Data Collection, EHRs, and Poverty Determinations

Craig Konnoth
University of Colorado Law School

Follow this and additional works at: https://scholar.law.colorado.edu/faculty-articles

Part of the Health Law and Policy Commons, Privacy Law Commons, and the Social Welfare Law Commons

Citation Information

Copyright Statement
Copyright protected. Use of materials from this collection beyond the exceptions provided for in the Fair Use and Educational Use clauses of the U.S. Copyright Law may violate federal law. Permission to publish or reproduce is required.

This Article is brought to you for free and open access by the Colorado Law Faculty Scholarship at Colorado Law Scholarly Commons. It has been accepted for inclusion in Publications by an authorized administrator of Colorado Law Scholarly Commons. For more information, please contact lauren.seney@colorado.edu.
Collecting and deploying poverty related data is an important starting point for leveraging data regarding social determinants of health in precision medicine. However, we must rethink how we collect and deploy such data. Current modes of collection yields imprecise data that is unsuited for research. Better data can be collected by cross-referencing other sources such as employers and public benefit programs, and by incentivizing and encouraging patients and providers to provide more accurate information. Data thus collected can be used to provide appropriate individual-level clinical and non-clinical care, and, to systematically determine what share of social resources healthcare should consume.

Key Words: Poverty, EHR, social determinants, precision medicine, privacy

Today, patients emanate increasing quantities of health information. However, not all health data are created equal. Questions regarding technology and innovation have been well explored when it comes to certain kinds of medical data. Genetic data have received the most attention. Less prominent, but nonetheless commonly analyzed, are the more typical kinds of biometric data that are already present in electronic health records (EHRs). These data include the bread and butter of medical practice—heart rate, blood pressure, and the like. When scholars write about clinical and non-clinical technologies that capture health information and deploy it for research and treatment, they usually are considering this kind of data.¹ Work that focuses on deploying data relating to social determinants of health is rarer.

Social determinants of health are environmental factors that help determine health outcomes in individuals and populations. Research suggests that on average, these factors are more determinative of health outcomes than medical interventions. It is therefore important to collect data both to understand how these determinants affect health, and to figure out how to prevent these effects—either by curing the health conditions, or by changing the underlying social conditions. However, because these determinants lie beyond the control of the healthcare system, and often, the expertise of medical professionals, information about them is lacking.
As part of the *Medicalization of Poverty* Symposium, this article is a first step in considering how data that relates to social determinants of health can be collected and analyzed by focusing on the collection, analysis, and deployment of poverty-related data to provide better rounded healthcare. There are benefits to considering poverty as an entry point to understanding other social determinants of health. First, there is already research on how poverty should be measured that takes into account parameters that are both financial and non-financial. I lay out some of those below.

More importantly, what exactly a social determinant of health *is* remains unclear. Many aspects of human activity have some bearing on health. But without data on those determinants, it is unclear how serious their effects are. Poverty escapes this classic chicken and egg problem because we *know* that poverty affects health in clear and plausible ways, some of which have already been researched. Creating a template for collecting information about poverty can then be used as a springboard for data collection regarding other determinants.

Finally, poverty is a useful place to begin precisely because it is connected to or comprises so many other social determinants of health. To understand a given individual’s poverty, we must have some information regarding his or her employment, family structure, housing situation, and welfare status. All of these are key candidates for social determinant status. Collecting data regarding poverty is a good starting point for determining where to go next. And collection and analysis of poverty data raises many of the same challenges and questions as data regarding other determinants.

The article is divided into two parts. The first examines questions of data collection, and considers the kind of data needed, sources of such data, and how to incentivize collection. It also considers privacy objections. The second considers how the data can be used for research to achieve interventions at both the individual and systemwide levels.

I. Data Collection

A. Problems

The first and possibly biggest hurdle to utilizing income data is that it is neither generated nor collected as part of regular medical practice. And even when income-related information is collected in medical settings, it is not clear how it is or should be recorded, classified, and accessed. Further, the measures currently used for assessing such data are wanting, lacking the granularity necessary for patient care and research.

