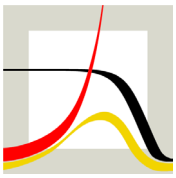

Max-Planck-Institut für demografische Forschung

Max Planck Institute for Demographic Research

Konrad-Zuse-Strasse 1 ▪ D-18057 Rostock ▪ Germany ▪ Tel +49 (0) 3 81 20 81 - 0 ▪ Fax +49 (0) 3 81 20 81 - 202 ▪ www.demogr.mpg.de

MPIDR Working Paper WP 2017-010 I May 2017

An ordinal measure of population health

Héctor Pifarre i Arolas I pifarreiarolas@demogr.mpg.de
Christian Dudel I dudel@demogr.mpg.de

This working paper has been approved for release by: Mikko Myrskylä (sekmyrskyla@demogr.mpg.de), Head of the Laboratory of Population Health and Head of the Laboratory of Fertility and Well-Being.

An ordinal measure of population health

Abstract

We propose a population-level health index that addresses two of the main concerns of existing health measures: i) the biases in self-assessed health measures, and ii) the difficulties associated with making comparisons across populations afflicted by a variety of conditions in a context in which multi-morbidity is high. Starting from a set of general axioms, we derive a partial order index that ranks health across populations based on the relative prevalence of various conditions within these populations, and that can be readily applied to many existing health surveys. We illustrate the use of our health measure by applying it to data from the National Health Interview Survey in order to examine health differences across racial groups in the U.S.

1 Introduction

Comparisons of health across populations are often made in the academic literature, and are frequently cited in political discussions. From cross-country evaluations (König et al., 2009), to studies on the socioeconomic gradient of health (Marmot, 2005), to research on changes in health over time (Anderson et al., 2000); relative health assessments are among the most studied and hotly debated topics across disciplines. However, despite the importance of this issue, there is still an active debate about how best operationalize population-level health comparisons (Murray, 2002; Etches et al., 2006).

In this work we focus on the measurement of morbidity and introduce an approach specifically designed to tackle two of the main challenges faced by health indicators, as discussed by Burgard and Chen (2014). Namely, differences in disease distributions across groups and problems with self-assessed health measures. Health comparisons are often made across populations who are not necessarily afflicted by the same types diseases.This issue of comparability is perhaps most obvious in cross-country comparisons among countries in different stages of the epidemiological transition and in health comparisons over time (Burgard and Chen, 2014), but it is also present in subnational health assessments (Rivera et al., 2002; Weimann et al., 2016). A further potential problem that can arise when using existing measures is that some of the most frequently collected measures of health rely on self-assessments of health, even though a growing body of evidence suggests that there is systematic bias in how respondents rate their health. While self-assessed health measures have been validated by their association with objective health indicators (such as mortality Idler and Benyamini 1997, 1999; Van Doorslaer and Gerdtham 2003; Jürges 2008, or health utilization Humphries and Van Doorslaer 2000), more recent research has found evidence that certain socioeconomic groups suffer from reporting bias (Dowd and Zajacova, 2010; Van Doorslaer et al., 2002; Etilé and Milcent, 2006). Thus, the reliance on self-rated health measures when making group comparisons has been called into question.

Our strategy for addressing these challenges is to develop an analytical framework that is based strictly on relative prevalences. Comparisons based on disease prevalences (and

on other indicators of health) have a long tradition in the public health literature, and comparisons that use this approach remain common Braveman et al. (2010). The appeal of focusing on relative disease prevalences lies in the assumption that these indicators are more objective than self-rated health measures, and are more directly comparable in cross-population studies. Furthermore, as was noted in Burgard and Chen (2014), measures based on disease prevalences are more closely related to the origins of health disparities, and can therefore provide more direct guidance for public health interventions than other methods that aggregate health states in complex ways. We recognize that despite these advantages prevalence-based analyses are not without problems. As was pointed out by Burgard and Chen (2014), diagnosis and reporting bias can lead to an over- or an underestimation of the relative disease burden across socioeconomic groups Johnston and Propper (2009). While we acknowledge the importance of such issues, addressing such data collection concerns is beyond the scope of this work.

Setting aside data related challenges, a difficulty that often arises in comparative health assessments based on prevalences is that no group is shown to be universally healthier than the other groups across all health outcomes. For example, it has been documented that in the U.S. white elderly individuals have higher cancer prevalences than Hispanics and blacks, but that the latter groups are more likely to report having diabetes (National Research Council of the National Academics, 2004). In such a situation, figuring out how to aggregate this information into a single measure without losing the benefits of prevalence-based indicators can be a challenge.

Our approach addresses this technical difficulty. The basic logic behind our proposal which we formalize later on in a set of axioms is as follows. Suppose we want to compare two individuals who might be suffering from cancer and cardiovascular disease. If individual A has both cancer and cardiovascular disease and individual B only has cancer, then, barring differences in the severity of the conditions, we say that individual A is in a worse health state than individual B. However, if individual A has cancer and individual B has cardiovascular disease, we say that they are not comparable. While we are focused on diseases in this example, the idea can also be applied to other prevalence-based measures,

such as limitations in activities of daily living or health behaviors.

Based on this reasoning, which we aggregate to population- level comparisons, we take the following three steps. First, we construct an index of comparability. In a context of multi-morbidity, this indicator captures to what extent the populations can be ranked using our axioms, and is influenced by the overlap in the profiles of conditions affecting the comparison groups. Second, we take the comparability information into account using our main indicator, the partial ordering disparity index (PODI). The PODI is an ordinal measure for population-level pairwise comparisons that ranks groups on the basis of relative prevalences. Third, we show how the PODI can be extended to generate a health ranking in multi-group comparisons. Our aim is to provide an intuitive health index that can be readily implemented using the information commonly collected in health surveys. Given the concerns that have been raised about self-assessed health measures, especially in the presence of multi-morbidity, our health indicators provide conservative estimates of the health disparities between populations that can be used to complement existing analyses. We illustrate the applicability of our analytical framework by drawing on data from the National Health Interview Survey (NHIS) to study racial health disparities. Those wishing to apply this approach in the future can refer to the user-friendly R-code and examples presented in the online appendix of this work.

