Data Governance Network
Anchored by IDFC Institute

October 2021

Working Paper 19

Health Data as Wealth: Understanding Patient Rights in India within a Digital Ecosystem through a Feminist Approach

Radhika Radhakrishnan

internet democracy project

Data Governance Network

The Data Governance Network is developing a multi-disciplinary community of researchers tackling India's next policy frontiers: data-enabled policymaking and the digital economy. At DGN, we work to cultivate and communicate research stemming from diverse viewpoints on market regulation, information privacy and digital rights. Our hope is to generate balanced and networked perspectives on data governance — thereby helping governments make smart policy choices which advance the empowerment and protection of individuals in today's data-rich environment.

About Us

The Internet Democracy Project works towards realising feminist visions of the digital in society, by exploring and addressing power imbalances in the areas of norms, governance and infrastructure in India and beyond.

Disclaimer and Terms of Use

The views and opinions expressed in this paper are those of the authors and do not necessarily represent those of the organisation.

Design

Cactus Communications

Suggested Citation:

Radhakrishnan, Radhika. (2021). Health Data as Wealth: Understanding Patient Rights in India within a Digital Ecosystem through a Feminist Approach. Working Paper 19, *Data Governance Network, Mumbai*.

Supported by a grant from Omidyar Network

Abstract

Over the past few years, there has been a drive towards the digitisation of healthcare in India. Policy frameworks in the country are incentivising further datafication by considering health data to be a commodity. In the context of big data, I argue that when health data is viewed as a disembodied resource, access to people's health data becomes a form of power, giving those with such access the unparalleled and unprecedented power to influence the governance of people's bodies and lives. Recognising the interconnections between our bodies and data from within an alternative feminist framework, this paper analyses the datafication of health in India through emerging developments proposed under the National Digital Health Mission (NDHM) ecosystem, and its implications for the bodies and rights of people. This work seeks to understand how datafication contributes to the disembodiment of health data in policy frameworks; the consequences of disembodiment for how people's health data is understood to have value and who can benefit from that value, with a focus on health insurance companies; and how acknowledging the relationship between health data and bodies within policy frameworks can empower people to safeguard their right to equitable healthcare.

Table of Contents

List of Abbreviations

ANM: Auxiliary Nurse Midwife
DEO: Data Entry Operator
DISHA: Digital Information Security in Health Act
EHR: Electronic Health Record
EMR: Electronic Medical Record
HIP: Health Information Provider
HIU: Health Information User
NDHE: National Digital Health Ecosystem
NDHB: National Digital Health Blueprint
NDHM: National Digital Health Mission
NHA: National Health Authority
NHM: National Health Mission
NHS: National Health Stack
PAN: Permanent Account Number
PDP: Personal Data Protection
PHR: Personal Health Record

Part I: Overview

Introduction

> Much of clinical...knowledge is embodied knowledge – knowledge sensed through and with the body.
> –Gordon (1988: 269)

Health data has been collected for decades in India in various forms, such as patient data collected within clinical settings, data collected for clinical trials and research, and census data. Such datafication of health is therefore not a new phenomenon. The collection of this data can also potentially lead to better health outcomes. For instance, data from X-ray scans give a more granular understanding of a patient's physiology and help health practitioners better diagnose diseases. Over the years, critical thinking about such data within the domain of healthcare has led to the development and implementation of medical legislation and codes of ethics to ensure that patient rights remain safeguarded (Indian Medical Council (Professional Conduct, Etiquette and Ethics) Regulations, 2002).

However, in the age of big data, we are at a crucial stage where such regulation is falling short and newer ways of thinking about health data are urgently needed. This is because the past few years have seen key changes in the manner in which the datafication of health is taking place. With the advent of big data, there is a rise in how much and what kinds of data is generated and consumed within the domain, blurring the boundaries of what constitutes health data and how it should be regulated. For instance, data about people's health is no longer confined to medical records generated within clinical settings for the use of healthcare practitioners to treat patients. Rather, it unfolds on a number of different scales extending, but not limited, to the use of Internet and social media platforms for medical consultations

and continuous patient monitoring through a vast array of wearable fitness and health devices and smartphone applications (Ruckenstein & Schüll, 2017). While some forms of this data may have been collected in the past, the quantitative explosion in the health data collected today is leading to a shift in the qualitative experience of healthcare for patients. Further, this data may in turn be used by private actors to serve business interests over the patient's best interests, with significant consequences for people and their rights, as this paper will explore.

Policy frameworks in the country are incentivizing and driving this increasing datafication of health, as will be discussed later. The framework they provide is an opportunity to critically engage with these developments more closely. At the same time, they don't fully capture the risks of datafication upon the bodies and rights of patients. These policy frameworks conceptualise health data as a disembodied resource and an enabler for economic progress. Data is predominantly understood as a resource (like oil), available for human extraction, and existing independently from the bodies producing it. This dominant framework can be traced back to the field of cybernetics which conceptualised data as a layer permeating everything while existing independently from the medium carrying it, making it possible to transfer it from one medium to another (Hayles, 1999). Till date, this understanding of data as a resource finds its way into various domains, including the emerging policy framework governing the digital health ecosystem in India, as I will explore in greater detail later in this paper.

This is concerning because in a datafied world, disembodiment of data opens it up to possibilities of human exploitation and manipulation (Couldry & Mejias, 2019). To account for this harm, feminist scholarship has led the way in foregrounding the relationship between data and bodies by showing that they are intimately interconnected, thus calling for a deeper understanding of data as embodied (Kovacs & Ranganathan, 2019; Kovacs & Jain, 2020; Radhakrishnan, 2020; Van der Ploeg, 2012). Feminists argue that data is an extension of our bodies, and control over data is often experienced by people as control over their bodies. For example, victims of non-consensual sharing of intimate images often describe their experience in terms of physical violence, not in terms of a data protection violation (Patella-Rey, 2018). When viewed through such experiences, some feminist scholars argue that the line between our physical bodies and our virtual bodies is becoming irrelevant because of the extent to which data is used to determine and control our bodily experiences (Van der Ploeg, 2012). Applying this understanding to health data raises important questions about patient rights as envisioned by policy frameworks that this paper will highlight and attempt to respond to. Therefore, it is imperative to engage directly with regulation around the datafication of health from feminist perspectives.

Moreover, there is a need to ground this discourse within emerging Global South ecosystems in the age of big data. Existing literature on the datafication of health focuses almost exclusively on advanced industrialised demographics where there is a relatively broad embrace of digital health technologies by citizens along with robust public debate around the social implications of this embrace. Further implications of datafication are expected to emerge when contextualised in demographics grappling with constraints such as lack of Internet access and digital literacy, weak public health systems, and powerful medical-industrial assemblages.

Responding to these gaps in the literature, in this paper, I examine the ecosystem of the National Digital Health Mission (NDHM)—a government of India initiative to develop an integrated digital health infrastructure for the country—as an analytical framework. The developments studied as part of this ecosystem are very recent, with some proposed less than a year ago at the time of writing. This paper offers an early-stage, grounded critique of these emerging developments by evaluating possible outcomes and scenarios and flagging areas of concern, based on fieldwork data and policy analysis.

I argue that in the age of big data, when health data is viewed as a disembodied resource, access to people's health data becomes a form of power, giving those with such access the unparalleled power to influence the governance of people's bodies and lives.

Part I of the paper provides an outline of this research study in Section 1, and of the key components, stakeholders and data flows that are part of the National Digital Health Mission ecosystem in Section 2.

Part II of the paper focuses on the (dis)embodiment of health data. Section 3 analyses the factors that contribute to the disembodiment of health data during the movement through its life cycle. Section 4 examines the shifts that are observed in healthcare through the datafication of our bodies, in terms of how datafication changes how people's health can have value and who can benefit from that value in the age of big data.

Part III focuses on the embodied impact of datafication through the NDHM in the lives of patients. Section 5 analyses how patients are impacted by disembodied datafication, how their rights are supported or undermined by it, and the varied ways in which such datafication influences the governance of their bodies. Lastly, Section 6 charts a way forward, by unpacking how the recognition of the embodied nature of health data can empower patients to affirm their rights.

1.1 Research Methodology

This research adopts a mixed methodology approach, primarily relying upon interviews conducted during ethnographic fieldwork and desk analysis of relevant government policies and documents. The process of desk review involving a critical analysis of policies on the subject is useful to understand how the state views the datafication of health and to examine the regulatory framework within which this datafication is happening. Interviews with key stakeholders help to fill the gaps in publicly available information and provide a deeper understanding of the implications of policies in the everyday lives of people.

I conducted nineteen semi-structured, in-depth interviews, of which sixteen were conducted in-person and three were conducted telephonically or online. With the exception of four interviews, which were conducted in English, all interviews were conducted in Hindi. Eight days of ethnographic fieldwork was carried out in the Union Territory of Chandigarh, the capital of the northern Indian states of Punjab and Haryana. This site was chosen because it is one of the seven union territories where the NDHM programme has been piloted by the government of India before rolling it out to the rest of the country, and thus one of the few places where these policies can be seen in action at the time of conducting this research.

Stakeholders interviewed include grassroots health workers such as ANM workers (Auxiliary Nurse Midwife workers or female health workers based at health sub-centres or Primary Health Centres) and Anganwadi workers (community-based frontline workers of the Integrated Child Development Services program of the Government of India); data entry operators in civic hospitals; Senior Medical Officers, medical interns, pharmacists, multi-purpose workers, and other staff workers at civic dispensaries; members of the National Health Mission (NHM) Employees Union; persons enrolled in the digital Health ID programme of the NDHM; and subject matter experts on the NDHM ecosystem.

These participants were chosen because they are key stakeholders in the NDHM ecosystem. Some participants are involved in the process of enrolling citizens in the NDHM programme, some are responsible for promoting the programme among local communities, some are themselves enrolled in the programme, and others have a working knowledge of the on-ground realities of its implementation.

Purposive and snowball sampling methods were used to identify research participants during fieldwork. Members of health worker unions were initially contacted, along with independent visits to community healthcare centres, civic hospitals, and civic dispensaries. From here, other participants were contacted by snowballing. Subject matter experts were contacted telephonically or online through purposive sampling.

Prior to as well as following fieldwork, this study has undergone an independent, rigorous ethics review, to ensure that the dignity and well-being of participants is respected, and has been approved by the Anusandhan Ethics Committee.

A major challenge for data collection during fieldwork was gaining institutional access to government health facilities. A long-drawn procedure was prescribed to obtain official government permission to conduct interviews within public health facilities, which was stalled indefinitely. In light of this limitation, in consultation with the Anusandhan Ethics Committee, I strove to protect the rights of all those I spoke to as part of this research. I built relationships with union workers who were helpful not only as key data sources but also as intermediaries in snowballing to more accessible data sources and sites.

Some names used in this paper have been changed as per the request of the research participants in their informed consent forms, and this has been indicated in footnotes for their first usage in the paper. Names and affiliations that have been used as is have been mentioned after seeking explicit written consent from research participants through informed consent forms.

As part of the mixed methodology adopted for this research, I carried out desk analysis of key state policy documents, bills, and consultation white papers relating to health and health data in India. I also analysed academic literature and news reports to support the theorisation in this paper and triangulate data.

2. The National Digital Health Mission (NDHM) Ecosystem

2.1 Evolution

Plans for digitising India's health ecosystem have been in the works since at least 2017, as laid out in various policies by the government. The National Health Policy, 2017, proposes "extensive deployment of digital tools for improving the efficiency and outcome of the healthcare system" (Ministry of Health & Family Welfare, 2017, p. 25). The National Health Stack (NHS), 2018, further envisages the deployment of a "powerful technology arsenal [for]... a complete redesign of the flow of people, money, and information, as well as a layered approach to providing comprehensive foundational health functions" (NITI Aayog, 2018, p. 6). The implementation framework for the NHS is proposed in the National Digital Health Blueprint (NDHB), 2019, which details the building blocks and institutional mechanisms for data integration of the health system in India to mediate the generation, collection, exchange and standardisation of health data (Ministry of Health & Family Welfare, 2019). The NDHB also recommends the establishment of a specialised program - the National Digital Health Mission (NDHM). The NDHM aims to "create a national digital health ecosystem that supports universal health coverage in an efficient, accessible, inclusive, affordable, timely and safe manner, that provides a wide-range of data, information and infrastructure services, duly leveraging open, interoperable, standards-based digital systems, and ensures the security, confidentiality and privacy of health-related personal information" (National Health Authority, 2020, p. 1). Lastly, the draft Digital Information Security in Healthcare Act (DISHA) 2018 seeks to provide privacy, confidentiality, security and standardization of health data (Ministry of Health & Family Welfare, 2018).

On 15 August 2020, Prime Minister Narendra Modi officially launched the NDHM. On 26 August 2020, the National Health Authority (NHA)—which will implement the NDHM—put out a draft Health Data Management Policy for public consultation. The policy was finalised and passed in December 2020 (Ministry of Health & Family Welfare, 2020a). In August 2020, the NDHM laid out a framework to run a sandbox to test new products and services to form a digital health architecture (Ministry of Health & Family Welfare, 2020b). In March 2021, the NDHM published a Draft Implementation Strategy for this ecosystem (National Health Authority, 2021).

2.2 Key Components

A reading of the major policy documents pertaining to health data in the country indicates that the National Digital Health Ecosystem (NDHE) under the NDHM comprises of the following key components:

1. **Health ID** refers to a voluntary unique identification number or identifier allocated to individuals to whom the health data relates, to enable them to participate in the NDHE (Ministry of Health & Family Welfare, 2020, Chapter IV: 10).
2. **Health Registry** contains the master data of all the entities in the ecosystem, including doctors, hospitals, clinics, laboratories, pharmacies, and insurance companies, providing the basic information about these entities (National Health Authority, 2020, p. 15).
3. **Consent Manager** is an electronic system that interacts with the data principal and obtains consent from them for any intended access to their personal data (Ministry of Health & Family Welfare, 2020, p. 3).
4. **Health Claims Platform** will provide the building blocks required to implement any large-scale health insurance program in an automated, data-driven manner, by both public and private actors (NITI Aayog, 2018, p. 23).
5. **Health Data Analytics Platform** will enable the creation of anonymised and aggregated datasets that assist in the creation of statistics leading to data-driven decisions and targeted policymaking in the health sector (NITI Aayog, 2018, p. 32).
6. **Open Telemedicine and e-Pharmacy Network** will expand access to healthcare services via a model enabling public and private sector participation (National Health Authority, 2020, p. 18).
7. **Software** used to manage health data in the ecosystem, including (National Health Authority, 2020, pp. 9-10):
    a) **Personal Health Records (PHRs)** to enable patients to compile, update and keep a copy of their own health records, to help them better manage their care.
    b) **Electronic Medical Records (EMR)** used within a hospital or a clinic to support patient diagnosis and treatment and transaction focused.
    c) **Electronic Health Records (EHR)** containing records for a patient across multiple doctors and providers and used within a healthcare system (such as across a state government or other hospitals) to provide better care for patients.