Income-related data can be collected in two ways. First, the International Classification of Diseases (ICD)-10, which comprises standard medical diagnosis categories, also captures socio-behavioral information through its z-code category. However, the z-code standards relating to income are relatively vague. The Z59 code comprises all “problems related to housing and economic circumstances.” The subcodes include, *inter alia*, homelessness and inadequate housing (Z59.0 and Z59.1 respectively). There is only one subcode for income—Z59.6, “low income” (though extreme poverty, Z59.5, and insufficient welfare support, Z59.7, are closely related). An unspecified “low income” category is hard to use for precision research. The value of such research lies in its ability to produce
algorithms that combine hundreds, even thousands of data points about an individual to make predictions about them. Such predictions become harder the less precise measures are.

Relative to the z-ode standards, the Systematized Nomenclature of Medicine—Clinical Terms (SNOMED-CT) is clearer. Its hierarchical approach, designed for EHR use, presents somewhat more granular categories. For example, it distinguishes between “variable income,” and “stable income,” among high, low, moderate, and no household income, between income that is insufficient to buy necessities, to “meet needs,” to buy anything more than necessities, and income that is sufficient to meet needs. It also sets apart “low” income and “no” income. And yet, the distinctions are vague. What is, for example, “low income”? How is one to distinguish between “needs” and “necessities”?

Most important, perhaps, is the impressionistic nature of these measurements. However granular our approach, concepts such as these are subjective, especially for clinical staff that lack training in them. More importantly, in the era of telemedicine, can a doctor in New York City make these judgments about someone living outside Rochester?

B. Approaches to Collection

Rather than creating new income data categories specific to the clinical context, it makes more sense to use existing poverty measurement approaches to collect income information. This would allow linking clinical research to existing literature and understandings of poverty. Such analysis would require obtaining actual income numbers, and combining those numbers with the demographic data that poverty studies have historically used to ensure proper analysis. Much like the Meaningful Use” program, which set out “stages” of data collection and employment for Medicare providers, each stage harder than the next, poverty data can be collected in increasingly sophisticated steps.4

1. Data Required

Poverty measures can take many forms. One reasonable approach would be to rely on measures regularly used by experts in determining poverty thresholds. These measures collect data in some instances, and make educated estimates in others. Although somewhat dated, a seminal National Research Council Report from 1995 provides a well-respected and oft-cited set of measures for the field.5

The Report first seeks to account for the disposable income a person enjoys. This includes, first, “the sum of money income from all sources together with the value of near-money benefits (e.g., food stamps) that are available to buy goods and services in the budget.”6 This comprises benefits on which money would otherwise have been spent, including “housing, school lunches, and home energy assistance.” Next, the Report deducts “expenses that cannot be used to buy … goods and services…[s]uch [as] … income and payroll taxes, child care and other work-related expenses, child support payments to another household, and out-of-pocket medical care costs, including health insurance premiums.”7
One important measure that the Report does not include is expenditures on housing. Families who own their homes are better off than those with identical incomes who do not. But, as the Report notes, such a measure is difficult to make, not least because (as of 1995), 62% of low income homeowners without mortgages have housing costs in the form of property taxes, insurance, and utilities that exceeded 30% of their income. The Report calls for further research on this issue.

How well-off someone is depends on criteria other than net income. Thus, the Report considers various factors, including estimates of “patterns of household behavior” to measure “the differential needs of adults and children as well as economies of scale.” However, after identifying various approaches, the only two non-income factors the Report relies on are the size and composition of the family (with different values for adults and children), and geographical adjustments based on differential housing costs, which, the Report notes, is the biggest geographic variable.

2. Automatic Collection and Categorization

At a basic level, much of this data can be collected or estimated automatically. Half of the United States population receives health care through an employer that has easy access to income information. Income data should be passed from the employer’s databases to the provider, via any insurer. Further, those partially or fully dependent on the patient, such as unemployed or non-full time spouses, and children under the age of 26, will, in many cases, be on the plan with them. Payers will therefore also have rough information about household size.

