When data drive health: an archaeology of medical records technology

Colin Koopman[1,2] · Patrick Jones[1] · Valérie Simon[1] · Paul Showler[1] ·
Mary McLevey[1] · for the Critical Genealogies Collaboratory

Abstract Medicine is often thought of as a science of the body, but it is also a science of data. In some contexts, it can even be asserted that data drive health. This article focuses on a key piece of data technology central to contemporary practices of medicine: the medical record. By situating the medical record in the perspective of its history, we inquire into how the kinds of data that are kept at sites of clinical encounter often depend on informational requirements that originate well outside of the clinic, in particular in health insurance records systems. Although this dependency of today's electronic medical records on billing requirements is widely lamented by clinical providers, its history remains little studied. Following the archaeology of medicine developed by Michel Foucault in *The Birth of the Clinic* and expanding his methodology in light of more recent contributions to the field of media archaeology, this article excavates some of the underexplored technological conditions that help constitute today's electronic medical record. If in some contexts, it is true that data drive health, then an archaeology of medical records helps reveal how health insurance records often impact clinical care and, by extension, health and disease.

Keywords Medical records · Information technology · Data systems · Formats · Philosophical archaeology · Media archaeology

✉ Colin Koopman
koopman@uoregon.edu

1 University of Oregon, Eugene, USA

2 Department of Philosophy, 1295 UO University of Oregon, Eugene, OR 97403, USA

Introduction

The very idea that health could be data-driven kindles skepticism in light of the common sense that the secrets of health and disease are buried deep within the tissues of our bodies. Yet if any questions about the matter remained, the SARS-CoV-2 pandemic has laid to waste lingering suspicions concerning how crucial data technology is for health. Consider the very definition of the pandemic, which has depended from the very beginning on probabilistic forecasting and other data-intensive epidemiological techniques that calculate not only anticipated rates of transmission but also quantitative measures of social interaction and estimated supplies of critical medical infrastructure. Dashboard aggregations of enormous data-collection enterprises summarize for their users, both public health experts and concerned citizens, the specific shapes of the pandemic.

In a dispatch for *The New Yorker* written from the front lines of the pandemic in the spring of 2020, the physician and author Sidhartha Mukherjee highlights with surgical accuracy the multifarious causes of the extremely fragile condition of the U.S. health system, among which he includes the frailty of the country's medical records and health surveillance networks. Mukherjee notes, with a casualness indicative of a taken-for-granted truth, that "Clinical medicine is, among other things, an information system" (Mukherjee 2020, p. 29). This echoes an observation by the French historian of medicine Georges Canguilhem, who suggested three decades ago that medical practice had recently changed such that "consultation now consists of a combination between a computer search through semiological and etiological databanks and a formulation of a probability diagnostic supported by the evaluation of statistical information" (Canguilhem 2012 [1989], p. 38).

To make vivid his point, Mukherjee recounts a poignant episode that began with a burst of anecdotal reports about blood clots associated with Covid-19 patients. Hoping to research the prevalence of clotting, he contacted a fellow physician who, despite recognizing the obvious benefits of those data, had to inform him that, "It's not easy at all to search the E.M.R. [electronic medical record] for any of this information" (Mukherjee 2020, p. 29). Mukherjee quotes another doctor, from Mass. General Hospital in Boston, lamenting on social media, "Why are nearly all notes in Epic [a major software platform for EMRs in the U.S.] … basically *useless* to understand what's happening to patient during hospital course?" (Mukherjee 2020, p. 29). To this another doctor replied, "Because notes are used to bill, determine level of service, and document it rather than their intended purpose, which was to convey our observations, assessments, and plan" (2020, p. 29). Another doctor, whom Mukherjee does not quote though the comment is archived on the relevant Twitter thread, went even further in replying that, "EMR has become mostly fluff with no substance, it exists for billing and coding—not for patient care!" (Jeyaraj 2020).

Mukherjee's observations neatly frame the three central domains of medicine in the United States today: clinical care, public health, and health insurance. Reflecting on this triad, his article poses a question that no patient, doctor, or epidemiologist will find unfamiliar: How is it that one of the world's richest

democracies would find its health systems driven, even if only at times, by the medical informatics requirements of health insurance? Those who pose this question often perceive the entanglement of EMRs in insurance to be a relatively recent occurrence that must have originated in the supposedly recent events of digitization and computerization. But as with most things digital and computable, the central technical structures underlying EMRs and their connections to insurance protocols actually had their beginnings nearly a century ago.

Across all of its domains, very specific and tightly specifying data infrastructures are at the heart of the contemporary health system. What are the historical conditions that have made possible such a reliance on informational infrastructure in the practice of both public health policy and personal health provision? And what are the specific shapes, requirements, and formats for medicine specified by these information systems? We here show that the enormous information infrastructures on which health so heavily depends today are built atop crucial-yet-overlooked technologies of medical informatics that were first put into widespread circulation in the first decades of the twentieth century. Some of these first-generation information architectures were developed by a nascent health insurance industry and from there loaded into clinical practice.

Taking a general perspective, these recordkeeping technologies include, among other elements, specific formats for structuring and standardizing data, particular protocols for data storage and transmission, and precision procedures for auditing the accuracy of data systems. For example, consider that every public health database relies for its data points on a highly curated specification of how data are to be formatted such that those data can be implemented for any purpose at all. Making use of the idea of the data structure (or what we refer to as the "format") as an analytical lens, we here seek to explore how health data began to be subjected to wide-scale formatting decades ago. Returning to earlier moments of emergent formatting helps make plain some of the technical and conceptual choices that over time have become so deeply entrenched in data formats that it can now be hard to acknowledge that these formats embody decisions at all. Data are often presented to researchers, and those who rely upon researchers, as facts that are simply given. But no data just are. All data are curated. All data are first shaped by formats.

Data formats bear moral and political choices that are often neglected by the time they are baked into operable technology. Formats, and the data they make possible, have implications that are difficult to see when viewed through the impressive computational edifices of contemporary data analytics techniques. By excavating earlier histories of data technologies, we can observe the basic (indeed almost rudimentary by comparison) decisions that are often soldered into the circuits of later complex computational apparatus that depend upon them. We can, for example, observe how particular technical formats emerged in early health insurance records systems in such a way that those formats could come to influence clinical medical records in later decades. We can observe, in other words, how it has come to pass that so many clinical care providers find their work lives neatly summarized by the title of another *New Yorker* essay by another doctor, Atul Gawande's "Why Doctors Hate Their Computers" (Gawande 2018).

Our analysis proceeds as follows. We begin in the first section by situating our inquiry with respect to significant precedents for the methodology we employ here. Most prominent for us are the contributions of Michel Foucault to the philosophy of medicine. We aim to expand some suggestive leads offered, but never really developed, in Foucault's work. In the second, and main, section of the paper, we turn to chronologically locating the takeoff of robust health informatics in public health (which occurred in the U.S. in the 1920s), then health insurance (achieved in the 1940s) and clinical medicine (occurring later in the 1960s). We use this chronology to consider how formalized health insurance forms were a predecessor technology to the formalization of clinical medical records which the latter had to confront in their development. We conclude by discussing the implications of our study for today's polarizing ethical and political debates about the relationship between insurance and care in medicine.

How should we do the archaeology of medical data?