Since this study aims to understand patient rights, I will focus upon the main patient-facing component within this ecosystem which is already being piloted in the country: the Health ID, which in turn facilitates access to a patient's Personal Health Records. With the launch of the NDHM on 15 August 2020, Health IDs were rolled out in six union territories as part of Phase 1 of the implementation, and are expected to be expanded nation-wide in upcoming phases (National Health Authority, 2020, p. 21).

2.3 Key Stakeholders

The stakeholders most directly relevant to this ecosystem include:

1) **Data Principal** is an individual to whom the health data relates (Ministry of Health & Family Welfare, 2020, p. 3)

2) **Health Information Providers (HIPs)** are hospitals, diagnostic centres, public health programs, labs, health apps, or other such entities which act as information providers by generating, storing and distributing health records in the digital health ecosystem (Ministry of Health & Family Welfare, 2020, p. 4). Another example of an HIP that will be discussed later in this paper is the Health Locker, which is an interoperable specification that can be implemented by multiple players to enable the creation and exchange of EHRs (National Health Authority, 2021, p. 16).

3) **Health Information Users (HIUs)** are entities that are permitted to request access to the personal data of a data principal and can access this data with the consent of the data principal (Ministry of Health & Family Welfare, 2020, p. 4). These could include hospitals, doctors, insurance providers and personal health apps.

4) **Data fiduciaries (trustees)** shall facilitate consent-driven interaction between entities that generate health data and entities that want to obtain access to PHRs for delivering better services to the individual (NITI Aayog, 2018, p. 29).

Given the focus of this paper on patients, the main standpoint here will be that of the data principal. From the perspective of data principals, I will unpack their data flows with respect to private health insurance providers as HIUs within this ecosystem. This HIU has been particularly chosen for this study because in the age of big data, the power wielded by health insurers through their access to patients' health data raises new and important questions that require urgent attention, as will be analysed later. Moreover, health data policies have created many provisions and incentives for insurers to be a part of this ecosystem (IRDAI-NHA Joint Working Group, 2019). For instance, the NHS specifically recognises the challenges of insurance claims processing and fraud management as areas where a data-driven health infrastructure would be beneficial, thereby making the provision of a coverage and claims platform to implement insurance programs (NITI Aayog, 2018). It further mentions that its analytics platform "will initially focus on Health Insurance" (NITI Aayog, 2018, p. 33).

2.4 Data Flows

Potential relevant data flows between stakeholders are represented in Figure 1. In this paper, I will focus upon the data flows relating to the data principal and their interactions with HIPs and HIUs through the Health ID. The NDHM ecosystem will provide data principals with a unique Health ID, and then use that Health ID to share personal health data among various stakeholders for different purposes, facilitated by their consent.

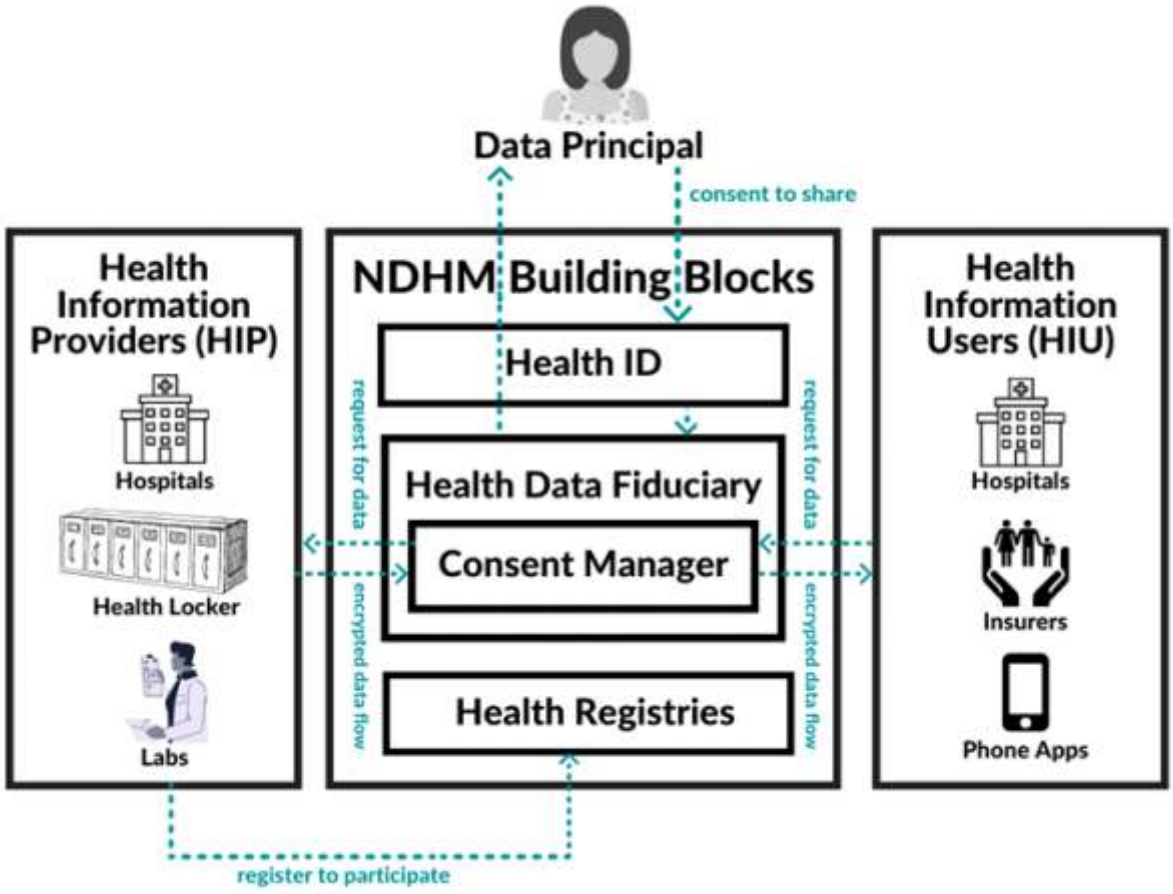

Figure 1: Data flows in the NDHM ecosystem

Part II: (Dis)embodiment of Health Data in Policies

3. Disembodiment of Health Data through Datafication

What is happening to a patient's health data as it moves through its life cycle - starting at the point it is generated from a patient's body till it is shared with different stakeholders for various purposes? In this section, I will analyse the key stages in the life cycle of health data (as illustrated in Figure 2) to understand how its relationship with a patient's body undergoes changes at each step, and how this relationship is reflected, and often invisibilised, in health data policies within the NDHM ecosystem.

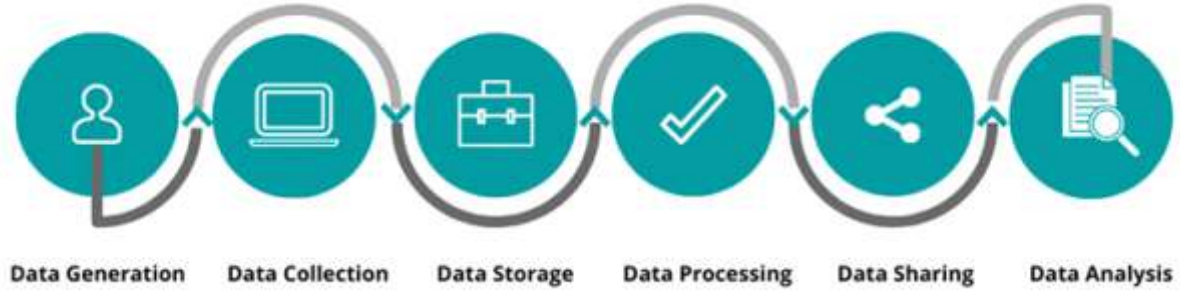

Figure 2: Life cycle of health data

3.1 Data Generation and Collection

The first three steps of the data lifecycle i.e. generation, collection, and storage often occur simultaneously. For example, a blood test generates data, entering the test results in an electronic record collects the data, and saving the results stores it. The storage stage is represented separately here to permit a more granular analysis.

Data generation refers to data coming into existence. Health data can be generated through any kind of medical encounter with people, devices or processes - from completing an admission checklist to non-clinical activities such as financial transactions. Here, I focus upon the health data generated from the bodies of patients. In this regard, data may be generated through clinical encounters such as visits to medical practitioners or clinical labs, as well as non-clinical self-reporting such as step counts or sleep patterns.

Health data is collected when the generated patient information is recorded or entered into either a clinical system such as a paper document or an electronic record, or a non-clinical device such as a wearable fitness tracker. In the NDHM ecosystem, entering information about the data principal in a PHR/EHR amounts to collecting their data. The NDHM also enables data principals to add readings from devices like wearables to their PHRs (National Health Authority, 2020, p. 12).

Though the NDHM ecosystem is said to be designed along the principle of “collection limitation” (Ministry of Health & Family Welfare, 2020, p. 15), its political economy enables more and more data to be structurally generated as well as collected. In the NDHM ecosystem, the availability of increasing amounts of data is framed as an unquestionable state of affairs (Kovacs & Ranganathan, 2019), but in reality, such a rise in the amount of data generated and collected is a key feature of surveillance capitalism, which creates a market where there is both demand for more data and a promise of profit from this data (Zuboff, 2019).

In many of these cases, data is generated through some degree of technological mediation. For instance, the data generated when a patient visits a medical practitioner in a clinic may include symptoms that are checked through devices such as stethoscopes and test results from scans done using X-ray machines, in combination with a doctor's sensory judgements of the patient's body (what can be directly seen, touched, smelled and heard). Similarly, the data generated when a patient is running, such as step counts or heart rate, may be mediated through wearable fitness tracking devices. The extent of this mediation may vary in different cases. A health check-up wherein a doctor inspects the discoloration of skin involves mostly a doctor's sensory judgment, whereas telemedicine almost entirely relies upon digital visualisations involving patient self-reporting which come to stand in for flesh-on-flesh responses (Lupton & Maslen, 2017).

Datafication leads to an increase in the extent to which technological mediation plays a role in the generation of health data. The term 'digital phenotype' conceptualises the data that people generate as an extension of their individual phenotype through interactions with networked digital technologies. Yet, irrespective of the degree to which technologies are used for medical mediation, the link between a patient's body and the data generated by it is strongly recognised within healthcare. For example, an X-ray is a literal scan of a person's body, representing an intuitively uncontested link between the body and data generated by the body. Moreover, errors in data-mediated diagnosis are acknowledged to potentially cause bodily harm to patients, and are accordingly regulated by legally requiring adherence to a standard of reasonable care while performing any acts that could cause foreseeable harm to patients (Ministry of Health & Family Welfare, 2010).

Despite the relationship between data and bodies being recognised within healthcare at this stage, a degree of disembodiment is already taking place here within policy frameworks. When data about patients is generated in the form of EHRs, its digital nature allows it to be alienated from patients' bodies through an erasure of its material basis, even though the data's connection with the body still remains. In other words, "data's digitality makes it easier for this centrality of our bodies, of the material world, to be obscured" (Kovacs & Jain, 2020). More importantly, while the techno-materiality of data distances it from bodies to some extent, the link between data and bodies gets further obscured by our conceptual and metaphorical frameworks of viewing data as a disembodied resource. Thus though data is originally generated from a patient's body, this link which is centrally recognised within healthcare, starts becoming tenuous within policy frameworks.

3.2 Data Storage

Collected health data may be stored manually in filing cabinets and folders or digitally in databases and data warehouses. The NDHM ecosystem is based on the principle of 'federated' architecture (Ministry of Health & Family Welfare, 2020). In other words, health data would be stored on interoperable, independent and decentralised information systems. This means that each interaction that a patient has with the healthcare system will generate many data points which are now likely to be held by several different institutions on the EHRs of independent systems. As a result, data is distributed across many institutions and data systems.

But while a person can physically exist in only one place at a given time, this need not be true of their data, which can co-exist in multiple places and systems. In fact, it is this ability of data to be distributed across spatial dimensions and dispersed over multiple locations that lends digitality its power while also blurring its links to the body. Moreover, data may be duplicated in ways that human bodies cannot be: data about one person may exist in not only different places, but may also end up corresponding to different digital identities. Duplication of data is considered key to ensuring that it can be distributed across digital networks (O'Brien & Marakas, 2008).

What does duplication of data mean for the bodies corresponding to this data? Consider the case of the NDHM Health ID to understand this. A data entry operator (DEO), Ms. Nirmala,[1] who is responsible for enrolling Health IDs for patients at a civic dispensary in Chandigarh noted (translated from Hindi):

> Let's say today you generate the [Health] ID for someone and then they come to me. That person does not know that I also did it [generated a Health ID for them]. Then it becomes different IDs... It gets duplicated... I have done this 1-2 times... Once, in the duplicate, a different [Health ID] number came... They did not know [that their Health ID had been made earlier]... They came back next time... That time I found out. Earlier they forgot, and next time when they came, they said that I had already made it [Health ID] elsewhere and again you also made it. When I saw it, both ID numbers were different... Now what can I do? I told them 'keep it, use whatever you want'... Duplication [of Health IDs] will be happening a lot I feel.

The duplication of Health IDs was observed in various other civic dispensaries and hospitals in Chandigarh. Having multiple Health IDs means that PHR data linked to a person gets fragmented across multiple systems. It may get linked to one PHR on one EMR system at a hospital when its corresponding Health ID is provided, and may get linked to another Health ID on a different system. This means that each of these duplicates has some information about the patient but is missing other details, and the

[1] Name changed

patient may receive treatment on the basis of incomplete data. This defeats the purpose of the Health ID, which is meant to provide a longitudinal view of a patient's health history in one place.

The distribution and duplication of stored data about a digital identity, meant to correspond to a single individual's health, points to a form of disembodiment enabled partly by the view of health data as a commodity; while the body is indivisible, data is dispersed. The commodification of data fuels these processes by incentivising different actors to further duplicate and store this data in their databases for their own purposes, which is where the greatest threat lies.