Based on this information, estimated taxes can be calculated and deducted. The accuracy of this approach depends on the income level of the individual. Those who are under the age of 65, and who earn less than $50,000 earn approximately 90% of their income from wages. In families where no other adult is claimed as a dependent, one might assume that the beneficiary is a single parent, or that the other parent works. Accordingly, child care expenses might be imputed in such cases. And a flat amount can be deducted, as the National Research Council Report suggests, for work expenses. Adjustments can be made based on geography.

Next, a third of the population is covered through public programs that also have access to income data. The Centers for Medicare and Medicaid Services has the ability to cross-reference data with beneficiary financial information. For example, Medicare Parts B and D have income sensitive premiums. Medicaid similarly collects detailed information about income and family size, though it does not seek that of unmarried partners. One shortcoming of Medicaid data is that it fails to collect various kinds of public benefits information, such as veterans’ benefits, beneficiary payments, TANF (Temporary Assistance for Needy Families) payments, and SSI (Supplemental Security Income) payments. However, both in the case of public and private beneficiaries, over time, technology should be developed that can link medical databases, through centralized health information networks, to public benefits databases that would provide such income information.

3. Incentivized Collection
The key objection to the previous approach is that it remains inexact. Income can come from additional sources besides employers and welfare programs. Such automatic collection also does not clearly identify household size, and makes extremely rough adjustments in other cases. And while the payer and even the provider might have information regarding most medical expenses, housing related expenses are harder to determine.

Thus, ideally, the patient would be involved in collecting or at least correcting the data collected. Either at the point of intake or in their own homes, patients should be able to privately review and edit income and expense information and basic data related to family size—for example, by using electronic patient portals. Patients might be incented to complete this task by receiving income sensitive discounts, similar to the insurance payment discounts they receive when they participate in wellness programs. Payers would have to measure the cost of providing these discounts against the benefits of better population health among their patient panel due to the availability of income data. Further, although the individual providers (i.e., doctors, nurses, social workers, etc.) should not have access to the income-related data by default, patients should be informed that letting their doctors and other clinicians see the data might lead to improved and cheaper care.

To be sure, these incentives are unlikely to incent those who are better off to provide data. However, poverty data about them would (based on current knowledge) probably have less of an impact on the care they receive.

Providers should also be incentivized to ask patients for data. Even under an automatic collection regime, doctors should ask about family size and structure. Incentive programs such as the former Meaningful Use program and the current Merit Based Incentive Payment System program, which pay or penalize Medicare providers based on how they collect, deploy, and report health data, should include social determinants of health among these measures.

C. Privacy Objections

The main objection to this approach comes from privacy considerations. Just as commentators and litigants raise concerns about employers’ gaining access to personal health information through wellness programs, some may raise concerns about payers gaining access to private salary information. Entities may then use the information to discriminate against individuals based on their income, shuttling them into lower value plans, or refusing to reimburse them, confident that lower income individuals will lack the resources to challenge denials. Finally, many might have similar concerns about providers seeing the information—at worse, low income patients might receive inferior treatment, at best, they might receive less courtesy when providers or their staff see that they are low income.

At the same time, using data segmentation techniques, neither the payer, nor even, if the patient so chooses, the provider, should be able to see the content of the transmitted data. As discussed below, even if not seen, the information might be valuable for patient care through diagnostic support systems for providers. Further, data segmentation
techniques can also be used to ensure that only some individuals within the practice and not others can see the data.

Further, questions about privacy should be separated from questions about discrimination. Providers and payers already discriminate based on a range of patient characteristics, such as race, sexuality, gender identity, diagnosis, and disability, some of which may be learned through access to health information that might be considered more sensitive than income data. Discrimination is therefore a larger problem that should be addressed through education, attitudinal change, and enforcement.