In 1963, Georges Canguilhem's former student Michel Foucault published *The Birth of the Clinic: An Archaeology of Medical Perception* (Foucault (1994 [1963]). Of all Foucault's books, the central concerns of *The Birth of the Clinic* are surely the most timely for our contemporary pandemic-stricken moment. Foucault's later work in the mid-1970s exhibited a resurgence of interest in health and medicine, including his Rio lectures on "social medicine" (Foucault 2002 [1974], 2004 [1974], 2007 [1974]) and his widely discussed theorization of biopolitics in *The History of Sexuality, Volume 1* (1990 [1976]).

One of the most interesting, but also under-discussed, themes found throughout Foucault's forays into the archaeology and genealogy of medicine is his attention to the role of information technology in health. Despite the decade separating them, this theme resonates across Foucault's mid-1970s writings on medicine and his earlier 1960s work on the clinic. In both periods, Foucault called attention to the crucial role of information, documentation, and recordkeeping in the constitution of modern medical practice.

In *The Birth of the Clinic*, Foucault traced the need for an organized and far-reaching system of medical information gathering and recording to the 18th-century "medicine of epidemics" (Foucault 1994 [1963], p. 25). During this period, epidemics were understood not so much in terms of contagion, but as the result of a complex interplay of environmental and social factors which needed to be investigated in their "historical individuality" (Foucault 1994 [1963], p. 25). An adequate medical knowledge of epidemics would, therefore, require "a complex method of observation" (Foucault 1994 [1963], p. 25) facilitated by a body of health inspectors who would "collect information about the various domains related to medicine, as well as about physics, chemistry, natural history, topography, and astronomy, would prescribe the measures to be taken, and would supervise the work of the doctor" (Foucault 1994 [1963], p. 26).

Just over ten years later, in the third of his 1974 Rio lectures, titled "The Incorporation of the Hospital into Modern Technology," Foucault again sounded the theme

of information, and this time more loudly (Foucault 2007 [1974]). He noted that, "a collection of documents is formed in the heart of the hospital, and thus is constituted not only a place of cure but also a place of record and the acquisition of knowledge…. It is no longer what was written and printed, but what every day was recorded in living, active and current actions which the hospital represents" (Foucault 2007 [1974], p. 151). This amounts to, said Foucault, "a permanent and as far as possible complete records system, which registers whatever occurs" (2007 [1994], p. 150). Foucault here marked the emergence of a new role for medical records in the production of medical knowledge. This suggests the possibility that medical records form part of what Foucault would come to describe, in an essay on the politics of health two years later, as "a common global strategy" that establishes continuity across different forms of medicine, which for our inquiry below includes most notably clinical practice, public health, and health insurance (Foucault 1984 [1976], p. 274).

All of this attests to how Foucault's writings on medicine explicitly feature a concern with recordkeeping, documentation, and registration. Yet it must also be admitted that, despite having noted this theme, Foucault left the crucial role of information in medicine relatively underexplored. We here take Foucault's tantalizingly obscure nods to statistical and individual health data as an invitation to build out those dimensions of his inquiries concerned with the role of informational records in the history of modern medicine. For Foucault's glances toward beta-version health data systems raise an important question for us today in a moment when the scope and stakes of public health measures like the response to the Covid-19 epidemic are increasingly defined in terms of big data, data science, and data capital. How does a health information apparatus help shape health care? How do informational accoutrement contribute not only to the formation of the many domains of medicine (sprawling from the clinic to the insurer), but also how do they contribute to medical practice in a way that defines or, as we put it here, "formats" the multiple subjects of medical care (from patient to insured)?

In posing such questions in light of Foucault's contributions to the history of medicine, we are by no means entering uncharted territory.[1] But where prior scholarship on medical data in a Foucauldian light tends to borrow his concepts to make a point about the role informational records play in the systems that organize medical practice, we rather aim to retrieve and extend the archaeological method at work in *The Birth of the Clinic*. Stated succinctly, we mobilize Foucault's methods rather than relying on his concepts.[2]

Our understanding of Foucault's archaeology is influenced by Ian Hacking, who describes his own archaeological efforts in what he calls "historical ontology" as a project of "analyzing *our* concepts, but not in the timeless way… in the finest

---

[1] See Hess and Ledebur (2011) and Hess and Mendelsohn (2010) on early iterations of paper forms in medicine; see Berg and Bowker (1997) and Bowker and Star (2000) on classificatory artifacts in medicine.

[2] See Koopman and Matza (2013) on concepts versus methods as two different instruments for mobilizing the work of a philosophical predecessor.

tradition of philosophical analysis," since it is crucial to observe, says Hacking, that "concepts have their being in historical sites" (Hacking 2002, p. 25).[3] In line with Hacking's observation, we observe in Foucault's archaeology not the analysis of "the mere surface of contact between the knowing subject and the known object" but an inquiry into "the more general arrangement of knowledge" (Foucault 1994 [1963], p. 137; cf. Foucault 1989 [1969]) that functions as epistemological "depth" (Hacking 2002, p. 77).

In pursing an archaeological analysis in the style of Foucault and Hacking, we find a need to amplify their archaeology by way of methodological interventions in the emerging approach of "media archaeology" (Kittler 1999 [1986]; Huhtamo and Parikka 2011). If Foucault's archaeology examines the deeper conditions of knowing (*savoir*) that silently structure the surface statements of scientific knowledge (*connaissance*), then media archaeology posits among those depth structuring conditions their technological apparatus. For instance, media archaeology provides a method for excavating the structuring role played by information systems (contrasting a book ledger with a file cabinet with an electronic database) or communications channels (contrasting radio with internet). If most users of Foucault, such as Hacking, have understood how central is the domain of the conceptual to archaeology, then media archaeology helps us further understand how central is the domain of the technical to an archaeology of the conceptual.

Two precedents help illuminate the media-archaeological approach we adopt. One is Lisa Gitelman's discussion of "how blanks, and job-printed documents more generally, may have worked to structure knowledge and instantiate culture" (Gitelman 2014, p. 24). In short: a printed blank can show us what can even count as allowable information, and hence possible, or at least storable, knowledge. Another is Cornelia Vismann's examination of the history of the power of files in law as they migrate from antiquity to modernity along a trajectory that runs from parchment documents (which authenticate information) to early-modern book registries (which store information) to modern filing systems and "file plans" that "give birth to a transcendental order of files prior to all content," that is, that constitute information itself (Vismann 2008 [2000], p. 142). In short: a file plan organizes data before it can ever become a dossier, a document, or a record in a database. Gitelman and Vismann both attend to the structuring role of mediating information technology. Combining this kind of attention with prior work on "formats" undertaken by one of us (Koopman 2019, 2021), our inquiry here proceeds through a primary focus on how data formats mediate medicine.[4]

How does one investigate the way that a format structures information so that it can be stable, storable, sortable? Here is where media archaeology is indeed deeply archaeological. Records, be they in the form of files (Vismann), or blanks (Gitelman), or formats (our focus), cannot simply be examined for what they say, that is,

---

[3] See Hacking (2002, pp. 2–5) on the importance of Foucault to his own conception of historical ontology.

[4] See also Leonelli and Tempini (2018) for another resonant approach, focused in this case on environmental and health data.

for their meaning, but must also be examined for their technical functionalities, that is, for their actual effects. The focus for archaeology is always pragmatics rather than semantics. The archaeological question of 'what do documents do?' replaces a hermeneutical pondering over 'what do these texts mean?'. The pragmatics of records, when archaeologically excavated, can be found in their formats, file plans, and protocols for inscription, storage, processing, and reproduction. The media archaeology we develop therefore aims to explicate how the formats internal to health records have contributed to the formation of medical practice across overlapping, and sometimes competing, domains.

