



Unjust Health Inequalities: An Analysis Based on Canada's National Population Health Survey and the HealthPaths Microsimulation Model

Michael Wolfson (University of Ottawa, Canada)

Geoff Rowe (Statistics Canada)

Paper Prepared for the IARIW 33rd General Conference

Rotterdam, the Netherlands, August 24-30, 2014

Session 8A

Time: Friday, August 29, Afternoon

Unjust Health Inequalities: an Analysis Based on Canada's National Population Health Survey and the HealthPaths Microsimulation Model

Michael Wolfson and Geoff Rowe¹

IARIW Session 8A, August 2014

Measuring and Explaining Individual Attainment of Economic and Social Justice

Organizer: Gordon Anderson (University of Toronto) anderson@chass.utoronto.ca

“A model is a lie that helps you see the truth”, Howard Skipper

“action to reduce health inequalities should be taken for moral not economic reasons”, Michael Marmot²

Abstract

Discussions of the fairness or justice of health inequalities has tended to cluster into two relatively separate literatures – one empirical, the other philosophical. In the empirical literature, the preponderance of concerns regarding health inequalities is with the social patterning of health, the ubiquitous observation that higher socio-economic status, however measured, is associated with better health, also measured variously. The (often implicit) judgment is that this correlation is intrinsically unfair. The philosophical literature, in contrast, looks to the sources or causes of health inequalities – with some sources judged unfair, such as those which are clearly remediable, and others not. However, the empirical literature tends to be weak on explicit moral reasoning, while the philosophical literature typically does not include any empirical analysis.

In this paper, we endeavor to bring these two strands together. The key innovation enabling this bridging of the two literatures is a microsimulation model, HealthPaths. HealthPaths is based on and tightly coupled to a very intensive multiple equation statistical analysis of the dynamics of individuals' health and key covariates based on the longitudinal Canadian National Population Health Survey with linked mortality follow-up. This analytical machinery therefore embodies best current quantification of many of the main elements of the “web of causality” for Canadians' health. In turn, by simulating counter-factual scenarios, HealthPaths can generate health distributions “as if” a given source of health inequalities were absent.

¹ We are indebted to the members of the NIH-funded Network on Inequality, Complexity and Health for valuable discussion. This work has been funded by the Canadian Foundation for Innovation.

² Skipper quoted in Mukherjee (2010); Marmot (2013)

We have focused on three potential sources of health inequalities – smoking, socio-economic status (SES) differences, and pain – all of which give rise to unfair health inequalities based on moral reasoning. As might be expected, eliminating smoking reduces health inequalities as we have measured them. But surprisingly (including to the authors), putting everyone in the top SES group is to some extent disequalizing, and eliminating pain is substantially disequalizing. These results suggest the need for a fundamental rethinking of the concept of health inequalities.

Introduction – It is obvious that individuals within a population vary in their health. To what extent is this variation fair or just? Insofar as it is unfair or unjust, should that portion of variations in health be the focus of a polity's public policy? This paper addresses these questions.

Broadly speaking, the literature on health inequalities falls into two substantially different groups – empirical and descriptive on the one hand, and philosophical on the other. The empirical literature is generally weak compared to philosophical discussions in terms of the detail of the moral and ethical reasoning. But the philosophical literature, while rich in ideas and arguments, is generally not at all empirical.

In turn, within the empirical literature on health inequalities, there are two main strands. One treats health analogously to income and examines its univariate distribution, however measured (le Grand, 1987, 1989; Gakidou et al., 2000; Smits and Monden, 2009). The other focuses on health in relation to covariates, especially measures of socio-economic status (SES) – the ubiquitous SES gradient in health (Marmot, 2013; Mackenbach et al., 2008). This latter gradient essentially involves a focus on the bivariate distribution of health and SES.

One strand of the philosophical literature is principally concerned with which of these two main ways of thinking about health inequalities – which we have called univariate and bivariate respectively (Wolfson and Rowe, 2001, Asada 2010), while others have called them individualistic and group (e.g. Lippert-Rasmussen, 2013) and individual and “health correlations” (Hausman, 2013) – is best suited to constructing judgements about the injustice associated with observed health inequalities. For example, Braveman et al. (2011) focus on group inequalities such as those across race / ethnic lines as clearly unjust. From this vantage, income and wealth, even though they are essentially continuous variables, are often lumped in as another form of group inequality.

Further, much of the moral reasoning about health inequalities rests on the sources of observed differences in health amongst the members of the population. For example, Hausman (2013) argues that *remediable* sources of health differences may be unjust, while those for which no remediation is possible are not. While he does not discuss how to define remediable sources of health differences, this presumably includes kinds of ill health that can be cured, or avoided by reducing exposure to a risk factor, or more broadly in Marmot's sense of “the causes of the causes”, by reducing the likelihood of exposure to a risk factor by intervening on determinants of risk factor exposure – i.e. proximal and distal risk factors. On the other hand, Deaton (2013) argues that “self-inflicted” health differences, such as injuries due to freely chosen risky behaviours, are not unjust.³

³ “I treat health inequalities as important to the extent that they involve inequalities in overall well-being, and treat them as unjust when they are not compensated for by other components of well-being, when they do not play an essential part in some other good outcome, or when they cannot plausibly be attributed to freely undertaken personal choices.” Deaton

These philosophical arguments pose major challenges for any empirical grounding of judgements about the fairness or justice of observed health inequalities. It is not enough simply to observe the univariate distribution of health (however measured), or the bivariate distribution of health and SES (also however measured). It is necessary to partition, in some way, the observed variations in health, and its correlates such as SES, between various sources or causes. Then the portion of health variations attributable to one set of causes can be considered unfair or unjust, while the balance of observed health variations is not.

In this paper, we take on the ambitious task of developing a significantly new approach to the discussion of health inequalities. We bring together, in a novel manner, both the empirical and philosophical approaches to discussions of just versus unjust health inequalities. Our approach builds on a new empirically-based quantitative description of the sources of health inequalities, the HealthPaths microsimulation model. We then use this novel empirical work, in combination with the philosophical literature, to begin partitioning sources of health inequalities between those giving rise to unjust health variations, and all others. In the following section, we set out the way we will “observe” health inequalities, and then discuss the moral principles to which we appeal for partitioning sources of health inequalities between the clearly unfair and the rest. (We do not, at this stage, claim that any sources of health inequalities which are not “clearly unfair” must therefore be “fair”.) Following this, the next main section describes our method for generating the needed empirical results. This method is a complex combination of statistical estimation, in turn tightly coupled with a specially designed microsimulation model, HealthPaths. The following section then uses the HealthPaths model to simulate an initial series of counter-factual scenarios designed to support a partition of the sources of health inequalities between “clearly unfair” and other, and then to estimate the degree of health inequality in Canada that can therefore be considered unjust.

Observing Health Inequalities At any point in time, we can in principle assess individuals’ health at the moment, or over a recent relatively short time interval – say the past year. Typically, these data come from periodic population health surveys, and we shall refer to these kinds of data as cross-sectional health status.

A very large portion of health inequality research (and health research more generally) focuses on health defined in terms of diseases, a biomedical approach. But some diseases like hypertension are asymptomatic, while pain may not have any obviously associated disease. As a result, this analysis focuses on functional health status, for which there are appropriate question sets such as the McMaster Health Utility Index (HUI) (Feeny et al., 2001) and the EQ-5D (Kind et al., 2006). With these question sets, health status can be represented by a scalar in the unit interval.⁴

(2013) We discuss Deaton’s other criteria (health as a component of overall well-being, whether compensable, and health as instrumental) below.

⁴ This approach is in line with Daniels (2001) quoted below, and the frequent references, e.g. in the papers in Eyal et al. (2013) to QALYs (quality-adjusted life years), notwithstanding the predominant use of diseases in the health research literature. This latter is an artefact of the dominance of the biomedical perspective, and the way most health-related data are collected. In lay open-ended questions on what people think about when thinking about health, functional health status is the more frequent response than disease (van Dalen et al., 1994).

Similarly, we can in any given year examine the data on deaths, which are typically compiled in relation to population counts, both broken down by age and sex. The resulting data are cross-sectional age- and sex-specific death counts and mortality rates. And for each sex, we can readily tabulate the distribution of ages at death, as well as the average age at death, both of which can be “age standardized” using conventional life table methods.

In the first instance, both of these cross-sectional data sets – health status among the living, and death counts by age – can be considered as univariate distributions. Health inequalities in these cases – which we refer to here as univariate health inequalities – are evident to the extent that there is any variance at all in the distributions. This kind of thinking is analogous to the widespread literature on economic inequality where the univariate distribution of annual income, for example, is typically characterized by summary inequality measures such as the Gini coefficient. Some (e.g. LeGrand, 1987, 1989, Smits and Monden, 2009) have applied widely used income inequality measures to variations in ages at death distributions. From this vantage, health inequality – measured in terms of mortality patterns – is zero when everyone in the population dies at exactly the same age, hence the survival curve is perfectly rectangular. However, as noted by Atkinson (2013), there are crucial differences between income distributions and age at death distributions such that naïve application of summary income inequality measures (e.g. Gini coefficients) is ill-conceived.⁵

An alternative to cross-sectional health inequalities is lifetime or life course health inequalities. In the case of mortality, it is conventional to posit a hypothetical birth cohort that is exposed to a given year’s vector of age-specific mortality rates, and then compute its survival curve. This curve is the distribution of life lengths (LLs) of this hypothetical (and steady-state) birth cohort. The average age at death for this cohort, the average of the LLs, is conventionally referred to as (period) life expectancy (LE).

Further, the two aspects of health just mentioned, health status while alive, and life length = age at death, can be combined into a composite lifetime or life course measure which we can call health-adjusted life length (HALL). Instead of simply counting the number of years of life (LL), we weight each year or period of life by the (unit interval) index of health status based on a measure like the HUI or EuroQuol, and then take the sum (or integral if we are in continuous time). If we had a suitably long-lasting longitudinal health survey, for a century at least, we could observe HALLs for a representative population sample.