3.3 Data Processing

Stored health data undergoes processing through many methods. Most commonly, health data is first 'cleaned' as a quality control measure to ensure that the stored data is reliable and accurate. It is also 'normalised' to ensure that reporting formats are consistent and missing data is accounted for. After this, most relevantly for this paper, health data may be anonymised, to remove personal identifiers, and aggregated, whereby individual patient data is consolidated into pools of patients based on various criteria.

The NDHM ecosystem allows HIPs to “make anonymised or de-identified data in an aggregated form available...for the purpose of facilitating health and clinical research, academic research, archiving, statistical analysis, policy formulation, the development and promotion of diagnostic solutions and such other purposes as may be specified by the NDHM'' (Ministry of Health & Family Welfare, 2020, p. 18). The goal here is to gain the benefits of correlative analysis while reducing privacy risks associated with having so much data on individuals.

It is assumed that the anonymisation or de-identification and aggregation of data would provide enough protection to individuals whose bodies generate the data. However, aggregate data can reveal personal details of particular individuals with the right kind of analysis (Tisné, 2020), and anonymised data can be re-identified through various techniques (Rocher et al., 2019), as will be explained in section 3.5. The Personal Data Protection Bill, 2019, makes re-identification of de-identified personal data without consent a punishable offence with imprisonment of up to three years, or fine, or both.

Moreover, using anonymised aggregation of data as a form of privacy control puts the focus on individual identifiability as the sole source of privacy violations. This ignores an entire class of other collective and embodied harms. For instance, some scholars argue that processes of aggregation move the focus away from individual bodies onto collective groupings, where the individual body becomes less and less relevant though it is affected by being part of the group for which decisions are made. This leads to the conceptualisation of “group privacy” (Taylor et al, 2016) which is an attempt to protect against the potential harms that may arise from being part of a group. Since such harms can remain considerable, they also require protection.

However, the NDHM framework stipulates guidelines for data processing based on personally identifying data, where personal data relates to an identifiable individual, not groups. In this light, the concept of group privacy raises questions about the underlying assumptions behind the NDHM framework in two primary ways.

First, limiting the focus of privacy protections to the possibility of individual identification invisibilises the vulnerability that people can suffer even when their health data is anonymous but part of a group, as discussed previously.

Second, it ignores the fact that when individual health data is aggregated, it paints a collective social picture of a population's health; a move from “biopolitics of the self to the biopolitics of the population” for the purpose of decision-making (Lupton, 2016). This move makes decisions taken on the basis of anonymised aggregated datasets seem disembodied, generic and impersonal. This is because with aggregation and anonymisation comes decontextualisation. While this may partly be a consequence of any kind of datafication, particularly stripping the data of its identification allows us to forget that this data is produced as intimate and immediate, tied to individual projects and contexts (Nissembaum, 2011). What is directly observable is data without a source or identity, though as mentioned earlier, it is possible to re-identify such data through various techniques. This specific manner of decontextualisation that comes with anonymised aggregation makes it easier for data to be controlled and manipulated; once the data is defined as an available de-identified resource, its processing and eventual monetisation are less likely to be contested. Within policy frameworks that permit for such processing as a form of privacy control, de-identification thus justifies and legitimises the processing of patient data through the perceived disembodiment that comes with aggregation and anonymisation.

3.4 Data Sharing

Health data may be made available for sharing with different stakeholders. As mentioned earlier, the Health Data Management Policy states that de-identified data in an aggregated form may be made available for sharing with the consent of the individual.

This is an important stage for critical analysis because sharing of data with various stakeholders is part of what constitutes the unprecedented threats and harms of the datafication of health. A large part of the problem here is that data collected in one context is shared and then used to make decisions in another context. To illustrate this problem, consider the case of insurance providers in the NDHM ecosystem who can collect various kinds of data about a patient's life and history. For example, data on a patient's socio-economic living conditions could be used to predict their health, which in turn could determine their insurance premiums, as illustrated by Mr. Murali, Principal Lawyer at Amicus, who has previously worked with the National Health Service in the United Kingdom (UK):

> [Insurers] may use data sources from the local municipal corporation about population density, existence of slums and low-cost housing and disease prevalence in my municipality... They could conclude from that data that there is a very high risk of contagious diseases here because there are slums all over this place. They don't have good sanitary facilities, they don't have adequate water supply in this whole area... Now I live in a high rise building where we have green space, we have an RO purifier for the whole of the building, we have proper healthcare facilities close by... But we, my neighbours and I, are only a few hundred in this whole ward.

The data that the insurance company is collecting about policyholders in this example is 'proxy data', and it is one of the many kinds of proxy data that can be collected about patients within the NDHM ecosystem, which provides a digital infrastructure to HIUs to collect data that they deem relevant to their decision-making. Health insurers could use this data to justify a higher premium for policyholders based on the socio-economic data of their neighbourhoods, even if that data is not an accurate representation of a particular individual's health. In some demographics, this is termed as “postcode lottery” where postcodes can directly affect the services an area can obtain, such as insurance prices. In India, beyond demographic location, other socio-economic predictors of health used to determine insurance already include an individual's occupation, education and income (Yellaiah & Ramakrishna, 2012).

Proxy data does not necessarily say anything meaningful about people's health. It alludes to a digital self or profile put together from collected data points that may not be representative of people's corporeal realities, as observed in the example above. Thus, digital profiles of people are produced from interconnected datasets, irrespective of whether they fit the profile or not, becoming “a person's shadow: hard to fight, impossible to shake” (Van der Ploeg, 2012). Yet these data points and profiles, defined by flawed proxies, are legitimised in the name of health, and can be collected on a mass scale through the NDHM digital infrastructure.

Further, one of the main objectives of the NDHM is to “ensure portability in the provision of health services'' (Ministry of Health & Family Welfare, 2020, p. 2) by making data shareable across platforms and stakeholders. A patient's longitudinal history of diagnosis, treatment, and reports will be available and accessible in its entirety to HIUs through a patient's Health ID. Portability enables data to be in places that its corresponding bodies are not. The digital nature of data contributes to this portability, allowing for data to be shared across systems. This means that a patient does not have to physically take the results of a lab test to a HIU; their data can be transported on its own through the digital infrastructure provided by the NDHM. However, our bodies are not shareable in the same way that portable data is. More importantly, with anonymisation and the possibilities for sharing that allows, people don't actually know where their data (or their bodies) go. Thus, portability tends to hide from view the relationship between data and bodies.

To some extent, the fundamental connection between a patient's body and their data is also recognised by the NHA because in principle, it makes provision for consent-based mechanisms to facilitate data sharing. This means that the authorities implicitly realise that the person whose body generates the data must have a say in how and with whom it is shared. Section 5 discusses how the model consent-based mechanism adopted by the NDHM does not empower patients, and thus any principled recognition of the relationship between bodies and data does not translate into praxis.

3.5 Data Analysis

The sharing of health data among various stakeholders leads to possibilities for data analysis that did not exist before. Data analysis is used to understand and identify trends or draw conclusions from data for making evidence-based medical decisions. Apart from manual analysis done by the medical professional who has access to the patient's health data, this stage may be carried out through computational and statistical techniques and programming. A key component of the NDHM ecosystem is proposed to be a Health Data Analytics Platform or National Health Analytics Framework, to enable the creation of anonymised and aggregated datasets that assist with statistics leading to data-driven decisions and targeted policymaking in the health sector (NITI Aayog, 2018, p. 32). According to NITI Aayog, this National Health Analytics Framework will initially focus on decision-making within the health insurance sector.

Discrete data is often analysed in conjunction with other information, and therefore, datasets are usually merged or linked with one another at this stage. Consider the NDHM Health ID, which would be used to share personal health data of patients among various stakeholders. The Health ID is linked to other data points. For instance, fieldwork indicated that it has been made a system-level requirement to create an Electronic Medical Record (EMR) in civic dispensaries in Chandigarh: without a Health ID, a person's EMR cannot be generated. Most Health IDs are also linked to Aadhaar, India's 12-digit unique identity number, either directly or indirectly. There are two ways of registering a Health ID at present: through a person's phone number or through their Aadhaar number. However, even many people's

phone numbers are already linked to Aadhaar. Moreover, some civic dispensaries are accepting only Aadhaar to register people. The reasons for this, as stated by Mr. Jadhav,[2] a multi-purpose worker at a civic dispensary in Chandigarh are the following (translated from Hindi):

> If you register through an Aadhaar card, the registration form is auto-filled. Because your name, photo, date of birth, address, email ID, mobile number - everything will be saved on Aadhaar, and it will automatically show up in the registration form. But if we register using a mobile number, all these details have to be manually entered, and many people don't know their own details... It takes me 15 minutes per registration through phone number, and just 5 minutes through Aadhaar. Also, with phone number registration... their Health ID card will not have any photo.

This preference for Aadhaar was echoed by data entry operators in other civic dispensaries and hospitals in Chandigarh. One of them, Ms. Namrata,[3] offered another reason for why the Aadhaar number was a preferred mode of registration (translated from Hindi):

> The reason many people don't do it [Health ID enrolment] through mobile is because they need to generate a password for it. So some people find the password difficult to generate... You have to make your own password... They don't do it then... After filling it [the form] in fully, there is an option below to provide a password... [For Aadhaar] it does not come. That's why it's easier through Aadhaar.

The NDHM Strategy Overview document itself makes a distinction between the creation of Health IDs using Aadhaar and those using a phone number. It mentions that "for those individuals intending to seek benefits of Government subsidy schemes... Unique Health ID will be generated based on Aadhaar" whereas "for those individuals not intending to seek any benefit of Government subsidy schemes, the Health ID may be generated...using email, mobile number, or any reliable government-issued proof of identity." (National Health Authority, 2020, p.14)

In this way, there are many system-based and structural incentives that have been designed to link a person's Aadhaar number with their Health ID, and therefore Aadhaar details eventually get linked to their PHRs. These incentives can be conceptualised as structural 'nudges' (Susser et al., 2019). For instance, in some cases, people are being explicitly required to link their phone numbers with Aadhaar if they choose to enrol for the Health ID using their phone numbers. This happened to Mr. Rizwan,[4] a caterer in Chandigarh who had visited a civic dispensary to get his Health ID made (translated from Hindi):

> Mine [my Health ID] did not get made... Mobile number was not linked to this. [Shows Aadhaar card]... So I did it, ma'am, I went and got it linked just now... When I asked them [health workers], they said if you link your Aadhaar to your mobile number, only then you will get the message for the Health ID on your mobile number.

The Health ID has also been integrated into the government's CoWIN (Covid Vaccine Intelligence Network) vaccinator portal to manage the inoculation drive for COVID-19 (Sheriff, 2020). There are many reported instances of Health IDs being registered for people without their consent through the CoWIN portal at the time of vaccination (Dogra, 2021). Further, Health IDs are expected to be linked to

[2] Name changed
[3] Name changed
[4] Name changed

to the government's e-Sanjeevani telemedicine service under the NDHM: the platform will be capable of generating Health IDs, which could then be used to catalogue health records and allow for them to be shared digitally (Porecha & Singh, 2021).

The PHRs that are finally made available at the analysis stage are thus likely to be linked to various other databases and systems, such as the EMR, Aadhaar, CoWIN, and e-Sanjeevani. Each of these databases is in turn connected to other databases. For example, it was recently mandated for all Aadhaar numbers to be linked to the holder's PAN (Permanent Account Number) card issued by the Indian Income Tax Department (Sharma, 2021). All this data put together contains a lot of personal information, including a person's name, age, date of birth, phone number, individual and family medical history, financial details, etc. Moreover, since data exists across decentralised proprietary systems that are used by different HIUs such as insurers, they may include other data points in their own databases that are relevant to their decision-making. For example, in their own databases, insurers may include data about a person's sleep patterns, family history etc.

Databases are linked in this manner for analysis primarily because 'raw' data has little monetisation value without meaningful analysis performed upon it (Mandel, 2017). But in conjunction with other data points, this data can feed algorithms to analyse trends and slot people into risk categories. Through anonymised aggregated health data, health insurers can identify patterns such as 'women over the age of 40 who sleep less than 6 hours a day are more prone to the risk of heart disease'[5] and use these patterns to raise premiums for policyholders that fit this risk description at the time of issuing the policy. When machine-learning algorithms are used to identify such patterns, their complexity often also makes these trends unexplainable and therefore uncontestable (Radhakrishnan & Sinha, 2020).

A key assumption within the framework of conceptualising data as a disembodied resource is that data is valuable only for human consumption and extraction, as is characteristic of commodities, but may not be valuable in itself. The notion of 'raw data' is based upon similar assumptions. Raw data has strong metaphorical parallels to natural resources such as raw metal ores. By themselves, these resources have no real monetisation value. Moreover, just as natural ores need to be mined to unearth valuable minerals, raw data is considered to require human or technological interventions to be transformed into something useful. Techniques of data analysis facilitate this transformation to 'unearth' meaning and money from data, and in this process, also cement the construction of data as disembodied.

This section has analysed how the relationship between health data and people's bodies that generate this data is invisibilised at each stage of its life cycle in the NDHM policy framework. This happens through digitalisation when data is generated and collected; duplication of data when it is stored in a decentralised digital ecosystem; anonymisation and aggregation when data is processed; the use of proxies and portability when data is shared with different stakeholders; and dynamic interlinkages with other datasets when this shared data is analysed.

4. What changes with the disembodiment of health data?

This section analyses three major shifts that are observed within healthcare in the age of big data when health data becomes disembodied within policy frameworks: 1. what data can be collected to predict a

[5] Illustrative example

person's health; 2. who can access this data ; and 3. what these entities can know and do through such access to this data.

