In order to enhance the "structural competency" (Metzl and Hansen 2014) of medicine—the capability of clinicians to address social and institutional determinants of their patients’ health—physicians need a theoretical lens to see how social conditions influence health and how they might address them. We consider one such theoretical lens, fundamental cause theory, and propose how it might contribute to a more structurally competent medical profession. We first describe fundamental cause theory and how it makes the social causes of disease and health visible. We then outline the sorts of "fundamental interventions" that physicians might make in order to address the fundamental causes.

Keywords Structural competency · Fundamental cause theory · Fundamental interventions · Social determinants · Population health
and health visible. To construct a response to “fundamental causes” we propose “fundamental interventions” in which clinicians can play a key role. We review preliminary research in support of such interventions and indicate the types of actions physicians might take to support them.

Fundamental Cause Theory

Fundamental cause theory makes sense of the fact that an inverse association between socioeconomic status (SES) and mortality has been remarkably persistent across places and times. The association is re-expressed even in the face of enormous historical changes in the risk factors and diseases afflicting humans. For example, there was a strong association between SES and mortality in Rhode Island in 1865; people of greater means (“taxpayers”) were much less likely to die than people of lesser means (“non-taxpayers”) (Chapin 1924). The main causes of death at that time were cholera, tuberculosis, and smallpox, and the main SES-related risk factors were contaminated water, substandard housing conditions, and poor sanitation. These risk factors are no longer prominent, given the development of modern sanitation systems, improved housing conditions, and cleaner water systems. Moreover, vaccines and new medications ensure that people no longer die of cholera, tuberculosis and smallpox in Rhode Island. Nevertheless, as a robust set of findings tells us, the association between SES and mortality remains strong in the current era.

Fundamental cause theory also predicts that the introduction of new health technologies, in the absence of interventions to address social inequalities, leads to larger health inequalities, given that social stratification drives unequal access to these technologies. Examples of new technologies to which unequal access has meant increasing health disparities include life-saving cancer screenings, medications that lower cholesterol, new regimens for diabetes care, and antidepressants that reduce suicide risk (Chang and Lauderdale 2009; Clouston et al. 2014; Link et al. 1998; Lutfey and Freese 2005). Thus, clinicians who strive to improve health through the introduction of technologies without attending to their social and institutional settings may end up exacerbating, rather than alleviating, inequities in health.

In order to account for the reproduction of SES gradients in different places and at different times, Link and Phelan argue that SES implies differential access to resources such as knowledge, money, power, prestige, and beneficial social connections that can be flexibly used in different places and at different times to garner a health advantage (Link and Phelan 1995). Consequently, fundamental causes affect health even when the profiles of risk and protective factors and diseases change radically. If the problem is cholera in the nineteenth century, for example, a person with greater resources is better able to avoid areas where the disease is rampant, and highly resourced communities are better able to prohibit entry of infected persons. If the problem is heart disease in the current era, a person with greater resources is better able to maintain a heart-healthy lifestyle and get the best medical treatment available. The flexible nature of the resources leads to the re-expression of SES gradients at different places and at different times. It is this re-expression that lead Link and Phelan to deem such causes “fundamental causes.”

Mechanisms of Action in Fundamental Cause Theory

Understanding how SES-related resources such as knowledge, money, power, prestige, and beneficial social connections reproduce SES health gradients requires attention to a “massive multiplicity” of health relevant circumstances (Lutfey and Freese 2005). When we step into a car we enter one that is either well-appointed with respect to safety features or not so well-appointed, we either hitch up our seat belt or choose not to do so, and we proceed on our way either driving defensively or aggressively. Our workplaces vary according to whether there are physical dangers, pollution, noise, pace demands, access to fitness rooms, and good health insurance plans.

Many of these health-relevant circumstances do not have immediate health consequences (eating two Big Macs, for example, or choosing fruit instead of French fries), have lagged health consequences (for example, working in a polluted environment), are associated with only a small elevation in risk (for example, failing to secure a seat belt), or are only relevant in a health emergency (for example, knowing people who can help you find the best doctor). As such, at any particular time these health relevant circumstances may not be prominent in our minds. Across populations and over time, though, they affect who gets sick and who succumbs.
when sickness strikes. Fundamental cause theory puts forward the proposition that flexible resources shape our exposure to and the consequences of this massive multiplicity of health relevant circumstances. But how are these fundamental causes visible in patterns of health and disease?