We refer to the distributions of LLs or of HALLs as indicative of lifetime health inequalities, as compared to cross-sectional health inequalities. Naturally, lifetime health inequalities are much more challenging to observe. Indeed, their averages are typically approximated or “indicated” by life table measures like life expectancy (LE)

⁵ Among his arguments are that: the shapes of the distributions of income and ages at death are very different, not only in being skewed in opposite directions, but also in there being an effective biological upper limit to life length (though it has continued to increase steadily over the past century in most developed countries). This is unlike income where it makes no difference to measured inequality whether it is measured in dollars or cents, it is widely agreed that economic inequality it is purely relative. For health, however, absolute levels (e.g. years) as well as relative changes (percentages) typically do matter. Finally, income is purely instrumental for welfare more broadly defined (its value is in what it can buy), while health is itself directly a major component of welfare (as well as being instrumental for other components of welfare broadly defined).

for mortality alone, or by health-adjusted life expectancy (HALE) for mortality and health status combined. HALE is analogous to LE, where LE is the average of LLs, and HALE is the average of HALLs. The idea of HALE was originated by Sullivan (1971) and has been increasingly adopted as the preferred summary measure of population health (e.g. WHO, 2014). An early estimate of HALE for Canada is Wolfson (1996); while McIntosh et al. (2009) have produced high quality estimates of the distribution of HALE by income for Canada.

These life table measures are usually based on patterns observed over a relatively short time interval, most often one year. As a result, the underlying individuals' lifetimes are hypothetical, based as they are on a sequence of age-specific rates (i.e. mortality and health status) observed in a given year, and assumed to apply every year as the cohort's members age. It is as if the cohort were aging outside of calendar time, i.e. going through its entire life cycle all in a given short period of time. Hence, these are commonly referred to as *period* life table results. (The main alternative, true *cohort* life table results, is much more demanding in terms of data, requiring both detailed historical data plus detailed projections of mortality and health status.)

We have now described two general ways of characterizing health inequalities – cross-sectional and lifetime. We have further noted that “health” in each of these cases can be based on mortality data alone, or for cross sectional data also in terms of health status, or for lifetime data (really estimates or indicators based on hypothesized data) in terms of health status integrated with mortality, in particular using health-adjusted life lengths or HALLs.

In addition, there is a crucial difference between univariate and bi- (or multi-) variate health inequalities. Analytically, univariate health inequalities are easy to understand: there are differences in (cross-sectional) health status or in mortality rates dispersed over ages, or in (lifetime) life lengths or health-adjusted life lengths (LLs or HALLs). Given such univariate distributions, there is a well-established toolkit of summary inequality measures that can be built upon from the economic inequality literature to describe the extent of inequality in a compact manner, albeit subject to the caveats noted above.

However, these univariate approaches fail to capture what is most commonly of concern in broader discussions of health inequalities – namely the correlates of health. The most noteworthy correlation is the ubiquitous gradient in health with socio-economic status (SES) (“Black Report”: DHHS, 1980; also Davey-Smith et al., 1990; Wolfson et al. 1993). However, in the U.S., the focus of bivariate health inequalities is often the correlation of some measure of health with a categorical or “group” variable, such as defined by race / ethnicity. In any case, one key strand in the discussion of health inequalities considers any gradient or correlation with a socio-economic (continuous) factor or (categorical) group to be the core of injustice associated with variations in health – if there were no strong correlation, any remaining variations in population health would not be of ethical concern.

Judging Fairness The starting point in judging the fairness of any distribution of health is the definition of good health. Most would agree with Daniels' (2001) definition as “normal functioning ... (which) preserves for people the ability to participate in the political, social and economic life of their society. It sustains them as fully participating citizens ... in all spheres of social life.” In this way, health is both an intrinsic component of any broader measure of well-being, and a means for achieving or maintaining other major elements of well-being

such as social and economic participation. In this latter sense, health is *instrumental* for, as well as intrinsic to, well-being. The existence of inequalities in health means that some members of society must have less than good health, hence a major derogation from their potential well-being. Thus, unless there are good (i.e. just or fair) reasons for poor health, the presumption must be that any (univariate) inequalities in health must be unfair or unjust.

Implicit in many of the expressed concerns about and analyses of observed SES gradients in health (i.e. a bivariate health inequality and correlation) is a moral judgment that health should be independent of SES; or relatedly that lower SES should not be a source of poorer health. As Daniels puts it, “we should view health inequalities that derive from social determinants as unjust unless the determinants are distributed in conformity with (Rawlsian) principles.” (2001, p8 col 2) Similarly, Braveman et al. (2011) state, “Health disparities ... warranting focused attention ... are a specific subset of health differences of particular relevance to social justice because they ... are likely to reinforce social disadvantage and vulnerability. Disparities in health and its determinants are the metric for assessing health equity...; (and) health equity is social justice in health.” (p2 box)

On the other hand, health differences that arise from choices freely made, that are in some sense self-inflicted rather than beyond the individual’s control, may not be considered unjust. For example, one might argue that smokers or those engaging in extreme sports should take responsibility for the adverse health effects of these choices, and no injustice should be ascribed to their premature disability or death (Deaton, 2013). However, Hausman (2013 p101) gives the example of a person who takes care of an individual suffering from a contagious disease, who then contracts the disease herself. Most would think it morally offensive to consider her illness fair because it was the result of a free choice. Marmot (2013) further criticises Deaton’s position explicitly by observing that, “In Britain, as in many countries, smoking prevalence has declined, and obesity prevalence has increased. If increasing fecklessness accounts for the obesity epidemic, what accounts for our apparent outbreak of responsible behaviour in relation to smoking? The decline in smoking has occurred because of concerted social action ...”

This latter observation regarding smoking points to another central distinction, namely between causes where public policy interventions can have some impact – i.e. the health inequalities are remediable by policies directed to their causes, and those that are beyond the control of collective actions. For example, Braveman et al. (2011, p4 col 2) state, “Society will generally be more motivated to address health differences that appear to result from modifiable circumstances over which individuals may have little control; for example, the quality of local schools, exposure to pollution or crime, or absence of stores selling nutritious food in one’s neighborhood.” Marmot (2013) states, “I regard as unfair health inequalities that could be avoidable by reasonable means.” Hence ameliorable or remediable sources of health inequalities are widely viewed as generally unjust.

In contrast to modifiable sources of health inequalities, in many cases (though possibly a decreasing number in future), there is nothing that publicly provided health care can do about infants born with (many kinds of) congenital anomalies, or the incidence of many kinds of cancer where there are no major known risk factors. These individuals’ health status is just “the luck of the draw”; and is therefore not usually considered in any society-wide sense to be unjust or unfair. (It may be considered unfair in the same sense that “Nature is

inherently unfair”, though it would be more sensible to say, “Nature is simply indifferent to or unaware of what is fair or just.”)

There is a third broad group of ethical arguments regarding the fairness of observed health inequalities. Health, while very important in and of itself, is generally only one component of the broader concept of overall well-being. Other components include SES, social relationships, and liberty. To some extent, it could be argued that health inequalities may be offset by, for example, higher income. Thus, we might be less concerned about individuals living with disabilities if there are substantial income transfers to these individuals compared to those in otherwise similar circumstances but living without disability. Simplifying in order to illustrate the point, if wellbeing is a function of only health and income, then we may judge inequalities in health to be less unfair if the joint distribution of health and income, after some income redistribution, is such that the distribution of wellbeing is less unequal. In this situation, we could say that income redistribution is *compensating* for health inequalities, and thereby reducing their injustice or unfairness (Hausman, 2013; Deaton, 2013).

While this is an important aspect for judging the fairness of health inequalities, it is generally beyond the scope of this analysis. We focus only on health inequalities without any explicit account of the (multivariate joint) distribution of other components of overall wellbeing. The key factor we will explore is the ameliorability or remediability of selected sources of health inequalities. Public policy interventions designed to reduce health inequalities may involve reductions in income (e.g. via taxes to fund programs) or the range of choice (e.g. via regulations limiting access to noxious behaviours like smoking). As a result, interventions to reduce health inequalities may reduce other components of wellbeing more broadly conceived, or act to increase their inequality. In this analysis, however, we generally avoid this third group of ethical argument by restricting our consideration of “ameliorable” sources of health inequalities to only those without major impacts on other dimensions of wellbeing.

Given this background, we shall focus on lifetime rather than cross-sectional health inequalities, and on HALLs rather than LLs. While these choices are more challenging empirically, they are conceptually and ethically (Daniels, 2001) superior. Most individuals have phases in their lives when they are healthy, and others, usually when older, when they are not. Using a life course perspective takes this general age-related downward trend in health explicitly into account; whereas using a cross-sectional (e.g. annual) approach inevitably confounds intrapersonal variations in health at different ages with inter-personal health inequalities – something to be avoided. Further, the prevalent focus on LE rather than HALE is no more than an artefact of the widespread availability of standardized mortality data, but the absence of rigorous and internationally comparable health status data – a major data gap in serious need of remedy. (see Washington City Group: CDC, no date; Budapest Initiative: UNECE, no date).

We further accept the importance of multivariate distributions of HALLs, in contrast to univariate distributions, in line with the common understanding that it is the “patterning” of variations in health that give rise to injustices, rather than dispersion per se. However, we wish to move beyond simply observing the extent of these correlations, the ubiquitous SES gradients in health. Rather, we draw on the philosophical discussion of health inequalities, especially insofar as it is premised on the causal pathways by which variations in health status emerge and play out over the life course. While this focus is most challenging from an empirical

perspective, it is essential if we are to consider the claims that the extent to which observed variations in health are fair or just depends on *why* they have arisen, on their sources. In doing so, we are not ignoring SES gradients and the social patterning of health. Rather, we seek to enrich this discussion by embedding the observation of SES gradients within a causal framework.

