PUBLIC HEALTH 175 (2019) 68–78

Available online at www.sciencedirect.com

Public Health

journal homepage: www.elsevier.com/puhe

Original Research

A robust health equity metric

F.J. Zimmerman

Department of Health Policy and Management, Jonathan and Karin Fielding School of Public Health, UCLA, 650 Charles E Young Drive South, Los Angeles, CA 90095-1772, USA

ARTICLE INFO

Article history:
Received 1 November 2018
Received in revised form
5 June 2019
Accepted 19 June 2019
Available online 9 August 2019

Keywords:
Health equity
Health measurement
Health disparities

ABSTRACT

Objectives: Progress on health equity will require a robust metric. The aim of this article is to propose a new health equity metric that is distinct from existing measures and that allows meaningful comparisons across time and place, is calculable using health data typically available, and measures health equity across all major forms of social exclusion.

Study design: A cross-sectional study.

Methods: The new health equity measure was calculated using data included from all 50 states and the District of Columbia in the 2017 Behavioral Risk Factor Surveillance Survey, collected by the US Centers for Disease Control and Prevention. The total sample size was 287,602. State-specific sample sizes ranged from 2269 (Alaska) to 14,685 (Kansas) with a median of 4452. A Healthy Days measure was calculated as the mean number of days that the respondents reported being physically healthy and mentally healthy out of the previous 30 days. The proposed measure defines individual health disutility as the distastefulness associated with one's health falling short of optimal achievable health, instrumentalized as the median health of the most socially privileged category, that of upper-income white men. The value of the health equity metric in a population is the mean value of this distastefulness over the entire population and has a theoretical range of $-\infty$ to 1.

Results: There is substantial variation across states (mean: 0.13; standard deviation: 0.15), with the District of Columbia (0.48), Minnesota (0.37), and Connecticut (0.30) showing the greatest health equity, and West Virginia (−0.26), Arkansas (−0.18), and Kentucky (−0.13) exhibiting the least. Across states, the value of the health equity metric is not correlated with the size of black-white health disparities.

Conclusions: It is feasible to use a single health equity metric for consistent and objective measurement of health equity. Doing so may facilitate more rapid progress toward health equity.

Introduction

Health equity is widely proclaimed to be an important goal of public health policy and practice. Yet, while there is an emerging consensus on what health equity means, the measurement of health equity is fractious, with no one measure emerging as a satisfactory empirical correlate of the central definition. Often, health equity is proxied by health disparities, which are a distinct concept. Improving health equity

E-mail address: fredzimmerman@ucla.edu.
https://doi.org/10.1016/j.puhe.2019.06.008

will require accountability, which will in turn require a strong empirical measure of health equity because comparisons across time and place are essential to finding effective policies for improving health equity.[1]

Concepts of health equity

Social disadvantages are consequential for health. Such disadvantages include the social marginalization that comes with low income, minority racial/ethnic status, female gender, as well as certain religions, immigration status, geographic location, sexual orientation, gender presentation, and so on. Reducing the health risks associated with these forms of social marginalization has become a major goal for public health.

To assist this goal, measurement of health equity must clear and be objective so that researchers can identify what interventions are associated with better or worse levels of health equity and so that policy-makers can gauge whether health equity in their jurisdictions is improving or declining over time and whether they are performing well relative to their peers. Without this kind of concrete measurement, there is a risk that successful interventions will not be recognized and that policy-makers will be unable to see clearly what works and what does not. To fulfill the role of a useful metric for policy and research, a health equity measure must accordingly have three attributes: it must allow meaningful comparisons across time and space, even when the specifics of social marginalization differ; it must be calculable using health data commonly available; and it must measure health equity across all major forms of social marginalization.

After an extensive process and building on years of previous work,[1–8] the following definition of health equity was developed:[9]

> *Health equity means that everyone has a fair and just opportunity to be as healthy as possible. For the purposes of measurement, health equity means reducing and ultimately eliminating disparities in health and its determinants that adversely affect excluded or marginalized groups.*

This definition makes clear that health equity involves minimizing adverse health outcomes that disproportionately affect socially marginalized groups or as they will be termed here, societal categories.[10]

Existing measures of health equity

A debate has existed contesting whether health equity measures should be defined over societal categories or over individuals without regard to societal categories.[3] Those who argue for measures oriented around societal categories argue that the very concept of equity necessarily implies social marginalization, brought about precisely because of people's social identities.[7,8] They argue for measures that specify *a priori* particular societal categories and measure the disparities in health outcomes between these societal categories. Common examples include black and white or two or more income categories. On the other side, those who have favored measures of variation across all individuals point out that the reliance of the disparities measures on *a priori* judgments about what kinds of social exclusion matter conflate empirical observations of health variation with theoretical hypotheses about the causes of this variation. Moreover, they say, measures of variation across individuals without regard to societal categories place more limited demands on the available data and therefore facilitate meaningful comparisons of health equity across time and place.[1,6]

These data limitations are real and can be severe. For example, an emerging literature has identified the role of intersectionality in the social determinants for health.[11,12] The experiences of gay African Americans and straight African Americans may in fact be quite different; yet, such distinctions are not captured by societal category–based disparities measures. One could address this problem by identifying a larger number of societal categories, each with a more internally homogenous experience, but this approach would create very small category sizes and make reliable estimation impossible. In this sense, the promise of disparities-based equity measures is never quite achieved—they capture certain forms of social disadvantage while leaving others unexplored.

This article presents a metric of health equity that is distinct from both the disparities approach and the individual-variation approach, while also recognizing and incorporating the advantages of each. This measure meets the criteria listed above for a health equity metric that is useful in both policy and research: comparability across time and space; feasibility with common health data; and completeness in incorporating of all major forms of social marginalization.

Methods

This study has two parts. In the first part, a new measure of health equity is presented, and its properties briefly described. The second part applies this measure to 2017 data from the Behavioral Risk Factor Surveillance System to show the variation in health equity performance across US states. This second part illustrates how the measure meets the criteria of comparability, feasibility, and completeness described above.

Part I: a new conceptual approach to health equity measurement

The new approach begins with the above definition of health equity as being about social exclusion—however defined.[3,9,10] The key insight implied by this definition is that while the particularities of different types of social exclusion may differ, the moral urgency to redress them does not. This might be called the solidarity principle that—whether by income, education, race/ethnicity, immigration status, sexual orientation, religion, gender, or any other social attribute—societal categories should not translate into systematic differences in health outcomes.