Formalizations of health records in the twentieth century

We of course do not propose to exhaustively survey all aspects of medical record-keeping that over the course of the past century have imprinted our very bodies with the stamps of informational infrastructure. Such recordkeeping imprints are found across clinical medicine, medical research, pharmaceutical engineering, health insurance, epidemiology, hygiene research, nutrition science, and more. A comprehensive survey of all these fields would clearly be well beyond the scope of a single essay, or even a single book. Rather, we selectively excavate one trajectory along which data came to drive health by attending to relations among two specific domains of medicine: health insurance and clinical care. To frame our chronology, however, we begin with a third domain closely related to these other two: public health.

A central category in our analyses of these domains is that of the *degree of formalization* of a records system. Formalization involves, primarily, formatting—structuring data. Such structuring is always simultaneously technical and conceptual. Through formats are data born. Formatting occurs for all data from legal names to email addresses to the gender you select on a drop-down list. As we use the term here, information systems are increasingly formalized as their formats are given greater technical and conceptual specificity.

Formalization is easy to achieve locally (anyone can specify anything they wish for themselves) but exceedingly complex to implement at scale. Insurance, by definition, is a big-scale project. Formalization in insurance systems is always significant for wide swaths of patients and providers. We might locate that significance, for example, in the way that formalized insurance documents require that providers in clinical settings collect (or even constitute) certain kinds of information on behalf of their patients. A historical archaeology of this clinic-insurer relationship, thus, promises one way of responding to the question with which we began: what are the technological conditions that have made possible today's electronic medical records systems whose formats are shaped at least in part by insurance and billing?

We here locate the emergence of formalized health informatics in, first, public health in the 1920s, and then, second, in health insurance in the 1930s and 1940s, finally turning to the formalization of medical records in clinical practice in the 1950s and 1960s. Attending to this chronology, we here show how health insurance data systems helped constitute informatic conditions for clinical medical records

systems. In medicine as elsewhere, priority in technological development often produces established pathways that function as grooves into which subsequent technologies can be efficiently fit.

The broader pattern we discern is well captured in recent work by media theorist (and part-ways media archaeologist) Tung-Hui Hu's concept of the "graft" (2015, p. 1). Hu's work convincingly details how new networks are often grafted onto pre-existing ones. His clearest example is that of today's fiber-optic networks which were in some cases physically laid atop older networks of railroad tracks and telegram lines. He notes along the way that the telecom company Sprint is a spandrel of the Southern Pacific Railroad for which it is named ("SPRINT" being an acronym for Southern Pacific Railroad Internal Networking Telephony). Translating Hu's concept to our case enables us to discern how health insurance records were formalized decades prior to clinical records in such a way that the former created conditions for the latter. In short, later-evolved clinical records systems would often need to be grafted onto the shapes first cut by insurance data systems. First-generation health insurance forms instituted information-gathering requirements for care providers. Insurance records are a part of the story of how it came to pass that clinical providers today would find themselves exasperated at the insurance-driven documentation requirements of their electronic medical records systems.

Public health records, 1900s–1930s

We begin with public health even though its data systems are not central to our story. Our account in this section is preliminary. It helps set the stage by returning to a period in which medical records were not yet subject to a high degree of formalization, except perhaps in very restricted local settings. Public health offers clear sight of a medical domain in which we can witness the entry onto the scene of a robust data reporting apparatus. This apparatus was sporadic as late as 1918 but soon came into fruition in the early 1920s.

Consider the data apparatus constituting the core of the initial response in spring 2020 to the ongoing coronavirus pandemic. All those data are central to the agile social-distancing mandates that helped to mitigate the first wave of impacts of the virus. A century earlier, the response to the 1918–20 global influenza epidemic suffered from a notable lack of data. Indeed, public-health officials at the time did not even have at their disposal the kind of health surveillance technology needed to capture the most basic of epidemic data. This would soon change.

Historians have marked the inter-war period as the birth moment of formalization in epidemiology. Alfredo Morabia refers to these years, specifically the 1920s, as the beginnings of "early epidemiology" in both the U.S. and U.K. contexts (Morabia 2004, p. 112). Charles-Edward Winslow's classic *The Conquest of Epidemic Disease* argues that modern quantitative studies of public health did not even come into clear view until 1910 (Winslow 1980 [1943], p. 366). And what first came into view in the 1900s and 1910s would not be implemented on a wide scale until the 1920s and 1930s, a chronology made painfully obvious by the fallout of a new strain of influenza that began circulating in 1918. Historian Alfred Crosby observed in his

1989 study of the 1918 pandemic that, "Many boards of health didn't make influenza a reportable disease" while most "physicians and nurses had much more compelling demands to answer than the call to be accurate clerks" (2003 [1989], p. 204; Spinney 2017). That was in the U.S.; globally, the data on the pandemic were even worse. Crosby concluded in his book that the task of estimating the global death toll from the pandemic is such that "[a]t the present state of research no such thing as even an educated guess on the matter exists" (2003 [1989], p. 207; Spinney 2017).

The disease surveillance systems so markedly absent in 1918 would be developed in the next decade. By 1925 every state in the U.S. was contributing data to national morbidity reports and in 1935 the first national health survey was conducted (Spinney 2017, p. 241). By the middle of the 1930s, most states were like Oregon, where "[l]ocal health officers regularly reported public health data to the Oregon State Board of Health" including case counts of communicable diseases (Oregon Health and Science University Historical Collections and Archives 2020). These early improvements in recordkeeping apparatus contributed to a redefinition of the very field of public health by enabling the use of newly developed statistical techniques in epidemiology. Nearly thirty years later, modern health data protocols were solidified when, in 1960, the National Health Survey and the National Office of Vital Statistics merged to form the National Center for Health Statistics. Epidemiological science today continues to retain at its core the kinds of data, statistics, and records apparatus first set up in the 1920s (Hardy and Magnello 2004; Broadbent 2013).

Although our focus here is on conceptual and technical relations between health insurance and clinical care, the foregoing consideration of the role of data in public health establishes a crucial point for our argument. For it shows that by the 1920s, and evidently not much before then, it was both conceptually and technically possible to implement highly formalized health records systems on a mass scale. This is surely of independent interest for the history of epidemiological public health. But for our quite different purposes here it shows how such formalized medical records technology was available to, but not widely taken up in, settings for clinical care until decades later, and indeed long after it had been implemented by insurance firms. We show this in the following two sections.

Health Insurance Records, 1930s–1940s

In turning to health insurance records we begin with some preliminary orientation about the health insurance industry in the U.S. According to historian John Murray, it was in the 1930s that private insurance companies began to use the kind of actuarial technology that made it possible for them to compete with other approaches to health insurance that had been predominant in earlier decades, chief among them industrial sickness funds (Murray 2007, p. 231). Industrial sickness funds, or establishment funds, were forms of employment-based insurance typically administered by companies, unions, or fraternal organizations, which aimed to insure workers for lost wages due to illness or injury. Due to their proximity to those whom they insured, sickness funds enjoyed certain advantages over early forms of commercial insurance. In addition to being able to more closely monitor claimants and to

"inspire feelings of personal investment in the funds among their members" (Murray 2007, p. 219), sickness funds typically required lower administrative costs associated with fee collection and distribution (Murray 2007, p. 7). But as private insurers came to use increasingly sophisticated actuarial techniques, they were able to offer more competitive rates that enabled them to eventually eclipse sickness funds (Murray 2007, p. 220).