One further point concerns the easily made but facile analogy between health and economic inequalities. There is a crucial difference because public policy can directly affect economic inequality, both through regulation (e.g. minimum wage laws) and income or wealth redistribution (e.g. income or estate taxation and cash transfers). However, health status and longevity cannot be directly redistributed. Public policy may attempt to induce such redistribution; but it can do so only indirectly, by seeking to manipulate (with more or less effectiveness) factors affecting health – ranging from provision of health care services to anti-smoking policies to income redistribution. This restriction of public policy to indirect approaches reinforces the salience of considering the “web of causality” (Krieger, 1994) associated with health, or in other words the determinants of health.

It is important to be clear what we mean by causality in the context of this analysis. The causes or sources of health inequalities are central to much of the moral reasoning as to whether health inequalities are fair or not. But establishing causality based on empirical observation is challenging and contentious. Not least, the growing availability and increasingly sophisticated analysis of multivariate longitudinal studies has clearly shown that health status at later ages is heavily conditioned by an individual’s biography, their prior experiences and exposures extending not only back to childhood but even prenatally. “Opportunities (regarding health) accumulate throughout the life course.” (Marmot, 2013). While many of the philosophical discussions recognize this fundamental point, the complexities of the empirical analysis entailed, typically spanning many published studies, has been too much to incorporate.

A major advance in this discussion of the fairness or not of observed health inequalities therefore depends fundamentally on being able to characterize quantitatively the morally relevant “web of causality” – the network of epidemiological relationships among health, risk factors and the “causes of the causes” in order to separate out unjust sources of health inequality. This is the essence of the HealthPaths microsimulation model. Still, it is important to appreciate that we are not basing our analysis on causality that has been established with the rigour of the physics underlying a light switch, or the randomized controlled trial. Rather, as described below, we have estimated a complex series of statistical relationships which offer a plausible quantified network of causal pathways, and then drawn out their joint and interacting implications. In doing so, we recognize that there are myriad caveats and limitations. But in the spirit of real world policy analysis, we are avoiding having the perfect being the enemy of the good.

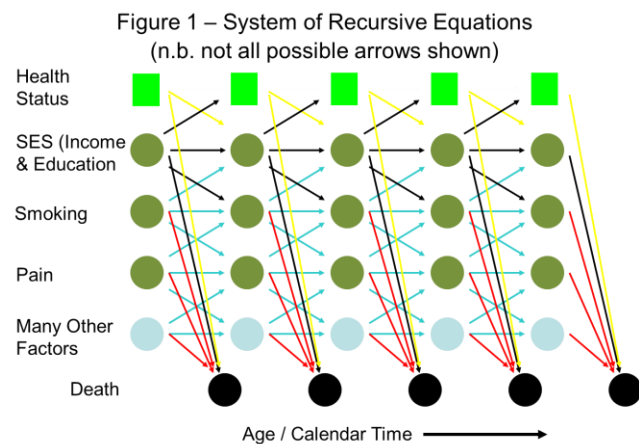
Finally, this analysis focuses on Canada. This necessarily excludes an important range of health inequalities which are difficult to describe as anything but unjust, such as the huge gulf in infant mortality rates between poor and rich countries.

Methods – This analysis is based on the HealthPaths model, an initial version of which is described in Wolfson and Rowe (2014). The version used in this analysis is a substantial extension.

The basic idea of the HealthPaths model involves two distinct but tightly coupled phases: statistical estimation and then microsimulation modeling. The statistical analysis is intensively based on the National Population Health Survey (NPHS; Statistics Canada, 1998) with an added mortality follow-up. Table 1 shows the NPHS variables included.

The empirical relationships for both phases of the analysis – statistical estimation and then microsimulation – form a network, as illustrated in Figure 1, where each “blob”

(coloured circle or rectangle) represents a measure of each of the indicated constructs, and each arrow



represents a “causal pathway”, or more concretely coefficients representing (in epidemiological parlance) effect size(s). The circles for SES represent educational attainment and income decile, which along with smoking (current, former, never, or occasional on a daily basis), and pain (measured on a 5-level scale), will be the focus of the analysis to follow. All the other variables in Table 1 are also included, and are represented in Figure 1 collectively by the “many other factors” series of gray circles. Death is shown as a black circle, where the measure in this case is the exact time of death.

Health status, in contrast, is shown as a rectangle, in order to indicate that this construct is actually a composite, an eight-dimensional vector of health states in each of the domains forming the McMaster Health Utility Index (HUI; Feeny et al., 2001) – vision, hearing, speech, mobility, dexterity, pain, cognition, and emotion. For each of these domains, the measure is an ordered categorical variable with 5 or 6 levels. Additionally, there is a valuation or scoring function that maps the full vector of eight ordered categorical variables into an index where 1 is full health and 0 is dead.

Note that in order for the diagram not to be too cluttered, only a few of the underlying arrows are actually shown. For example, while at each period there are arrows from SES to health, to smoking, and to SES in the following time period, and to death during the intervening interval, no arrow is shown from smoking or the “many other factors” to pain, for example, though there is likely some relationship.

While all the possible arrows indicating causal pathways are not drawn in Figure 1 (to avoid clutter), implicitly our “theory” is that potentially everything can affect everything else; the individual characteristics co-evolve. Thus, each variable must be a dependent variable in one equation, and it is given the chance to be an independent variable in the equations for all the other co-evolving state variables, both by itself and combined with other variables via a set of interaction terms.

Table 1 – Risk Factors / Events and Health States Included

Ordinal Variables

- Vision
- Hearing
- Speech
- Mobility
- Dexterity
- Emotion
- Cognition
- Pain
- Income Decile
- Leisure Activity
- Daily Activity
- Smoking

Functional
Health
summarized
via HUI

Binary Variables

- Employed this Year
- Family Member
- Institutional Resident
- High School Graduation
- Community College
- University Graduation
- Mortality

Quantitative Variables

- Body Mass Index
- Sense of Mastery
- Sense of Coherence
- Years of Daily Smoking

While Figure 1 shows only arrows for one period lags (a conventional first order Markov assumption), the actual statistical analysis of the NPHS allows up to two period lags (i.e. second order Markov). The result is a coherent and complete network of recursive equations. Since every equation in this network of recursive relationships is estimated from the same NPHS data, the equations are coherent both in terms of definitions, and their measurement and error structures.

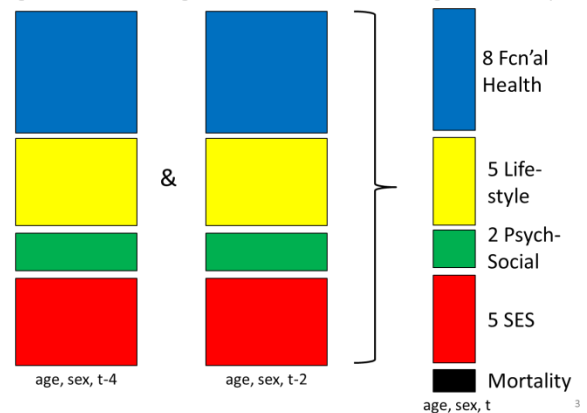
Given the conceptual framework indicated in Figure 1, in effect a system of simultaneous equations describing the (stochastic) co-evolution for any given individual of all the variables shown in Table 1, a unique approach to statistical estimation was afforded by the existence of bootstrap weights in the main data set used, the National Population Health Survey (Yeo et al., 1999). In order to allow users to account for the complex multi-stage clustered design of the NPHS sample, each individual observation in the NPHS has been given a set of 500 bootstrap sample weights. Ordinarily, these are used to estimate more accurate sampling errors for any given statistic derived from the sample. These sampling errors are typically larger than those derived under the assumption that the data had come from a simple random sample.⁶ For example, the sampling error of a number in one cell of a cross tabulation, or of a coefficient in a regression, can be simply generated (albeit via intensive computation) by running the tabulation or the regression n times, once using each vector of bootstrap weights, and then *post facto* computing the variance of the n estimates.

In this analysis, we have used these bootstrap weights in a highly novel fashion, as described in Rowe and Binder (2008). Not only have we used the bootstrap weights to determine the variances of the estimated coefficients in each of the regressions in the system of equations outlined above, but also to estimate a form of specification error and ultimately the variances of our simulations as well. Each regression equation in the coherent network of relationships was estimated 40 times, once with each of 40 bootstrap weights (sampled from the set of 500 available). Further, the fitting was done with the elastic net method with penalized likelihoods using out of sample prediction error (see Appendix).

The overall regression specification is shown in Figure 2. Since waves of the NPHS were every two years, the equations to predict the values at time t for for a given age and sex were based on the observed values at times $t-2$ and $t-4$. The set of coherently estimated transition dynamics relationships forms the first phase of the analysis. Including interaction terms, there were typically about 200 right-hand side variables available in each regression. These regressions were done by by single year of age.⁷

This richly empirical description of individuals' life course

Figure 2 – Estimating the Network of Co-Evolving Relationships

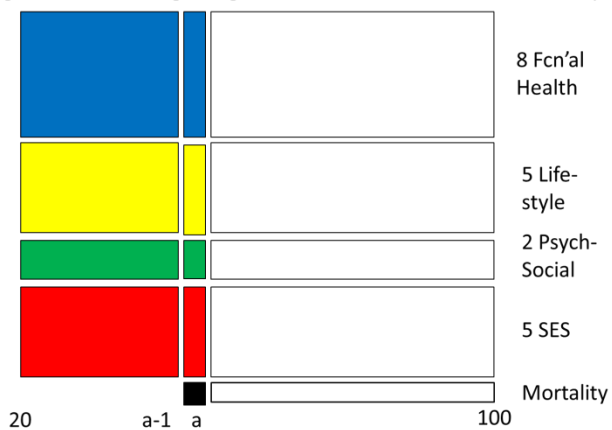


⁶ The ratio of the variance of the bootstrap replicates to the simple random sample variance is termed the “design effect” of the complex sample, and for the NPHS is generally well over two.