The paper is organized as follows. In section 2, we introduce our axioms and formalize how they can be used to provide population-level comparisons. Building on the pairwise relations discussed in section 2, we present in section 3 a new health measure, the partial ordering disparity index. We then derive the formal properties of the PODI and illustrate how it can be extended to compare any number of populations. In section 4, we use data from the NHIS in applying the PODI to examine racial differences in health in the U.S. In section 6, we conclude.

2 Definitions and concepts

2.1 Axioms

Within the scope of our work are health comparisons based on data from health surveys, such as the NHIS, the Health and Retirement Study and related surveys, and the World Health Survey. In addition to collecting data on socioeconomic variables, these surveys generally gather some information on the burden of disease at an individual level, such as self-reports on the prevalence of diseases or on limitations in daily life activities. In the analysis that follows, we assume that we have information on the prevalence of relatively common diseases, even though our approach can be implemented for any prevalence-based variable (see section 5). The following axioms provide the foundation for our analysis.

Axiom 2.1. *Individuals who have no condition are healthier than individuals with at least one condition.*

Axiom 2.2. *Individuals with the same combination of conditions are all equally (un)healthy.*

Axiom 2.3. *Individual j is less healthy than individual k if j has the same conditions as k and at least one additional condition.*

We expect that these axioms will provide a reasonable basis for comparisons at the population level, assuming that the diseases and the limitations covered in the health survey data used accurately reflect the conditions that afflict the population in question. We wish to emphasize that our goal is to make comparisons at the population level, as we are well aware that the judgments derived from our axioms might not hold in all individual-level comparisons.

In the discussion, we explore the question of how these axioms can be extended to cover more informed assessments. The basic analytical framework is, however, unchanged by those modifications.

2.2 Partial ordering

Axioms 1 through 3 provide a partial ordering of any two individuals based on pairwise comparisons. We define those comparisons using two relations, denoted *dominance* and *comparability*. Let $\mathcal{D} = \{D_1, \ldots, D_k\}$ be the set of all conditions under study. The power set $\mathcal{C} = \mathcal{P}(\mathcal{D})$ gives the set of all possible combinations of conditions, including the case of no condition as the empty set, $\emptyset$. Elements of $\mathcal{C}$ will be denoted by C_j. Define $\succ$ as the non-reflexive, non-symmetric, and transitive relation of dominance and $\sim$ as the reflexive, symmetric, and transitive relation of comparability.

Definition 2.1 (Comparability)**.** $\forall C_i, C_j \in \mathcal{C} : C_i \sim C_j$ *iff* $C_j \subseteq C_i \vee C_i \subseteq C_j$.

Definition 2.2 (Dominance)**.** $\forall C_i, C_j \in \mathcal{C} : C_i \succ C_j$ *iff* $C_j \subset C_i$.

While these relations are defined in terms of conditions, they can be easily translated into comparisons between individuals defined by condition profiles. The comparability relation captures the idea that individuals with different medical profiles might not be comparable. In turn, among comparable individuals, the dominance relation establishes a health ranking. Coming back to the example in the introduction, we can say based on definition 2.1 that if individual A has cancer and individual B cardiovascular disease, they are not comparable. However, if individual A has both cancer and a cardiovascular disease, then A is comparable to B. Furthermore, since individual A's conditions are a strict subset of individual B's conditions, we can say that, barring differences in the severity of the conditions, individual A's health state dominates that of individual B (definition 2.2).

The relations 2.1 and 2.1 imply a partial ordering of the elements of C which can be represented as a graph. A dominance relation is a directed, acyclic graph $\mathcal{G} = (N, E)$, with nodes N given by $\mathcal{C}$ and edges E given by the elements of $\mathcal{C} \times \mathcal{C}$ for which the conditions stated in definition 2.2 apply. $\mathcal{D}$ is the maximal element of $\mathcal{G}$ and $\emptyset$ is the minimal element of $\mathcal{G}$. In a similar way a non-directed graph $\mathcal{G}'$ can be constructed that represents comparability relations (relation 2.1). An example for $\mathcal{D} = \{A, B, C\}$, given in figure 1, represents the dominance relation. For instance, the node $\{A, B\}$ dominates both nodes $\{A\}$ and $\{B\}$ and also the minimal element $\emptyset$ because of transitivity. In turn,

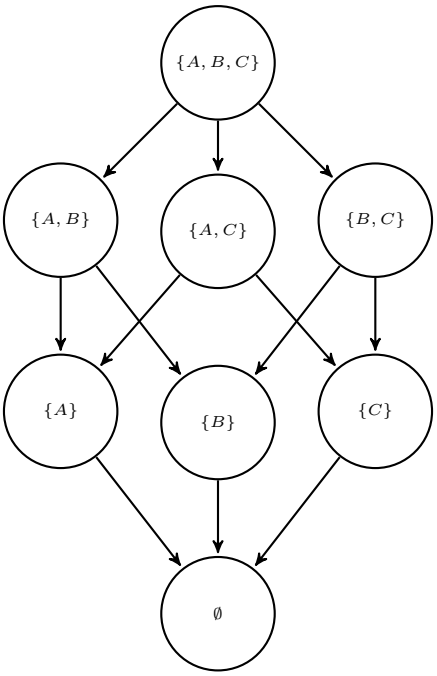

Figure 1: Example of partial ordering with three conditions A, B, and C. Nodes (circles) represent all of the possible combinations of conditions. Edges (arrows) show the dominance relation, e.g., $\{A, B\}$ dominates $\{A\}$ and $\{B\}$ and because of transitivity, also $\emptyset$.

it is dominated by the maximal element of the graph. A graphical representation of the comparability relation would look similar, by directed edges replaced by undirected edges.