Consider next a few landmarks in the development of the kind of large-scale private health insurance in the U.S. that displaced industrial sickness funds. In 1936 the American Hospital Association established the Commission on Hospital Services, which would later become the Blue Cross Association, an organization that by 1937 had 894,000 members and then soon after in 1943 had a staggering almost 12 million members (Jost 2014, p. 391). The chronology of Blue Cross is typical of an industry that on the whole had, according to an early observer of "hospitalization plans" writing in 1940, experienced "phenomenal growth" in the latter years of the 1930s (Faulkner 1940, p. 283). The expansion of commercial insurers was relentless in the next decades: "from 1940 to 1960 the number of Americans with relatively simple medical coverage for physician visits rose from 3 million to nearly 88 million. Similarly, the number of those with hospital insurance increased from 12 to 132 million" (Murray 2007, p. 233). By the time President Lyndon Johnson signed the 1965 bill establishing what would later become Medicare and Medicaid, 138.7 million Americans were covered by private hospital insurance (Jost 2014, p. 392).

The common health insurance forms, records, and files we are accustomed to today were first formalized in the midst of this period of remarkable growth. Before we turn to charting the specificities of some of this health insurance information, a preliminary distinction is needed. There are (at least) two kinds of very distinct formalized data in modern health insurance information systems. First, there are population-level statistics that feed into actuarial estimates used by insurance companies to calculate costs, premiums, and other incomes. Second, there are individual-level records about insured customers, some of which are standard blank forms supplied to (and filled out by) the insured, others of which are standard blanks supplied to billing departments and physicians (or their assistants) at sites where health care is provided. It is the latter that is our primary interest here.

As to the first of these data types, statistical-actuarial information, health insurance on the model of large-scale private providers that we remain accustomed to today emerged in part because of the advances in actuarial techniques noted above. This occurred in the 1920s and 1930s, at the very time that statistical techniques were beginning to redefine possibilities for scientific field after scientific field, including epidemiology (Porter 1986). In deploying these new techniques health insurance firms actually lagged slightly behind other insurance enterprises, most notably life insurance (Bouk 2015). On the whole the statistical dimensions of formalization in health insurance is a story that is already well known—it can be considered as yet one more chapter in the history of biopolitical statistics.[5]

---

[5] See Ewald (1991) and Hacking (1991) on the history of biopolitical statistics.

In contrast to copious scholarship on statistics and its role in insurance enterprises, the second type of formalized data in modern health insurance, individualized records and forms, remains largely underexplored. It is also more pertinent to how health insurance information directly interfaces with clinical medical practice. Just as the statistical dimensions of health insurance data share technical procedures with the statistically driven field of epidemiology, the individualized documents of health insurance are procedurally interfaced with clinical practice records systems. Patient insurance records interface directly with the billing dimensions of clinical health provision, and thus, exercise an influence on clinical practice through the back door of patients records systems. This interfacing happens at two levels. The first and more obvious is at the level of the contents of the records systems (e.g., the patient's name and their billing allowances for certain procedures). The second is at a more technical or formal level akin to the ways that statistical methods in public health and health insurance intersect. At this technical level, we can pose questions not just about the contents of this patient's records from hospital to insurer, but more crucially about the formats (that is, the very data fields and allowable data variables) of the records systems employed by hospitals and insurers. We can ask, to return precisely to the question with which we opened, how these information systems are designed such that the types of information allowed into, and required by, a patient record database are determined at least in part by the types of information demanded by insurer information systems.

From their very inception, commercial health insurers made extensive use of documentary apparatus to keep detailed records on their customers. Of course, insurance of any kind is an information technology enterprise as well as being an economic (and some argue, social) enterprise. So what kinds of records were kept by early health insurers?

One of the most obvious standard documents employed by health insurers were records sourced from questionnaires used in applications for insurance by would-be customers. In the 1930s and 1940s when health insurance began to boom, these documents began to boom too. They were among the many kinds of patient-facing health insurance forms that we are quite accustomed to today.

In addition to these kinds of patient-facing documents, there were other types of formal records that insurers regarded as essential to their business but which few patients would ever see. Of particular interest are the physician-facing forms that any medical facility caring for an insured patient would need to fill out. As these forms eventually came to accompany more and more patients, hospital personnel from administrators to doctors and nurses would adjust clinical care protocols (as well as informal office practices) to keep up with this information technology that was increasingly an essential part of any clinical interaction.

To see how these self-adjustments came to seem necessary, consider the long history within which insurer records systems have been delivered to the site of clinical care with the clear expectation that their formats and protocols will be abided. These expectations were from early on backed not just by individual insurance companies, but also by the heft of insurance industry professional organizations. In an early study of health insurance in 1942, C.A. Kulp observed that industry professionalization beginning in the 1930s was characterized by

more objective and uniform standards for rate-making as well as more uniform policy contracts (Kulp 1942, p. 376). These tendencies toward formalization were implemented through standard forms first developed by professional organizations and then promulgated through such of their professionalizing publications as textbooks.

Two professional organizations stand out as the most important: the Health and Accident Underwriters Conference ("the Conference") and the Bureau of Personal Accident and Health Underwriters ("the Bureau"). The Advisory Forms prepared by the Bureau "represent the closest approach to uniformity" at the time (Faulkner 1940, p. 293). Their Revised Program released in 1932 included six different policy schedules for different levels of insurance (Kulp 1942, p. 379). The members of the Conference were "less homogenous" than those in the Bureau and, thus, offered a "wider range of forms" (Faulkner 1940, p. 294; Kulp 1942). But here too the entire point of the Conference's 1931 effort to create standard advisory policy forms was to "encourage greater uniformity in contract phraseology" (Faulkner 1960, p. 553).

With these general trends in view, consider just one specific landmark of the formalization of health insurer records. Edwin Faulkner's 1940 *Accident-and-Health Insurance* (Faulkner 1940) was published in the *McGraw-Hill Insurance Series*, a collection targeted to a newly professionalizing (and rapidly expanding) class of insurance workers, as well as those looking to train and educate them. Faulkner's book was, in the words of the series editor, "an educational textbook"—it intended to instruct on how to conduct the new commerce in health insurance (Gordon 1940, p. vii). The textbook as a genre offers an illuminating window. A professionalizing textbook (and especially one from the mid-twentieth century) does not seek controversy and even shies away from argument. Its rhetoric is that of the plain fact. It asserts only what is eminently assertible. What is stated in a textbook, we can be sure, is widely taken as fact (whether, in fact, it is fact).

Faulkner's textbook offers an exemplary instance, in its Appendix VI-D, of a "Specimen Physician's Claim Report" (Faulkner 1940, p. 358; see Fig. 1). The specimen report is titled "Certificate of Attending Physician." The document consists of a one-page form to be filled out, dated, and signed by an "Attending Physician" who must also supply their credentials (in the form of the medical school from which they graduated) and their "Address in Full." Twelve questions (in the form of blanks) are presented to the physician completing the form. The first blank identifies the patient (by name and address). The bulk of space on the form is given to blanks 2–5 in which the physician is asked to, "Describe and locate accurately [the] character and extent of injury" as well as any "complications," and the "present condition" and "prognosis" of the diagnosed injury. A further series of blanks requests details specific to the impairments and disabilities caused by the injury as well as the physician's opinion on how long it would have prevented the patient from working (Faulkner 1940, p. 358). This particular specimen form is meant to function as part of an accident and injury policy reimbursing insureds for occupational losses like missed wages due to injury. The narrow specificity of its intended use is of interest. But what is most notable about this document is simply the way it indexes the need for a physician attending to a patient to produce a series of medical observations that have been pre-formatted by an insurance company.