⁷ To mitigate sample size concerns, a kernel including adjacent ages was used in each regression.

dynamics was next used to construct a Monte Carlo microsimulation model, HealthPaths. The basic idea of this microsimulation is to start by creating one fully synthetic individual biography, and then to repeat this process many times.. To synthesize one biography, an individual is first “born” (*in silico* in the memory of the computer) by being endowed with the values for all their state variables, including their correlation structure, in this case at “birth” at age 20. (The NPHS data were not sufficient to support estimates of transition dynamics below this age.) The set of equations indicated in Figure 2 is then applied recursively to generate all the individual’s state space values each successive time period as illustrated in Figure 3. The given individual’s synthetic biography is then completed, in as realistic a manner as possible, moving to the right in Figure 3 until the individual is simulated to die.

Figure 3 – Simulating Using the Network of Estimated Relationships



Pictorially, the values for each row in the rightmost coloured column of Figure 3 are determined by the values for the two preceding time periods (recall Figure 2), plus the coefficients of the estimated network of equations, plus many random number draws to determine whether a given transition occurs.

This second order Markov assumption, indicated in Figure 2, is a compromise. On the one hand, higher order terms (i.e. variables’ values lagged three or more periods) would likely be statistically significant (e.g. Wolfson et al., 1993). Still, it is a notable

improvement on analyses which assume the relevant processes are only first order Markov. Further, the second order Markov assumption, given the eight waves of data from the NPHS from 1994 to 2008 that have been used, enables multiple observations of three time period sequences (3-tuples) of events for each individual, thereby allowing direct estimation of individual heterogeneity terms.

Note, however, that Figures 1 and 3 are misleading in our context because they give the impression that the simulation (as opposed to the estimation) uses discrete time. In fact, the simulation is considerably more sophisticated in that it uses continuous time. Correspondingly, instead of the model generating changes in status (i.e. changes in the values of any of the state variables in Table 1) that occur only at fixed annual (say) time intervals, it generates event histories with discrete events such as an improvement or decline in health status that can occur at any time t where t is a continuous variable. As a result, the statistical analysis forms a bridge between the discrete time data from the NPHS and the continuous time microsimulation modeling.

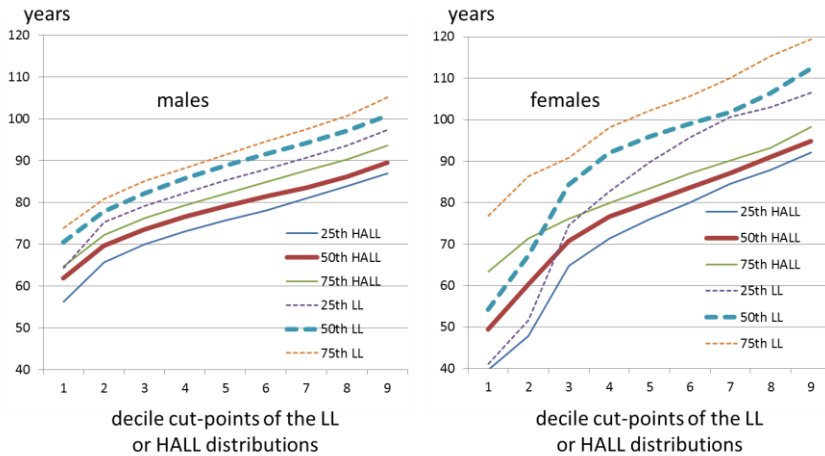
Given a complete set of empirically-based stochastic descriptions of an individual’s state transition dynamics, used to generate a single individual’s synthetic but realistic biography, the next step is simply to repeat the process shown in Figure 3 many times in order to generate a complete representative sample of the Canadian population. In this analysis, we focus on individuals in a hypothetical 1976 birth cohort⁸, and generate samples

⁸ The cohort starts with an observed multivariate joint distribution of the state variables in Table 1 for 20 year olds in 1996 (i.e. born in 1976). The trajectories of all their state variables are then simulated based on the regressions from the NPHS

on the order of 1 million individuals in order to assure Monte Carlo error is negligible (the sampling variability of LE was explicitly examined to verify this). Finally, the eight functional health status variables are aggregated using the McMaster HUI weighting function to produce for each individual during the (continuous time) interval between state changes a summary health status index with values in the [0,1] interval. The sum (actually integral, given the model's use of continuous time) of this summary health index provides each individual's entry into the distribution of health-adjusted life lengths (HALLs), our main indicator of health inequalities.⁹

Initial Simulation Results – Given the extensive statistical estimation process using the NPHS, and its incorporation into the HealthPaths microsimulation model, Figure 4 shows baseline results for a hypothetical 1976 period birth cohort (assuming survival to age 20). The dashed lines show the simulated baseline distribution of LLs, while the solid lines are the HALLs.¹⁰ In order to give an indication of the uncertainty of these results, the bold lines are the medians of the distributions over the 40 replicates, with the inter-quartile ranges

Figure 4 – Quartile Distributions of Baseline LLs and HALLs by Decile Cut-Points for Men and Women (years)



are shown by the thinner lines. The horizontal axis is the nine decile cut-points in the distribution of LLs or HALLs.¹¹

If these LL curves were rotated clockwise 90 degrees, they would correspond to conventional life table survival curves, as shown to the left in Figure 5. Since the survival curve in Figure 5 is an actual recent survival curve, the curves in Figure 4 are plausible in their shapes. However, they do show considerable uncertainty, based on the 40 bootstrap replicate statistical estimates, and their

corresponding 40 replicate simulations. This is indicated by the sometimes wide inter-quartile ranges.¹² The range of uncertainty narrows for women at higher deciles for both their HALL and LL distributions, while it widens somewhat for men. Also, the variability tends to be somewhat skewed: for both men and women, and

over the 1994 to 2008 period. As a result, the simulation generates a “thick” period birth cohort, one whose transition dynamics derive from data spanning 14 years.

⁹ Note, that the average of these HALLs is *cohort* HALE based on full individual or micro-level life course trajectories. This is different from semi-aggregate *period* HALE as typically measured using the Sullivan (1971) method, for example in Wolfson (1996) and McIntosh et al (2009).

¹⁰ While the LLs may appear high, they are in line, for example, with the projections of Canada's Chief Actuary (OSFI, 2014)

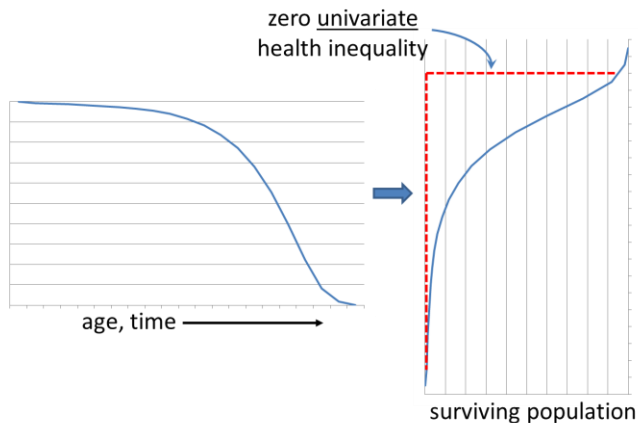
¹¹ Note that individuals are ordered separately in each of the four distributions shown (males or females x LLs or HALLs), so a simulated individual in the *k*th percentile of HALL need not be in the *k*th percentile of LL.

¹² Recall that with samples of 1 million individuals, Monte Carlo error is negligible. What remains is a mix of sampling and specification error.

both LL and HALL, the uncertainty is skewed toward lower values in the lower deciles, and shifts toward being skewed to higher values in the higher deciles (seen by comparing the distances between the 25th and 50th percentiles with those between the 50th and 75th percentiles).

Univariate and Bivariate Health Inequality – As noted in the earlier discussion of the moral reasoning regarding judgments of whether health inequalities are unfair or not, there is a fundamental distinction between what we have called univariate and bivariate health inequality (Wolfson and Rowe, 2001; Asada, 2010). On the one hand,

Figure 5 – Standard Survival Curve and Univariate Health inequality



there is a considerable tradition of using curves like those in Figures 4 and 5 as the basis for measuring health inequalities (Le Grand, 1987, 1989).

Technically, full univariate health equality is illustrated in Figure 5 by the dashed lines on the right hand part of the diagram. To the extent that the survival curve becomes “rectangular” in this sense, health inequality in this population is reduced to zero (ignoring the small tail for the longest living survivors). However, as noted by Atkinson (2013), (rotated) survival curves, or the distribution of ages at death (= first derivative of the survival curve) do not look at all like income

distribution densities. Hence there is no reasonable parallel with the economic inequality measure literature, especially to the extent that it is axiomatized in terms of Lorenz dominance (Atkinson, 1970).

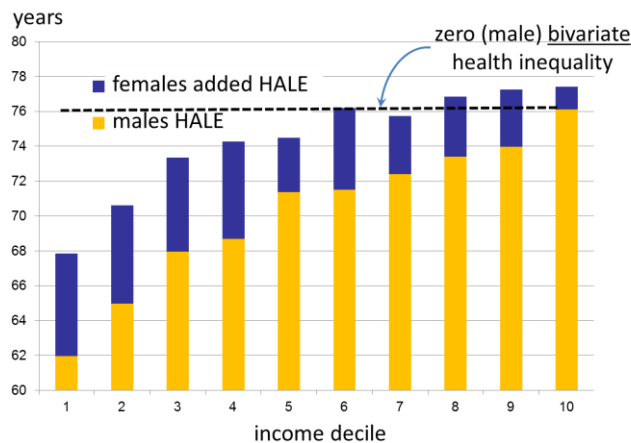
In terms of the baseline HealthPaths simulation results shown in Figure 4 above, there would clearly be no (univariate) health inequality at all if these curves were all perfectly horizontal straight lines. So the steeper the slopes of these curves, the more unequal the distribution of LLs and HALLs. Females in particular show almost a bifurcation in health inequality, with a considerably steeper slope for the first four decile cut points.