When disease is a matter of cruel fate, untouchable by human knowledge and action, people cannot use flexible resources of knowledge, money, power, prestige, and beneficial social connections to gain a health advantage. But as successful strides are taken in preventing disease or addressing it once it develops, these same resources have substantial health utility (Phelan et al. 2004). This reasoning leads to the prediction that, when health-relevant knowledge or technology emerges, people with more resources will benefit more and people with fewer resources will benefit less—producing health inequalities. If this is true we should be able to “see” fundamental causes at work in epidemiological trends over time for specific diseases for which health relevant knowledge has emerged. Health relevant knowledge and technology has developed for several major killers over the past fifty or so years and can also be observed in declining age-adjusted death rates—for heart disease since the 1950s, for lung cancer since the mid-1990s, for colon cancer since the late 1970s to early 1980s, and for breast cancer since the early 1990s (National Center for Health Statistics 2013).

As this knowledge and technology has developed, who has benefited more and who less? Figures 1a and 1b show age-adjusted rates of heart disease and breast cancer mortality between 1950 and 2010 in the United States by race.

In both instances we note strong declines in age-adjusted mortality with the decline starting somewhat later for breast cancer mortality. The reasons for the overall declines are multifactorial but include changes in behavioural risk factors (e.g., smoking), changes in our capacity to address risk factors such as high cholesterol (e.g., with statins), a capacity to identify disease earlier when it is more treatable, and an enhanced capacity to forestall mortality through improved medical intervention. Because of the U.S. social context of white racism and resulting discrimination, blacks are expected to be disadvantaged with respect to whites in the use of this beneficial knowledge and technology, both because they are blocked from the attainment of flexible resources and because of direct discrimination above and beyond differences in socioeconomic status. As Figure 1a shows, in the 1950s and 1960s, whites had mortality rates either equal to or higher than blacks. Then mortality rates declined dramatically and unequally so as to create a large black-white mortality disparity in the current era. As Figure 1b shows, black women had lower rates of breast cancer mortality than whites in the 1950s and 1960s but then showed climbing rates that began to match those of whites. In the 1990s, breast cancer mortality rates began their first decline and, while both white and black women benefited, a sizeable health disparity favouring whites emerged and is prominent still in 2010.2

Figure 1c and 1d show age-adjusted death rates in colon and lung cancer mortality in the United States between 1968 and 2005 (colon) and 2009 (lung) according to county-level SES (Saldana-Ruiz et al. 2013; Rubin, Clouston, and Link 2014). Individual-level indicators of SES were not recorded on death certificates over this long span of time, necessitating the use of county-level measures gauged by five indicators: the proportion of persons twenty-five years and over with less than nine years of education; the proportion of persons twenty-five years and over with at least twelve

---

1 There are some circumstances, of course, in which resources are negatively associated with health outcomes—what Lutfey and Freese (2005) call “countervailing mechanisms.” For example, goals considered even more important than health may cause those with more resources to use them in ways that negatively influence health—as when men fail to use sunscreen as part of their achievement of masculinity (Courtenay 2000). Nevertheless, as Lutfey and Freese observe, such countervailing mechanisms do not threaten the validity of fundamental cause theory since they are cumulatively smaller than the “massive multiplicity” of mechanisms that support the association between resources and health (see also Phelan, Link, and Tehranifar 2010).

2 Mortality data might be biased with respect to the accurate designation of the cause of death for minorities (Wailoo 2011). Moreover, the cause of death of minorities may not have been investigated as thoroughly as the mortality of whites, especially in earlier decades. Nevertheless, if this were true we would expect other causes of death to show similar patterns by race. Yet mortality rates due to diseases such as pancreatic cancer—for which we have not yet developed specific prevention strategies or effective treatment options—fail to show crossover like lung cancer and colon cancer (Rubin, Clouston, and Link 2014). More recent data concerning HIV/AIDS mortality demonstrates a similar pattern—the spread of antiretroviral drugs in the mid-1990s corresponds with a substantial increase in the relative risk of mortality associated with race over time (Rubin, Colen, and Link 2010). Thus, while there may be problems with mortality data, the preponderance of the evidence is suggestive of processes associated with fundamental causes.
years of education; the proportion of persons sixteen years and over currently employed in a white collar occupation; the proportion of families at or above the federally-defined poverty level; and the proportion of households with access to a telephone.

As shown in Figure 1c, during the earliest period depicted (1968–1971) age-adjusted death rates for colon cancer were highest among people living in the top SES tertile of counties, next highest in those in the middle tertile, and lowest among people living in the lowest SES tertile of counties. Then sometime after 1980, mortality rates began to decline after guidelines for screening and subsequent treatment were implemented (Saldana-Ruiz et al. 2013). But the rates declined substantially more rapidly for people living in high versus low SES counties such that the SES gradient actually reversed direction by the last period examined. In 2001–2005 the highest rate is in the lowest tertile of SES, followed by the middle tertile, with the lowest rate occurring in the highest tertile of SES counties.