Consider the following scenario to illustrate these shifts within the NDHM ecosystem:

- Data Principal, Kranti, has registered for a Health ID through which she can access her Personal Health Records (PHRs). The longitudinal health data in her PHRs is stored in a federated manner across different HIPs (such as hospitals and diagnostic labs), and is unified through her Health ID.
- Under the NDHM ecosystem, there is a provision for a health locker cloud storage facility, as introduced in Section 2. Kranti can set up a health locker account to act as a gateway to all her data stored across different HIPs. She may use it to store her one-time issued health documents (such as health certificates and reports) and her transactional data (such as PHRs).
- Kranti has a family history of diabetes and gets regular blood-sugar tests done, the results of which she shares with her doctor. Through the NDHM infrastructure, Kranti can create a consent artefact to share her blood-sugar test results stored in her Health Locker (HIP) directly with her doctor (HIU) on an ongoing basis, without having to personally visit the doctor each time to update her diet plan.
- Kranti and her doctor may also get creative. Under the NDHM ecosystem, “users can add reading from IoT and other devices like wearables to their PHR. The data will be stored in the Health Locker which can act as a HIP for the user” (National Health Authority, 2020, p. 12). So her doctor may ask Kranti to use a GoQii wearable fitness tracker to track her glucose level and the number of calories she is burning everyday. In this case, Kranti may set up a consent artefact to automatically share data from her fitness tracker with her doctor as well.
- Kranti's doctor is not the only HIU who is getting creative in this ecosystem. Another HIU, Kranti's health insurance provider, Max Bupa, is looking to improve their market share. They have collaborated with GoQii, the company designing Kranti's fitness tracker. GoQii shares user data from their fitness trackers with Max Bupa, which enables Max Bupa to assign a health score to its users. Based on this score, they offer discounts on insurance premiums to select policyholders (Subramanian, 2018).

4.1 What data is collected about health?

Earlier, health data was largely confined to analogue medical records which were the sole means of determining an individual's health. In such cases, patients were more likely to be aware of what data of theirs was collected by medical professionals and what was being done with this data for their treatment. Analogue health data has also historically undergone the stages in the life cycle analysed in Section 3. For example, to contain the tuberculosis epidemic in India since the mid 1980s, data has been collected on a mass-scale, aggregated and analysed, and thus must have undergone a degree of disembodiment that accompanies these processes. Thus, the disembodiment of health data is not by itself an altogether new phenomenon.

However, in the age of big data, wider means of data collection are fueled, in particular by enabling the collection of data that may not necessarily signify the health of the body. An example of this, in the case of proxy data about socio-economic conditions to predict health, was mentioned earlier. In this respect, health data may now include both clinical data and non-clinical data. Clinical health data includes all data that is collected during a clinical encounter within traditional clinical boundaries, such as lab test results, patient diagnosis, etc. Non-clinical health data could include self-reported health data, data from wearable devices (step count, sleep patterns), environmental data (air quality), social media data (user posts), behavioural data (smoking, alcohol use, diet, drugs, sexual history), socio-economic data

(education, employment, housing conditions) etc. (Golembiewski, 2019). There are also other kinds of data that may come into the picture, such as data about a person's shopping history, driving habits, etc. Though such data points are not intuitive predictors of health, given the increasingly networked databases that health data is stored in (as analysed previously) this data can also enter health databases. Thus the boundaries of what kinds of data can be used to predict an individual's health have been blurred.

With the emergence of newer, non-clinical predictors of health, our digital phenotypes have now extended to data beyond our direct bodies, and include social and environmental factors. Some scholars argue that in contemporary digitised contexts, health data can be considered as any data relating to health and well-being and defined by the characteristics of being ubiquitous, comprehensive, personalised, and measurement-based (Ada Lovelace Institute, 2020), while other scholars claim that 'all data is health data.' One of the major shifts that thus come about through the use of big data in healthcare and the associated disembodiment of health data is what kinds of data can be collected to predict a person's health.

Consider the “Quantified Self Movement” which is aimed at quantifying the self by attempting self-knowledge through numbers. One way this is achieved is through the digitisation of self-tracking of health, such as the use of non-clinical health data gathered from wearable fitness trackers. Scholars argue that such quantification represents an instance of biopolitics of the self whereby the body is made amenable to digitised health management and monitoring techniques (Ajana, 2017; Lupton, 2016). Such self-tracked data has initiated new relationships between people and their bodies, with bodily intuition now outsourced to non-clinical data in a way that was previously not possible through analogue techniques of self-analysis (Smith & Vonthethoff, 2017).

To a limited extent, policies governing the NDHM ecosystem recognise some forms of non-clinical data to be a form of health data, such as data from fitness trackers (National Health Authority, 2020, p. 12). That is how, in the illustrated use case, Kranti is able to add data from her fitness tracker to her PHRs. But the current policies do not make any attempts to respond to challenges that arise from different forms of health data being treated differently within broader policy frameworks. For example, in India, most fitness trackers are not classified as medical devices (Drugs Controller General, 2017) and are thus not subject to the same regulations as medical devices, though the data they collect is used to predict a person's health. These regulations are primarily aimed at the manufacture, import, sale and distribution of medical devices in the country. This creates a governance gap, and many newer applications of health data are at risk of falling through this gap.

4.2 Who can access health data?

Not only is a lot more data about health being collected in the age of big data, this data is now increasingly accessible to stakeholders outside of traditional clinical boundaries. Earlier, health was primarily a matter of discussion between a patient and their doctor. This was partly due to the limited availability of technologies to collect data at a mass scale as well as the limited technical ability to produce meaningful insights from it.

When health data is treated as a disembodied resource, it legitimises the sharing of data among more non-clinical stakeholders who can derive value for themselves from the resource. But the reason private actors want to access patient data in the first place is because in reality, health data is embodied and says something about the bodies that generate this data, something that private actors have a business interest in knowing and monetising. Legibility is defined as “the way in which citizens become visible...

to authorities through data collection and analysis" (Taylor, 2016). This increased legibility of patients to private actors, irrespective of whether it is accurate or not, is the second major shift that arises from the disembodied datafication of health.

Private actors—health insurance providers, fin-tech companies, employers, advertisers, and data analytics companies—can now access the most intimate data about people that is collected under the broad ambit of health. Such companies monetise data that has been voluntarily generated by individuals by extracting and combining it with other data to draw correlations and inferences that hold value in the market of health data (Nissenbaum & Patterson, 2016).

For example, in India, private companies such as HealthifyMe (http://www.healthifyme.com/corporate-wellness/), ekincare (http://www.ekincare.com/about-us/who-we-are), and InnerHour (http://www.innerhour.com/corp-features), have begun offering data-driven corporate workplace wellness schemes.[6] Such schemes incentivise employees with the prospect of bonuses or discounts on health insurance premiums to engage in self-tracking activities such as exercise, thus generating data that employers or insurance carriers can analyse when the schemes are offered as part of employment health insurance packages (Christophersen et al., 2015). Since health insurance costs of employees often get borne by employers through such offerings, companies have a business incentive to increase their potential cost savings through such datafied tracking of their employee's health. Thus, datafication of health in this case becomes a method of increasing control over employees by mitigating economic risk for employers and shifting this risk on to employees, who now are obligated to maintain a 'healthy lifestyle' as prescribed by their employers even beyond the workplace (Ajana, 2016).

At present, India doesn't yet have a national digital database of clinical health records of patients. Due to this, insurers do not have much insight yet into their policyholders beyond the preliminary health-checkup at the time of issuing policies. As Kranti's case suggests, this has led to some insurers mining alternative, non-clinical forms of data to attempt to paint a more holistic health profile of policyholders. For example, Kranti's insurance provider is able to access the data collected by her wearable fitness tracker, such as her sleep patterns, calorie counts, step counts, heart rate, etc. Solving their problem of the unavailability of health data, the NDHM particularly incentivises insurance providers to integrate within its ecosystem. According to Mr. Srikanth, an independent public interest technologist:

> This entire datafication is happening because the state wants private sector led healthcare using a health insurance model, and the insurance industry needs data. This is probably the reason why we are having the health datafication... The architecture and thinking behind the National Digital Health Ecosystem is in line with this vision of the private sector led healthcare and the state playing the role of an 'enabler.'

Patient data within the NDHM ecosystem may be available to private insurance providers in a variety of ways. One way this sharing could take place is through formal agreements between HIPs and HIUs, which will now be possible through the NDHM digital infrastructure, as illustrated in Kranti's case. It has been reported that more than 20 companies have already offered their technology and solutions to the NDHM, including GOQii and some insurance companies (Singh & Porecha, 2020). Cyber leaks and hacks could also make patient data vulnerable to access by private entities. According to the draft Digital Information Security in Health Act (Ministry of Health & Family Welfare, 2018, p. 25), any person who breaches digital health data shall be liable to pay damages by way of compensation to the owner of the

[6] I reached out multiple times to these companies for interviews for this study, but did not receive any response from them.

digital healthcare data in relation to which the breach took place. In cases of a 'serious' breach, the Act mentions punishment with imprisonment or fines. However, despite these provisions, various cases of serious data breaches have occurred in the past few years relating to health data indicating that sufficient precautions are not being taken to safeguard patient data (Kurian, 2021).

4.3 What can those with access to health data do with such access?

When individuals become legible to private actors who can access massive amounts of personal data about their lives, it changes how people's health can have value, and who can benefit from that value. A major incentive for datafication is to collect personal data of people to gain granular insights into their lives and experiences and make profitable predictions about them. This is possible because our personal data contains insights into who we are, what we do, and what we want (Greenwood, 2014).

Health insurers use personal data to sort, rank, and differentially charge their customers for services. Consider Kranti's use case of Max Bupa's collaboration with GOQii to determine insurance premiums.[7] This is not a singular case. HDFC Ergo[8] has a “Stay Active” scheme that encourages policyholders to increase the number of steps they take every day by rewarding them with discounts on their policy premium at the time of renewal, based on the average daily step count they record on their Health Jinn App throughout the policy year (“Stay Active,” n.d.). Aditya Birla uses a health assessment to assign a Healthy Heart Score to policyholders, wherein they are categorised as Red, Amber, and Green to receive discounts on premiums, based on their probability of developing heart disease (Aditya Birla Health Insurance Company Pvt. Ltd., 2016). This may not yet be a popular model in India, but the NDHM infrastructure incentivises more takers for it.

In these examples, proxy data such as step counts and sleep patterns is used to singularly predict a person's health, and consequently their insurance premiums, even though there is no evidence to suggest that people who take a certain number of steps a day are less likely to make an insurance claim. The low-cost mass-market tracker devices and apps used for this data collection further put into question the accuracy of the data collected in this manner. While the insurer doesn't directly raise premium rates, less 'healthy' policyholders who do not meet their targets on the app (or may not even be able to afford using a smartphone app) end up paying more than their 'healthier' counterparts, with health here being defined by flawed data proxies.

This affects not only individuals, but also their families. In Kranti's case, she is not herself diabetic but has a family history of diabetes, as mentioned. Because health insurance coverage is often offered at the level of families in India, insurers have a business interest in mining data about the health of families. Though insulin for diabetes is itself not covered by health insurance in India, related health conditions such as hypertension, retinopathy or diabetic foot ulcers, are covered. In Kranti's case, insurers would now be able to predict if she is likely to get diabetes because of Kranti's PHRs which would be linked to her family's PHRs in the NDHM ecosystem. This would be possible even without Kranti's explicit consent. During fieldwork, I observed that when an individual is enrolled for a Health ID, they are asked for the details of their entire family so that their families can also be enrolled, as evidenced in the following conversation with Ms. Narima, an ANM worker in a civic dispensary in Chandigarh (translated from Hindi):

[7] I contacted Max Bupa for comments but received no response to my emails.

[8] I contacted HDFC Ergo for comments but received no response to my emails.

[Ms. Narima]: Everyone in the family won't come [for Health ID registration]. One family member comes with the date of birth of all members and with one family phone.
[Me]: So then you make IDs for the full family through that one person and one phone?
[Ms. Narima]: Yes.

In this way, a person's health data gets linked to their family's health data through names, phone numbers, addresses, or other common identifiers, irrespective of whether the family has consented to it, or is even aware of it. With all of this interlinked data available to insurers under the NDHM, they are in a powerful position to access health records of people who have not consented to share this data.

In Kranti's case of hereditary diseases such as diabetes, her own chances of inheriting the disease are high. While earlier, insurers would not be able to know this unless it is declared by the individual, they would now be able to access this data at ease. As a result of this, insurers can either exclude individuals like Kranti from their insurance policies, or differentially treat them by pricing premiums to account for that risk, based on data obtained through their PHRs as well as the aggregated data of others related to the individual. Thus, an individual may find that they have been denied health insurance on the grounds that they are at a high risk of a particular disease, while this risk may have been indicated by their digital phenotype rather than their self-declared medical history (Ada Lovelace Institute, 2020). These predictions also attain credibility and normative authority because of their medicalised nature, which may be used in turn to further justify the practices of categorising and profiling individuals.

All this undermines the original purpose of insurance, which is to balance risk in society and protect those most in need. Insurance is predicated on the understanding that everyone deserves medical care without a strain on their finances. However, as private insurance providers gather more data about people and are able to make predictions about people's lives, they can pinpoint the riskiest people and either increase their premiums or deny them coverage, hitting those who can least afford them the highest (O'neil, 2016).

In this way, access to people's health data becomes a form of power, giving those with such control the unparalleled power to influence the governance of people's bodies and lives. Private actors can use biopolitical categorisations of 'fit' vs. 'unfit', 'healthy' vs. 'unhealthy', 'risky' vs. 'risk-free' to discriminate and differentiate more decisively between people (Ajana, 2017). This commodification and exploitation of personal health data is accompanied by new forms of techno-scientific objectification of bodies which are “sliced and diced into decontextualized parts, and bought and sold” in the health data marketplace (Nafus & Neff, 2016, p. 62). The disembodiment of data within the NDHM ecosystem thus creates a scenario where these actors controlling its digital infrastructure can now exploit patients from afar, rather than having to control their bodies in person (Van der Ploeg, 2012). This happens without patients necessarily knowing what is going on and who has access to information about them.

Part III: Embodied Implications of the Datafication of Health

5. Patient Rights

So far, the paper has analysed the disembodiment of health data in the age of big data - how it takes place at various stages in the life cycle of data (section 3), as well as the changes that come about from it (section 4). By focusing on the implications of datafication through the NDHM, Part III now shows that the disembodiment of health data undermines patients' right to healthcare, and that recognising this embodiment can empower patients to affirm their rights. This section unpacks the impact of current trends regarding the disembodied datafication of health on the rights of patients and foregrounds the threats to these rights that come to light when we put bodies back into the picture.

Before analysing various patient rights in detail, it should be noted that though some degree of disembodiment has historically taken place with health data, as mentioned earlier, medical legislation and codes of ethics have been developed and implemented over the years to ensure that patient rights remain safeguarded (Indian Medical Council, 2002). However, with the advent of big data and the changes that have come about with it as analysed earlier, the patient is no longer the primary focus of the health data ecosystem. Therefore, this section centers the embodied lived experiences and needs of patients within the NDHM ecosystem to better understand what is new in this moment of time that requires our urgent attention.

