: https://doi.org/10.1038/s43586-021-00056-9 (cit. on p. 329).

[UJM19] A. Ukil, A. J. Jara, and L. Marin. “Data-driven automated cardiac health management with robust edge analytics and de-risking”. In: Sensors 19.12 (2019), p. 2733 (cit. on p. 331).

[Uni] United Nations. “COVID-19 and Human Rights: We Are All in This Together”. In: () (cit. on p. 320).

[USF13] C. Uhlerop, A. Slavković, and S. E. Fienberg. "Privacy-preserving data sharing for genome-wide association studies". In: The Journal of privacy and confidentiality 5.1 (2013), p. 137 (cit. on p. 329).

[VEJ20] D. A. Vyas, L. G. Eisenstein, and D. S. Jones. Hidden in Plain Sight—Reconsidering the Use of Race Correction in Clinical Algorithms. 2020 (cit. on p. 339).

[Vel+19] D. Velmeshev, L. Schirmer, D. Jung, M. Haeussler, Y. Perez, S. Mayer, A. Bhaduri, N. Goyal, D. H. Rowitch, and A. R. Kriegstein. "Single-cell genomics identifies cell type–specific molecular changes in autism". In: Science 364.6441 (2019), pp. 685–689 (cit. on p. 328).

[VS09] D. Vu and A. Slavkovic. "Differential privacy for clinical trial data: Preliminary evaluations". In: 2009 IEEE International Conference on Data Mining Workshops. IEEE. 2009, pp. 138–143 (cit. on p. 326).

[VSB12] S. A. Vinterbo, A. D. Sarwate, and A. A. Boxwala. "Protecting count queries in study design". In: Journal of the American Medical Informatics Association 19.5 (2012), pp. 750–757 (cit. on p. 323).

[Wan+19] S. Wang, M. B. A. McDermott, G. Chauhan, M. C. Hughes, T. Naumann, and M. Ghassemi. "MIMIC-Extract: A Data Extraction, Preprocessing, and Representation Pipeline for MIMIC-III". In: arXiv:1907.08322 [cs, stat] (July 2019). arXiv: 1907.08322. URL: http://arxiv.org/abs/1907.08322 (visited on 11/04/2019) (cit. on pp. 317, 331).

[Wu+19] D. Wu, H. Kobayashi, C. Ding, L. Cheng, and K. G. M. Ghassemi. "Modeling the Biological Pathology Continuum with HSIC-regularized Wasserstein Auto-encoders". In: arXiv:1901.06618 [cs, stat] (2019). arXiv: 1901.06618. URL: http://arxiv.org/abs/1901.06618 (visited on 05/22/2019) (cit. on p. 331).

[Wu+20] X. Wu, M. R. Khosravi, L. Qi, G. Ji, W. Dou, and X. Xu. "Locally private frequency estimation of physical symptoms for infectious disease analysis in internet of medical things". In: Computer Communications 162 (2020), pp. 139–151 (cit. on p. 331).

[Zil+21] A. Ziller, D. Usynin, R. Braren, M. Makowski, D. Rueckert, and G. Kaissis. "Medical imaging deep learning with differential privacy". In: Scientific Reports 11.1 (2021), pp. 1–8 (cit. on p. 333).

[ZWB20] Y. Zhou, Z. S. Wu, and A. Banerjee. "Bypassing the Ambient Dimension: Private SGD with Gradient Subspace Identification". In: arXiv preprint arXiv:2007.03813 (2020) (cit. on p. 333).