Figure 1d shows that during the initial periods of observation (1968–1971 and 1972–1975) there was little difference in lung cancer rates by SES of county of residence. But then subsequent to the discovery of the harmful effects of smoking and the lag time one might expect between changes in smoking behaviour and disease occurrence the situation began to change. As mortality rates continued to rise, they rose more rapidly for people living in lower SES counties and when they finally began to decline they did so earlier and at a faster pace in higher SES counties thereby producing a substantial disparity in the current era (Rubin, Clouston, and Link 2014). What each one of these charts shows is that, when new knowledge or life-saving technology emerges, the benefits accrue more rapidly for those who are better situated with respect to SES-related

**Fig. 1** a Age-adjusted rates heart disease mortality 1950 to 2010 by race. b Age-adjusted breast cancer mortality 1950 to 2010 by race. c Age-adjusted colon cancer mortality 1968 to 2005 by SES. d Age-adjusted lung cancer mortality 1968 to 2009 by SES.
resources of knowledge, money, power, prestige, and beneficial social connections thereby creating disparities over time. In this way we can also see fundamental causes at work in epidemiological trends over time.

From Fundamental Causes to Fundamental Interventions

Fundamental cause theory is a useful lens for physicians when they seek to understand and act on the social determinants of health, but they also need to do something about the problems that the theory brings to light. We propose that fundamental causes call for physicians to make “fundamental interventions”—interventions that address the array of social issues involved in the fundamental causes of disease. Such interventions range from physicians’ participation in social movements and policy changes, to more local efforts physicians might make to improve access to the resources on which many health outcomes depend—education, housing, jobs, safety, and food security, among others.

This is a new formulation of an old idea. In the late nineteenth century, the Prussian doctor Rudolf Virchow argued that the physician was “the natural attorney for the poor” (Mills 2000, 248). Several contemporary physicians have continued this legacy by connecting their clinical practice to broader social movements or advocacy campaigns. For example, H. Jack Geiger was an advocate for Community Health Centers in the 1960s when he received national attention for writing prescriptions for malnourished children in the Mississippi Delta that were filled in local grocery stores. He points out that it was not his prescriptions that remedied the problem of poor nutrition in the region but rather his collaboration with unemployed former agriculturalist community residents to start a cooperative farm on land purchased with Community Health Center funds. In this case, clinical needs mandated a broad-based community development strategy (Geiger 2014). Geiger went on to develop a ground-breaking curriculum at the Sophie Davis School of Biomedical Education of the City University of New York, in which medical students are taught public health interventions alongside clinical medicine by undertaking neighbourhood health assessments in order to practice Community Oriented Primary Care, an approach developed in South Africa in the 1950s (Woolf and Aron 2013). More recently, public health scholar Len Syme has discussed the process by which his and his colleagues’ attempts to ameliorate hypertension among San Francisco bus drivers progressed from a focus on individual-level risk factors to the realization that the “cause of the situation was to be found … in the job itself” (Syme 1998, 501). In turn, then, Syme and his colleagues supported the drivers as they worked through their union to advocate for changes in the bus schedule that would reduce stress and increase opportunities for exercise and nutrition (Syme 1998, 2008). At a more macro level, there is increasing evidence that major health advances can be achieved via social and economic policy (Schoeni et al. 2008).

As one example, in the United States following the Civil Rights Movement of the 1960s there were dramatic improvements in both African American infant survival in the South (Almond, Chay, and Greenstone 2006) and widespread reductions in black men’s and women’s adult mortality nationwide (Kaplan, Ranjit, and Burghard 2008; Masters 2012).

We concur with Schoeni et al. (2008) that economic and social policy is health policy. This view is gaining support at multiple levels in the policy-making community, as evidenced by the recent Rio Political Declaration on Social Determinants of Health that was adopted by 125 WHO member states in October 2011 (Marmot et al. 2008). In fact, international comparisons reinforce this view. Wilkinson and Pickett (2009) found that countries with better distribution of wealth had better mental and physical health indicators and greater life expectancies. This held even for the wealthiest quartile of the population in the countries studied. A U.S. Institute of Medicine report (Woolf and Aron 2013) presented evidence that the United States has the poorest health outcomes among the industrialized nations despite spending the most on healthcare per capita. The report argued that institutional and social inequalities explained these patterns, including unequal healthcare provision; behavioural health risks that are related to social conditions such as smoking, unprotected sex, and the use of firearms; the lack of safety net programs for the poor; and the low quality of public services in poor neighbourhoods. As a counter example to the U.S. case, low to middle income countries that emphasize human development, universal healthcare, and public health oriented social policy, even with otherwise divergent political systems, such as Costa Rica and Cuba, demonstrate health indicators and life expectancies that
exceed or equal those of high income countries (Keck and Reed 2012; Unger et al. 2008).