2.3 Population level comparisons

Having formally defined health relations at the individual level, the next logical step is to formulate a strategy to aggregate these relations for the purposes of making group comparisons. To that end, we build on the concept of expected dominance, which can be traced back to Lieberson (1976) and has also been applied to health comparisons in recent work by Herrero and Villar (2013). This intuitive measure reflects the probability that a given relation (2.1 or 2.2) can be established between any two randomly selected individuals from two different groups. Computationally, this measure corresponds to the proportion of pairwise comparisons between individuals of two groups that constitute a relation of comparability or dominance. A possible drawback of using a measure based on the "average" (expectation) is that it is blind to the presence of particular subgroups for observations for which we might observe different patterns in health disparities. However, in

section 3.2 we show that it is possible to decompose our health index into the contributions of different subgroups.

In this section we build on this reasoning to define two population level measures: an index of dominance and an index of comparability.

2.3.1 Comparability

There is no guarantee that any two groups are composed of highly comparable individuals. If relations of comparability can be established across a relatively small number of individuals only, then a health ranking might not be informative. We acknowledge this potential shortcoming in two ways. First, we construct an index of comparability following the reasoning outlined above. This index explicitly measures to what extent the disease profiles of the individuals in the two populations can be ordered using our relations. Based on the value of the index of comparability, the analyst can judge whether the results are meaningful. Second, we use the index of comparability to adjust our health disparity measure. In the following, we formalize the index of comparability; the health disparity measure is introduced later.

Let us define the index of comparability formally. Let Ω_a and Ω_b denote the sets of individuals from two groups. The proportion of observations individual ω is comparable to can be written as

$$J_i(\omega) = \sum_{C_j : C_\omega \sim C_j} \mathrm{P}_i[C_j], \tag{2.1}$$

where C_ω is the set of conditions that afflicts ω and $\mathrm{P}_i[C_j]$ is the proportion of individuals with conditions C_j in group Ω_i, the reference group, which can be either Ω_a or Ω_b. That is, the comparability of ω can be assessed both relative to her own group or to the comparison group. The choice of reference group is not important for our health disparity measure (see section 3), but is needed to define aggregate comparability and dominance.

Having defined J_i, we only need to aggregate across individuals to arrive at an aggregate measure of comparability.

Definition 2.3. *The comparability index* $\mathrm{CI}_i(\Omega_j) \in (0,1)$*, with reference group* Ω_i*, is given by*

$$\mathrm{CI}_i(\Omega_j) = \frac{1}{n_j} \sum_{\omega \in \Omega_j} J_i(\omega).$$

The interpretation of this equation is straightforward; i.e., high CI values mean that most observations can be compared to each other, while the opposite is true when the CI is close to zero. By construction, every observation is comparable to observations with no conditions (minimal element of $\mathcal{G}$) and to observations that suffer from all conditions (maximal element of $\mathcal{G}$); and vice versa. Thus, the CI is bounded below as follows:

$$\mathrm{CI}_i(\Omega_j) \geq \mathrm{P}_i[\emptyset \vee \mathcal{D}](2 - \mathrm{P}_i[\emptyset \vee \mathcal{D}]), \tag{2.2}$$

where $\mathrm{P}_i[\emptyset \vee \mathcal{D}] = \mathrm{P}_i[\emptyset] + \mathrm{P}_i[\mathcal{D}]$. For example, if $\mathrm{P}_i[\emptyset \vee \mathcal{D}] = 0.5$, the CI is equal to or higher than 0.75, even if the observations with at least one but not all conditions are not all comparable to each other. This means that large proportions of healthy individuals in both groups will increase comparability, even if the conditions of the unhealthy individuals differ substantially between groups.

To account for this, we can decompose the CI into two parts. The first part, $\mathrm{CI}_i^D(\Omega_j)$, given on the right hand side of equation (2.2), gives the proportion of comparability due to the presence of observations with no or all conditions (see above). The second part is given by:

$$\mathrm{CI}_i^N(\Omega_j) = \frac{1}{n} \sum_{\omega \in \Omega_j : C_\omega \in \mathcal{N}} J_i^N(\omega), \tag{2.3}$$

where $\mathcal{N} = \mathcal{C}/\{\emptyset, \mathcal{D}\}$ is the set of conditions that includes all combinations of conditions except the minimal element (healthy) and the maximal element (all conditions) and

$$J_i^N(\omega) = \sum_{C_j : C_\omega \sim C_j} \mathrm{P}_i[C_j] \quad \text{for} \quad C_j, C_\omega \in \mathcal{N}. \tag{2.4}$$

This gives the contribution to the overall comparability of individuals with at least one but not all conditions; in other words, to the comparability among the sick.Before turning to dominance, we wish to make a final, important remark regarding the interpretation of comparability. All of our measures are aggregations of individual-level comparisons. For this reason, even if two populations have identical distributions of prevalences at the population level, they might have a low comparability index if the individuals constituting the groups are not comparable. Ultimately, this can be the case because our measures of comparability are all based on the concept of dominance, and not on dissimilarity. One individual is not comparable to another individual, not only because each has a different affliction, but because it is not possible to establish a relation of dominance (or weak dominance) between them. The clearest example of such a situation is a comparison between a completely healthy individual and an individual suffering from several conditions. While these individuals may satisfy the comparability criteria, they clearly have very different medical profiles. At the same time, as a necessary but not sufficient condition for achieving high levels of comparability, populations have to suffer from similar diseases.