The converse is that there are some privileged individuals who belong to a societal category that is not subject to

substantial social exclusion. How narrowly should this societal category be defined? After all, at some point every individual has had the experience of feeling like an outsider. Yet, their social exclusion matters to health only when it is sustained, substantial, and unbuffered by other advantages that typically come with higher income, such as good access to health care and to health information.[13] Defining such a non-marginalized societal category does not imply that individuals in this category never experience any disadvantages, nor is it to ignore other dimensions of systemic exclusion that may affect them. It says only that whatever disadvantages they experience are not sustained, substantial, or unbuffered enough to systematically worsen their health.

One recent study, for example, finds that higher-income gay men are substantially less likely to experience discrimination on the basis of sexual orientation than lower-income gay men.[14] Higher-income men were also less likely to report symptoms of depression or anxiety than lower-income men and their high income significantly moderated the link between perceived discrimination and psychological distress.

Comparing outcomes to the single most-privileged societal category is similar to the approach taken by the US Centers for Disease Control and Prevention in tracking progress toward Healthy People goals, which defines the concept 'as healthy as possible' as the outcomes for the best-achieving societal category, whatever that category may be for a given outcome.[15]

The reason to instead define a privileged category by its social attributes (e.g. the class of wealthier white men) and to use this category's health as the indicator of best possible health across all outcomes is to make it clear that optimal health is being achieved by a particular, socially defined category. As Braveman has noted, 'The health of the most advantaged social stratum indicates a minimum level that should be biologically possible for everyone'.[3]

The first feature of the proposed health equity metric is therefore to define two societal categories: the privileged and those who suffer at least some sustained, substantial, and unbuffered exclusion because of their social identity. The definition of two societal categories comes with an important advantage. As long as the privileged societal category is defined to include a sufficiently large number of people to permit statistical inference, the two-category approach avoids problems of sparse data.

The second feature of the new measure involves a calculation of health disutility, in which shortfalls from the best possible health matter more when they are larger and by an amount that is more than proportionate to the size of the shortfall. For example, the health-equity harm of having a life expectancy 10 years below that of the privileged category would be more than double the health-equity harm of a life expectancy 5 years below that of the privileged category. This is a common assumption in economics that a given increment to life circumstances (typically income or consumption in economics) has a greater impact on the well-being (or utility) of the poor than for the rich.

This utility feature makes the health equity measure sensitive to within-category inequality; in two societies in which the disadvantaged category has the same average health, the society with more inequality within the disadvantaged category will have higher disutility—that is, less health equity because of the people with very poor health outcomes. The advantage of this approach is that it recognizes the importance of multiple, overlapping forms of social exclusion and reduces the health equity score when these overlapping social exclusions result in severe health shortfalls for some.

In the debate between those who insist on societal category-based measures of health equity and those who prefer the clarity and comparability of non-societal category-based measures, the health equity metric proposed here takes a little from both sides. Like the societal category-based measures, this metric is sensitive to the processes of social exclusion that separate the health outcomes of the privileged from those who are not privileged. Like the non-societal category-based measures, this measure places more-than-proportionate emphasis on the experiences of the least well off, and—as long as adequate data are available for the most-privileged societal category—it is comparable across time and place. In effect, this measure follows Braveman in recognizing that 'Virtually everywhere in the world, social position varies according to economic resources, power or control, and prestige or social standing,' and assumes that these advantages will be clear in the data, even if the breadth of social processes that create disadvantage are unmeasured.

Formal definition of the utility-based approach to health equity measurement

The assumption here is that in the US, wealthy white men make up the privileged societal category. Although women sometimes have better health outcomes, in most domains it is men who have the social privilege, and therefore it is their experience that is normatively important to this analysis.[2,3]

$$\mathrm{HEM} = 1 - \frac{1}{N}\sum_{i=1}^{N}\left(\beta.max\left\{\frac{\bar{y}^{*} - y_i}{\bar{y}^{*}}, 0\right\}\right)^{\alpha}$$

A health equity metric (HEM) compares the experience of individuals within societal categories to the median experience of this privileged category. Assume that there is a continuous measure of health, y, for a large number of individuals in a population. The Health Equity Metric is defined as: where N is the total number of individuals in the sample, y_i is an individual's health, y^* is the median health in the most-privileged category, and β and α are parameters, with $\alpha > 2$ and $\beta > 0$. The value of HEM ranges from $-\infty$ (maximum inequality) to 1 (perfect equality), with higher values indicating greater health equity. Appendix 1 elaborates the derivation of this measure.

This measure is scaled so that greater values imply greater health equity. The value of 0 has a special significance: an individual with half of the health of the median of the most-privileged category would have a value on the health equity metric of 0. This interpretation does not of course directly translate to population values, but it provides a useful heuristic anchor.

This formulation of the metric shows how it meets the criteria articulated above for a health metric to be useful in research and policy: those of comparability, feasibility, and completeness.

Comparability means that the measure should be useful in making meaningful comparisons across time and space. The major constraint for this criterion is that the social meaning of categories is different in different times or places. To be a high-school dropout or an Asian American in California has a different social meaning and therefore different health implications in 2018 than in 1942.[16,17] The health equity metric here meets this criterion because only the most privileged societal category is explicitly identified. And while the identities of socially marginalized groups have changed over time, as have the ways in which marginalization translates into health outcomes, the privileged status of upper-income white men has been remarkably stable for decades.

The precision and consistency of this health equity metric is limited by the overall sample size and by the sample size of the most-privileged category only. Unlike the other health equity measures, this metric does not require large samples for all the subcategories. Avoiding the curse of dimensionality is a considerable advantage, even in large data sets and makes this health measure much more feasible to estimate for many jurisdictions than the disparities measures currently in use.

Finally, and for similar reasons, the measure here is comprehensive in that it incorporates the health experiences of all those who are not members of the privileged group as part of the health equity measure. Because it measures and places a premium on equality within the non-privileged group, it recognizes the possibility of disparate experiences within that group, including in ways that are not theorized *a priori*. For example, differences within the African American experience that would be missed by a black-white disparities measure are incorporated in this measure. In this sense, the measure has some of the advantages of the individual health variation measures discussed in the introduction.