5.1 Consent

The NDHM proposes a consent-based framework for the collection, processing and sharing of personal data of patients (Ministry of Health & Family Welfare, 2020, pp. 6-10). According to this framework, data fiduciaries can collect or process personal data only with the consent of the data principal, and this consent must be free, informed, specific, clearly given, and capable of being withdrawn (Ministry of Health & Family Welfare, 2020, p. 6). The consent framework also requires all data fiduciaries to give a clear and conspicuous privacy notice to data principals (Ministry of Health & Family Welfare, 2020, p. 7).

Scholars have argued that such mechanisms through which consent is sought are inadequate in many ways (Solove, 2013). Consider the case of consent fatigue which is caused by requiring individuals to read through lengthy privacy policies and manage consent for every application they use, as would be required under the NDHM. An example of how this may work in this ecosystem is explained by Mr. Murali:

> Once the [NDHM] system comes in place, it will be a condition of the health insurance policy that they [HIUs] will constantly get the right to update their information with my medical records...So during the duration of this policy, I consent that they [HIUs] will have access to my medical health records... They will add this at the bottom of the page somewhere. And when there are five pages of terms and conditions, you would say 'I agree'... So I don't know that I'm giving consent to...because I have not read five pages of terms and conditions.

Moreover, seeking consent through individualistic privacy notices cannot empower people within disempowering structures and ecosystems. For instance, as this paper analysed previously, even if a person denies consent to their health data being used, an insurer can use the health data of other people such as their family members to make statistical extrapolations and set premiums in accordance with family health risk for the person. This is applicable even if that data has not been sought with the consent of the family members in question, as was noted earlier.

Thus, though the NDHM states that data principals should at all times have decision-making power over the manner in which their personal data is collected and processed (Ministry of Health & Family Welfare, 2020, p. 6), an individual's own decisions are not always sufficient to maintain control over their personal health data. In addition, individuals often lack the ability to withhold or withdraw consent vis a vis actors such as health insurance companies because these companies wield a lot more power than they do (Tisne, 2018). Patients are always at the disadvantaged end of power and information asymmetries and may also be in a vulnerable state of mind at the time of giving consent.

As seen here, the consent framework proposed by the NDHM does not empower patients in reality. This is largely because consent is being obtained here within an ecosystem that fundamentally views health data as a resource and structurally incentivises businesses to monetise this resource to generate value for their businesses, as analysed earlier. By proposing to introduce consent within such inherently disempowering structures that the NDHM operates within, the NDHM effectively invisibilises the power relations that keep patients from consenting meaningfully to the usage of their health data (Kovacs & Jain, 2020).

Due to the many failings of such consent-based mechanisms, some scholars have proposed to do away with consent altogether in data protection (Matthan, 2017). However, Kovacs and Jain (2020) warn that such proposals further legitimise the construction of data as a resource. In order to emphasise the relationship between bodies and data, what is needed is a feminist re-envisioning of consent frameworks which takes into account people's social realities in accessing healthcare. An implementation framework for the same in the context of health data will be discussed in Section 6.

5.2 Choice

The Health Data Management Policy mentions that “participation of an individual in the NDHE will be on a *voluntary basis* and where an individual *chooses* to participate, he/she will be issued a Health ID... by the NDHM” (Ministry of Health & Family Welfare, 2020, p. 1, emphasis mine). However, fieldwork indicates that the Health ID has been made mandatory in Chandigarh. The Chandigarh Health Department authorities sent a WhatsApp message in August 2020 to all health workers mentioning that “The registration for generating Health IDs is mandatory for all the citizens of our country” (see Figure 3). All the health workers I spoke to in Chandigarh, across multiple civic dispensaries and hospitals, confirmed that they had received this message.

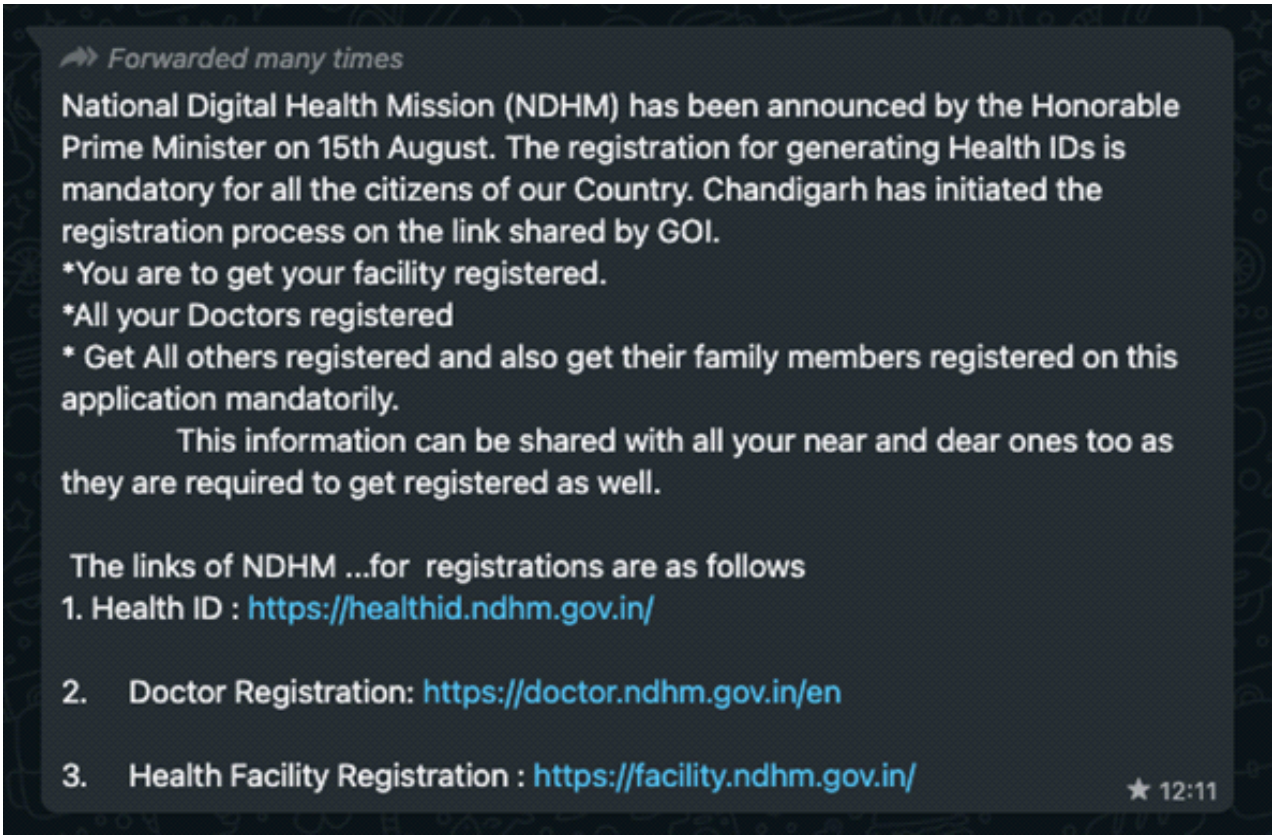

Figure 3: A WhatsApp message sent to health workers in Chandigarh by the Health Department

A member of the NHM Employees Union, Mr. Irfan,[9] said (translated from Hindi):

> [Health ID] has been made compulsory. There were directions from the top that Health ID has to be made for everyone... For any hospital service, when they [patients] have to make the visit card [appointment card for hospital visits], that time the first thing they are being asked for is the Health ID.

I verified that registration counters at civic hospitals are requiring patients to register for a Health ID before issuing them an OPD card to see a doctor. A pharmacist, Mr. Arun, also mentioned his sister's experience with this: “Around 4-5 months ago, my sister, she was pregnant, she went to the hospital...Her [Health] ID was on-the-spot generated at the registration desk” (translated from Hindi). At a civic dispensary I visited, Dr. Amit,[10] a medical intern who was given the responsibility of enrolling patients with the Health ID, said that they were not providing medication to patients unless they got a Health ID made:

> We write on the side of the prescription "health ID and EMR" and...**we tell our staff that till I don't sign this [saying] that they [patients] have made a health ID and EMR, you don't give the medication.** This is one of the ways to positively influence people. **This is a mandated thing for them**... It's mandatory because they asked us specifically how many [Health IDs] we made in a day... So **they give us targets that this is how much you should achieve in a day. So that's why we have to... force patients or people to make it.** So you have to innovate different techniques to tell them this is important. (emphasis mine)

[9] Name changed
[10] Name changed

Most patients in India don't have a meaningful choice when it comes to accessing medical care which has historically been an "imposed preference" for the poor-sick (Prasad, 2007). Their choice is always already assumed. Thus, patients end up participating in the NDHM ecosystem not because they freely choose to do so, but because they do not have a meaningful choice in rejecting it (Radhakrishnan, 2021).

While this has a more adverse impact on marginalised communities, it also impacts other classes of society. For example, if one health insurance company decides to introduce a clause in their policy that enables them to continually access a person's health records and profit from it, as pointed out earlier, then it is expected for other insurance companies to follow suit, since the NDHM incentivises such monetisation. Mr. Murali explains:

> I will have no choice because every insurance company is going to have the same terms and conditions. So I can't do anything about it. Either I sign up to it and get insurance or I don't sign up and don't get insurance.

It is important to look at policies not just on paper, but also in terms of their implementation, especially since Indian regulation has a history of setting up digital infrastructures that are voluntary on paper but mandatory in practice (Khera, 2017). In practice, one of the reasons for why the NDHM is able to mandatorily enrol people into its digital ecosystem is that it most likely regards health data to be a 'public good.' The Economic Survey 2018-19 specifically pitched for data to be treated as a public good in India (Ministry of Finance, 2019). Framing data to be a public good is an emerging trend under 'data philanthropy' that has been gaining momentum in recent years after it was introduced by the United Nations United Nations Global Pulse (2009). At the heart of this trend is the belief that data sharing through partnerships between private and public entities is a positive act that can be beneficial to the public and can enhance policy action. Since then, various campaigns have been launched globally to promote the philanthropic sharing of personal health information (Patients Like Me, n.d.). In a localised context, the NDHM considers health data to be an asset or a resource that could potentially be beneficial not only to the individual but to society at large, thus trumping individual interests with the larger national interest of digitising health data in collaboration with private entities such as health insurance companies. This framework of viewing data as a disembodied public resource is therefore at least partly responsible for its mandate in compulsorily enrolling people into its digital ecosystem.

5.3 Privacy

One of the guiding principles of the NDHM is "privacy by design" for the protection of a data principal's personal digital health data privacy (Ministry of Health & Family Welfare, 2020, p. 14). Yet, many privacy issues exist within this ecosystem which are concerning due to the nature of sensitive health data involved. Mr. Arun, a member of the NHM Employees Union, said (translated from Hindi):

> [Registering Health IDs] is not the work of any particular post...If you go and tell people that there is a scheme of NDHM, and if you promote it, and if the person is willing to enrol or if you convince him somehow to enrol, then you can also do it... [but] nowadays on social media, TV, radio, FM etc. they say that if someone asks you for your OTP, don't give it. So this [cybersecurity] fraud can happen. We have to be alert about it.

This risk is compounded by the low levels of digital literacy in India. For instance, as mentioned in Section 3.5, the Health ID registration requires a password to be generated when the enrolment is being done through a phone number instead of Aadhaar. Since patients often find it difficult to generate their own passwords, the health workers doing their registration provide passwords for them. In most civic dispensaries I visited, a specific format is being used for the passwords, as explained by Dr. Amit, a medical intern:

> We provide a password.... We write the name of the person with the [redacted] as the capital letter and then we use the symbol '@' and then we write the [redacted] of that person as the password. So we tell them this or we write it on the prescription as well that 'this is your password'. Also once they have their Health ID number they can access it later and change the password according to their convenience.

None of the patients I spoke to said they knew how to change this password. They did not know they were even required to. This means that most of the Health IDs being generated for those lacking digital literacy are prone to easy cyber attacks.

In principle, the need to protect data from intrusions such as cyber attacks is the understanding of privacy that the NDHM adopts. However, feminists argue that privacy violations are not just violations of data itself. Privacy violations have embodied social consequences which lead to violations of people's bodily integrity, autonomy, and dignity (Radhakrishnan, 2020). In fact, privacy is an important component of bodily integrity (Debatin, 2011). For example, sensitive patient history can now be accessed by multiple private stakeholders through the NDHM digital infrastructure. Mr. Arun, a member of the NHM Employees Union, said (translated from Hindi):

> Privacy issues may be there. If someone has some major illness like TB [tuberculosis] or some contagious disease where the patient does not want to share their history with anyone, then this disease history will go to the doctors. If the doctor's laptop is accessible to anyone, then that information can be passed on to someone else.

There is precedence for why this is dangerous in the health care system in particular. In 2015, the National Aids Control Organisation (NACO) began urging states to collect the Aadhaar numbers of people living with HIV (PLHIV) to link them with their patient identity cards issued by antiretroviral therapy centres. Given the stigma surrounding HIV/AIDS in India, many PLHIV started dropping out of treatment programmes across India for fear of being identified through a breach of their privacy (Tomar, 2017). One of the fears then was that databases being formed on the basis of the Aadhaar numbers could be seeded by private actors and be used by insurance companies (Rao, 2017). The takeaway from this is that when states insist on people identifying themselves at the cost of their privacy to receive treatment, those living with stigmatised conditions are more likely to refuse healthcare. As Mr. Arun points out, this is likely to repeat with the NDHM because health data would be stored in databases accessible to many actors.

This example also indicates that an increased legibility of people's health and a loss of control over who can access one's data blurs the contextual boundaries that structure a patient's privacy expectations around health data (Ada Lovelace Institute, 2020). A PLHIV may be comfortable sharing their HIV status with their doctors within the context of clearly defined clinical boundaries, such as patient-doctor confidentiality, which prevent that data from being shared with others. But increased legibility to private actors such as insurers enables this patient data to now flow into proprietary databases that patients may not be comfortable with. This is concerning not only because such legibility hampers one's ability to live a life free of capital and state control, but also because it fails to account for the boundaries of privacy management within one's own home. For example, PLHIV don't always share their HIV status with family members. If Health IDs are known to all family members, they would lose control over the decision of whom to share this information with and when.

Similar concerns also apply in the case of abortion, which is highly stigmatised culturally despite being legal until 24 weeks of pregnancy. Privacy is a key concern among women accessing abortions and it is already under attack due to mandatory submission of digital ID documents required to curtail sex-selective abortion. Currently, if a woman needs a female contraceptive, she can privately approach an

ASHA worker in her locality for it without her family or others finding out about it. But if this history is made available in an EHR which is updated with her medical history and may be accessible to various stakeholders, it hampers not only women's privacy but also their autonomy to make decisions about their bodies.