APPENDIX VI-D

SPECIMEN PHYSICIAN'S CLAIM REPORT

Commercial Accident Claim

Certificate of Attending Physician
(To be furnished without expense to the Company)

THE BLANK INSURANCE COMPANY
100 Main Street   St. Louis, Missouri

1. Patient's name and address.   Name   Address

2. Describe and locate accurately character and extent of injury.

3. Complications, if any.

4. Describe the present condition of injured parts.

5. Prognosis.

6. If fracture or dislocation, state whether complete or incomplete. If fracture of **long** bones, state whether fracture is through head or shaft.

7. If injury necessitated surgical operation, please give date and describe fully.

8. Give details of any history of physical impairments which may have contributed to lengthen period of disability.

9. Give date of first and last consultation or treatment.   First Date............................ Last Date..............

10. How many times did you see claimant for these injuries?

11. In your opinion how long was he disabled from performing any and every kind of duty pertaining to his occupation?   Total Loss of Time....................weeks....................days

12. In your opinion how long was he disabled from performing one or more important daily duties pertaining to his occupation?   Partial Loss of Time....................weeks....................days

Date...   ....
                                                                                              Attending Physician.

Graduate of...   ....
                          Medical School.   Address in Full.

Fig. 1 Specimen Physician's Claim Report, from Faulkner 1940, p. 358

The specimen form asserts itself only modestly. Yet it is remarkable in what it demands. It is a record produced by an insurance form for commercial purposes which requires of a licensed clinical practitioner that they specify certain medical facts, and conduct an examination for their discovery, whether the doctor themselves finds these particular facts relevant. The insurance form does not just ask for data, but rather it explicitly and precisely demands the provision of certain medical information.

Whenever such a form would appear in a clinical setting where patient records remained relatively informal or idiosyncratic, they would bring with them powerful incentives toward adjusting clinical practices toward the more efficient provision of the requested medical information. Consider a hypothetical mid-century family doctor presented with an increasing panoply of forms like that shown in Fig. 1.[6] An increasing majority of the doctor's patients began arranging payment through insurance and soon the forms became obligatory for the financial solvency of their practice. Our doctor was typical of the time in that they took patient notes on informal record forms they had revised from those they inherited from senior physicians or professors in medical school. Those forms were used in idiosyncratic ways, since they were mostly meant to stay within the doctor's practice. But now our doctor suddenly has to send records of specific medical observations to a third party who would approve or deny payment. This was a clear incentive to reformat recordkeeping practices in ways that make them more legible for users outside of the practice, and specifically for insurance industry professionals. The insurance forms perhaps even requested statements of medical fact that doctors on their own would not have considered pertinent to their clinical examination. But rather than refusing to record a fact they know they will have to produce later on for insurance purposes, the efficient physician redesigned their recordkeeping practices so as to observe and inscribe precisely those facts that they would need to report later. In short, they reformatted their patient records. Not all doctors and hospitals undertook these changes. But a direct incentive to reformat was there for them all nonetheless.

Faulkner's specimen form was a prototype for a paper technology constituting precisely such an incentive. After the first edition of his textbook met with success, Faulkner revised his textbook twenty years later, when he gave it the simpler title of *Health Insurance*, publishing it again in the same series (Faulkner 1960). Located again in the appendices are two specimen forms titled "Attending Physician's Statement"—one a template for individual insurance policies, the other a template for group insurance plans (Faulkner 1960, p. 616–7). These two forms offer small, but notable, revisions of their 1940 predecessors that bring them closer to the health insurance forms we know today. Whereas the 1940 form concerned only injuries,

---

[6] Though our example is hypothetical, it fits in the context of 1930s debates about the role of insurance in medical practice. With the cost of medicine rising, many patients could no longer afford to pay for expensive medical procedures, and their doctors were forced into either offering payment plans or simply donating their services. Meanwhile, organizations like the American Medical Association opposed the move towards insurance because it violated patient confidentiality. See for one account of these debates Brenner (1934) reporting on a clash between the A.M.A and the Michigan State Medical Society.

by 1960 health insurance was decidedly oriented by both sickness and injury, as attested by the first blank on the form: "Nature of Sickness or Injury." The seventh blank (following the enumeration on the individual form) requests the physician to describe any surgical procedure performed and then introduces an innovation from the 1940 form: a precise specification of the "charge for this procedure" followed by a blank with a leading dollar sign. The next blank requests specific dates of treatments and then again requires the exact "charge per call" with a dollar-sign blank. Again, the attending physician must sign and date the form, supply their address, and now their phone number (but no longer the medical school from which they received their form, though the words "M.D." are printed onto the form following the blank for the signature).

Such forms as these directly exhibit some of the ways in which insurance companies loaded data specifications onto clinical providers. Since clinical practice in 1940 had not yet achieved the same level of formalization as insurance, not only would most providers have had to simply comply with the requests of the insurance companies but the later formalization of clinical records systems in the 1960s would occur in conditions where there was already an entrenched incentive to accommodate requirements for insurance-mandated data. The result, in sum, would be the acquiescence to billing requirements widely observed today.

Commercial health insurance forms of the 1940s not only created conditions for the incorporation of insurance data into clinical care settings, but they also instituted important departures in recordkeeping from the industrial sickness funds that preceded them. While the business form of these two kinds of entities clearly differ, also notable is a sharp difference with respect to the underlying information technologies through which each conducted their operations.[7] Viewed through a media-archaeological lens, the rupture between commercial health insurance and industrial sickness funds points to a wider tension between the roles of information technology in two different styles of business organization. Industrial sickness funds had often depended on more local, intimate, and informal relationships between workers, doctors, and companies. Beginning in the first decades of the twentieth century, but then scaling up considerably in the 1930s, such community-centered initiatives were increasingly displaced by managerial efforts throughout private industry and public government. According to historians, a key quality of the new administrative style was its focus on systematicity (Yates 1989; Sklar 1988). Efforts in systematization required a more extensive formalized information apparatus. That formalization was achieved in health insurance at scale well before it would be achieved at scale in clinical settings.

---

[7] The documentary formalization developed by commercial insurers was not needed by the early industrial sickness funds. These funds were primarily aimed at supporting workers for lost wages but not for expenses of health care (Jost 2014, p. 390). And in cases where funds employed medical examination audits, they often "required that workers be examined by a physician chosen and paid by the fund or that the personal physician's decision be approved by a fund-employed physician" (Murray 2007, p. 8). It was not until insurance expanded to include hospital and clinic visits that those providing insurance would have an interest in receiving records from care providers.

Clinical medical records, 1950s–1960s

Consider now clinical medical records. Here our account begins with historical surveys conducted by others. Of specific interest are those written from an insider perspective by clinical or teaching doctors interested in the history of their own profession, or those by the technicians who helped develop early standardized medical informatics systems in an era when computers were first being brought into the clinical setting.

Richard Gillum, an M.D. at Howard University Hospital, argues that clinical medical records did not congeal as a standardized piece of professional equipment until the late 1950s and early 1960s. Gillum notes earlier partial attempts, such as an American College of Surgery program in 1918 that required hospital records for all patients with a specific protocol for types of information to be contained on the record. Clearly by 1918 the idea of a formalized clinical record was in view. But it remained only that: a vision. Less than twenty percent of physician offices kept adequate records (Gillum 2013, p. 855). Hospitals sought to remedy the lack of standard data but their efforts met with mixed success. Full success would not come until the development and uptake of information systems like Lawrence Weed's problem-oriented medical information system to which we return below (Gillum 2013, p. 855; Weed 1971).[8]

One way Gillum evidences his chronology is by counting the volume of contributions to the study of medical records by professionals. He shows that there was little attention to these topics prior to World War II and that "the number of citations [in professional journals] with 'medical records' in the title changed little [after the war] (41 in 1945–1949 and 42 in 1950–1954, 43 in 1955–1959)." But then there was "a rapid increase in varied journals after 1960 (in title/in all fields 60/687 in 1960–1964; 113/1763 in 1965–1969; 171/3098 in 1970–1974; 260/3737 in 1975–1979, compared with 414/37,950 in 2005–2009)" (Gillum 2013, p. 855).