In addition to these functional criteria, a strong metric should have certain mathematical properties. The Box and Fig. 1 discuss the desirable mathematical properties of any health equity index and illustrate them with examples. The health equity metric shown here has all of the properties.

Box 1
Properties of health equity measures

To permit meaningful comparisons across time and space, a health equity metric should have certain properties, which can be more easily understood in reference to Fig. 1. In this Figure, several different distributions are displayed, with different levels of health for different groups displayed as box-and-whisker plots, with width proportional to the group's share of the population.

A health equity metric is higher for each of the distributions than it is for the baseline distribution if the metric displays the relevant property, including:

- **Inequality sensitivity.** The health equity metric increases as disparity in the population declines. (Because this property is of course the most fundamental and obvious one, it is not represented in Fig. 1.)
- **Solidarity.** The health equity metric increases when health becomes more equally distributed across two or more disadvantaged groups. Equity metrics that display this property will have higher values for Distribution 2 than for Distribution 1 in Fig. 1.
- **Subpopulation inequality sensitivity.** The health equity metric increases as disparity within any subpopulation declines. Equity metrics that display this property will have higher values for Distribution three than for Distribution 1.
- **Vulnerability sensitivity.** When there is health inequality within two disadvantaged groups (e.g. groups A and D in Distributions 4a and 4b of Fig. 1), health equity should be greater when the health inequality is smaller in the more disadvantaged group (group A) than when it is smaller in the less disadvantaged group (group D).
- **Health sensitivity.** Holding some reference value constant, this property implies that health equity increases when all groups see their health improve. Compare Distribution 5 to Distribution 1. Note that without this property, a health equity measure is a disparity measure.
- **Group-size sensitivity.** With this property, a health equity metric increases when a smaller proportion of the population is in a disadvantaged group. See Distribution 6.

The health equity measure proposed here displays all of these properties. Most existing measures display only a subset of them. See Appendix 2 for details.

Part II: Application of the new health equity metric to data from Behavioral Risk Factor Surveillance System

This section applies the health equity metric to data from the Behavioral Risk Factor Surveillance System (BRFSS) from 2017. The health outcome used is the measure of healthy days of the Centers for Disease Control and Prevention's Health-Related Quality of Life Scale.[18,19] The scale is constructed by summing the answers to two questions about how many days in the previous 30, the respondent felt that their mental or physical health was not good. This value is truncated at 30 and subtracted from 30 to create a measure of healthy days. This measure has psychometric properties.[18]

Data from all 50 states and the District of Columbia are included, for an initial sample size of 287,816 adults aged 18–64 years. Of these, 214 observations were dropped because of missing data for a final sample size of 287,602. State-specific sample sizes ranged from 2269 (Alaska) to 14,685 (Kansas) with a median of 4452.

For each state the median number of healthy days was calculated for White men with incomes in the highest category: $75,000 annually or more.

The Health Equity Metric was calculated according to the equation on page (8) with $\alpha = 2.5$ and $\beta = 2$. Appropriate survey weights were applied in making these calculations,[20] which were also age-adjusted using the following age strata: 18–29; 30–39; 40–49; 50–56; and 57–64 years. These strata were

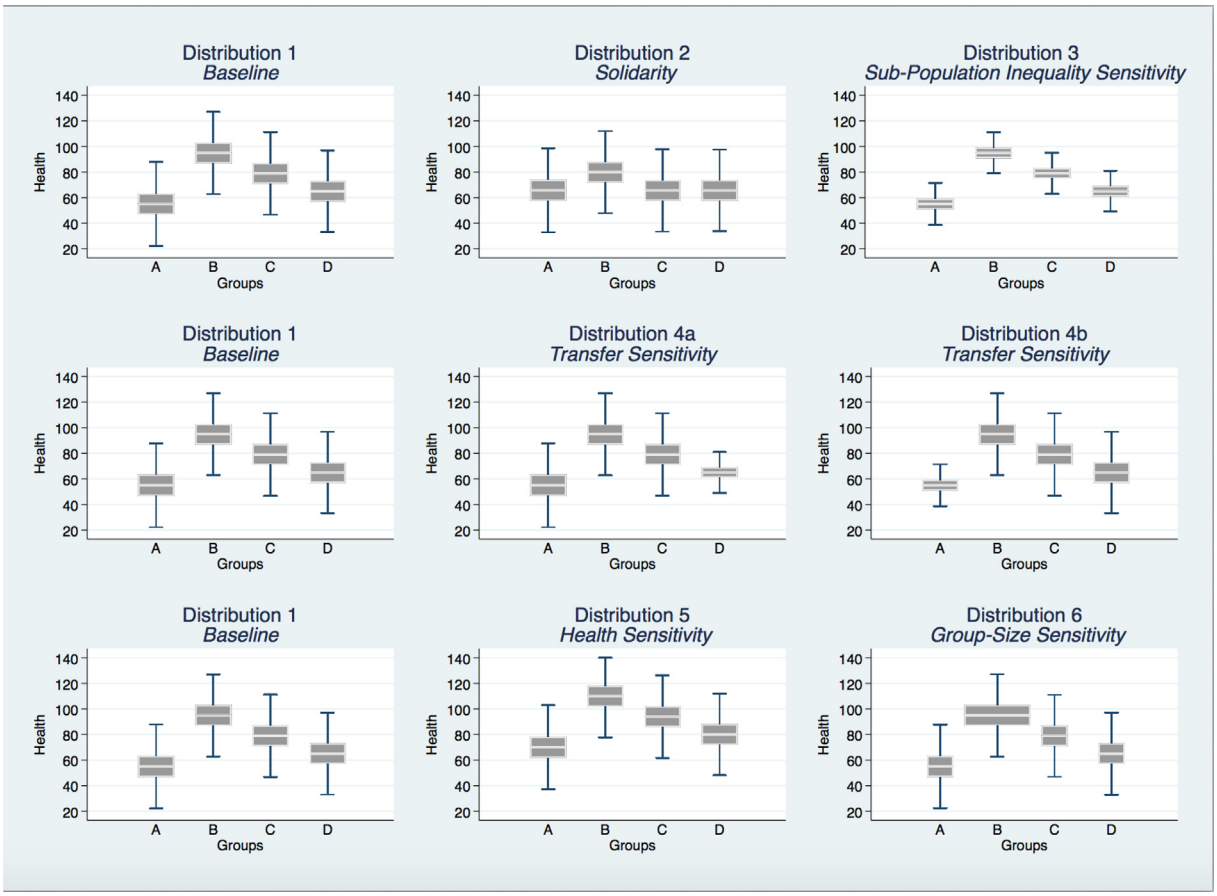


Fig. 1 – Hypothetical distributions of health outcomes across four social groups (A, B, C, D).