While it is clear, as shown in Figure 5, how we should define univariate health equality, there is no settled definition of when health inequalities are “unchanged” in this context. Unlike Lorenz dominance and the condition of transfers (Atkinson, 1970), which form the foundation for measures of economic inequality (though see Wolfson, 1994 with regard to the concept of polarization), there is no corresponding foundation for judging when a change in the shape of the (rotated) survival curve is either equalizing or disequalizing. Even the very widely accepted economic inequality axiom of mean independence (i.e. that proportional changes in income should have no effect on measured inequality) is not necessarily accepted in the health inequality literature.

Arithmetically, it is straightforward, when comparing two survival curves, to consider either absolute or proportional changes. Specifically, we could consider an equal absolute shift (i.e. a linear translation) as leaving inequality unchanged, or alternatively an equi-proportional shift. These two definitions are certainly not mathematically equivalent. However, the results shown below yield roughly similar qualitative interpretations for both approaches, so for simplicity we focus on translational changes, measured in years. Further, we see no

need to use or develop any summary inequality measure in for health inequalities, analogous to the Gini coefficient in the realm of economic inequality, since it will be sufficient to inspect graphical results, including those similar to Figure 4.

Figure 6 – Male and Female HALE (years) in Canada by Income Decile and Bivariate Health Inequality



Notwithstanding this discussion of measuring univariate health inequality, the majority of the literature on health inequalities, as already noted, focuses on the bivariate distribution of health and socio-economic status (both measured in many ways, as discussed above). As an illustration of bivariate health inequalities, Figure 6 shows the distribution of HALEs across income decile groups (MacIntosh et al., 2009). This is an instance of the health gradient, where every step up the socio-economic spectrum is associated with an improvement in health. Gradients like this are ubiquitous, since virtually everywhere one cares to look, health status is positively correlated with SES. In this case, essentially zero health inequality (via “levelling

up”) would correspond to the horizontal dashed line in Figure 6 where for every income decile group, HALE for males would be at that level.

As a digression, it is important to note that there are clear arguments that a majority of these social gradient associations is causal – income determining health rather than the reverse (e.g. Wolfson et al., 1993). However this is contentious. Marmot (2013) notes that there is a very strong correlation between being an economist and believing the contrary, that there is “reverse causality” where the most important causal path is that poor health leads to low income.¹³

Our analysis does not take sides in this debate. Instead, we have used sophisticated statistical analysis that makes no a priori judgments, and then combined it with agnostic and far more detailed microsimulation modeling, to transcend the extant simpler analyses. We have endeavored to capture the co-evolution, over an individual’s life course, not only of their health and SES, but also all the other factors in Table 1 above. As a result, our analysis is completely open to both social determinants as causes, and to the possibility of reverse causality, as well as to indirect effects via confounders, and lagged effects; with the results depending on the actual coefficients that have been estimated (over multiple replicates).

¹³ Interestingly, Deaton (2013) argues that this gradient is no older than the 18th century. His argument, though, is poor. Just because he cites some fragmentary evidence about the European aristocracy not living longer than everyone else, this is insufficient to conclude that there was no SES gradient at all in earlier societies. In contrast, there is evidence of a gradient in health by social status among Rhesus Macaques (Stephen Suomi, personal communication). This suggests that biological mechanisms like stress operate in all primates, including pre-18th century homo sapiens. Interestingly, Suomi’s various observations of wild Macaques, and experiments with his laboratory Macaques, suggest that the steepness of their gradient is variable, not unlike differences in the SES – health gradient between Sweden and the UK.

Returning to the measurement of bivariate health inequality, note that there could still be variability in HALLs within each income decile. To be really certain that there was truly zero bivariate health inequality of HALLs, one could break the deciles down into percentiles, but going to ever finer SES groups would soon encounter methodological issues, since HALE is a synthetic indicator that can only be estimated based on the characteristics of a group. In the limit, as group size gets ever smaller, the concept of an individual HALE converges to HALL. And a sample of individual HALLs is precisely what we are exploring with the HealthPaths simulation results. However, instead of simply describing the bivariate distribution of a measure of SES and HALLs, and using this to characterize health inequalities, we are instead building on the moral reasoning arguments surveyed at the outset. Our focus regarding bivariate health inequalities, therefore, will be on SES as a source of individual variations (i.e. univariate inequality) in HALLs, a source widely considered to be unjust.

In the U.S., much of the discussion of bivariate health inequalities is couched in terms of disparities amongst groups (Braveman et al., 2011), defined for example by race / ethnicity. In this case, instead of income deciles along the horizontal axis of Figure 6, there would be a small number of groups defined in terms of other social characteristics. But the measurement principle is the same: health inequalities exist when the (average) levels of health are not the same across the groups, and they would not exist if there were no inter-group differences. However, as Asada (2010) has argued, this is somewhat simplistic, because it ignores variations in health within each of the groups. Instead, she argues that a more sophisticated approach is needed, where one examines both between and within group variations in health (again variously measured). To this end, she briefly sketches an empirical approach for addressing these concerns. It is precisely the Asada (2010) kind of empirical approach that the intense statistical analysis of the NPHS and the tightly coupled construction of the HealthPaths microsimulation model enable us to do.

Constructing Counterfactuals – The core of our analysis of health inequalities draws on HealthPaths simulation results for a number of counterfactual scenarios. Recalling Figure 1, Figure 7 schematically indicates how this is done. The X's indicate where the estimated connections between co-evolving factors and the factor of interest are "snipped" or "knocked out". In other words, the simulation proceeds exactly as in the baseline scenario, but the role of one set of factors is changed: the counterfactual scenarios we simulate snip all the pathways leading into the variable being examined, but leave all the pathways emanating from this variable as in the baseline scenario. These selective "snips" apply from "birth" throughout each of lives of the millions of individuals simulated, not simply at the beginning of each life.

Figure 7 – Counterfactual ("Knockout") System of Recursive Equations
(n.b. not all possible arrows shown)

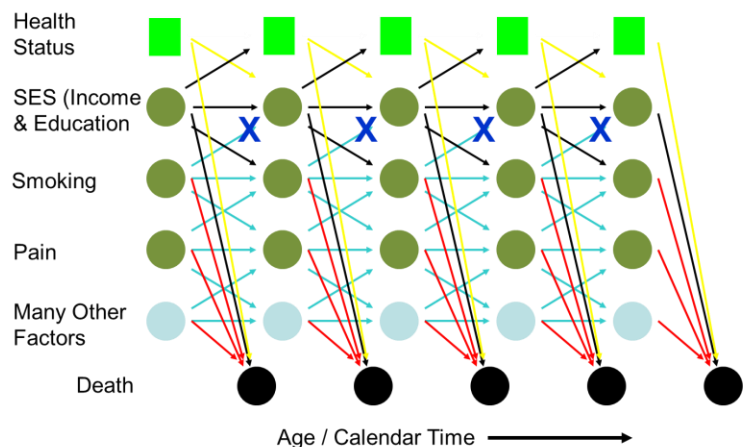


Figure 8 – LE and HALE for 40 Bootstrap Counterfactual Simulations for Each of Eight Components of Functional Health Status¹⁶

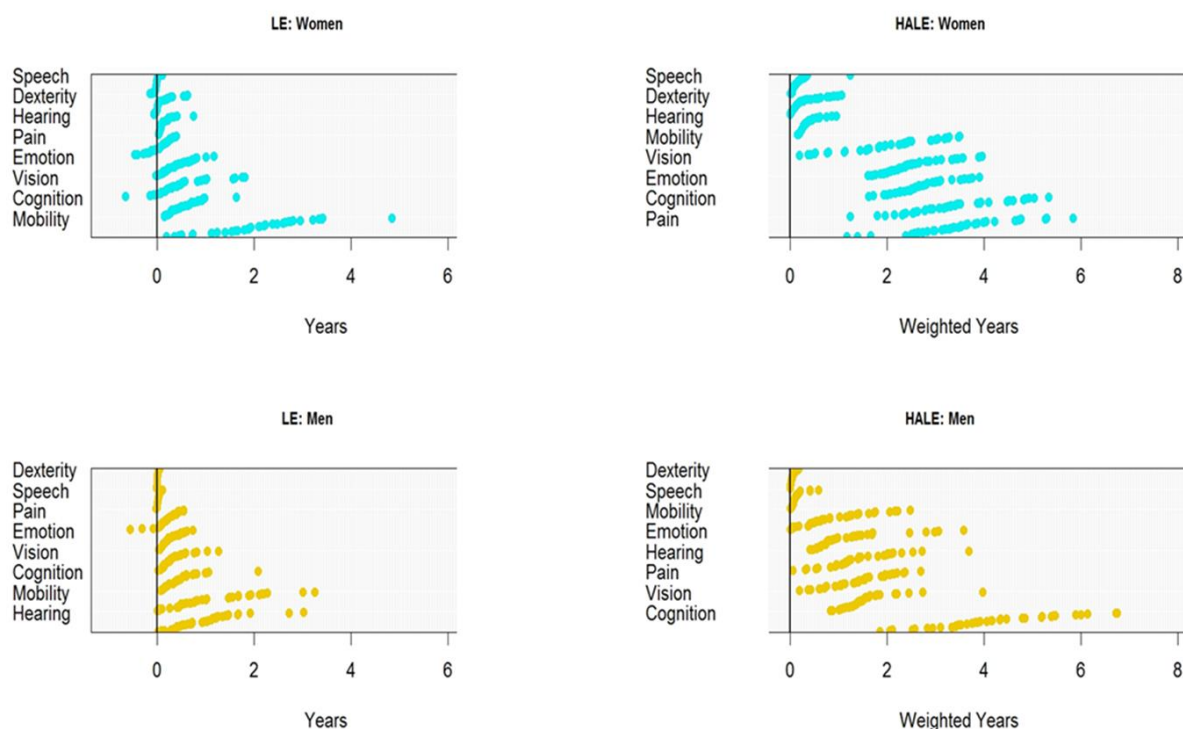


Figure 8 introduces this kind of counterfactual simulation by showing the impacts of simply eliminating any deficits, one at a time, in each of the eight underlying health domains of the McMaster HUI. In other words, whatever the baseline simulation entails, by virtue of the inputs to a given factor such as mobility, in the counterfactual scenario that factor is repeatedly reset, over the entire lifetime, to full mobility. The results are shown separately in four panels for LE and HALE, and for men and women.¹⁴

The uncertainty in these results is shown explicitly: each of the dots is the result from one of the 40 replicate simulations (in turn based on one of the 40 replicate estimations from the NPHS). The horizontal spread of these dots indicates considerable uncertainty – not only in the coefficient estimates, but also in the specification of the each equation (i.e. which variables have non-zero coefficients at all). Perhaps not surprisingly, the uncertainty is greater for HALE (on the right) than for LE (on the left). Note the different horizontal axis ranges for LE (up to 6 years) and HALE (up to 8 “weighted years”).