Thus, threats to privacy under the NDHM not only affect data, but have far-reaching social consequences for the bodies and lives of people, with the worst affected being the marginalised with stigmatised health conditions who are most in need of healthcare.

5.4 Ownership and Control

According to the NDHM, "true ownership and control of the personal data will remain with data principals" (Health Data Management Policy, 2020, p.14). This sounds empowering and appealing because it suggests that the NDHM is giving patients power over their health data.

However, ownership is an inappropriate and counterproductive framework for data protection because it is embedded in extractive market logics and ignores concerns of how data shapes societies. A person cannot really own most of their data, and even if they could, it would not protect their rights. For instance, the practice of "postcode lottery" was discussed earlier in this paper to highlight how proxy health data such as socio-economic conditions of a person's neighbourhood can be used to determine the provision of services to a person, including justifying a higher health insurance premium for them. A person can never really "own" their demographic profile, and doing so would not stop insurers from making statistical inferences about their health based on the demographic profile of their neighbours.

Fundamentally, data ownership is a popular framework because policies view data as a resource, a capital commodity, property to be owned and traded. This fails to capture the relationships between people's bodies and their data, and the power relations governing these relationships.

We need to rethink ownership in a way that has little to do with capitalist relations of exchange and more to do with feminist thinking of the body. A popular feminist slogan in sex education to teach children about physical boundaries is "my body belongs to me." (Interface Children & Family Services, n.d.) Similarly, there is a long history of feminists using the slogan "my body is mine" ('My Body Is Mine', 2019) in the fight against sexual violence. These mottos could be understood to indicate a form of ownership. But when feminists support such mottos, it is important to note that they centre the notion of bodily integrity within this framework. For instance, when individuals engage in shared sexual experiences, even within market logics such as commercial sex work, the inviolability of the body is still central to that experience. But such concerns of bodily integrity are not covered by the framework of data ownership.

5.5 Clinical Care

The datafication of health offers potential to improve clinical care. As mentioned earlier, one of the main objectives of the NDHM is to ensure portability of health services. This means that with the advent of the NDHM, "all individuals will be able to conveniently access their personal health records'' digitally (National Health Authority, 2020, p. 27). According to a member of the NHM Employees Union, Mr. Arun[11] (translated from Hindi):

[11] Name changed

> India is right now very backward. In a town like Chandigarh, we don't have computerised systems. In the future, there is a vision that all doctors will have a laptop, all pharmacists will have a digital record of transactions. Through NDHM, the [Health] ID that is being created will help you to see all this data online. If a patient comes, we register his data online. If the patient tomorrow goes to... any other nearby facility, then the doctor can see all the data online.

For the general public, this vision holds promise for clinical care, as explained by Mr. Maajid,[12] a BPO worker in Chandigarh (translated from Hindi):

> We won't have to carry our medical documents. Everything will become digital. Our entire health history will be there in it. Sometimes our documents get lost, the [visit] cards they make for us also get lost. Digitally there will be past 10 years history also - what had happened earlier, all that.

The NDHM thus plans to improve clinical care by provisioning for data portability. As analysed earlier, among other factors, digital portability in the age of big data contributes to blurring the relationship between data and bodies by enabling data to be in places that its corresponding bodies are not.

At the same time, the datafication of health also shifts focus away from the care of bodies towards the care of datafied bodies through increasingly data-driven methods of patient engagement beyond traditional clinical boundaries. One way this happens is through the promotion of evidence-based treatments. While evidence-based treatment is not in itself bad, the way it plays out in private healthcare requires attention. One of the expected outcomes of the NDHM is to "enable evidence-based interventions" (National Health Authority, 2020, p. 27). Mr. Srikanth explains what this could mean for clinical care:

> The insurance-led model already has made treatment, particularly in secondary and tertiary care, diagnostic-heavy to justify course of treatment, in some cases solely for the benefit of insurance providers...[thus] increasing cost of care. Health datafication will only lead to more of it, where more and more parts of care will be preconditioned... on the data itself, and access, affordability of care will be affected. Denying or delaying treatment preconditioning on access to data... could lead to suboptimal outcomes.

Such datafication is driven by the twin motivations of HIPs wanting to modernise patient care and technology companies and private HIUs wanting to gain a bigger share of the lucrative health data market. Electronic personal health records, a key component of the NDHM, are one example where these motivations converge. They are understood within the ecosystem to be necessary for improving patient care and for patient empowerment. Dr. Amit, a medical intern at a civic dispensary in Chandigarh, explained how the EMR systems designed in India currently work:

> EMR has a lot of details.... You get the option to first choose what kind of patient came to you, [whether] it's an outpatient or inpatient or somebody who just came in emergency. So a lot of... options are being provided, so we choose an option first. Generally here it's a dispensary...so we have to choose an... outpatient... Then for the outpatient, a different window comes, in which there are a lot of options... Whatever has been written in the prescription, we've tried to record all the data in it. So we type in allergies... to any drugs or some form of allergens that they know of. And then the option comes of history... So, the long term chronic histories, we note down in that, and

[12] Name changed

> then after that we have to write the symptoms of the patient, and then what diagnosis did we make... Then treatment and then pharmacologically whatever drugs we gave, we do that. So we fill all of this in and we save the data. That data is supposedly saved in that patient's record.... The future concept of it is, the next time a different doctor sees it, he can access all the records, and he can know that maybe the patient has some allergies to some drugs, and he should not prescribe those drugs.

While this sounds promising as an ideal, it should be noted that EMR systems are being implemented in India based on their success in high-income contexts, without taking into account the reasons for clinicians' inertia in the adoption of EMRs in low-income contexts (World Health Organisation, 2021). In India, doctors tend to be more time-strapped and pay less attention to the elaborate documentation of clinical notes that is required for an EMR (Kandhari, 2017). According to the most recent OECD data, India has a ratio of 1.9 doctors for 1,000 inhabitants, while Germany has 4.5 doctors per 1,000 inhabitants (OECD Data, n.d.).

Moreover, the use of EMR systems can take decision-making away from medical experts (Gawande, 2018). It is another way in which 'evidence based interventions' may become predetermined by data. Mr. Murali explains how this can happen:

> When you go to visit a doctor, he will open up the screen, put in the symptoms... and the symptoms will all be drop-down boxes. So if I say I have a headache... that will give a choice of four things. And then depending on which one of those I pick... it will give me the next ones... The doctors are just going to tick boxes. So rather than them having a medical view of how they will diagnose or come to conclusions, that process will be usurped by drop boxes. So you tick the box, and it gives you two choices. If you hadn't got those two choices, you would have probably made choice three.... So it's taking away decision making from the doctor, and the pharmacist and all the experts.

He further explains how this can lead to an erosion in clinical care and exclusion of those who do not fit patterns identified by EMR systems:

> It [EMR system] is throwing up options based on patterns from the past. But individual human beings don't fit patterns... And once we go down this path, then everybody who doesn't fall within the pattern will be excluded or will be treated badly. And every time you treat somebody badly... it affirms the pattern. So then more and more people get excluded, because the pattern is reconfirmed by these exclusions. And so you create a pattern which is not true at all. But we'll have lots of data to point you and push you towards a course of treatment... And then you will show this and say 'listen, you can't blame me because 99% this is the situation.' Whereas today 99% is irrelevant.... The reason I'm going to the doctor is because I might be in that 1% and I don't have enough sense to figure out if I'm that 1%, so I need to go to a specialist. But now you're reducing the doctors to just ticking boxes. And as soon as you finish ticking the boxes, it'll automatically say what tests you need and what pills you need to be given. So where's the doctor's role in this whole exercise? We will employ more and more data entry operators and fewer doctors. Data entry operators are cheaper than doctors, anyway.

This analysis is based on a drop-down model of implementing EMR systems, which is the model used in countries like the US, and is likely to be adopted by the Indian government as well. The other possible way to implement an EMR is a free-text model where doctors can input free text into the system instead of selecting drop-down boxes. However, this model is much harder to implement for data sharing. This

[12] Name changed

is because with free text, each doctor would record data differently and there would be no consistency or mapping across different records, making it hard to analyse and share in a standardised and understandable format.

What this analysis also suggests is that at least a part of the decision-making is happening at the level of the system design. Yet, corporations who design these technological systems do not have a legal 'duty of care' towards patients, their data, and their bodies, thus challenging their right to equitable clinical care (Ada Lovelace Institute, 2020).

5.6 Accountability

Recognising the relationship between bodies and health data also means recognising that violations of health data can lead to violations of these bodies and their rights, as this section has analysed. In such a context, accountability must be determined keeping in mind potential harms to the body, not just to health data. This calls for a higher standard of accountability.

However, all the challenges to patients' rights described in this section are exacerbated due to the lack of accountability placed upon the state and private actors within the NDHM ecosystem. The National Health Data Management Policy mentions that data fiduciaries will be held accountable for "complying with measures which give effect to the privacy principles while processing any personal data by it or on its behalf" (Ministry of Health & Family Welfare, 2020, p. 14). However, the policy makes no mention of what these specific accountability measures would be, who would frame them, and how they would be enforced. This lack of accountability for the people whose health data is under consideration points to the NDHM's understanding of data as a legitimately tradeable resource that does not require the protection of associated bodies.

For example, if an electronic health record is inaccurate, and the resulting treatment given to a patient is wrong, who will be held responsible for such a scenario? Mr. Murali provides an example:

> You could have eye drops and also something that helps psychiatric issues - one letter is the difference between these two drugs. So if there is a mistake - who is responsible for the consequences of it? Doctor? DEO [data entry operator]? Pharmacist? In the offline world, you'd have a paper prescription, you'd have a pharmacist, and you can figure it out - if the handwriting is bad, the pharmacist is expected to ask what issue the patient has: 'do you have problems with eyes or psychiatric issues', and give the right medicine. But if it is all done online, and the pharmacist is dispensing it according to the system, then there is no human check on this to pick even such basic errors.

In another example, if a patient follows a doctors' advice but falls sick, the doctor can be held accountable. But if a person follows the data-driven recommendations of a fitness application and over-exercises or under-eats and consequently falls sick, it is not clear who is liable or what recourse they might have. Since healthcare is a state responsibility, the involvement of private actors in the ecosystem, such as technology companies who are not directly responsible for healthcare provision, should call for an even higher standard of accountability. But even the ethical codes and legal duties adhered to by health professionals are not applicable to them (Ada Lovelace Institute, 2020). Mr. Srikanth explains:

> The State... now acts as an enabler which will collect data or mandate data collection, and then auction out that data or share the data to market to build that infrastructure... I see this datafication, collection and sharing of data as a way to let more and more functions away from the government, which means that those functions now have far

> less accountability because that's not in the hands of the state anymore and private interests will prioritise return of investments over well-being of the population.

This means that there are very few safeguards and means of recourse for embodied harms to patients that may arise through the NDHM digital infrastructure.

For the rights discussed in this section, the NDHM's underlying framework of health data as a resource and source of capital is core to why these rights are undermined. In reality, as this section has shown, harms from data violations impact people's bodies and lives intimately. The violations of these rights come into picture only when we analytically put bodies back into the framework and question not only how data may be harmed, but how bodies may be harmed through their data, and how this harm undermines patients' right to healthcare. Failing to recognise the relationship between health data and bodies therefore risks the exclusion and exploitation of patients.

6. Way Forward: Recognising the Embodiment of Health Data to Empower Patients

A new, feminist framework is needed to reconceptualise how we fundamentally understand the nature of health data and the rights pertaining to it. This framework must be grounded in the notion of embodiment and bodily integrity so as to safeguard patient rights. Based on the challenges pertaining to the datafication of health identified in the earlier sections, I propose three levels of changes to implement such a framework: 1. regulatory and legal changes; 2. system-design and structural changes; and 3. ground-level changes.

6.1 Regulatory and Legal Changes

1. Enactment of the Personal Data Protection (PDP) Bill, 2019

The PDP Bill, which provides a data protection framework for the country, has not yet been enacted. The regulation of health data comes within the ambit of 'sensitive personal data' within this Bill. Since the PDP Bill is not yet enforced, the Information Technology (Reasonable security practices and procedures and sensitive personal data or information) Rules, 2011 (Ministry of Communications and Information Technology, 2011) currently govern health data. However, various limitations of these Rules have been pointed out (Acharya, 2013) indicating that they are insufficient to tackle the challenges highlighted in this paper. In such a scenario without strong regulation governing health data, private stakeholders within the healthcare ecosystem have already begun collecting such data and modelling insurance products around it. Examples of such initiatives, in particular collaborations between health insurance companies and technology companies producing fitness tracking devices, were discussed earlier. Even if the PDP Bill is enforced in the near future, it must be noted that the Bill has many drawbacks and legitimises the data-as-resource framework (Internet Democracy Project, 2019). What is most urgently needed is a strong law that takes into account the challenges with the datafication of health in the age of big data (as pointed out in this paper) since current legislation fails to do so.

2. Revision of what counts as health data

As analysed earlier, the disembodiment of health data has led to a broadening of what data can be used to infer a person's health. Legal frameworks pertaining to people's data rights thus need to be amended to account for the expansion of what counts as health data. For instance, proxy data and data pertaining to non-clinical predictors of health need to be regulated keeping in mind that they are often used to predict a person's health. Yet, such data would likely not fall under the category of sensitive personal data as defined in the PDP Bill. A consensus must thus be established to decide what kinds of data can accurately

be used to determine a person's health and well-being, either singularly or in conjunction with other data points, and what the risks are of doing so.

3. Imposition of a duty of care upon corporations

Tools that collect health data to make health-related decisions, such as wearable fitness trackers, must be regulated similar to how medical devices are regulated. This will ensure that companies which use this health data are regulated similar to other stakeholders that collect health data. This may be done by establishing a 'duty of care' for corporations towards individuals whose health data they collect and process. This would function as an accountability measure for the protection of people's health data.