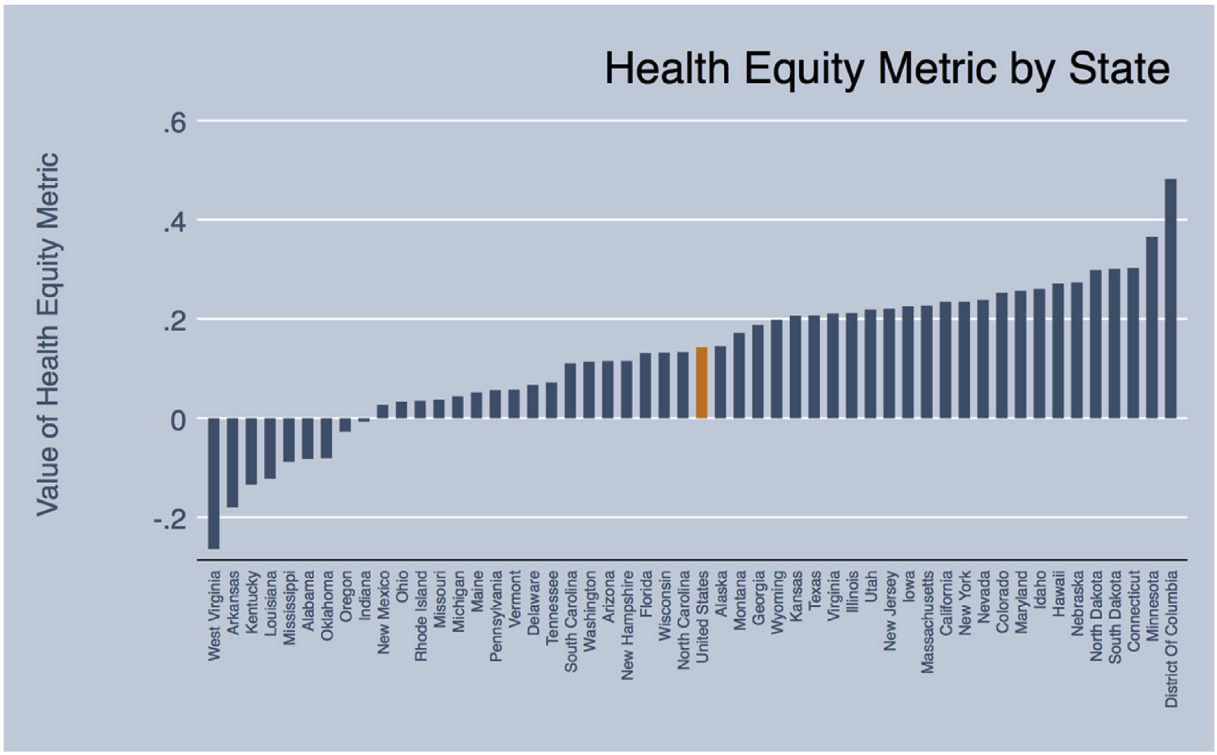


Fig. 2 – Health equity metric values for each state.

chosen to be large enough for adequate state-specific sample size within age strata (and racial/ethnic and income categories) and small enough for within-strata homogeneity of outcomes.

Results

There is substantial variation across states, with a mean value of 0.13 and a standard deviation of 0.15. Fig. 2 shows the values of the Health Equity Metric for each state and for the US as a whole. The District of Columbia, Minnesota, and Connecticut take the top three slots, and West Virginia, Arkansas, and Kentucky exhibit the least health equity. The US average is 0.14.

Health equity and health disparities

Health equity and health disparities are distinct concepts. Empirically, too, at least in these data, they are quite different. As an example, Fig. 3 presents a scatterplot of state scores on the Health Equity Metric against a common measure of health disparities: the gap between white average healthy days and black average healthy days. As can be seen, there is little correlation between these measures; the correlation coefficient is 0.04 (*P*-value 0.79). An example is illustrative. Both Minnesota and Kentucky exhibit almost no racial health disparities, yet Minnesota scores among the highest states on the health equity metric while Kentucky is among the lowest. One reason (cf. Appendix 3) is that both black and white middle-income residents fare much better in Minnesota than in Kentucky.

This difference is consequential because so much of the measurement of health equity to date has been to use health disparities as a proxy for health equity. The sharp divergence of health disparities and health equity shown here suggests that in fact racial health disparities are a poor proxy for health equity.

The figure also highlights the small-numbers problem that arise in health-disparities measures. In Alaska, whites are having an average of four more healthy days than the black average. Yet, although Alaska has 2269 observations in total, the African-American experience is represented by only 39 observations. Of course, the experience of Latinx residents is represented by even fewer respondents and therefore cannot be compared across states in these data.

It is possible to disaggregate the health equity metric to identify particular disparities across specific categories that are driving these differences. Appendix 3 shows many other measures of disparities. While the analysis of particular disparities raises interesting questions, it can suffer from the data problems of small group sizes described above.