2.3.2 Dominance

The procedure we used to count the proportion of comparable individuals can also be applied when assessing the proportion of individuals linked by a relation of dominance. In addition, we build on this reasoning and include a comparability correction: we propose weighting the dominance relations not by the total possible pairwise comparisons, but rather by the total comparable relations only. In essence, our axioms provide us with tools that can be used to establish rankings among the comparable, while staying neutral among the non-comparable. For this reason, while we still report information on the total proportion of comparable relations, we exclude those individuals for whom we cannot establish an ordering from dominance measures.

Let Ω_a and Ω_b denote the sets of individuals from the two groups to be compared; let $\mathrm{P}_i[C_j]$ be the proportion of individuals with conditions C_j in group Ω_i. For individual ω

with conditions C_ω the proportion of dominated individuals is given by:

$$I_i(\omega) = \sum_{C_j : C_\omega \succ C_j} \mathrm{P}_i[C_j]. \tag{2.5}$$

For a given reference group, $I(\omega)$ can be computed for each individual $\omega \in \Omega_a, \Omega_b$.

Note that since $J_i(\omega) \geq I_i(\omega)$ the expectation of $I_i(\omega)$ is bounded by comparability. To account for this, we count the number of relations of dominance over the comparable pairwise relations instead of the total. This correction defines the comparability weighted health dominance index:

Definition 2.4. *The comparability weighted health dominance index* $\mathrm{CHDI}_i(\Omega_j) \in [0, 1]$, *with reference group* Ω_i, *is given by*

$$\mathrm{CHDI}_i(\Omega_j) = \frac{1}{n_j} \sum_{\omega \in \Omega_j} \frac{I_i(\omega)}{J_i(\omega)}.$$

There are two issues that remain unresolved. First, the values of the comparability index and of the CHDI are dependent on the choice of reference group. Second and more importantly, we need a way to compare health across multiple groups since, as we show in the next section, dominance measures are not always transitive. That is, such measures do not necessarily lead to a unique ranking of populations. Our proposed health ranking indicator addresses both concerns.

3 The PODI index

Having formally defined the relations of dominance and comparability at the population level, we propose the partial ordering disparity index (PODI). The PODI is constructed through a simple modification of the CHDI that resolves the issue of having to choose a reference population. In the first part of this section, we present the PODI and show that, despite its simplicity, the index has a number of desirable properties: i.e., it is not altered by the sample size or the choice of reference group, and it properly reflects the

health improvements within groups. Then, in the remainder of the section, we illustrate how the PODI can be used to establish a health ranking between any number of groups. In conclusion, we demonstrate how it is possible to capture the essence of the axioms we presented in the first part of the paper in a health measure capable of ranking across multiple groups or populations.

3.1 Definition

We define the PODI between two groups as the difference between the CHDIs resulting from alternating the reference group. This essentially captures the difference in expected domination.

Definition 3.1. *For two groups* Ω_a *and* Ω_b *the* $PODI \in [-1, 1]$ *is defined as*

$$P_{ab} = \text{CHDI}_a(\Omega_b) - \text{CHDI}_b(\Omega_a).$$

A value of 1 means that every member of group Ω_b is less healthy than every member of group Ω_b; the reverse is true when $P_{ab} = -1$. A value of zero implies that both groups are equally (un)healthy. Intermediary values can also be interpreted directly; values above zero indicate the disadvantage of Ω_b relative to Ω_a, while the opposite is the case for values below zero. Intuitively, the PODI measures the relative expected dominance of one group over the other. For example, $P_{ab} = 0.3$ means that on average, members of Ω_b dominate 30percent more members of Ω_a than members of Ω_a dominate members of Ω_b. A value of 0.3 again indicates a reversed situation.

3.2 Properties

The PODI has a number of interesting properties. The choice of the reference group, the sample size, and the definition of the pairwise relations do not alter the results in unexpected ways.

Property 3.1 (Symmetry)**.** *If the distributions of conditions are changed for groups* Ω_a *and* Ω_b *the sign of the PODI is reversed, i.e.* $P_{ab} = -P_{ba}$*, otherwise it remains unchanged.*

Property 3.2 (Sample size independence)**.** *The PODI does not depend on sample size, only on relative frequencies. That is, if* n_a *and* n_b *denote the sizes of populations* Ω_a *and* Ω_b *and* n_c *and* n_d *are the sizes of two other populations* Ω_c *and* Ω_d *for which* $n_a \neq n_c$ *and* $n_b \neq n_d$ *hold, but the distribution of conditions is the same for* Ω_a *and* Ω_c *as well as for* Ω_b *and* Ω_d*, then* $P_{ab} = P_{cd}$*.*

Property 3.3 (Mirror property)**.** *If the dominance relation as given in definition 2.2 is reversed, i.e.* $\forall C_i, C_j \in \mathcal{C} : C_i \preceq C_j$ *iff* $C_j \subset C_i$*, calculating the "reversed" PODI yields* $1 - P_{ab}$*.*

Furthermore, the PODI is well behaved with respect to changes in the prevalence of conditions that result in an improvement in the health of one of the groups. We consider two changes that unambiguously improve the health of a group and demonstrate that those changes also improve the PODI. First, let one condition be removed from a member of group Ω_a; call this a cure. The second change consists of adding an individual with no conditions to group Ω_a; we call this a group improving addition. Let Ω_{a_C} be the group resulting from curing one member of a, and Ω_{a_A} be the group resulting from a adding a healthy individual to group a. Then, the following properties hold (proofs are given in the appendix).