Morris Collen's field-defining studies of the history of medical informatics hews to the same timeline as Gillum. Dr. Collen was Medical Director of the West Bay Division of Kaiser Permanente and their Physician in Chief in San Francisco as well as chair of the Executive Committee of the Permanente Medical Group from 1949 to 1973. In those positions, he led the development of electronic medical records, hospital information systems, and made other contributions to the field of medical informatics (American Medical Informatics Association 2020). In his 1995 *A History of Medical Informatics in the United States, 1950 to 1990*, Collen quickly dispatches with the first decades of the twentieth century before turning to the 1950s as the time at which medical informatics began (Collen 1995). In a companion 1986 essay titled "The Origins of Medical Informatics," Collen details the first glimmers of formal medical records systems in the late-nineteenth century, but argues that these remained only inchoate (Collen 1986). The essay is then almost entirely

---

[8] Though clinical records formalization was largely restricted to smaller-scale ad hoc settings such as family doctor practices ad specific hospitals, one larger-scale context where records were subject to a particularly idiosyncratic formalization was that of the armed forces, and specifically that of the American Expeditionary Forces during World War I, as discussed in a forthcoming media-genealogical study by Packer et. al. (forthcoming, chapter 3).

silent about the following decades (except to note advances in the development of electronic computers outside of any medical context) until we get to the introduction of computerized medical records system in the 1960s and the birth to medical informatics proper. Collen traces this history across medical contexts as varied as clinical practice, health research, and early bioengineering. In each he finds the same starting point of the late 1950s to early 1960s. As Collen describes it in his book, the era beginning in the 1950s was one that was witness to the "creation" of a "new discipline" (Collen 1995, p. 37).

A skeptic might object that Collen and Gillum overstate the absence of medical records systems in the early-twentieth century. At a localized level, a number of physicians and recordkeepers were indeed working toward the production of a more information-intensive medical practice. For just one example, neglected by both, consider "the card system" proposed in a 1903 publication by Carl E. Black, M.D., Surgeon to Passavant Memorial Hospital and Our Savior's Hospital in Jacksonville, Illinois (Black 1903, p. 103). Despite Black's introduction of index-card information management into the art of medicine, and despite his continued refinements to his system in the next decades, Black's approach received almost no uptake in the profession at large (judging from the paucity of subsequent citations). Such attempts at early medical informatics in patient-care settings notwithstanding, what stands out is that these were attempts only, rather than large-scale initiatives meeting with wide success. This lack of wide-scale formalization (for example, at the level of the profession) was no doubt in part a function of the relative decentralization of clinical medicine in the U.S., especially in contrast to state-sponsored public health, but also even in contrast to the growth of large managerial enterprises in the insurance industry.

Collen and Gillum's account, thus, clearly holds with respect to the specifically scaled-up use of medical records in clinical settings. Their point is that wide-scale usage of formalized records systems did not begin in clinical settings until the late 1950s. Concurring with this assessment is a recent state-of-the-field article in *International Journal of Medical Informatics* locating the birth of robust medical informatics in a 1959 paper in *Science* (Haux 2010, p. 601). That same paper is cited in historian Joseph November's recent account of the rise of computation in biomedicine, a development he also locates in the late 1950s (November 2012, p. 8, p. 7). Although November's account is focused mostly on biological and medical research, he does briefly consider clinical-setting informatics. He here points to a mid-1960s effort in screening records formalization at Kaiser spearheaded by Collen himself (November 2012, p. 196) and a 1966 NIH report by Mass. General Hospital doctor G. Octo Barnett (who later developed the MUMPS medical programming language) where medical records are explicitly discussed (November 2012, p. 201).

We concur with this historiographical consensus that the late 1950s were an important moment for the first formalization (via computerization) of medical records. But we also believe this observation can be enriched by attending to the role that non-clinical medical records were already playing in clinical medicine. Standard histories of medical informatics do not often focus on the precedents other kinds of records would have set for formalization and formatting in clinical care. But already implicit in the standard chronology just recounted is the possibility of an

investigation into how clinical records, as they were first formalized amidst 1950s computerization, would have to be grafted onto already-existing insurance records.

Formalization Across Health Insurance and Clinical Medicine, 1960s

To briefly recap, we have thus far excavated a chronological lag between two moments of the formalization of records in medicine. We showed that health insurance records were formalized in the 1930s and 1940s. We then showed that similar formalization did not take root in clinical medical settings until the late 1950s and especially early 1960s. With this lag in view, we turn now to exhibiting some continuities that are evident in the grafting of clinical medical informatics onto health insurance records systems.

Two kinds of medical information systems, both at the heart of contemporary health care, met in a heady moment of culture-wide computerization in the 1960s. As in so many such cases, those records systems that came first instituted tendencies and dependencies that later technicians often needed to respect. The designers of computerized clinical medical informatics systems in the 1960s would have had to deal with the fact that insurance records systems dating back to the 1940s were already circulating in clinical settings. Said differently, the relatively informal clinical recordkeeping apparatus that new computerized systems sought to formalize in the 1960s had in fact already been subject to at least some formalizing requirements due to the presence of insurance records in the clinic going back to the 1940s.

The continuities between these two kinds of formalized medical records systems were only ever partial—credit ratings do not appear in a doctor's records nor are certain specificities of medical observation ever needed by insurers. But it is indeed partial continuity that interests us, for were these two kinds of records systems completely continuous they could hardly be considered separable. The continuities in formatting and data structure we deem important are often both plain to see and yet obscured from attention. Consider, for just one example, the formats that must be in place in both insurance records and clinical informatics such that a patient can be tracked across both domains. It is a truism of records clerks that no data system implemented at scale can rely on the name as a unique identifier (for too many names are shared by many individuals). Such systems thus use identifiers. These create the need to establish continuities, or translations, across data systems. If a hospital assigns a patient an alphanumeric identifier, then this needs to be reconciled with a numeric identifier assigned by an insurer. This is almost never a trivial task.

To give a fuller sense of the continuities here, we turn now to one early and important electronic records system. We do not posit this system as a paradigm for all subsequent medical records systems. Rather, it is of interest because (recalling our methodological commitments stated above) it presents a clear case of technical and conceptual operations where insurance and billing requirements were a direct concern in designing a clinical records system. In later decades, specifically with the

widespread adoption of electronic medical records systems, these kinds of operations requirements became even more visible.

One early pioneer of medical records computerization was physician and researcher Lawrence Weed, a member of the National Academy of Sciences and the recipient of numerous awards for contributions to medical informatics. In his landmark 1968 essay "Medical Records that Guide and Teach," Weed proposed an approach to medical records relying on "new technics" of data collection and curation (1968a, p. 595). Weed's approach included standardized interview questions, the use of precision data-collection instruments, and paramedical personnel who enter data collected from patient narratives and physical examinations into standardized electronic forms. The central idea was that a computer could then generate a report featuring "only the positive findings printed out under the appropriate system heading" (Weed 1968a, p. 596). The data from these medical records could then be translated into a set of treatable "problems" for each patient. These data and their problems have significant implications for both treatment and extra-therapeutic concerns. Concerning the latter, Weed was explicit that, "A medical record maintained by the technic described will make possible a fiscal management audit" such that the use of "resources and services involved in care" can be measured directly in relation to the specific problems outlined in the medical record (1968b, p. 657). It was clearly understood that this audit would have been imbricated with the data requirements that health insurance had already loaded into clinical practice.