Finally, studies show that a seeming desire to keep information private actually reflects embarrassment about discussing certain information in face-to-face interactions. Automatically and electronically collecting patient information on the sensitive topic of income might better comport with preferences. It is plausible that many patients recognize the importance of information such as income to their medical care, but do not want the awkwardness of discussing the information with the doctor themselves.

II. DATA USE

Once obtained, data on income can be used for various purposes, both at the individual and system level. The data can be used for interventions in clinical contexts and in non-clinical contexts. Research developed through the learning health system, in which the experiences with respect to each patient will recursively improve clinical decision-making when it comes to future patients, will determine which interventions are likely to be successful.

A. Individual Interventions

First, data can be collected to determine how best to intervene to help patients—at least those patients below a certain level of income. Financial stability can have a range of implications on outcomes. For example, studies show that low income individuals have problems adhering to medication due to costs, and a lack of transportation. Low income patients are also disproportionately affected by surgical follow-up costs, including those involving travel, childcare, and lost wages. Studies have therefore shown less post-discharge follow up and hospital utilization among this group. Other studies show that low income individuals are more subject to cardiac events; the causes merit further investigation.

Collecting and analyzing data would allow a more precise calibration between poverty and health. The risks of particular health events might plausibly be linked with different levels of income. More importantly, by analyzing—in the long run, hopefully, in real time—the effects of various kinds of interventions, providers and others might more precisely apply solutions in order to address poverty’s harms.

First, providers themselves might adopt clinical solutions, that is, medical solutions to physiological problems caused by social conditions such as poverty. One way to do this would be to incorporate income-based approaches in Clinical Decision Support systems (CDS). CDS systems offer providers ranges of recommended options based on the
patient’s customized profile (which might include their income among other characteristics). CDS is used, for example, to warn clinicians of potential complications arising from drug interactions based on other medications that the patient is taking. To be sure, providers might ignore these recommendations, and other commenters have considered the liability issues (or lack thereof) this would raise.\textsuperscript{33}

CDS, or, if available, the poverty information itself can offer a provider guidance in determining, for example, the kinds of follow up questions they should ask. Whether the patient has access to regular meals might determine what kind of medication is prescribed. Whether the patient has access to electricity, or running water will determine what kind of care a provider can reasonably expect a patient could administer for herself at home. Whether a patient has access to straightforward transportation or childcare options might also determine what kind of follow up the provider can reasonably expect—and again, might affect the care she provides today. In a fee for service regime, it might affect the kinds of procedures recommended. To be sure, the physician might get a sense of the patient’s income from the kinds of recommendations the CDS makes, but that again is far more limited information than if the physician collected the data herself, and may not exceed the kinds of hints a physician already gathers by discovering that a patient receives Medicaid.

Biological or clinical solutions do not address the root of the problem—poverty. Rather, they seek to address the harm ex post. However, the information might also be used to offer structural solutions, what I refer to elsewhere as a social approach, to solving the problems of poverty itself.\textsuperscript{34} These social interventions involve interventions outside the clinic, such as ensuring access to resources that help ameliorate or eliminate poverty. These interventions can take place in contexts that are closely related to the clinic. For example, the information can be used with payers to get telemedicine waivers if transport is hard. It might be used to waive co-payments; this, in the short term, might lose a payer money, but in the long run, might ensure better care (and lower costs) by preventing catastrophic events.

Income data can allow providers to connect individuals with other services such as counselors and social service providers to ensure that they remain adherent to expensive or hard-to-administer medication. The “cascades of care” approach that the HIV context has long used and community health care models offer blueprints—as soon as someone is diagnosed with HIV, for example, they are connected to social service providers and others to ensure adherence.\textsuperscript{35} Indeed, continued adherence might require access to other kinds of medical care—in the HIV context, for example, depression counseling has been shown to be critical.\textsuperscript{36} In the poverty context, we might see referrals to counselors, insurance co-pay programs, housing programs, and others, to ensure that patients maintain access to medication.