Property 3.4 (Cures)**.** *Cures can only increase the PODI, i.e.* $P_{ab} \leq P_{a_C b}$*.*

Property 3.5 (Group improving additions)**.** *Group improving additions can only increase the PODI, i.e.* $P_{ab} \leq P_{a_A b}$*.*

Finally, the PODI between two populations can be decomposed into the contributions of distinct subpopulations. For example, suppose we are interested in examining the relative health of the males and the females in a given country. We might suspect that these differences are not constant across, for example, socioeconomic groups. This hypothesis

can be tested directly since, as we show, the PODI is the weighted sum of the relative health across subpopulations.

Formally, suppose populations Ω_a and Ω_b can each be divided in S subgroups, denoted with the superindex s. For each subgroup s the PODI can be calculated as $P^s_{ab} = \text{CHDI}^s_a(\Omega^s_b) - \text{CHDI}^s_b(\Omega^s_a)$. Then, the aggregate PODI can be decomposed into weighted contributions of each subgroup comparison and an additional cross-group comparison, denoted by P^{-s}_{ab}.

Property 3.6 (Decomposability)**.** *The PODI can be decomposed as*

$$P_{ab} = \sum_{s=1}^{S} w^s P^s_{ab} + P^{-s}_{ab}, \tag{3.1}$$

where w_s is a weight given by $\frac{n^s_a n^s_b}{n_a n_b}$. P^s_{ab} equals

$$\frac{1}{n^s_b} \sum_{\omega \in \Omega^s_b} \frac{I^s_a(\omega)}{J^s_a(\omega)} - \frac{1}{n^s_a} \sum_{\omega \in \Omega^s_a} \frac{I^s_b(\omega)}{J^s_b(\omega)} \tag{3.2}$$

with $I^s_a(\omega)$ defined as $\sum_{C_j : C_\omega \succ C_j} \text{P}^s_a[C_j]$, where $\text{P}^s_a[C_j]$ is the distribution of conditions in Ω^s_a. $J^s_a(\omega)$, $I^s_b(\omega)$, and $J^s_b(\omega)$ are defined correspondingly. P^{-s}_{ab} is given by

$$\frac{1}{n_b} \sum_{\omega \in \Omega^s_b} \frac{I^{-s}_a(\omega)}{J^{-s}_a(\omega)} \frac{1 - n^s_a}{n_a} - \frac{1}{n_a} \sum_{\omega \in \Omega^s_a} \frac{I^{-s}_b(\omega)}{J^{-s}_b(\omega)} \frac{1 - n^s_b}{n_b} \tag{3.3}$$

with $I^{-s}_a(\omega) = \sum_{C_j : C_\omega \succ C_j} \text{P}^{-s}_a[C_j]$ and $\text{P}^{-s}_a[C_j]$ is the distribution of conditions in Ω_a / Ω^s_a. $J^{-s}_a(\omega)$, $I^{-s}_b(\omega)$, and $J^{-s}_b(\omega)$ are defined correspondingly.

A proof is given in the appendix. This property shows that a decomposition by subgroups as described above is relatively straightforward.

3.3 A health ranking

In applying any measure of health, we often aim to draw conclusions based on comparisons that involve more than two groups. Unfortunately, the pairwise measures defined in the

previous sections can not be applied directly to these scenarios since they are not always transitive.

Property 3.7 (non-transitivity)**.** *The PODI is non-transitive in the sense that if* $P_{ab} > 0$ *and* $P_{bc} > 0$ *(or* $P_{ab} < 0$ *and* $P_{bc} < 0$*) then* $P_{ac} \lesseqgtr 0$*.*

In our case, this issue can arise when there is a high degree of non comparability across groups (an example is provided in the Appendix).

This well-known shortcoming has been addressed in the literature on the optimal aggregation of pairwise comparisons (Gavalec et al., 2015), and on pairwise comparisons of health (Herrero and Villar, 2013). A common strategy for overcoming non-transitivity is to summarize all pairwise comparisons P_{ij} for each group i with all other groups j into a single index I_i. By construction, this index is transitive and thus provides a health ranking across groups.

Here we follow the approach developed in Crawford and Williams (1985) and Barzilai and Golany (1990). Intuitively, the strategy is to obtain individual group indexes that are as close as possible to the values of each of the possible pairwise comparisons. Since this cannot be fully achieved, the method specifies a penalty function consisting of taking the square of the differences between P_{ij} and the values for each group I_i.

Let $\Omega_1, \Omega_2, \ldots, \Omega_G$ be the groups for which we know P_{ij} and want to derive I_i. The individual I_i are normalized to sum to zero, i.e. $\sum I_i = 0$. Moreover, we want the individual I_i to be close to $P_{ij} = I_i - I_j$, i.e. pairwise differences of index values should mimic pairwise comparisons given by P_{ij} as closely as possible. To achieve this, we minimize

$$\min_I \sum_{i=1}^{G} \sum_{j=1}^{G} [P_{ij} - (I_i - I_j)]^2, \tag{3.4}$$

given the constraint that

$$\sum_i I_i = 0. \tag{3.5}$$

An optimal solution to the problem is given by:

$$I_i = \frac{1}{P} \sum_{j=1}^{P} P_{ij}, \tag{3.6}$$

which is simply the mean of all pairwise comparisons for group i.

This solution guarantees that higher values of I_i imply that a given group i has a relatively low health status. Furthermore, Barzilai and Golany (1990) prove that the solution is unique, and it does not depend on the ordering of groups. Note that the additive restriction (3.5) is required for a unique solution, but that the choice of normalization does not affect the ranking. In the next section we apply our concepts to a comparison of health across races in the U.S.

4 Empirical application

We use our health indexes to study racial health differences in the US. Drawing on data from the NHIS for the year 2014, we study the relative prevalence of 15 conditions across four races for 3570 individuals. The comparison is narrowed to ages 60 to 65 to control for the differences in age distributions across races. The objective of this exercise is to illustrate that while simple comparisons of prevalences across groups are appealing, they are often not sufficient to rank health across populations or groups.