Discussion

The health equity metric proposed here splits the difference between those who insist on measures that explicitly involve differences in specified societal categories and those who advocate for measures that track variation at the individual level only. It defines only the single privileged category and not requiring separate measurement—and the attendant

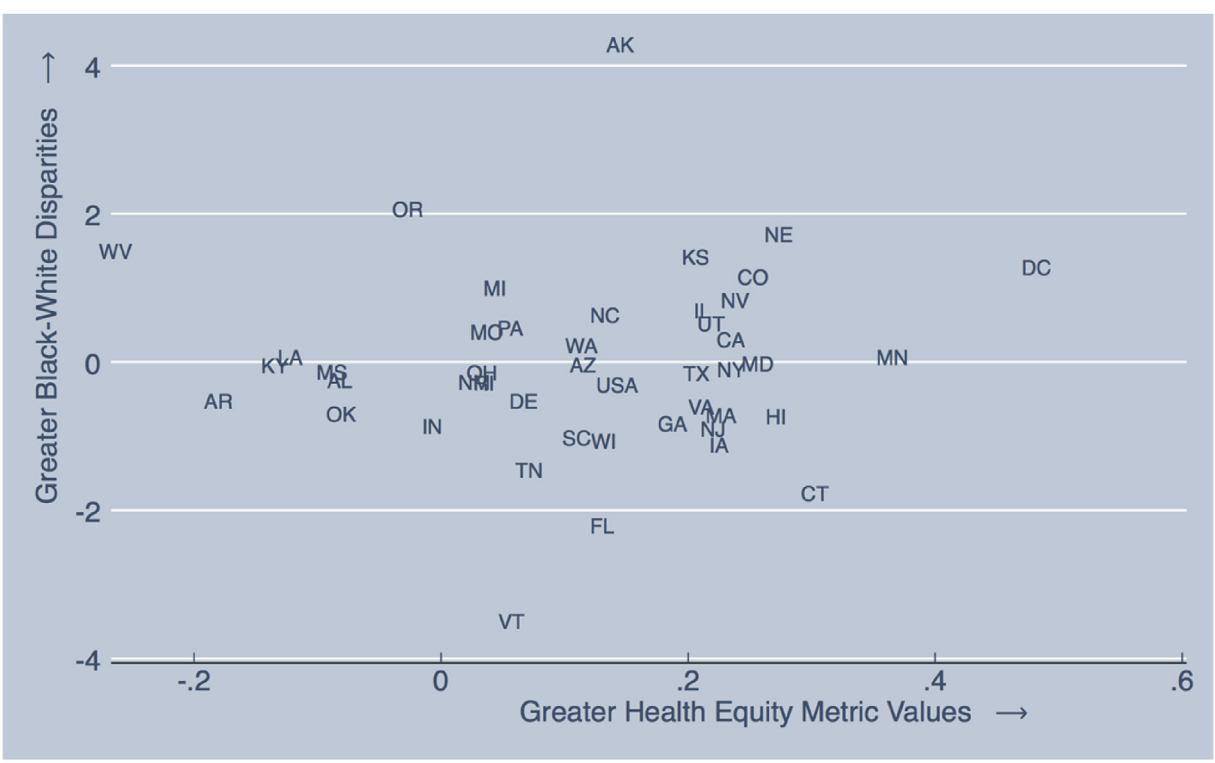


Fig. 3 – State comparison of health equity (health equity metric) and health disparities (white average healthy days minus black average healthy days), with the state abbreviations positioned at the state values.

large sample sizes—for each of the various disadvantaged categories in society. To the extent that the most privileged category is relatively stable—typically male and always those with the highest income levels—this measure facilitates meaningful comparisons of health equity over time and across place even as the social patterning of marginalization evolves.

The health equity metric has a particular advantage in that—unlike measures of health disparities—it can be calculated without deep data on every societal category. The analyses of BRFSS data here provide a compelling example of the problem. Defining societal categories by the interaction of gender, race/ethnicity, and three income categories, there were 30 states with fewer than 10 observations per societal category; only six states had at least 30 observations for each societal category. Even with over 1500 observations from each state, the curse of dimensionality is acute. Yet, without accurate estimates of societal-category averages, measures of health equity based on societal-category averages lose their usefulness. The health equity metric does not have this problem because only the sample size of the most privileged category matters. In these data, the state with the smallest cell size for upper-income White men had 193 observations, and the average across all states was 671.

The health equity metric is also distinguished by its focus on the most vulnerable. That focus reflects a particular value judgment[21] that is consistent with the mission of public health.

The limitations of this health equity metric include those common to all such measures—including a reliance on accurate health and socio-economic data. In addition, the health equity metric has its own limitations, chiefly that its scale is not in units that are easily understandable to policy-makers or the public. In addition, as the underlying health data become less normally distributed, the measure here tends to track average health and also the coefficient of variation. In the extreme case of binary health data, average health, the coefficient of variation of health, and the Health Equity Metric are all mathematical transformations of one another unless there are extremely large differences between the most privileged and other societal categories.

The results presented here for state performance on the health equity metric are intended as an example of its implementation, not as a thorough analysis of health equity in states. The underlying health measure—number of healthy days in the past 30—is not only defensible[18] but also limited. It is noteworthy that the metric's ranking is similar to other, similar efforts to measure health equity in states. For example, a recent report of education-related disparities in disability across US states[22] finds a similar pattern as that here, and the Spearman rank correlation coefficient between these two rankings is 0.60 ($P < 0.001$). These results, while preliminary, are consistent with existing research that suggests that state policies matter for population health and particularly so for the most vulnerable.[23]

Conclusion

The health equity metric proposed here is reliably measurable in samples of moderate size and is consistent over time and in different political jurisdictions. Regular use of a single health equity measure such as this one will enable public health advocates and researchers to track performance on health equity.

Author statements

Ethical approval

This study was deemed exempt from review by the UCLA Institutional Review Board.

Funding

The support of the Robert Wood Johnson Foundation is gratefully acknowledged. A brief residency at the Helen C. Whitely Center at the University of Washington is also gratefully acknowledged.

Competing interest

None declared.

REFERENCES

1. Murray CJ, Gakidou EE, Frenk J. Health inequalities and social group differences: what should we measure? *Bull World Health Organ* 1999;**77**(7):537.
2. Braveman P, Gruskin S. Defining equity in health. *J Epidemiol Community Health* 2003;**57**(4):254–8.
3. Braveman P. Health disparities and health equity: concepts and measurement. *Annu Rev Public Health* 2006;**27**(1):167–94.
4. Whitehead M. The concepts and principles of equity and health. *Int J Health Serv* 1992;**22**(3):429–45.
5. Mackenbach JP, Kunst AE. Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe. *Soc Sci Med* 1997;**44**(6):757–71.
6. Gakidou EE, Murray CJ, Frenk J. Defining and measuring health inequality: an approach based on the distribution of health expectancy. *Bull World Health Organ* 2000;**78**:42–54.
7. Szwarcwald CL. On the World Health Organisation's measurement of health inequalities. *J Epidemiol Community Health* 2002;**56**(3):177–82.
8. Anand S. The concern for equity in health. *J Epidemiol Community Health* 2002;**56**(7):485–7.
9. Braveman P, Arkin E, Orleans T, Proctor D, Acker J, Plough A. What is health equity? *Behav Sci Policy* 2018;**4**(1):1–14.
10. Beckfield J. *Political sociology and the people's health*. New York: Oxford University Press; 2018.
11. Smedley BD, Myers HF. Conceptual and methodological challenges for health disparities research and their policy implications. *J Soc Issues* 2014;**70**(2):382–91.
12. Bowleg L. The problem with the phrase women and minorities: intersectionality—an important theoretical framework for public health. *Am J Pub Helath* 2012;**102**(7):1267–73.
13. McGarrity LA. Socioeconomic status as context for minority stress and health disparities among lesbian, gay, and bisexual