The eight health domains underlying the McMaster HUI have been sorted from top to bottom in each of the graphs in Figure 8 in order of increasing median (over the 40 replicates) impact on LE or HALE. In general, these functional health domains have a much larger impact on HALE than on LE – since they are often burdens throughout long periods of life without necessarily being quickly fatal, or even fatal at all. Pain and cognition are particularly notable because they show the greatest discrepancies between their impacts on LE and on HALE. Yet these health problems tend to garner much less attention than heart disease and cancers because they are not visible as major factors in the mortality data. This is the key reason we are focusing on HALE and HALLs rather than on LE and LLs – we obtain a more complete picture of the burdens of various health problems.

¹⁴ Note that Figure 8 shows the results of 1,280 counterfactual simulations (8 health domains x male and female x LE and HALE x 40 replicates) compared with the corresponding 160 (male and female x LE and HALE x 40 replicates) baseline scenarios.

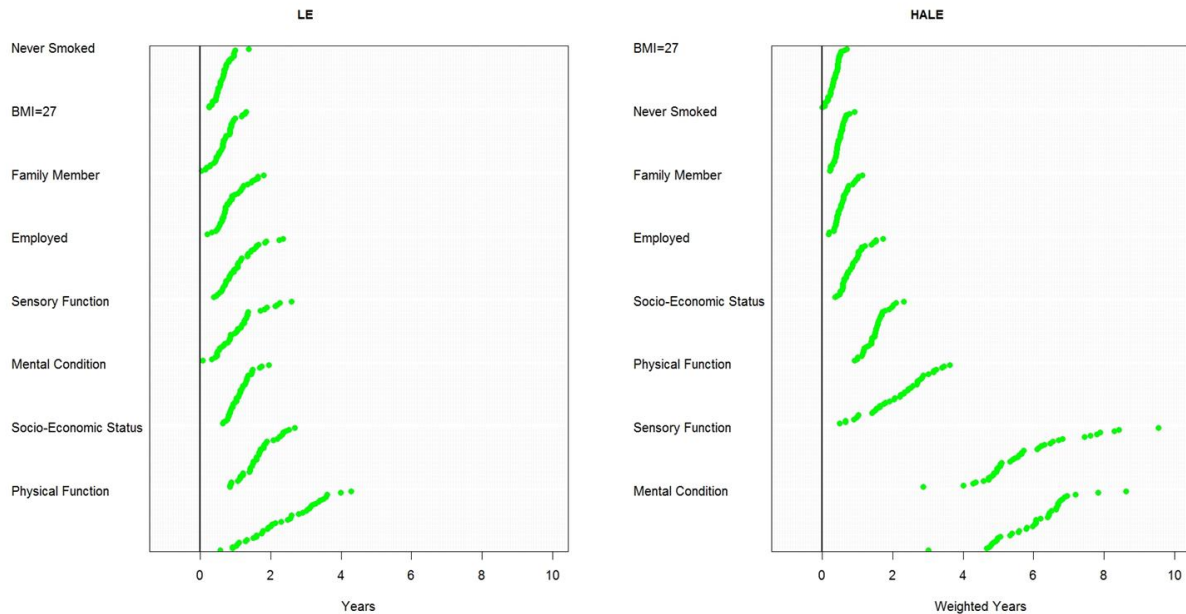
With Figure 9, we move to a fuller use of HealthPaths' capacity to simulate counterfactuals. In order to keep the results more compact, we have created several composite variables from those included in the statistical analysis of the NPHS (recall Table 1):

- Socio-Economic Status (SES) = Education + Income
- Physical Function = Leisure + Daily Non-Leisure time physical activity, + Mobility + Dexterity
- Mental Condition = Sense of Coherence + Sense of Mastery + Emotion + Cognition
- Sensory Function = Vision + Hearing + Speech + Pain

Figure 9 parallels Figure 8, but this time eliminates, one at a time, any adverse effects of each of the listed variables. And to avoid greater complexity, the results are shown for men and women combined.

Perhaps one of the most surprising results in Figure 9 is for BMI. The dots show the changes in LE (on the left) and HALE (on the right) if everyone was somewhat overweight according to the WHO categorization, specifically if everyone always had a BMI of 27 throughout their lives. Of course, in the baseline simulation, there are many individuals with BMIs above 27. But there are even more with lower BMIs, suggesting that normal (according to the WHO definition) BMI may actually be harmful for both LE and HALE. While this result runs counter to much of the current thinking, it is in line with our more detailed probing of this result (Rowe and Wolfson, forthcoming), and with several excellent published results in the epidemiological literature (Flegal et al., 2005; Flegal et al., 2013).

Figure 9 – Changes in LE and HALE for Grouped Risk Factors



The counterfactual scenarios shown in Figure 9 for Smoking Status are as if no one ever smoked; and the SES results posit that everyone is always in the top income decile and has completed post-secondary education (i.e. the top levels on both variables, again throughout life).

Eliminating smoking has a significant effect, adding on the order one year of LE and HALE. SES has an even greater effect. Forcing everyone to the top of the income decile and educational attainment groups for their entire lives, notwithstanding what their longitudinal dynamics would otherwise be in the baseline simulation, results in close to two years increase in both LE and HALE.

But especially for HALE, these BMI, smoking and SES effects are swamped by the components of sensory function and mental conditions, both of which increase HALE on the order of six years. As shown in Figure 8, pain is one of the most important factors for sensory function in Figure 9.

The counterfactual results in Figures 8 and 9 use LEs and HALEs averaged over all the LLs and HALLs of the individuals in our hypothetical 1976 birth cohort. In order to explore health inequalities, we have to delve into the “within group” inequalities. This is readily accomplished by returning to the LL and HALL survival curves in Figure 4, based on decile cut-points of the corresponding survival curves (n.b. not income deciles). But this time, we focus on *changes* in these distributions associated with specific factors, by exploiting the power of the underlying microsimulation capacity as estimated from the NPHS and incorporated into the HealthPaths model.

Moreover, we couch our discussion of health inequality in terms of the shapes of the (rotated) survival curves. In other words, we are using univariate health inequality as our normative framework. In so doing, we are not denying the fundamental importance of bivariate or group inequality, especially the SES gradient in health. Nor are we naively endorsing the univariate or individualistic approach to health inequalities. Rather, because we now have the analytical tools for exploring the “sources” of inequality – the key element in the moral reasoning about whether health inequalities are just or unjust – it is no longer necessary or even appropriate to focus directly on the SES gradient in health. Rather, in this case we can focus on the effects of a counterfactual scenario where there are no SES differences at all *ab initio*. In other words, our focus is on comparing two univariate health distributions – one a baseline, the other a hypothetical distribution where an unjust source of health inequality has been carefully removed.

Specifically, we focus on three counterfactual scenarios – where in turn each of smoking, SES differences, and pain are completely eliminated for everyone in the birth cohort over their entire lives. While there are myriad sources of health inequalities, and many arguments as to whether a given source gives rise to just or unjust health inequalities, based on the kinds of moral arguments reviewed at the outset, it is likely that there would be broad agreement that any health inequalities attributable to smoking, to SES and to pain are all unjust. The widespread moral argument against SES as a source of health inequalities has already been noted. One key strand of the argument in this case concerns equality of opportunity – it is simply unfair for someone born in poorer circumstances to be burdened throughout the rest of their life with poorer health than someone born in

more fortunate circumstances.¹⁵ In the case of smoking, as already noted by Marmot (2013), it can in principle be greatly reduced by social intervention. Pain is likely a more novel factor, but the widespread availability of inexpensive analgesics, for example, yet their pervasive under use, in part because of fears of creating addictions, suggests that pain is substantially ameliorable at no more than modest cost, hence an unjust source of health inequalities.

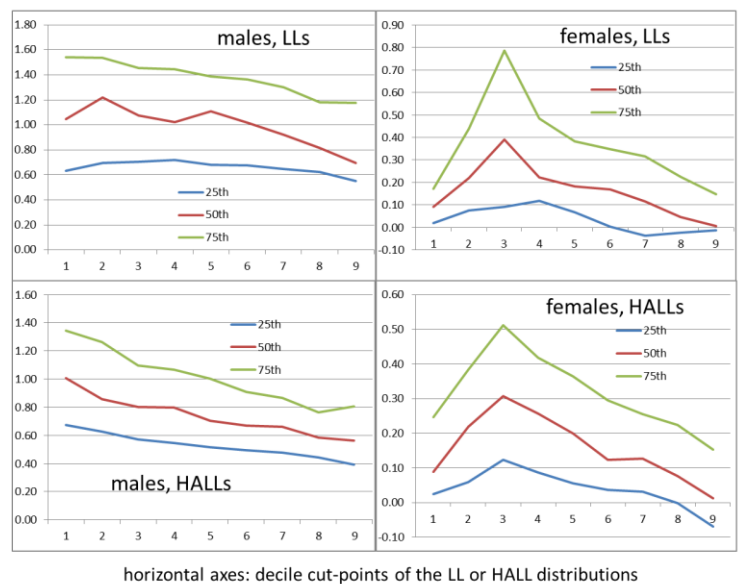
Figures 10 through 12 provide the results. In all cases, the inter-quartile ranges for the changes in the distributions of LLs and HALLs (i.e. the 25th, 50th and 75th percentile curves) indicate that there is considerable uncertainty.