4.1 Descriptive statistics

We report the prevalences of the full set of 15 diseases in the NHIS for our age and racial groups (Table 1). Prevalences are measured by survey questions that ask respondents whether they have ever been told by a doctor that they suffer from a given condition. Because they are relatively small in number, the individuals with missing values are removed from the sample.[1] In addition, we restrict our exercise to the individuals who reported belonging to one of the four main racial categories, and exclude those who

[1]There are a total of 153 missing values and no racial group has a disproportionate share of missing values.

reported belonging to the other category, as this category is difficult to interpret. The data should be read with some caution since the narrow age group results in small samples for some racial groups. For example, the Asian group is composed of 163 individuals. As a consequence, the number of cases of some of the less prevalent diseases might be subject to small sample bias.

Table 1: Prevalence by race

Diseases	White	Asian	Black	Hispanic	All races
Hypertension	48.1%	49.1%	69.6%	52.9%	51.6%
High cholesterol	48.8%	45.6%	49.1%	46.4%	48.5%
Coronary heart disease	7.0%	4.3%	8.0%	9.1%	7.2%
Angina pectoris	2.9%	1.2%	3.4%	3.0%	2.9%
Heart attack	6.1%	4.9%	5.2%	6.1%	5.9%
Heart condition/ disease	11.9%	7.4%	12.1%	7.2%	11.2%
Stroke	4.2%	1.8%	9.1%	6.3%	5.0%
Emphysema	3.9%	0.6%	4.0%	3.6%	3.7%
COPD	7.7%	3.1%	7.4%	3.6%	7.1%
Asthma	12.7%	12.3%	17.7%	11.5%	13.3%
Ulcer	10.0%	6.8%	9.9%	11.5%	10.0%
Cancer	17.4%	5.5%	9.5%	10.7%	15.1%
Diabetes	16.0%	17.8%	29.4%	29.1%	19.3%
Arthritis	40.8%	25.9%	47.2%	35.7%	40.5%
Chronic liver condition	2.7%	2.5%	2.8%	3.1%	2.7%

A first look at the relative prevalences across the races reveals that some conditions have a much steeper race gradient that others. For example, blacks have a much higher prevalence of hypertension (69.6 percent) than the other races (at around 50 percent). In contrast, the differences in the prevalence of heart attacks across races appear to be much smaller (between 4.9 percent for Asian individuals, and 6.1 percent for Hispanic and white individuals). Nevertheless, the table shows a wide range of prevalences across races for each disease. Given that no racial group has unambiguously lower prevalences than the other groups across all conditions, it is not possible to establish a relative health ordering without further analysis.

In addition, Table 2 highlights the empirical relevance of multi-morbidity. Only 15.3 percent of the sample reported having none of the 15 diseases, while 20.8 percent indicated

Table 2: Number of diseases per person by race

	0	1	2	3	4	5	6	Total
White	16.0%	21.2%	22.9%	16.0%	10.7%	5.9%	7.4%	100.0%
Asian	21.3%	25.8%	22.6%	14.8%	7.7%	4.5%	3.2%	100.0%
Black	9.2%	17.4%	21.9%	18.4%	16.6%	8.0%	8.4%	100.0%
Hispanic	16.8%	20.7%	23.3%	16.5%	7.7%	7.1%	8.0%	100.0%
All races	15.3%	20.8%	22.7%	16.3%	11.1%	6.3%	7.6%	100.0%

that they have one condition. The remaining 63.9 percent of the individuals across all races reported suffering from at least two conditions. While the prevalence of multi-morbidity differs across races, in each of the racial groups more than half the individuals said they have at least two conditions, and a high proportion indicated that they have several more. A further difference across that races is that among the individuals suffering from multi-morbidity, the most frequent combinations of conditions vary in prevalence as well. For example, among those individuals with two conditions, the combination of hypertension and cholesterol is twice as prevalent among blacks (30.8 percent) as among Hispanics (15.9 percent).

Table 3: Most prevalent combinations of diseases by race

Race	Hypertension & Cholesterol	Hypertension & Arthritis	Cholesterol & Arthritis	Rest
White	22.1%	19.3%	12.7%	45.8%
Asian	28.6%	11.4%	14.3%	45.7%
Black	30.8%	5.6%	13.1%	50.5%
Hispanic	15.9%	14.6%	13.4%	56.1%
All races	23.0%	16.5%	12.9%	47.7%

Our health measures are designed to provide a health ranking for precisely this type of complex information on prevalences. First, we assess the comparability of the group-specific profiles of diseases. Then, once we have established whether a prevalence-based comparison is reasonable, the PODI provides a health ranking.

4.2 Health measures

Comparability Overall, the comparability index across races (Table 4, panel A) ranges from .51 (between blacks and whites) to .59 (between Asians and Hispanics). This means that it is possible to establish relations of comparability for over half of the possible pairwise comparisons across individuals of differences races. In this example, we have purposely kept diseases at the lowest possible form of aggregation, which makes comparability more difficult. Hence, we treat this as a lower bound to the comparability levels. To illustrate the impact of the level of aggregation in conditions, we have calculated the comparability index after merging all of the circulatory diseases into one category (following the IC-10 category definitions). Grouping circulatory conditions has a very noticeable impact on comparability, raising the lowest comparability index to .69 and the highest to .79 (Table 4, panel B).