Weed's team took billing and insurance to be of such importance for a complete medical records system that they explicitly recognized the need to graft their new records systems onto extant pathways of health insurance information technology. The team, for example, featured the need to integrate billing into medical records as part of a 1967 grant application. Weed, along with lab co-director Jan Schultz and their team, submitted a grant proposal for the development of the computer system which would come to support Weed's Problem-Oriented Medical Record (POMR)—namely, the Problem-Oriented Medical Information System (PROMIS). According to Shultz's retrospective account, "One of the original goals for the automation of the POMR was the integration of the patient's bill with the clinical record" (1986, p. 163).[9] Another goal was the development of a "Problem Oriented List of Charges" which would function as a billing and insurance analogue to the "Problem Oriented Medical Record" (1986, p. 167). Overall, the four primary goals stated in the grant were "good patient care," improving "epidemiological studies and other research endeavors," facilitating a "medical audit" of "standards of care," and finally enabling "a business audit" (Shultz 1986, p. 159). The final goal concerned the efficiency and economics of medical approaches as they can be measured for "common medical and surgical disorders" (Shultz 1986, p. 159). Such an audit would have of course

---

[9] The team that became PROMIS Laboratories even moved their families in order to make feasible the exchange of clinical and insurance records. The team moved from Ohio to Vermont when the data processing department they worked with in Cleveland was reorganized so that "there would be no opportunity to integrate the electronic medical record with the patient's bill and the hospital's administrative operation" (Schulz, 1986, p. 163). At the University of Vermont they found an organization accommodating to such data integration needs.

included work in seamless data exchange between clinical settings and the insurance requirements that had already been installed long before 1967.

The medical records technology pioneered by Weed and his colleagues was, of course, only one of many such approaches. But we do not propose to generalize the evidence of insurer–provider interaction exhibited in Weed's work to all such systems. For our archaeology already shows that at mid-century it was possible, because in fact actual, for pre-existing health insurance records requirements to function as technical and conceptual conditions to which newer clinical records systems would have to be grafted. That grafting is evident as early as Weed's work in 1967. And it would continue for a few decades more. Octo Barnett, the MUMPS pioneer mentioned above, would observe in 1989 that, "Many medical practices have begun to use computer technology to assist in information processing. These systems have been introduced mainly because of the increasing complexity of reimbursement procedures." (Barnett 1989, p. 86). This was written on the cusp of the widespread uptake of the Epic EMR platform in the 1990s.[10]

The culmination of those three decades of grafting clinical data to insurance formats is the common observation among medical practitioners today that, as our opening scene by Mukherjee attests, health insurance data requirements do in fact shape significant swaths of electronic medical records. We have not sought to show precisely when, or even exactly why, this relationship between insurance data and medical data began. We have rather aimed to archaeologically clarify the conceptual and technical apparatus at the heart of this graft. At the center of that apparatus are found records that are highly formalized, that is, records that are tightly formatted in both technical and conceptual ways.

Conclusion

At the technological level of records systems, formalizations once in place often establish pathway tendencies (and even at times pathway dependencies) on to which later-developed formalizations are grafted. This is an important part of how it has come to be that today's clinical records systems could be shaped in part by insurance records requirements despite the widespread consternation of clinical providers. Put bluntly (as it is by numerous physicians commenting online and elsewhere), the dependence of medical records on insurance requirements often transforms patients into billable units of payment.

If it is widely understood that health insurance organizations play an outsize role in health care in the U.S. today, there is a technological dimension to this influence in addition to its widely acknowledged economic explanations.[11] Yet for all the debates that rage about health care, and all the explanations that aim to account for policy divides, too little attention is paid to the mundane data at the heart of contemporary

---

[10] See the company history at http://www.fundinguniverse.com/company-histories/epic-systems-corporation-history/, accessed January 22, 2021.

[11] See for more economistic explanations Gordon (2003) and Murray (2007), or consider the positions of many left-leaning politicians today.

medicine. We have here shown how there emerged a technological tendency, conditioned by a history of technological priority, for health insurance data to set a course for those data that populate clinical medical records. Thus is health beset by a technological inertia that conditions medicine across its many iterations. This is just one of the many ways in which data drive health, but surely also disease.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

American Medical Informatics Association. 2020. Morris F. Collen, MD, FACMI. https://www.amia.org/about-amia/leadership/acmi-fellow/morris-f-collen-md-facmi, Accessed 9 June 2020.

Barnett, G.O. 1989. The Application of Computer-Based Medical-Record Systems in Ambulatory Practice. In *Implementing Health Care Information Systems*, ed. H.F. Orthner and B.I. Blum. New York: Springer.

Berg, M., and G. Bowker. 1997. The Multiple Bodies of the Medical Record: Toward a sociology of an artifact. *Sociological Quarterly* 38 (3): 513–537.

Black, C.E. 1903. The Card System in Medical Practice. *Medical Fortnightly*, 26 October.

Bouk, D. 2015. *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual*. Chicago: University of Chicago Press.

Bowker, G., and S.L. Star. 2000. *Sorting Things Out: Classification and Its Consequences*. Cambridge: MIT Press.

Brenner, A. 1934. The Health Insurance Issue Stirs Doctors to New Debate; Need of Wider Medical Benefit and of Greater Security for Practitioners is Recognized, but Ideas on Methods Differ. *The New York Times*, 15 July: p. 3.

Broadbent, A. 2013. *Philosophy of Epidemiology*. New York: Palgrave Macmillan.

Canguilhem, G. 2012 [1989]. *Writing on Medicine* (trans: Geroulanos, S., and Meyers, T.). New York: Fordham University Press.

Collen, M.F. 1986. Origins of Medical Informatics. *Western Journal of Medicine* 145 (6): 778–785.

Collen, M.F. 1995. *A History of Medical Informatics in the United States, 1995–1990*. Indianapolis: American Medical Informatics Association.

Crosby, A. 1989/2003. *America's Forgotten Pandemic*: *The Influenza of 1918*. Cambridge: Cambridge University Press.

Ewald, F. 1991. Insurance and Risk. In *The Foucault Effect: Studies in Governmentality*, ed. G. Burchell, C. Gordon, and P. Miller, 197–210. Chicago: University of Chicago Press.

Faulkner, E.J. 1940. *Accident-and-Health Insurance*. New York: McGraw-Hill.

Faulkner, E.J. 1960. *Health Insurance*. New York: McGraw-Hill.

Foucault, M. 1984 [1976]. The Politics of Health in the Eighteenth Century. In *The Foucault Reader*, ed. P. Rabinow, 273–289. New York: Random House.

Foucault, M. 1989 [1969]. *The Archaeology of Knowledge* (trans: Sheridan, A.M.). London: Routledge.

Foucault, M. 1994 [1963]. *The Birth of the Clinic*: *An archaeology of medical perception* (trans: Sheridan, A.M.). New York: Vintage Books.

Foucault, M. 2002. [1974]. The Birth of Social Medicine. In *Power*: *The essential works of Foucault 1954–1984,* ed. J. Faubion, trans. R Hurley et al., 134–156. New York: New Press.