individuals. *Psychol Sex Orientation Gender Divers* 2014;**1**(4):383–97.
14. Gamarel KE, Reisner SL, Parsons JT, Golub SA. Association between socioeconomic position discrimination and psychological distress: findings from a community-based sample of gay and bisexual men in New York City. *Am J Pub Helath* 2012;**102**(11):2094–101.
15. Keppel KG, Pearcy JN, Klein RJ. Measuring progress in healthy people 2010. *Healthy People 2010 Stat Notes* 2004;(25):1–16.
16. Dowd JB, Hamoudi A. Is life expectancy really falling for groups of low socio-economic status? Lagged selection bias and artefactual trends in mortality. *Int J Epidemiol* 2014;**43**(4):983–8.
17. Ryan CL, Bauman K. *Educational attainment in the United States: 2015*. Washington, DC: US Census Bureau; March 2016.
18. Hagerty MR, Cummins RA, Ferriss AL, et al. Quality of life indexes for national policy: review and agenda for research. *Soc Indicat Res* 2001;**55**(1):1–96.
19. Hennessy CH, Moriarty DG, Zack MM, Scherr PA, Brackbill R. Measuring health-related quality of life for public health surveillance. *Publ Health Rep* 1994;**109**(5):665.
20. Centers for Disease Control and Prevention. *Behavioral risk factor surveillance System 2015 codebook report*. 2016.
21. Harper S, Lynch J. *Health inequalities: measurement and decomposition*. 2016.
22. Montez JK, Zajacova A, Hayward MD. Disparities in disability by educational attainment across US states. *Am J Pub Helath* 2017;**107**(7):1101–8.
23. Montez JK, Hayward MD, Wolf DA. Do U.S. states' socioeconomic and policy contexts shape adult disability? *Soc Sci Med* 2017;**178**:115–26.

Appendix 1. Development of the health equity metric

Assume that there is a continuous measure of health, y, for a large number of individuals in a population and suppose that there are J social groups within this population. The social groups might be defined by education, race/ethnicity, gender, and so on and assume that the J groups are mutually exclusive and collectively exhaustive. We start by defining a health deficit, h_{ij}^{-}, for each individual i in each group j as the amount by which his or her health, y_{ij}, falls below the average health in the most privileged group, y_{j*}.

$$h_{ij}^{-} = \max\left\{\frac{\bar{y}_{j*} - y_{ij}}{\bar{y}_{j*}}, 0\right\} \tag{1}$$

Simply falling below the average health of the most privileged group is distasteful; falling far below that average is considerably more distasteful. A standard assumption (for example in poverty studies[1] is that greater shortfalls are more-than-proportionately painful than modest shortfalls. We might dislike a cold or flu, but we abhor larger health problems. This non-linearity is a standard assumption in economics and is reflected in at least some health phenomena. For example, the odds of disability, poor physical function, and mortality all increase more than proportionally to BMI above the healthy range.[2–5]

Accordingly, with the parameter α taking on values greater than 1, the individual's distaste of poor health—or disutility, in the language of economics—can be represented as:

$$d_{ij} = \left(\beta h_{ij}^{-}\right)^{\alpha} = \left(\beta \cdot max\left\{\frac{\bar{y}_{j*} - y_{ij}}{\bar{y}_{j*}}, 0\right\}\right)^{\alpha} \tag{2}$$

At the population level, the health equity metric is then one minus the average of a non-linear transformation of this health deficit across all groups:

$$\text{HEM} = 1 - \sum_{j=1}^{J}\left(\frac{1}{N_j}\sum_{i=1}^{N_j} d_{ij}\right) \tag{3}$$

Here N_j is the population size of the j'th group, and α is a parameter greater than 1. With $\alpha > 1$ HEM has inequality sensitivity, and with $\alpha > 2$ HEM has vulnerability sensitivity. (See Appendix 2). The value of HEM ranges from $-\infty$ (maximum inequality) to 1 (perfect equality).

When β has a value of 2, the metric has a special interpretation. In this case, when an individual's health deficit is 0.5—half of the value as the average for the privileged group for whatever health measure is being considered—then the individual's contribution to the HEM is 0. Accordingly, 0 serves as a rough anchor for interpreting the metric, bearing in mind that because the function is non-linear, the average value is not the same as the value for the average.

This health equity metric can be rearranged as a sum over the entire population:

$$\text{HEM} = 1 - \frac{1}{N}\sum_{i=1}^{N} d_i \tag{4}$$

It may seem odd that the expression in Eq. (4) seems not to explicitly reference social groups. But of course, the health outcomes of the most privileged social groups are concretely embedded in d_i: when the most privileged group does better, then holding the health of all other social groups constant, the health equity metric falls, because they are excluded from the (presumably feasible) health gains of the most privileged.

This HEM has similarities to entropy indices and to variance-based measures, but unlike those other measures, this health equity metric has all of the properties enumerated in Appendix 2, including solidarity, subpopulation inequality sensitivity and vulnerability sensitivity.

A health equity metric should be distinct from a measure of health disparities, in that health equity should encompass the full array of social exclusion in a population. Health disparities, by contrast, focus on differences in health outcomes across specific groups defined by the researcher. Health equity and health disparities tend to be most similar when the number of groups defined by the researcher is large.

The health equity metric proposed here can be decomposed into a measure of health disparities, simply by calculating the expected disutility within each of any number of researcher-defined groups and comparing these group averages.