**Fundamental Interventions in Clinical Practice**

While medical professionals can have a significant impact as advocates for health-promoting national policy, this level of intervention may seem to some clinicians to be out of their purview; they may prefer to work on their patients’ immediate health needs by addressing local social conditions. Fundamental interventions need not involve broader institutional and social change; they may simply involve including basic needs and social support in medical care. For example, several studies have established that providing food, housing, and drug treatment improves the adherence to highly active antiretroviral medications among homeless HIV positive patients (Bamberger et al. 2000; Knowlton et al. 2006; Kushel et al. 2006), a population that some clinicians had previously argued should be ineligible for HIV medications because their low medication adherence rates led to drug resistant strains of HIV. Another example is a redesign of electronic medical records at a major children’s hospital to incorporate indicators of social determinants of health. Electronically prompting clinicians to ask about key social indicators led to increased referrals to social services and to integration of care among service providers (Gottlieb et al. 2015).

A few examples from our work, and the work of contributors to this volume, further illustrate how clinicians might participate in such fundamental interventions. Hansen (the second author) directs an elective that pairs psychiatry residents with community organizers in a community development network of twenty-seven local agencies and non-governmental organizations called The Brownsville Partnership. The Partnership works in Brownsville, Brooklyn, the neighbourhood with the highest concentration of public housing in New York City, as well as a high level of violent crime and emergency room visits for psychiatric crises. Rather than psychiatrically diagnosing individuals, trainees conducted a needs assessment alongside community organizers. They identified trauma from exposure to violence as a major community concern. They then collaborated with community organizers to form peer support groups for survivors of violence and to identify neighbourhood safe spaces. This effort has led to the creation of a wellness centre with youth programs and crisis drop-in services. Where clinical trainees might otherwise have treated PTSD in individuals, fundamental cause theory directed trainees to ameliorate and prevent trauma.

Celermajer and Saul (this volume) have also taken a systems approach to trauma, working with members of the security forces of Nepal to address their sense of entrapment in a professional culture of abuse and to create disincentives for the use of torture within their ranks. And Angoff et al. (this volume) describe an intervention in medical education to model and reward positive uses of power by clinical supervisors and residents who teach medical students. This is to counteract clinical supervisors’ verbal abuse of students, a form of abuse that medical students transfer to their patients. In these cases, fundamental cause theory predicts that individual level interventions to stem police violence and to reduce abusive behaviour of medical students toward patients will fail, because individuals within abusive systems do not have the resources to resist the broader institutional culture. The fundamental cause of such abuse lies further upstream, in the reward structure and hierarchy of policing and of medical training. Thus the distribution of resources and the reward structure in each system must be changed in order to reduce abuse.

**Next Steps for Clinicians**

There are many different lenses that physicians must use in the clinical encounter—and for good reason. In one respect it is important that the physician see the patient as a particular organ or a narrow molecular interaction, for this lets her apply her expertise with the most precision. Without dismissing the importance of this expertise, we argue here that the lens of fundamental cause theory offers another important set of insights to the physician, and opens up another set of possibilities for action.

But if fundamental cause theory provides a useful lens, fundamental interventions require that doctors expand their repertoires of practices and relationships. They must seek out and build relationships with the campaigns and organizations best positioned to take on fundamental causes. They must work to translate their stories of clinical practice into mandates for action and hone their messages so as to make them suitable for media consumption. In sum, in an age of medical specialization, physicians must at the same time become the
consummated generalists, connecting their clinical practices with the public good.

This is all the more urgent today. Dr. Paul Farmer has argued that a “focus on health offers a critical new dimension to human rights work and is a largely untapped vein of resources, passion, and good will” (Farmer 2005, 238). As economic inequality deepens and market mechanisms become even more fully taken-for-granted, a right to live—a right to health—may be one of the most important points of leverage for those seeking broader social reforms.

Our argument also raises questions for the field of bioethics—which until recently has focused more on the clinical encounter than on the social determinants of health or on health disparities (Daniels 2006, 2008). A fundamental cause perspective suggests that clinicians’ focus on individual-level interventions likely exacerbates health inequalities even as it may improve health in the aggregate. Thus, structural competency is not only necessary in order for clinicians properly to address their patients’ needs but also ethically significant in that it might result in greater health equity (Chapman 2010; Whitehead 1992).

References


Geiger, H.J. 2014. The political future of public health in a time of demographic change. Lecture presented as part of the Frank A. Calderone Prize at the Columbia University Mailman School of Public Health, October 28, in New York, NY.