Table 4: Comparability

(a) Comparability disagregatted

	White	Black	Asian	Hispanic
White	0.54	0.52	0.58	0.54
Black	0.52	0.51	0.58	0.53
Asian	0.58	0.58	0.63	0.59
Hispanic	0.54	0.53	0.59	0.55

(b) Comparability Aggregated

	White	Black	Asian	Hispanic
White	0.71	0.70	0.75	0.70
Black	0.70	0.69	0.75	0.69
Asian	0.75	0.75	0.79	0.75
Hispanic	0.70	0.69	0.75	0.70

PODI and dominance Table 5 shows both of the pairwise PODI indexes (3.1) across races and the summary index (3.6) that aggregates the pairwise comparisons. We include both the PODI for the full set of conditions and the PODI for the aggregation of circulatory diseases. In this particular case, we find that the pairwise PODIs satisfy transitivity, but

this need not be the case in the other scenarios. Across both scenarios, our results indicate that the Asian group is the healthiest, and the black group is the least healthy. However, the intermediate positions depend on the modeling of circulatory diseases. In our preferred model, with a single category for circulatory diseases, the white group is ranked higher than the Hispanic group. The opposite outcome is found when we use multiple circulatory disease categories; and, hence, when comparability is low.

Table 5: PODI

(a) PODI aggregated

Ranking		White	Black	Asian	Hispanic
2nd	White	-	0.12	-0.17	0.01
4th	Black	-0.12	-	-0.29	-0.10
1st	Asian	0.17	0.29	-	0.18
3rd	Hispanic	-0.01	0.10	-0.18	-

(b) PODI disaggregated

Ranking		White	Black	Asian	Hispanic
3rd	White	-	0.15	-0.16	-0.02
4th	Black	-0.15	-	-0.31	-0.16
1st	Asian	0.16	0.31	-	0.14
2nd	Hispanic	0.02	0.16	-0.14	-

Comments The results of our application are generally in line with existing findings on racial/ethnic disparities in health in the US (Braveman et al., 2010). It is well-established that morbidity levels are highest among the black population (Ward, 2013). In addition, the results highlight the importance of taking comparability into consideration. The Hispanic group is ranked as healthier than the white group in the low comparability case only. This finding corresponds with recent research indicating that the so-called Hispanic paradox [2] might not apply to morbidity. However, the prevalences found among the Asian population should be viewed with care given the relatively small sample size for this group.

[2] In the U.S., Hispanic individuals tend to have a lower socieconomic status than white individuals but also a relative mortality advantage. This finding has been called the Hispanix paradox(Abraido-Lanza et al., 1999)

For an in-depth study on racial differences in health among the elderly in the U.S. that includes such measures, see (Hummer et al., 2004).

5 Discussion

We believe that the main strengths of our approach are our relatively weak assumptions and our use of prevalences, which are arguably more objective than self-assessments. The PODI can also be easily used to compare populations on the basis of the distribution of other prevalence-based measures. In addition to being used to investigate the distribution of diseases, the PODI could, for instance, be used to compare populations based on the prevalence of limitations in activities of daily living. Beyond direct measures of health, other possible areas in which the PODI could be applied include health behaviors like smoking and alcohol consumption. Generally speaking, our indicator can be used to compare any distributions of binary measurements. In addition to the issues related to data collection (diagnosis bias and reporting bias) discussed above, a potential drawback of the simple criteria on which our approach is based is that they could result in low comparability; i.e., in health rankings that only capture a fraction of the population. Our recommendation for potential users is to avoid including too many individual conditions in the analysis. Aggregations across typologies of diseases are a straightforward way to increase comparability. Still, if the focus of the analysis is not on individual conditions, using coarser definitions for disease profiles might result in only a small loss of information. Ultimately, the best approach is clearly dependent on the application.

A distinctive feature of relative prevalence measures is that they require us to explicitly consider the comparability of the populations studied. The PODI relies on individuals having comparable disease profiles, while other measures, like self-rated health, can be used regardless of the underlying distribution of diseases. For example, consider a comparison between a population in an industrialized country and a population in a developing country. Self-rated health measures like the visual scale analogue [3] could be

[3]The visual scale analogue (VAS) simply asks individuals to rank their health on a zero-to-100 scale.

easily compared, whereas differences in the diseases afflicting the two countries would complicate a comparison based on prevalences. Because prevalence-based approaches have stricter comparability requirements, they can help to ensure the rigor of the results in a context in which comparability is suspect, as in the example above. Thus, we believe that the thorough treatment of comparability is a strength of our analysis.

Finally, while we did not incorporate into our framework any information on the relative severity of different diseases, it is possible to extend our analysis to include expert knowledge on this factor. Such an approach would entail modifying the definitions of the relations of dominance and comparability to include additional criteria, but the analytical framework would remain unchanged. The inclusion of severity could be a natural way to mitigate some of the shortcomings associated with ordinal measures that apply to the PODI. In our analysis, we seek to establish health rankings (a partial order), and avoid making judgments regarding the extent or the magnitude of the relative health advantages of individuals. Nevertheless, it is clear that there can be some degree of cardinality in relative health comparisons between individuals. An obvious example would be a comparison of an individual suffering from a life-threatening condition with an individual suffering from a minor chronic condition. In such a case, a severity adjustment could restore some of these cardinal judgments to an otherwise entirely ordinal measure.

6 Conclusion

We proposed a new indicator, the PODI, based on the relative prevalence of conditions across populations. Although no statistic can provide a universal solution for the methodological issues that arise when assessing health disparities, our method specifically addresses two of the main difficulties associated with health comparisons. Specifically, we attempt to provide a measure that is both robust to self-assessment biases, and that takes into consideration the differences in the types of diseases afflicting the comparison groups. We illustrated how our simple framework can be used to analyze complex data on disease prevalences by applying it to current health disparities across racial groups in the U.S.

References

Abraido-Lanza, A. F., Dohrenwend, B. P., Ng-Mak, D. S. and Turner, J. B. (1999). The latino mortality paradox: a test of the" salmon bias" and healthy migrant hypotheses. *American journal of public health* 89: 1543–1548.