The data can also enable individuals to be connected to services far beyond the clinic. Certain Accountable Care Organizations (ACOs) integrate social services into the care model. These ACOs link with social services organizations and engage in interventions such as providing transport for clinical visits, and helping to address environmental causes of disease.\textsuperscript{37} Similarly, medical-legal partnerships that have caught
the imagination of many law schools and community organizations, help address legal trouble that patients may face, ranging from landlord/tenant disputes to insurance denials. Indeed, many of these interventions can be achieved automatically with the right technology. At Boston Medical Center, for example, a medical-legal partnership used electronic health records to automatically generate electricity shut off protection letters to electric companies.

B. Systemwide Interventions

Apart from determining how interventions should be carried out with respect to specific patients, collecting income-related data is helpful for identifying system-wide changes to improve the health of individuals that do not involve direct clinical or non-clinical interventions in the lives of individuals.

Clinical or clinical-adjacent transformations may include changing how we provide care for patients, by changing staffing. We might seek to hire more individuals with social work expertise rather than those with clinical training for example. Another important area of consideration involves payment models. Recently, there has been a renewed push for moving away from fee for service models to performance based models. Objections to such changes include the fact that some patient panels have greater social hardships, including poverty, than others, and therefore have worse outcomes, all else (including provider performance) being equal. In such cases, it is unfair to penalize providers for taking care of more vulnerable patients. While payment adjustments exist today for such providers, collecting and analyzing data will help determine exactly how much to weight poverty in determining value based payments.

More importantly, perhaps, collecting and analyzing this data would give us a sense of how healthcare relates to other kinds of social services. It would give us a sense of how to allocate social resources. Commentators have questioned the value of an emphasis on medical research that benefits only the well off, but fails to increase health outcomes more widely. They question how much we allocate to clinical care and research relative to say, social services, income assistance, and other programs. Their claims arise from the fact that clinical and non-clinical methods are both plausible approaches to solving health problems. The clinical approach takes for granted the social determinant, such as poverty, and works on the back end, seeking to stem or mitigate the harmful effects of the social condition. To do so, it might suggest preventative clinical measures, such as regular vitamin tablet consumption. More often, the clinical approach works ex post, in the form of different clinical treatment approaches as outlined above. Non-clinical approaches work ex ante. They seek to nip the problem in the bud by ameliorating the social harm in the first place. Our research might therefore tell us what interventions would provide the most value for money—instead of cancer research, we may determine that outcomes would be improved by increasing TANF payments. Ultimately, such research can form the basis of determining how to allot social resources.

CONCLUSION

Collecting and analyzing poverty data is both possible and important. Much of the collection can be achieved automatically; the remainder of the data can be collected through
incentivizing physicians and patients to provide information. Well-established methods can be used to protect the privacy of this data. The data can be used for important interventions, both in clinical and non-clinical spaces as they pertain to individual well-being, and more broadly for structural solutions. More importantly, poverty data would prove an important pilot experiment for collecting social determinant data more generally and give us a sense of the additional information we need to ensure good health.

Acknowledgements: My thanks to the conference organizers for this thought-provoking conference, and to Kelly Ilseng for formatting assistance.


6 Id. at 11 and 219.

7 Id.

8 Id. at 244-46.

9 Id. at 169.

10 Id. at 181-82.

11 Id.


13 45 CFR 147.120 (2011).

15 Kaiser Family Foundation, supra note 12.


20 J. Pomeranz, “Participatory Workforce Wellness Programs: Reward, Penalty, and Regulatory Conflict,” Milbank Quarterly 93, no. 2 (2015): 301-318. Admittedly, these programs are not without controversy. Id.


27 See Konnoth, supra note 1 (explaining the learning health system).


32 See C. Konnoth, Medicalization and the New Civil Rights (Working Draft, on file with author).


34 C. Konnoth, Medicalization, supra note 32.


41 The kind of approach I seek to suggest is analogous to the one here. A. A. Markovitz et al., “Risk Adjustment may lessen Penalties on Hospitals Treating Complex Cardiac Patients under Medicare’s Bundled Payments,” Health Affairs 36, no. 12 (2017).