Figure 10 shows the changes in LLs and HALLs for men and women in the event that no one in our hypothetical 1976 birth cohort ever smoked, compared to the baseline scenario where there is a considerable amount of smoking. (Note that the vertical axis scales are different in each of the graphs. Further, individuals are sorted from left to right differently in each of the graphs, though within each graph the differences between the hypothetical scenario and the baseline always use the same replicate.)

Eliminating smoking has mostly an equalizing effect – since the *changes* in the survival curves get smaller as we move from left to right, i.e. from lower deciles of LLs or HALLs to higher deciles. This result likely follows from smoking rates being higher among those in the lower LL or HALL deciles.

Perhaps surprisingly, eliminating SES effects (Figure 11) – by forcing everyone always to be in the top income decile and in the top educational attainment group – is close to neutral for men across the health spectrum, and for women in the middle of the HALL spectrum. It is only clearly equalizing at the top of the univariate LL distribution for women, while it is substantially disequalizing for women in the lower 40% or so of both the LL and HALL distributions.

Figure 10 – No Smoking Scenario Compared to Baseline



¹⁵ This moral statement draws on the widespread evidence that birth and early childhood circumstances have powerful effects on health throughout the subsequent life course.

Figure 11 – Top SES Scenario Compared to Baseline

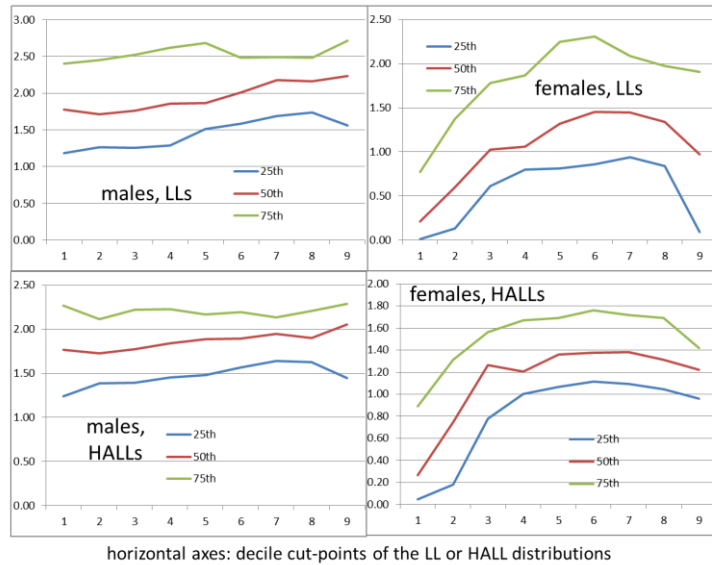
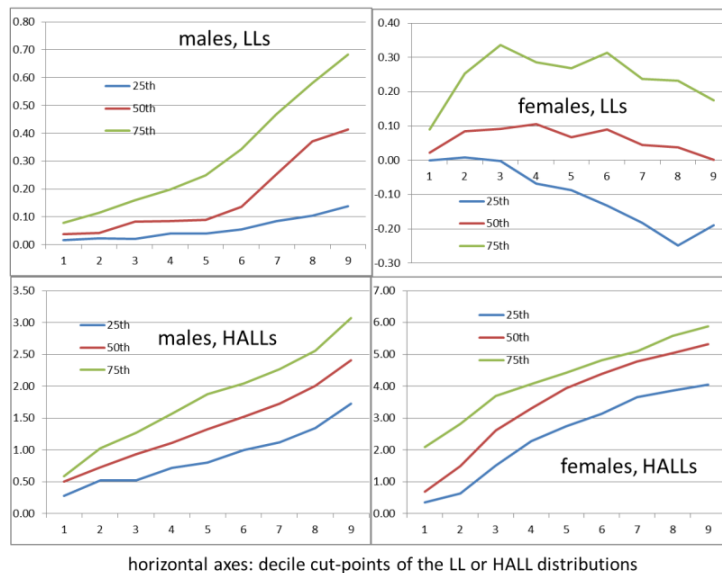


Figure 12 – No Pain Scenario Compared to Baseline



Still, there is no question that eliminating SES effects as posited in our counterfactual simulation increases both LLs and HALLs generally – on the order of two years for men and 1 1/3rd years for women in the middle decile range, as already shown in Figure 9 above.

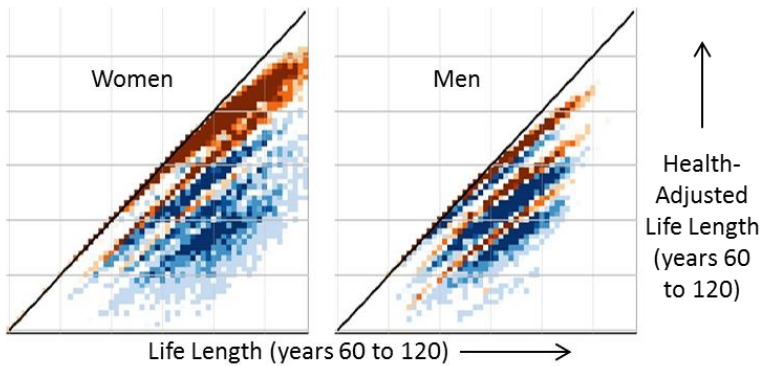
Finally, the most dramatic pattern is for pain (Figure 12). Its elimination would generally increase health inequalities for both men and women, but much more so for women's HALLs. (Again, note the differences in the vertical axis scales.) More specifically, eliminating pain has virtually no impact on women's LLs, less than 0.1 years, and less than this amount for over half of men. But eliminating pain has a much larger effect on HALLs – as much as 5 years for women in the upper range of HALL.

Figures 13 a to c provides another view of these results. These scatter plots show the bivariate distribution of *changes* in LLs and HALLs for the same three counterfactual scenarios, with LLs along the horizontal axis, and HALLs along the vertical. The more intense blue indicates a larger negative change; the more intense orange indicates a larger increase in the population in each LL x HALL cell.¹⁶

In general, it is apparent from Figures 13 a to c that pain has a more pervasive impact than SES, which in turn has more impact than eliminating smoking – indicated by the overall intensity of colour in each of the pairs of graphs. Further, elimination of pain appears to have a stronger impact for women than for men, while in the cases of the top SES and no smoking scenarios, the impacts are greater for men than those for women.

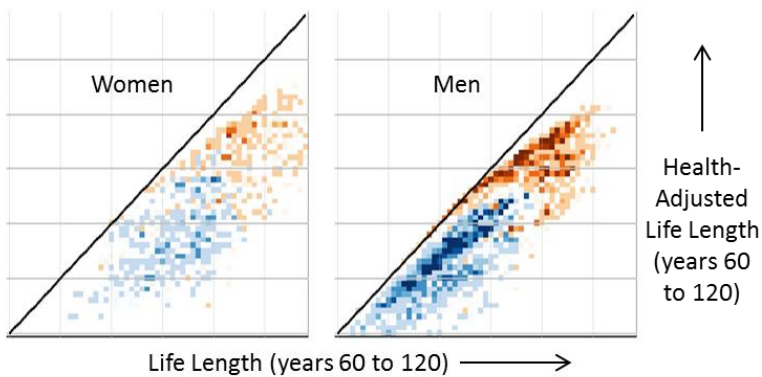
¹⁶ The horizontal and vertical axes are each 100 one year intervals, resulting in 10,000 cells. The colour coding is based on 19 equal sized intervals centered on 0.0 with an open-ended interval at the top and bottom. The "streakiness" visible in some of the scatter plots is due to a degree of clustering of HUI values.

Figure 13a – From Baseline to No Pain



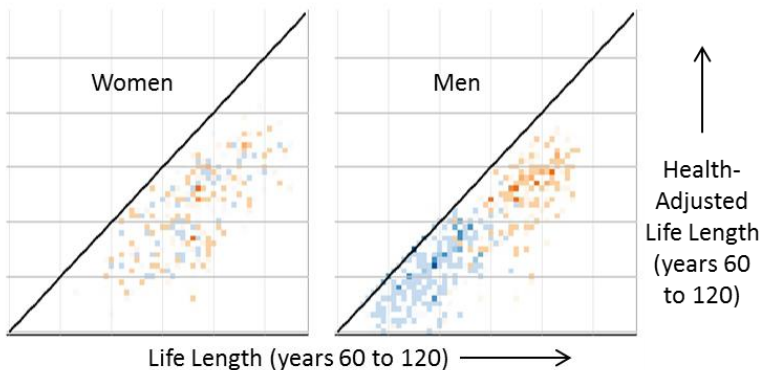
In all cases, the blue dots are toward the lower left, while the orange dots tend to the upper right – indicating that the counterfactual scenarios all increase both LE and HALE. However, for pain, the blue to orange movement in the graphs is much closer to vertical. In other words, elimination of pain is increasing HALLs much more than it is increasing LLs.

Figure 13b – From Baseline to Top SES



While the patterns in Figure 13 all show improvements associated with the counterfactual scenarios of no smoking, top SES and no pain, they are not as obvious for assessing the impacts of these scenarios on health inequalities. The key results in this regard are in Figures 10 – 12. And in the case of pain, it matters little what specific (univariate) health inequality measure is used. The upward slopes of the curves clearly indicate a widening of the HALL distributions for both men and women. Intuitively, this corresponds to pain being relatively more concentrated among those who have higher HALLs. This is indicated in Figure 13a especially for women by comparing the implicit contour lines for the concentrations of blue and orange areas, both of which are mostly elongated ovals. The slope of the most pronounced blue concentrations is shallower than that of the corresponding orange area. As a result, those with higher HALLs to start are experiencing larger increases in their HALLs from the elimination of pain. This is disqualizing.