Appendix References

1 . Foster J, Greer J, Thorbecke E. A class of decomposable poverty measures. Econometrica. 1984;52(3):761-766.
2 . Berrington de Gonzalez A, Hartge P, Cerhan JR, et al. Body-Mass index and mortality among 1.46 Million white adults. New England Journal of Medicine. 2010;363(23):2211-2219.
3 . Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. The Lancet. 2009;373(9669):1083-1096.
4 . Waidmann T, Freedman V, Himes C, Ahmad S. Examining The Relationships Between Excess Body Weight, Health And Disability. Office of Disability, Aging and Long-Term Care Policy U.S. Department of Health and Human Services 2008.
5 . Slagter SN, van Vliet-Ostaptchouk JV, van Beek AP, et al. Health-related quality of life in relation to obesity grade, type 2 diabetes, metabolic syndrome and inflammation. PloS one. 2015;10(10):e0140599.

Appendix 2. Properties of Health Equity Measures

Referring to the Box in the main text and Fig. 3, this Appendix briefly describes the extent to which existing measures of health equity satisfy the various properties described there. Because these properties are not relevant for non-group–based measures, these are not described here. The discussion does make a distinction between group-based measures based on a single attribute—such as income, education or occupational status—and those group-based measures that incorporate several categorical variables. Most of these attributes—excepting inequality sensitivity—are not defined for single-attribute, group-based measures.

Most existing group-based measures of health equity have inequality sensitivity and group-size sensitivity, but not all of the other attributes. For example, the absolute mean difference across groups has none of the other properties; the between-group variance measure exhibits solidarity and entropy-based measures such as the Rényi or Atkinson indices exhibit health sensitivity and solidarity.

None of the existing group-based measures exhibit subpopulation inequality sensitivity nor vulnerability sensitivity. These properties capture individual experiences of health that are influenced by social position but not reflected in group averages. For example, suppose there are two distinct social policies that benefit members of a racial/ethnic minority, but one targets the most well-off among this group (e.g. postdoctoral opportunities at elite universities) and the other targets more typical members (e.g. college scholarships). Supposing for the sake of exposition that both raise average outcomes for this racial/ethnic group by the same amount, one improves the group average by distending the distribution, raising the top end, while the other improves the group average by compressing the distribution, raising those below the average. It would be useful for a health equity metric to distinguish between these two kinds of effects, yet none of the existing group-based measures does. The single-attribute measures are sensitive to inequality throughout the distribution, so in a sense including among the disadvantaged, but only to the extent this inequality is correlated with the single metric.

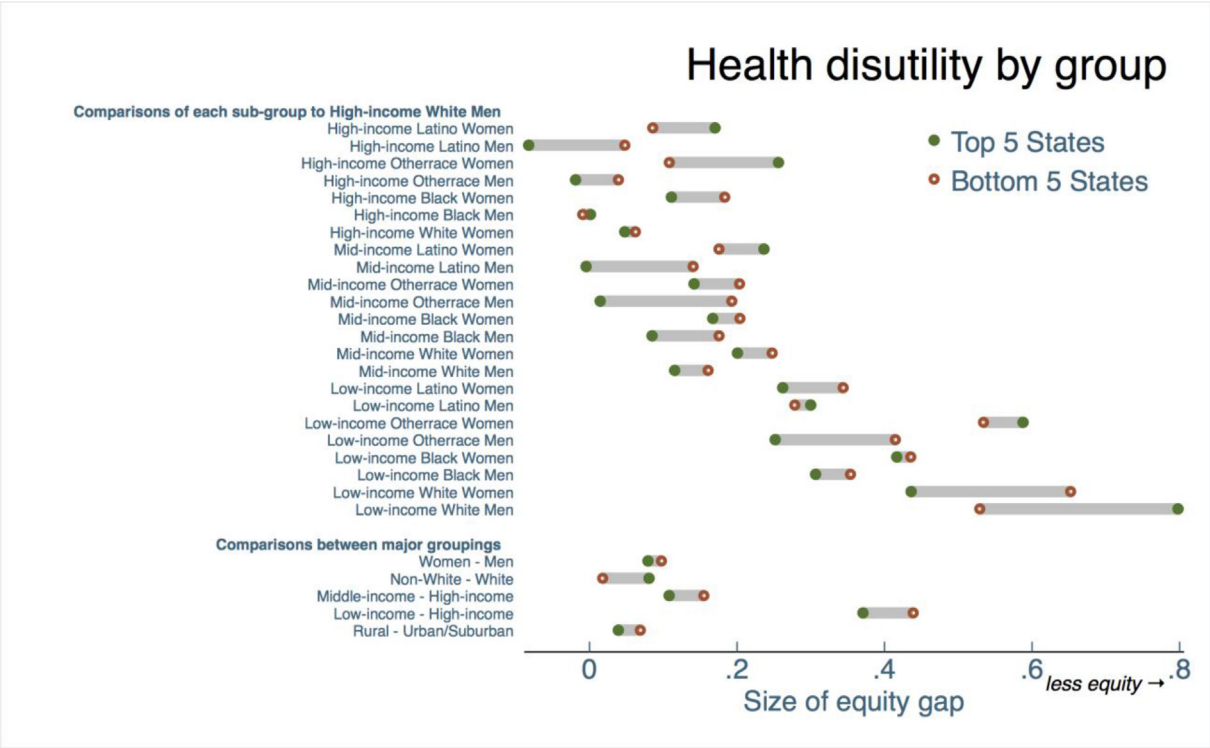
Vulnerability sensitivity (also called transfer sensitivity) takes the concern with the well-being of the least well-off within groups further. Suppose there are two interventions (e.g. college scholarships and nurse home-visiting) that both raise group-average health by the same amount and also compress within-group inequality by the same amount. The intervention (nurse home-visiting in this example) that achieves the reduction in within-group inequality by bringing up the most vulnerable as opposed to those who are less vulnerable would achieve higher health equity under this property. To the extent that one is concerned not just with social-group averages but in particular with the social processes that create particular vulnerability for the least well-off within social groups, vulnerability sensitivity is a useful property for a health equity measure.

It is tempting to ask how much impairment to health equity is caused by each social attribute independently: income, race/ethnicity, gender, et cetera. However, because health equity is defined only at a population level, this question has no answer. For most of the health equity measures in the literature it is possible to identify the contribution of each subgroup to health equity (e.g. low-income Latina men) against some benchmark (e.g. high-income White men). Because of the importance of intersectionality, it is not possible to calculate how much of this contribution is attributable to income as opposed to ethnicity. Separate comparisons by income or by ethnicity are of course possible, but these comparisons are descriptions of the population, not regressions in which it is possible to control for other variables to obtain an independent effect.