4. Regulation of anonymised or de-identified health data

The Srikrishna Committee's Data Protection Report acknowledges the possibility of failure of anonymisation or de-identification to protect an individual's identity (Ministry of Electronics and Information Technology, 2018). In addition, this paper earlier analysed why such processes do not provide a sufficient protection mechanism since trends identified based on aggregated and anonymised datasets can be used to target individuals by predicting patterns of behaviour. Thus, access to anonymised or de-identified health data must be strictly controlled by pre-determining a specific set of stakeholders who would be permitted to access such data and the purposes for which they can do so. This should be done keeping in mind the provision under the draft Digital Information Security in Health Act (DISHA) (2018) that "digital health data, whether identifiable or anonymised, shall not be accessed, used or disclosed to any person for a commercial purpose and in no circumstances be accessed, used or disclosed to insurance companies, employers, human resource consultants and pharmaceutical companies, or any other entity as may be specified by the Central Government" (Ministry of Health & Family Welfare, 2018). It should be noted that the Health Data Management Policy, on the other hand, makes anonymised or de-identified data in an aggregated form available for sharing for various purposes including but not limited to statistical analysis, policy formulation, the development and promotion of diagnostic solutions and any other purposes as may be specified by the NDHM (Ministry of Health & Family Welfare, 2020, p. 18). Thus, this contradiction between DISHA and the Health Data Management Policy needs to also be resolved to ensure the effective regulation of anonymised or de-identified health data.

5. Cybersecurity protections

This paper earlier analysed the lack of privacy protocols available on the ground to protect the data collected while enrolling citizens for the Health ID. Clear guidelines need to be devised to ensure that this data is protected, including provisions such as protocols for generating strong, randomised passwords during Health ID registrations in cases where those with low digital literacy are unable to do so themselves. Similarly, there need to be clear guidelines created for how to respond to potential cases of data breach.

6. Consent for data collection

The process of seeking consent for data collection can be improved. In the status quo, an individual's family members can be enrolled to participate in the NDHM ecosystem by generating Health IDs for them without their consent. As noted earlier, this partly enables insurers to use the health data of family members to make statistical extrapolations and set premiums for the individual in accordance with data collected about their family. Consent must always be obtained directly from able adults, and exceptions can be made to this only in the case of minors or persons with mental disabilities who may not be able to consent themselves. Moreover, if the health data of other individuals is used to determine an individual's health (such as aggregated data of a neighbourhood, as seen earlier) and make decisions pertaining to them, then the individual must be informed and their consent must be sought before doing so to ensure that control remains in their hands. For this consent to be meaningful, feminist scholars point out that consent should not be a binary yes/no decision, but an ongoing negotiation wherein each party may say

no as well as provide input on the terms of agreement (Kovacs & Jain 2020). Making privacy policies for seeking consent available in local languages and non-written formats would also help in making them more accessible.

7. Accountability
Consensus needs to be established on questions of who would be held responsible in case of an error in a data-driven decision, malfunction of a digital health tool, or the use of inaccurate or inappropriate data in the health system; how the level of liability of the stakeholders would be determined; and what the process for recourse would be. Higher accountability needs to be placed upon stakeholders collecting and processing health data, given the understanding that this impacts people's lives and bodies, as seen earlier. Ethical codes and legal duties pertaining to data that are adhered to by health professionals must also be applicable to anyone else collecting and processing health data.

6.2 System-level and Structural Changes

1. Alternative identification for enrolment of Health ID
As discussed earlier, a Health ID can be created using a mobile phone number or Aadhaar number but through either means, Aadhaar details are likely to get directly or indirectly linked to the Health ID. Moreover, the registration system has been designed in a manner that incentivises the use of Aadhaar, as analysed previously. This is problematic because Aadhaar has been shown to be an exclusionary form of digital identification (Khera, 2017). Thus, other valid digital identity proofs such as a person's driving license or passport number, which are not necessarily linked to Aadhaar, must be introduced for the registration of a Health ID, and system-level incentives or 'nudges' that make enrolment through Aadhaar easier or preferred must be removed.

2. De-duplication of the Health ID
A recurring challenge that health workers noted was the duplication of Health IDs wherein a single individual could make multiple Health IDs without their knowledge, thus fragmenting their health data across various systems. De-duplication must be implemented at a system-design level, so as to provide a longitudinal view of a patient's health history in one place for their ease of access.

3. Electronic Medical Records (EMR)
Some issues pertaining to clinical care that may arise from the use of EMRs have been pointed out earlier. Thus, before rushing into digitising health records, there needs to be a clear understanding of its challenges and its potential impact upon the practices of clinical care and the embodied experiences of health professionals and patients in low-income contexts. This may be achieved by first piloting EMR systems in some parts of the country, studying their impact in detail and making necessary adjustments in practice before rolling them out nation-wide.

4. Transparency
Under the NDHM, many of the decisions made about an individual's health are likely to be opaque and proprietary. For instance, given the increasingly networked and decentralised nature of databases as well as the use of algorithmic analysis noted earlier, an individual will not know what data points have been used by say, an insurance company, to determine an insurance premium for them. This hinders their ability to challenge these decisions. Thus, transparency must be introduced in decision-making drawing on health data so that individuals understand what they are signing up for and can get recourse when needed.

6.3 Ground-Level Changes

1. Internet access and digital literacy

The NDHM's assumption of digital literacy and availability of Internet access needs to be grounded in the realities of the country.

India has one of the highest digital divides in the world for access to communication technologies and the Internet, with this divide more starkly observed in some demographics, such as low-income groups, rural India, and for women (GSMA, 2020). This is compounded by challenges with digital infrastructure, which is observed even at the initial stage of digitisation, in the registrations for a Health ID, as mentioned by Mr. Jadhav, a Multi-Purpose Worker in Chandigarh (translated from Hindi):

> 10% registrations fail because of network issues. They have provided tabs [tablets] but they have BSNL SIM. Here in [retracted] village, we don't have a BSNL tower. I have a JIO sim on my phone that works well. So I do registrations through my phone only for others, using my own data.

A related challenge is the lack of digital literacy: nearly ninety percent of the population are not digitally literate ('A look at India's', 2018). Ms. Namrata, a data entry operator at a civic dispensary in Chandigarh, questioned how the NDHM would then benefit people (translated from Hindi):

> For those [patients] who have bigger phones... I take pictures of the [Health ID] card on their mobile phones and give them, saying, 'Keep this safe. Whenever you go to the hospital, you have to take this with you.' How will they be able to keep this safe? They will forget about it.... So I don't think this has any benefits... Those with smaller phones get a message. But so many people are illiterate, they don't know how to read the message. Then I write on the visit card and give it... If they keep that carefully, then it's good, otherwise what can I do?

The baseline digital infrastructure and digital literacy in the country needs to be strengthened for the proposed benefits of the NDHM to reach communities who are already underserved in the delivery of health services. Such proposals have already been made in the context of digital identification such as Aadhaar being required for accessing social benefits (Khera, 2017).

2. People-centric awareness drives, not data collection drives

Almost none of the people I spoke to in Chandigarh—Medical Officers, health workers, and patients—knew what the Health ID is and what benefits they would get from it. Asymmetries in knowledge are being sustained by asymmetries of power. Medical officers claim that ANMs have received training about the NDHM, but ANMs and Anganwadi workers deny this. Ms. Narima, an ANM worker, said (translated from Hindi):

> They should tell clearly what this work is for... We don't know absolutely anything. All this we're all doing by guesswork. That everything is becoming digital, in the future they'll only ask for your Health IDs, then only your treatment will start... That's what we told the public also, that's all we also know about it.

As a result of health workers being uninformed, no information about the proposed benefits of the NDHM has trickled down to communities. Exploiting this information asymmetry, some health workers have been using disinformation to influence people to get enrolled. Dr. Amit, a medical intern posted in a civic dispensary, said:

> People ask...'what is the use of it [Health ID]? Why am I wasting my time, we are here to just take medicine, let us go', but then we have to tell them something... So we tell them... in the near future for vaccination... this will be required, this will be a mandatory thing to have... I don't think there's any link between the vaccination and Health ID. It's just I think created by the staff to motivate patients to make health IDs... Because there's no reward [or] incentive to the people, or there's nothing that we can provide them for giving their time, so just to make them think that it is worth your time.

People cannot meaningfully exercise any control over their data if they aren't even aware of their participation in the digital health ecosystem. The state should initiate people-centric awareness drives and provide training for all stakeholders involved in the ecosystem instead of focusing on collecting the data of people who are unaware of their participation.

3. Voluntary participation and meaningful choice

As noted earlier in the paper, the NDHM stipulates that participation of individuals in the NDHE will be on a voluntary basis, but in reality, their participation is being mandated. These practices need to be stopped, and the voluntary nature of participation must be respected. Moreover, as seen previously, people are participating in the NDHM ecosystem not because of their willingness to do so, but because of a lack of meaningful choice in accessing quality healthcare if they do not cooperate. For people to have a meaningful choice in their participation, if they wish to get access to healthcare through a specific provider of their choice without their participation in the NDHM, that option must be available to them. Alternatives also need to be made available for accessing health services from other providers that do not require their participation in the NDHM programs.

4. Non-exclusion

Though the NDHM states that it will follow the principle of non-exclusion, whereby no individual will be denied health services if they do not participate in the NDHM ecosystem, this paper analysed how that is not the case in reality. For instance, people are being denied access to medication and hospital visits if they do not register for a Health ID. Thus, strict guidelines should be devised regarding non-exclusion and these should be displayed in prominent locations in all health facilities so people are informed of their rights. Cases where there is evidence of a denial of health services must be strictly and independently investigated and institutions found violating this principle must be held liable.

In all of the changes recommended in this section, people's experiences with accessing equitable healthcare and their rights take centre stage. Key to this framework is the fundamental shift in viewing data as embodied, as opposed to an exploitable resource. This enables us to pay attention to and prioritise people's corporeal realities and needs as opposed to satisfying capitalist market-driven wants.

Conclusion

We are at a stage in time when the datafication of health, as we have always known it, is undergoing crucial shifts. The critical perspectives, medical legislation and codes of ethics that have been developed over the years to safeguard patient health data and associated healthcare rights need to be re-thought in the age of big data. This paper has analysed the various ways in which key changes are being observed today in the manner in which the datafication of health is taking place, calling for urgent attention. Responding to this need, using the policy framework provided by the NDHM as an analytical starting point, this paper has analysed how health data is viewed in a disembodied manner within healthcare policies, how such disembodiment undermines patient rights, and how acknowledging the relationship between health data and bodies can empower people to safeguard their rights in the era of big data.

By outlining the key stages in the life cycle of health data, this paper unpacked how its relationship with a patient's body undergoes changes at each stage and how this relationship is reflected in health data policies. Building upon this life cycle, the paper then argued that three major shifts are observed within healthcare when data becomes disembodied within policy frameworks: 1. what data can be collected to determine a person's health; 2. who can access this data about health; and 3. what they can know and do through such access to this data. As a result of these shifts, when health data is viewed as a disembodied resource, access to people's health data becomes a form of power, giving those with such control unparalleled power to influence the governance of people's bodies and lives.

Next, the paper adopted a rights-based approach to highlight the on-ground implications of the disembodiment of health data, arguing that it undermines patients' right to healthcare, in particular their rights regarding consent, choice, privacy, control, clinical care, and accountability. The violations of these rights come into picture only when we analytically put bodies back into policy frameworks and question not only how data may be harmed, but how bodies may be harmed through their data, and how this harm threatens patients' right to healthcare.

Lastly, the paper proposed a feminist framework grounded in the notion of embodiment and bodily integrity to reconceptualise how we fundamentally understand the nature of health data and the rights pertaining to it. Under this framework, I proposed three levels of changes to empower patients to affirm their rights: 1. regulatory and legal changes; 2. system-design and structural changes; and 3. ground-level changes.

The feminist critiques and framework offered in the paper are meant to serve as a starting point for policy, public, and academic discourse around how patient interests may be best served and safeguarded at a time when they are most threatened in the age of big data. Since the developments discussed in this paper are very recent, some proposed less than a year before the time of writing, this paper serves as an early-stage blueprint for the direction in which we should move. The implementation of these ideas in practice may not always be straightforward, and may require further thinking and labour. There are also likely to be further developments in this domain that will require continuous engagements to keep the wheel in motion. With these caveats in mind, at the heart of this work remains the conviction that the intimate relationship between our bodies and health data fundamentally challenges our understanding of the datafication of health and shapes our responses to it. Moving forward, through the contributions made in this paper, I hope this crucial embodied relationship makes its way into policy frameworks governing health data in the country.

References

Acharya, Bhairav. (2013). Comments on the Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011. *Centre for Internet and Society.* https://cis-india.org/internet-governance/blog/comments-on-the-it-reasonable-security-practices-and-procedures-and-sensitive-personal-data-or-information-rules-2011

Ada Lovelace Institute. (2020). The data will see you now: Datafication and the boundaries of health. *Ada Lovelace Institute.*

Aditya Birla Health Insurance Company Pvt. Ltd. (2016). *Activ Health - Policy Terms and Conditions.* https://www.irdai.gov.in/ADMINCMS/cms/Uploadedfiles/HP1617/ABHI_ActivHealth.pdf

Ajana, B. (2017). *Digital health and the biopolitics of the Quantified Self.* Digital Health, 3, 2055207616689509.

A look at India's deep digital literacy divide and why it needs to be bridged. (2018, September 24). *Financial Express.* https://www.financialexpress.com/education-2/a-look-at-indias-deep-digital-literacy-divide-and-why-it-needs-to-be-bridged/1323822

Chandrasekhar, Ramya. (2018, April 24). Here are the consequences of linking women's medical records to their Aadhaar. *The Indian Express.* https://indianexpress.com/article/gender/here-are-the-consequences-of-linking-womens-medical-records-to-their-aadhaar-5139360/

Christophersen M, Mørck P, Langhoff TO, Bjørn P. 2015. Unforeseen challenges: adopting wearable health data tracking devices to reduce health insurance costs in organizations. In *International Conference on Universal Access in Human-Computer Interaction, ed. M Antona, C Stephanidis*, pp. 288–99. Berlin: Springer Int.

Couldry, Nick, & Mejias, Ulises A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media,* 20(4), 336-349.

Debatin, Bernhard (2011). "Ethics, privacy, and self-restraint in social networking". *Privacy Online:* 47–60. doi:10.1007/978-3-642-21521-6_5. ISBN 978-3-642-21520-9.