Figure 13c – From Baseline to No Smoking



Concluding Comments – There are two broad strands of literature on the fairness of health inequalities – one primarily empirical and descriptive, the other mainly philosophical. However there is not a great deal of overlap between the two. The analysis reported here provides a novel and provocative bridge. Based on a review of the moral reasoning with regard to health inequalities, we have built on the argument that a major criterion for judging fairness is whether or not the source of the inequality is fair. Further, sources of health inequalities that

are ameliorable by social action are widely judged to be unfair, as are sources associated with unequal opportunities – most often reflected in socio-economic status (SES) differences that are correlated with health.

Based on this moral reasoning, elimination of health problems attributable to any one of pain, SES and smoking would be eliminating a source of unjust health inequalities. Using the HealthPaths microsimulation model, we have built on these moral arguments and explored their implications empirically. The basic technique has been to construct explicit but hypothetical scenarios where each of these sources of health inequalities was absent. The key results are then based on the comparisons of these counterfactuals with the baseline scenario.

In so doing, we have focused on the univariate distribution of health, measured over the full life course of a representative Canadian period birth cohort using health-adjusted life length (HALL) as the health metric. By using a univariate approach to measuring health inequalities, we are in no way denying the importance of the social determinants of health. They exist, and have been incorporated into HealthPaths based on very extensive empirical analysis of Canada's longitudinal National Population Health Survey. Rather, we have endeavored to place these SES factors into a broader framework – moving beyond piecemeal empirical results.

The results are surprising (including to the authors). Elimination of pain, which seems on the face of it to be clearly an unjust source of health problems, would actually increase (univariate HALL) health inequalities. Also, elimination of SES effects on health and longevity, when dozens of other factors are explicitly included in the analysis, while clearly increasing average health, has essentially no effect on health inequalities among men, and generally increases health inequalities as here defined among women. Only the elimination of smoking reduces health inequalities, and then only modestly.

References

- Asada, Y (2013), "A Summary Measure of Health Inequalities", Chapter 3 in Eyal, N, S A Hurst, O F Norheim, and D Wikkler (Eds) *Inequalities in Health: Concepts, Measures and Ethics*, Oxford.
- Atkinson, A B (1970), "On the Measurement of Inequality". *Journal of economic theory*, 2(3), 244-263.
- Atkinson, AB (2013), "Health Inequality, Health Inequity, and Health Spending", Chapter 2 in Eyal, N, S A Hurst, O F Norheim, and D Wikkler (Eds) *Inequalities in Health: Concepts, Measures and Ethics*, Oxford.
<http://books.google.ca/books?hl=en&lr=&id=7B5pAgAAQBAJ&oi=fnd&pg=PA27&ots=tg312vVmdp&sig=b97snbYpisAqqvgwUuxlJ0kCM8Y#v=onepage&q&f=false> accessed on July 12, 2014
- Braveman, P. A., Kumanyika, S., Fielding, J., LaVeist, T., Borrell, L. N., Manderscheid, R., & Troutman, A. (2011). "Health disparities and health equity: the issue is justice". *American Journal of Public Health* 101
- CDC (no date) Washington Group on Disability Statistics, (http://www.cdc.gov/nchs/washington_group.htm)
- Daniels, N (2001), "Justice, Health and Health Care", *American Journal of Bioethics* 1, no. 2 (2001): 2-16.
- Davey-Smith, G, M Bartley, D Blane (1990), "The Black Report on Socio-Economic Inequalities in Health 10 Years On", *BMJ*, Vol 301,
- Deaton, A (2013), "What Does Empirical Evidence Tell Us About the Injustice of Health Inequalities", Chapter 17 in Eyal, N, S A Hurst, O F Norheim, and D Wikkler (Eds) *Inequalities in Health: Concepts, Measures and Ethics*, Oxford.
- DHHS (1980) Department of Health and Social Security. *Inequalities in health: report of a research working group*. London.
- Eyal, N, S A Hurst, O F Norheim, and D Wikkler (Eds) (2013) *Inequalities in Health: Concepts, Measures and Ethics*, Oxford Feeney et al.

- Feeny, D., W Furlong, G Torrance, C Goldsmith, Z Zhu, S DePauw, M Denton, M Boyle (2001), "Multi-Attribute and Single Attribute Utility Functions for the Health Utility Index Mark 3 System", *Medical Care*, Vol 40, No 2, pp113 – 128.
- Flegal, K. M., Graubard, B. I., Williamson, D. F., & Gail, M. H. (2005). Excess deaths associated with underweight, overweight, and obesity. *Jama*, 293(15), 1861-1867.
- Flegal, K. M., Kit, B. K., Orpana, H., & Graubard, B. I. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *Jama*, 309(1), 71-82.
- Gakidou, E. E., Murray, C. J., & Frenk, J. (2000). Defining and measuring health inequality: an approach based on the distribution of health expectancy. *Bulletin of the World Health Organization*, 78(1), 42-54.
- Hausman, D M (2013) "Egalitarian Critiques of Health Inequalities", Chap 7 in Eyal, N, S A Hurst, O F Norheim, and D Wikkler (Eds) *Inequalities in Health: Concepts, Measures and Ethics*, Oxford.
- Kind, P., Brooks, R., & Rabin, R. (Eds.). (2006). *EQ-5D concepts and methods:: a developmental history*. Springer.
- Krieger, Nancy (1994), "Epidemiology and the Web of Causation: Has Anyone Seen the Spider?" *Soc. Sci. Med.* 39(7). pp. 887-903.
- Le Grand, Julian (1987). "Inequalities in health: some international comparisons." *European economic review* 31, no. 1 (1987): 182-191.
- Le Grand, Julian (1989). *An international comparison of distributions of ages-at-death*. In J. Fox (Ed.) *Health Inequality in European Countries*, Gower Pub Co., 1989.
- Lippert-Rasmussen, K (2013), "When Group Measures of Health Should Matter", Chapter 4 in Eyal, N, S A Hurst, O F Norheim, and D Wikkler (Eds) *Inequalities in Health: Concepts, Measures and Ethics*, Oxford.
- McIntosh, C, P Fines, R Wilkins, M Wolfson (2009) "Income disparities in health-adjusted life expectancy for Canadian adults, 1991 to 2001", *Health Reports*, November 2009, Statistics Canada (<http://www.statcan.gc.ca/pub/82-003-x/2009004/article/11019-eng.pdf> accessed Feb 2012)
- Mackenbach, JP, I Stirbu et al. (2008), "Socioeconomic inequalities in Health in 22 European Countries", *N Engl J Med* 2008;358:2468-81.
- Marmot, M (2013), "Fair Society, Health Lives", Chapter 18 in Eyal, N, S A Hurst, O F Norheim, and D Wikkler (Eds) *Inequalities in Health: Concepts, Measures and Ethics*, Oxford.
- Mukherjee, S (2010), *The Emperor of All Maladies, a Biography of Cancer*, Scribner, New York.
- OSFI (2014) Mortality Projections for Social Security Programs in Canada, Actuarial Study No. 12, Office of the Chief Actuary, (<http://www.osfi-bsif.gc.ca/Eng/oca-bac/as-ea/Pages/mpsspc.aspx>)
- Rowe, Geoff and Binder, David (2008). "Can Bootstrap Replicates be used for Cross-validation?". 2008 *Proceedings of the Survey Research Methods Section*. American Statistical Association Annual Meeting.
- Smits, Jeroen, and Christiaan Monden. "Length of life inequality around the globe." *Social Science & Medicine* 68, no. 6 (2009): 1114-1123. – need pdf
- Statistics Canada (1998), "National Population Health Survey Overview, 1996/7" (<http://www.statcan.gc.ca/pub/82-567-x/82-567-x1997001-eng.pdf> accessed Feb 2012)
- Sullivan, Daniel F (1971), "A Single Index of Morbidity and Mortality", *HSMHA Health Reports* 1971;86(4):347-355.
- UNECE (no date), Task Force on Measuring Health Status (Budapest Initiative) (<http://www.unece.org/index.php?id=28480>)
- Van Dalen, H A Williams, C Gudex (1994) "Lay People's Evaluations of Health: Are There Variations between Different Subgroups?" *Journal of Epidemiology and Community Health*, Vol. 48, No. 3 (Jun., 1994),
- WHO (2014) *World Health Statistics 2014*, WHO, Geneva (http://apps.who.int/iris/bitstream/10665/112738/1/9789240692671_eng.pdf?ua=1)

- Wolfson, M, G. Rowe, J. Gentleman, and M. Tomiak (1993) "Career Earnings and Death: A Longitudinal Analysis of Older Canadian Men", Journal of Gerontology: Social Sciences, Vol. 48, No. 4, pp S167-S179.
- Wolfson, M (1994) "When Inequalities Diverge", American Economic Review, May 1994.
- Wolfson, M (1996) "Health-Adjusted Life Expectancy" Health Reports, Statistics Canada, Vol 8, No. 1, Summer 1996.
- Wolfson, M and G Rowe (2001) "On Measuring Health Inequalities", Bulletin of the World Health Organization, 2001, 79: 553–56
- Wolfson, M and G Rowe (2014), "HealthPaths: Using Functional Health Trajectories to Quantify the Relative Importance of Selected Health Determinants", Demographic Research, forthcoming
- Yeo, D, H Mantel and T-P Liu (1999), Bootstrap Variance Estimation for the National Population Health Survey, In *American Statistical Association: Proceedings of the Survey Research Methods Section*, pp. 49-57 (https://www.amstat.org/sections/srms/Proceedings/papers/1999_136.pdf, accessed July 2014)

Appendix – Statistical Estimation and Overview of the Microsimulation Algorithm

(available from the authors)