Measures that can be decomposed by political or geographic subjurisdictions will be more useful than those that cannot. It would be useful for researchers, advocates, and policy-makers to know, for example, whether their health equity is largely driven by rural or urban areas of a state, or is greater in certain Congressional districts, and so on. In principle, any of the health equity measures can be decomposed geographically, but doing so can invoke the curse of dimensionality.

Appendix 3. Disparities in Health Disutilities

In Appendix Fig. 1, the results reported in the main text are disaggregated into a health disparities measure, showing health disparities across various groups in the five most-equitable and five least-equitable states. Each dot represents as disparity, with dots farther to the right representing larger disparities. In the top panel, the disparity is that between high-education white men and each of the subpopulations labeled in the row headings. In the bottom panel the disparity is that between the two groups indicated in the row heading.



Appendix Fig. 1 – Disparities in health disutility across several groups.

As stated in the main text, some of the group sizes required to estimate these disparities may be very small, so particular disparities should be interpreted with caution.

The best-performing states have smaller expected disutility for most of the subgroups than do the worst-performing states. For example, relative to high-education white men, low-education white women have expected disutility of 0.65 in the five best-performing states, but 1.0 in the five worst-performing states. For some groups, the worst-performing states in fact perform better than the best-performing states. For example, Latino women with some college have lower health disutility in the worse-performing states than in the better-performing states. This difference may be because of small cell sizes, however.

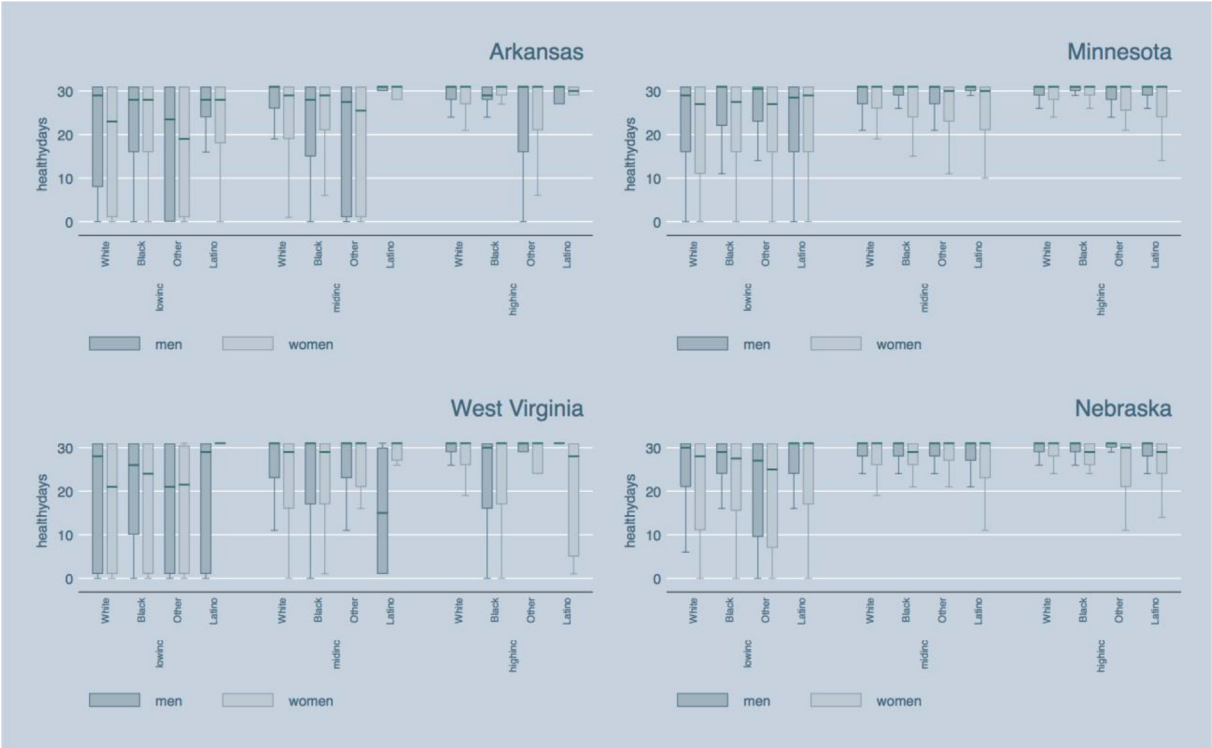
The differences shown in Appendix Fig. 1 suggest opportunities to improve health equity. The distance between dots on any row suggest opportunities for the poorly performing states to catch up to the more equitable states. For example, the relatively large gap between the equitable and non-equitable states for both white women and black women with some college suggests that policies to benefit this group are feasible and consequential. The distance of the dots on any row from the left axis suggests the magnitude of health equity shortfall for each social group. As the comparisons between the larger groupings show, the health equity gap is particularly large for low-income as opposed to high-income groups.

Appendix Fig. 2 provides a different way to explore disparities at the state level. This figure shows the distribution of the number of healthy days as box-and-whisker plots within states. For each state, the distributions are grouped first by gender, then by race/ethnicity, and then by threeincome groups, low, medium, and high. On the left of the Figure are two low-performing states—Arkansas and West Virginia—and on the right are two high-performing states—Minnesota and Nebraska. These states were chosen because they represent opposite ends of the equity spectrum, while being demographically similar.

Several interesting patterns emerge from Appendix Fig. 2. The most obvious difference between the low-performing and high-performing states is the middle-income group, which fares much more poorly in Arkansas and West Virginia than in Minnesota and Nebraska. This difference is more pronounced for the non-white than for white residents.

Gender differences can be observed throughout the income spectrum and across racial/ethnic groups and are especially strong for at low levels of income.

As with Appendix Fig. 1, Appendix Fig. 2 suggests feasible ways of improving health equity. Of course, the most important task would be to improve the health of the lowest income groups. At the same time, the empirical difference between high-performing and low-performing states is that the high-performing states do a somewhat better job of protecting the health of low-income residents and a substantially better job of protecting the health of middle-income residents.



Appendix Fig. 2 – Differences in healthy days in four states by income, race/ethnicity and gender.