Dogra, S. (2021, May 24). Took Covid vaccine using Aadhaar? Your National Health ID has been created without your permission. *India Today.* https://www.indiatoday.in/technology/features/story/took-covid-vaccine-using-aadhaar-your-national-health-id-has-been-created-without-your-permission-1806470-2021-05-24

Drugs Controller General. (2017). *Classification of medical devices and in vitro diagnostic medical devices under the provisions of the Medical Devices Rules, 2017 - Reg.* https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/medical-device/Classificationg1.pdf

Gawande, Atul (2018, November 5). Why Doctors Hate Their Computers. *The New Yorker.* https://www.newyorker.com/magazine/2018/11/12/why-doctors-hate-their-computers

Golembiewski, E., Allen, K. S., Blackmon, A. M., Hinrichs, R. J., & Vest, J. R. (2019). Combining nonclinical determinants of health and clinical data for research and evaluation: rapid review. *JMIR public health and surveillance, 5*(4), e12846.

Gordon, D. R. (1988). Clinical science and clinical expertise: changing boundaries between art and science in medicine. In *Biomedicine examined* (pp. 257-295). Springer, Dordrecht.

Greenwood, D, et al. (2014). 'The New Deal on Data: A Framework for Institutional Controls', In J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), Privacy, big data, and the Public Good: Frameworks for Engagement (pp. 192-210). Cambridge: Cambridge University Press. doi:10.1017/CBO9781107590205.012.

GSMA. (2020). *Connected Women: The Mobile Gender Gap Report 2020.* https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2020/05/GSMA-The-Mobile-Gender-Gap-Report-2020.pdf

Hayles, Katherine N. (1999). How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics. *Chicago: University of Chicago Press.*

Indian Medical Council (Professional Conduct, Etiquette and Ethics) Regulations (2002). India.

Interface Children & Family Services. (n.d.) *My Body Belongs to Me: A Child Personal Safety Education Tool Kit For Parents & Teachers.* https://www.icfs.org/wp-content/uploads/My-Body-Belongs-To-Me-Toolkit-English.pdf

Internet Democracy Project. (2019). Personal Data Protection Bill 2019: Submission to the Joint Parliamentary Committee.

IRDAI-NHA Joint Working Group. (2019). *Report of Network Hospital Management.* https://www.pmjay.gov.in/sites/default/files/2019-09/Report%20of%20Network%20Hospital%20Management_11-09-19.pdf

Jain, S. H., Powers, B. W., Hawkins, J. B., & Brownstein, J. S. (2015). The digital phenotype. *Nature biotechnology, 33(5)*, 462-463.

Kandhari, Ruhi. (2017). Why a backdoor push towards eHealth. *The Ken.* https://the-ken.com/story/why-backdoor-push-towards-ehealth/

Khera, Reetika. (2017). Impact of Aadhaar on Welfare Programmes. *Economic & Political Weekly. Vol. 52, Issue No. 50,* 16 Dec, 2017. https://www.epw.in/journal/2017/50/special-articles/impact-aadhaar-welfare-programmes.html

Kovacs, Anja, & Jain, Tripti. (2020) Informed Consent - Said Who? A Feminist Perspective on Principles of Consent in the Age of Embodied Data. Mumbai, *Data Governance Network.*

Kovacs, Anja, & Ranganathan, Nayantara. (2019). Data sovereignty, of whom? Limits and suitability of sovereignty frameworks for data in India. Mumbai, *Data Governance Network.*

Kurian, O. (2021, January 12). Data, Privacy, Pandemic: India just had the biggest medical records breach ever. *Observer Research Foundation.* https://www.orfonline.org/expert-speak/data-privacy-pandemic-india-just-had-the-biggest-medical-records-breach-ever/

Lupton, D., & Maslen, S. (2017). Telemedicine and the senses: a review. Sociology of health & illness, 39(8), 1557-1571.

Lupton, D. (2016). *The quantified self.* John Wiley & Sons.

Mandel, M. (2017). The Economic Impact of Data: Why Data Is Not Like Oil. *Progressive Policy Institute.*

Matthan, Rahul. (2017). *Beyond consent: A new paradigm for data protection* (Discussion Document 2017-03). The Takshashila Institution. https://takshashila.org.in/wp-content/uploads/2017/07/TDD-Beyond-Consent-Data-Protection-RM-2017-03.pdf

Ministry of Communications and Information Technology. (2011). *Information Technology (Reasonable security practices and procedures and sensitive personal data or information) Rules, 2011.* https://www.meity.gov.in/writereaddata/files/GSR313E_10511%281%29_0.pdf

Ministry of Electronics and Information Technology. (2018). *A Free and Fair Digital Economy: Protecting Privacy, Empowering Indians.*

Ministry of Finance. (2019). Economic Survey 2018-19. *Government of India.*

Ministry of Health & Family Welfare. (2010). *The Clinical Establishments (Registration and Regulation) Act*. http://www.clinicalestablishments.gov.in/cms/Home.aspx

Ministry of Health & Family Welfare. (2017). *National Health Policy.* https://www.nhp.gov.in/nhpfiles/national_health_policy_2017.pdf

Ministry of Health & Family Welfare. (2018). *Digital Information Security in Healthcare Act (Draft).* https://www.nhp.gov.in/NHPfiles/R_4179_1521627488625_0.pdf

Ministry of Health & Family Welfare. (2019). *National Digital Health Blueprint.* https://main.mohfw.gov.in/sites/default/files/Final%20NDHB%20report_0.pdf

Ministry of Health & Family Welfare. (2020a). *Health Data Management Policy.* https://ndhm.gov.in//assets/uploads/NDHM%20Health%20Data%20anagement%20Policy.pdf

Ministry of Health & Family Welfare. (2020b). *NDHM Sandbox Enabling Framework v1.0.* https://ndhm.gov.in/documents/sandbox_guidelines

'My Body Is Mine': tracing a feminist slogan. (2019, June 4). *Mama Cash.* https://www.mamacash.org/en/page/1166

National Health Authority. (2020). *National Digital Health Mission. Strategy Overview.* https://ndhm.gov.in/documents/ndhm_strategy_overview

National Health Authority. (2021). *NDHM Draft Implementation Strategy.* https://ndhm.gov.in/documents/draft_ndhm_implementation_strategy

Neff, G., & Nafus, D. (2016). *Self-tracking.* MIT Press.

New digital health ID will be used in Covid immunisation, says PM Modi. *The Indian Express.* https://indianexpress.com/article/india/new-digital-health-id-will-be-used-in-covid-immunisation-says-pm-modi-6795239/

Nissembaum, H. (2011). A Contextual Approach to Privacy Online, Daedalus 140(4): 32–48.

Nissenbaum H, Patterson H. (2016). Biosensing in context: health privacy in a connected world, pp. 79–100. In Nafus D. (2016). *Quantified: Biosensing Technologies in Everyday* Life. Cambridge, MA: MIT Press.

NITI Aayog. (2018). *National Health Stack. Strategy and Approach.* https://niti.gov.in/writereaddata/files/document_publication/NHS-Strategy-and-Approach-Document-for-consultation.pdf

O'Brien, J. & Marakas, G.M. (2008). *Management Information Systems* (pp. 185-189). New York, NY: McGraw-Hill Irwin.

O'neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy.* Crown.

OECD Data. (n.d.). Doctors: Total Per 1000 inhabitants. https://data.oecd.org/healthres/doctors.htm#indicator-chart

Patella-Rey, PJ. (2018). Beyond privacy: Bodily integrity as an alternative framework for understanding non-consensual pornography. *Information, Communication & Society,* 21(5), 786-791. DOI: 10.1080/1369118X.2018.1428653

Patients Like Me. (n.d.). http://news.patientslikeme.com/press-release/patientslikeme-launches-data-good-campaign-encourage-health-data-sharing-advance-medic

Porecha, M., Singh, P.V. (2021, March 19). eSanjeevani—The government-owned dark horse in India's telemedicine race. *The Ken.* https://the-ken.com/story/esanjeevani-the-government-owned-dark-horse-in-indias-telemedicine-race/

Prasad, Purendra (2007). Medicine, power and social legitimacy: A socio-historical appraisal of health systems in contemporary India. *Economic and Political Weekly:* 3491-3498.

Radhakrishnan, Radhika, & Sinha, Amber. (2020). Towards Algorithmic Transparency. *Centre for Internet and Society.* https://cis-india.org/internet-governance/algorithmic-transparency-pdf

Radhakrishnan, Radhika. (2020). *"I took Allah's name and stepped out":* Bodies, Data and Embodied Experiences of Surveillance and Control during COVID-19 in India. Mumbai, *Data Governance Network.*

Radhakrishnan, Radhika. (2021). Experiments with Social Good: Feminist Critiques of Artificial Intelligence in Healthcare in India. *Catalyst: Feminism, Theory, Technoscience* 7 (2): 1–29. http://www.catalystjournal.org | ISSN: 2380-3312

Rao, Menaka. (2017, November 17). Why Aadhaar is prompting HIV positive people to drop out of treatment programmes across India. *Scroll.in.* https://scroll.in/pulse/857656/across-india-hiv-positive-people-drop-out-of-treatment-programmes-as-centres-insist-on-aadhaar

Rocher, L., Hendrickx, J. M., & De Montjoye, Y. A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature communications, 10*(1), 1-9. https://doi.org/10.1038/s41467-019-10933-3

Ruckenstein, M., & Schüll, N. D. (2017). The datafication of health. Annual Review of Anthropology, 46, 261-278.

Sharma, D. (2021, March 28). Link PAN card to Aadhar in three days or pay a fine of ₹1,000. Here's why. *Hindustan Times.* https://www.hindustantimes.com/india-news/link-pan-card-to-aadhar-in-three-days-or-pay-a-fine-of-rs-1-000-here-s-why-101616913131068.html

Singh, S. & Porecha, M. (2020, September 2020). Behind the rush and hush of India's National Digital Health Mission. *The Ken.* https://the-ken.com/story/behind-the-rush-and-hush-of-indias-digital-health-mission/

Smith, G. J. D. and Vonthethoff, B. (2017) 'Health by numbers? Exploring the practice and experience of datafied health', *Health Sociology Review,* 26(1), pp. 6–21. doi: 10.1080/14461242.2016.1196600.

Solove, Daniel J. (2013). Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review,* 126(7), 1880-1903. https://harvardlawreview.org/wp-content/uploads/pdfs/vol126_solove.pdf

Stay Active And Get Rewarded With Optima Restore. (n.d.) *HDFC Ergo.* https://www.hdfcergo.com/blogs/health-insurance/stay-active-and-get-rewarded-with-optima-restore

Subramanian, A. (2018, February 14). Max Bupa launches 'GoActive' : A digitally enabled 'Everyday Use' Health Insurance Plan. *GOQII Blog.* https://goqii.com/blog/max-bupa-launches-goactive-a-digitallyenabled-everyday-use-health-insurance-plan/

Susser, Daniel, Roessler, Beat., & Nissenbaum, Helen (2019). Technology, autonomy, and manipulation. *Internet Policy Review,* 8(2). DOI: 10.14763/2019.2.1410

Taylor, L., Floridi, L., & Van der Sloot, B. (Eds.). (2016). *Group privacy: New challenges of data technologies* (Vol. 126). Springer.

Taylor, L. (2016). *Data subjects or data citizens? Addressing the global regulatory challenge of big data, Information, Freedom and Property.* Routledge. doi: 10.4324/9781315618265-13.

Tisné, M.. (2020). The data delusion: Protecting individual data isn't enough when the harm is collective. *Luminate, July.* https://cyber.fsi.stanford.edu/publication/data-delusion

Tisne, Martin. (2018). It's time for a Bill of Data Rights. *MIT Technology Review.* https://www.technologyreview.com/2018/12/14/138615/its-time-for-a-bill-of-data-rights/

Tomar, Shruti. (2017, April 3). Linking benefits for AIDS patients to Aadhaar triggers privacy concerns. *The Hindustan Times.* https://www.hindustantimes.com/bhopal/linking-benefits-for-aids-patients-toaadhaar-triggers-privacy-concerns/story-iR6HB8RmqPDaNwkX2Oj5EJ.html

United Nations Global Pulse (2009). Data Philanthropy. http://www.unglobal-pulse.org/blog/dataphilanthropy-public-private-sector-data-sharing-global-resilience

Van der Ploeg, I. (2012). The body as data in the age of information. *Kirstie Ball, Kevin & David Lyon (Eds.), Routledge Handbook of Surveillance Studies,* 176-183.

Warzel, C. (2019, August 13). All Your Data Is Health Data And Big Tech has it all. *The New York Times.* https://www.nytimes.com/2019/08/13/opinion/health-data.html

World Health Organisation. (2021). Global Health Workforce statistics database. https://www.who.int/data/gho/data/themes/topics/health-workforce

Yellaiah, J. and Ramakrishna, G. (2012), “Socio economic determinants of health insurance in India: the case of Hyderabad city”, *International Journal of Development and Sustainability,* Vol. 1 No. 2, pp. 111–119

Zuboff, Shoshana. (2019). *The Age of Surveillance Capitalism: the Fight for Human Future at the New Frontier of Power.* London: Profile Books.

Acknowledgements

All research–like all human labour–is social and collective. This research study has been shaped by the participation of many individuals, though all flaws in this paper remain my own. I am immensely grateful to all research participants for sharing their valuable inputs and experiences with me and I hope to have done justice to these through this work. I would like to thank the Ethics Review Committee of Anusandhan Trust for thoroughly and rigorously reviewing this research at all stages. I am thankful to Dr. Sunita Sheel Bandewar and Mr. Murali Neelakantan for their helpful inputs as discussants for the proposal and draft paper presentations of this research during two roundtables organised by the Data Governance Network respectively. I'd like to also thank Chahat Rana, journalist at The Caravan, for leads during fieldwork of this study, and members of the NHM Employees Union (who prefer to remain anonymous) for facilitating my access to fieldwork sites that were initially inaccessible to me. I am grateful to Dr. Anja Kovacs for guiding and supervising this research.

About the Authors

This research was carried out while the author was a Researcher at the Internet Democracy Project. Radhika Radhakrishnan is an interdisciplinary researcher from India currently at the World Wide Web Foundation as a Gender Research Manager. She researches the intersection of gender justice and digital technologies using feminist, qualitative methodologies. She has a Master's degree in Women's Studies and a Bachelor's degree in Computer Science Engineering. She has previously carried out research with the Internet Democracy Project and the Centre for Internet and Society, and has consulted with the Internet Governance Forum of the United Nations and the Women's Rights Programme of the Association for Progressive Communications. To make her research more accessible to a wider audience, she regularly writes for media publications, gives public talks, and hosts podcasts on digital rights from a gender lens.

Personal website: https://radhika-radhakrishnan.com/

datagovernance.org dgn@idfcinstitute.org

@datagovnetwork /datagovnetwork /datagovnetwork