Foucault, M. 2004. [1974]. The Crisis of Medicine or the Crisis of Antimedicine (trans: Knowlton, E. Jr., and King, W.J.). *Foucault Studies* 1: 5–19.

Foucault, M. 2007. [1974]. The Incorporation of the Hospital into Modern Technology. In *Space, Knowledge, and Power*: *Foucault and geography*, eds. J.W. Crampton and S. Elden, trans. E. Knowlton Jr, W.J. King, and S. Elden, 141–151. New York: Routledge.

Foucault, M. 1990 [1976]. *The History of Sexuality*, *vol. 1*: *An Introduction*. (trans: Hurley, R.). London: Penguin.

Gawande, Atul. 2018. Why Doctors Hate Their Computers. *The New Yorker*, 12 November, https://www.newyorker.com/magazine/2018/11/12/why-doctors-hate-their-computers. Accessed 13 April 2020.

Gillum, R.F. 2013. From Papyrus to the Electronic Tablet: A brief history of the clinical medical record with lessons for the digital age. *American Journal of Medicine* 126 (10): 853–857.

Gitelman, L. 2014. *Paper Knowledge: Toward a media history of documents*. Durham: Duke University Press.

Gordon, C. 2003. *Dead on Arrival: The Politics of Health Care in Twentieth Century America*. Princeton: Princeton University Press.

Gordon, H. 1940. Foreword, to Faulkner. In *Accident-and-Health Insurance*, ed. E.J. Faulkner. New York: McGraw-Hill.

Hacking, I. 1991. How Should We Do the History of Statistics? In *The Foucault Effect: Studies in governmentality*, ed. G. Burchell, C. Gordon, and P. Miller, 181–195. Chicago: University of Chicago Press.

Hacking, I. 2002. *Historical Ontology*. Cambridge: Harvard University Press.

Hardy, A., and E. Magnello. 2004. Statistical Methods in Epidemiology: Karl Pearson, Ronald Ross, Major Greenwood, and Austin Bradford Hill, 1900–1945. In *A History of Epidemiologic Methods and Concepts*, ed. A. Morabia, 207–221. Basel: Birkhauser Verlag.

Haux, Reinhold. 2010. Medical Informatics: Past, Present, Future. *International Journal of Medical Informatics* 79 (9): 599–610.

Hess, V., and S. Ledebur. 2011. Taking and Keeping: A Note on the Emergence and Function of Hospital Patient Records. *Journal of the Society of Archivists* 32 (1): 21–33.

Hess, V., and J.A. Mendelsohn. 2010. Case and Series: Medical Knowledge and Paper Technology, 1600–1900. *History of Science* 48 (3–4): 287–314.

Huhtamo, E., and J. Parikka, eds. 2011. *Media Archaeology: Approaches, Applications, and Implications*. Berkeley: University of California Press.

Hu, Tung-Hui. 2015. *A Prehistory of the Cloud*. Cambridge: MIT Press.

Jeyaraj, D. 2020. EMR has Become Mostly Fluff with No Substance, It Exists for Billing and Coding—Not for Patient Care. *Twitter*. 12 April, https://twitter.com/darwinjeyaraj/status/1249400228584796166.

Jost, T.S. 2014. Health Insurance in the United States, history of. In *Encyclopedia of Health Economics*, ed. A.J. Culyer, 388–395. Boston: Elsevier.

Kittler, F. 1999 [1986]. *Gramophone, Film, Typewriter* (trans: Winthrop-Young, G., and Wurtz, M.). Palo Alto: Stanford University Press.

Koopman, C., and T. Matza. 2013. Putting Foucault to work: Analytic and concept in Foucaultian inquiry. *Critical Inquiry* 39 (4): 817–840.

Koopman, C. 2019. *How We Became Our Data: A Genealogy of the Informational Person*. Chicago: University of Chicago Press.

Koopman, C. 2021. The Political Theory of Data: Institutions, Algorithms, and Formats in Racial Redlining. *Political Theory*. https://doi.org/10.1177/00905917211027835.

Kulp, C.A. 1942. *Casualty Insurance: An Analysis of Hazards, Policies, Companies, and Rates*. New York: Ronald Press Company.

Leonelli, S., and N. Tempini. 2018. Where Health and Environment Meet: The Use of Invariant Parameters in Big Data Analysis. *Synthese*. https://doi.org/10.1007/s11229-018-1844-2.

Morabia, A., ed. 2004. *A History of Epidemiologic Methods and Concepts*. Basel: Birkhauser Verlag.

Mukherjee, S. 2020. What the Coronavirus Crisis Reveals about American Medicine. *The New Yorker*, 4 May, 24–31.

Murray, J. 2007. *Origins of American Health Insurance: A history of industrial sickness funds*. New Haven: Yale University Press.

November, J. 2012. *Biomedical Computing: Digitizing Life in the United States*. Baltimore: Johns Hopkins University Press.

Oregon Health and Science University Historical Collections and Archives. 2020. *Modern Public Health Data: Local and state boards of health*. https://www.ohsu.edu/historical-collections-archives/modern-public-health-data. Accessed 31 May 2020.

Packer, J., et al. Forthcoming. *Prison House of the Circuit: A Media Genealogy*. Ms. on file with the corresponding author.

Porter, T.M. 1986. *The Rise in Statistical Thinking, 1820–1900*. Princeton: Princeton University Press.

Schultz, J. 1986. A History of the Promis Technology: An effective human interface. In: Proceedings of the ACM Conference on the History of Personal Workstations, January, 159–182.

Sklar, M.J. 1988. *The Corporate Reconstruction of American Capitalism, 1890–1916*. Cambridge: Cambridge University Press.

Spinney, L. 2017. *Pale Rider: The Spanish flu of 1918 and how it changed the world*. New York: Public Affairs.

Vismann, C. 2008 [2000]. *Files*: *Law and media technology* (trans: Winthrop-Young, G.). Palo Alto: Stanford University Press.

Weed, L. 1968a. Medical Records that Guide and Teach. *New England Journal of Medicine* 278 (11): 593–600.

Weed, L. 1968b. Medical Records that Guide and Teach. *New England Journal of Medicine* 278 (12): 652–657.

Weed, L. 1971. *Medical Records, Medical Education, and Patient Care: The Problem-Oriented Record as a Basic Tool*. Chicago: Year Book Medical Publishers.

Winslow, C.E.A. 1980 [1943]. The *Conquest of Epidemic Disease*: *A chapter in the history of ideas*. Madison: University of Wisconsin Press.

Yates, J. 1989. *Control through Communication: The rise of system in American management*. Baltimore: Johns Hopkins University Press.

Colin Koopman   is a Professor of Philosophy and Director of New Media & Culture at the University of Oregon. His writings have appeared in *Critical Inquiry, The Hastings Center Report, New Media & Society, and The New York Times*. His most recent book is *How We Became Our Data: A Genealogy of the Informational Person*.

Patrick Jones   received his Ph.D. degree in Media Studies at the University of Oregon in Summer of 2021. His research on media studies and media history focuses on the intersections of technology and democracy as located, for instance, in the ballot box.

Valérie Simon   is a Ph.D. candidate in the Department of Philosophy at the University of Oregon. Her research focuses on questions of history, technology, and queer archival activism.

Paul Showler   is a Ph.D. candidate in the Department of Philosophy at the University of Oregon. Their research is focused on ethics, pragmatism, and metaphilosophy. Their most recent article, on genealogy and pragmatism, is forthcoming in *Synthese*.

Mary McLevey   is a Ph.D. candidate in the Department of Philosophy at the University of Oregon. She researches topics at the intersection of philosophy of technology and social and political theory.