Anderson, G. F., Jeremy, H., Hussey, P. S. and Jee-Hughes, M. (2000). Health spending and outcomes: Trends in oecd countries, 1960-1998. *Health Affairs* 19: 150–157.

Barzilai, J. and Golany, B. (1990). Deriving weights from pairwise comparison matrices: The additive case. *Operations Research Letters* 9: 407–410.

Braveman, P. A., Cubbin, C., Egerter, S., Williams, D. R. and Pamuk, E. (2010). Socioeconomic disparities in health in the United States: what the patterns tell us. *American Journal of Public Health* 100: S186–S196.

Burgard, S. A. and Chen, P. V. (2014). Challenges of health measurement in studies of health disparities. *Social Science & Medicine* 106: 143–150.

Crawford, G. and Williams, C. (1985). A note on the analysis of subjective judgment matrices. *Journal of Mathematical Psychology* 29: 387–405.

Dowd, J. B. and Zajacova, A. (2010). Does self-rated health mean the same thing across socioeconomic groups? evidence from biomarker data. *Annals of epidemiology* 20: 743–749.

Etches, V., Frank, J., Di Ruggiero, E. and Manuel, D. (2006). Measuring population health: A review of indicators. *Annual Review of Public Health* 27: 29–55.

Etilé, F. and Milcent, C. (2006). Income-related reporting heterogeneity in self-assessed health: evidence from france. *Health economics* 15: 965–981.

Gavalec, M., Ramík, J. and Zimmermann, K. (2015). *Decision Making and Optimization. Special Matrices and Their Applications in Economics and Management*. Lecture Notes in Economics and Mathematical Systems. Springer.

Herrero, C. and Villar, A. (2013). On the comparison of group performance with categorical data. *PLoS ONE* 8: e84784.

Hummer, R. A., Benjamins, M. R. and Rogers, R. G. (2004). Racial and ethnic disparities

in health and mortality among the us elderly population. *Critical perspectives on racial and ethnic differences in health in late life* : 53–94.

Humphries, K. H. and Van Doorslaer, E. (2000). Income-related health inequality in canada. *Social science & medicine* 50: 663–671.

Idler, E. L. and Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior* : 21–37.

Idler, E. L. and Benyamini, Y. (1999). Community studies reporting association between self-rated health and mortality. *Research on Aging* 21: 392–401.

Johnston, D. W. and Propper, M. A., C.and Shieslds (2009). Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient. *Journal of Health Economics* 28: 540–552.

Jürges, H. (2008). Self-assessed health, reference levels and mortality. *Applied Economics* 40: 569–582.

König, H.-H., Bernert, S., Angermeyer, M. C., Matschinger, H., Martinez, M., Villagut, G., Haro, J. M., de Girolamo, G., de Graaf, R., Kovess, V. and ESEMeD/MHEDEA 2000 Investigators (2009). Comparison of population health status in six European countries: Results of a representative survey using the EQ-5D questionnaire. *Medical Care* 47: 255–261.

Lieberson, S. (1976). Rank-sum comparison between groups. *Sociological Methodology* 7: 276–291.

Marmot, M. (2005). Social determinants of health inequalities. *The Lancet* 365: 1099–1104.

Murray, C. J. L. (2002). Summary measures of population health: concepts, ethics, measurement and applications, World Health Organization.

National Research Council of the National Academics (2004). *Critical perspectives on racial and ethnic differences in health in late life*. National Academies Press.

Rivera, J. A., Barquera, S., Campirano, F., Campos, I., Safdie, M. and Tovar, V. (2002). Epidemiological and nutritional transition in Mexico: rapid increase of non-communicable chronic diseases and obesity. *Public Health Nutrition* 5: 113–122.

Van Doorslaer, E. and Gerdtham, U.-G. (2003). Does inequality in self-assessed health

predict inequality in survival by income? evidence from swedish data. *Social science & medicine* 57: 1621–1629.

Van Doorslaer, E., Koolman, X. and Jones, A. M. (2002). Explaining income-related inequalities in doctor utilisation in europe: a decomposition approach. *ECuity II Project* .

Ward, B. W. (2013). Prevalence of multiple chronic conditions among us adults: estimates from the national health interview survey, 2010. *Preventing chronic disease* 10.

Weimann, A., Dai, D. and Oni, T. (2016). A cross-sectional and spatial analysis of the prevalence of multimorbidity and its association with socioeconomic disadvantage in South Africa: A comparison between 2008 and 2012. *Social Science & Medicine* 163: 144–156.

A Proofs

Properties 3.1, 3.3, and 3.2 follow directly from the definition of the PODI. Property 3.5 follows from the proof of property 3.4, given below.

Proof of property 3.4. This proof uses the concept of paths from the maximal to the minimal node. A path is any sequence of unique nodes and edges from $\mathcal{D}$ to $\emptyset$ which follows the direction of the edges. For instance, in figure 1 going from $\mathcal{D}$ to $\{A, B\}$ to $\{A\}$ to $\emptyset$ is a path, while $\mathcal{D}$, $\{A, B\}$, $\{A\}$, $\{A, C\}$, $\{C\}$, $\emptyset$ is not. Using this concept one can show that $\text{CHDI}_{a_C}(\Omega_b) \geq \text{CHDI}_a(\Omega_b)$ and $\text{CHDI}_b(\Omega_a) \geq \text{CHDI}_b(\Omega_{a_C})$ hold, which proves the above property. Let ω_C be the cured individual from population a. We can distinguish four cases:

- For Individuals in Ω_b who are not on the same path as ω_C the contribution to the CHDI will not change.
- Individuals in Ω_b who are on the same path experience no changes to comparability. The contribution of dominance to the CHDI will stay constant or increase since dominance may increase, but not decrease.

- For Individuals in Ω_b who are on the same path as ω_C before the cure but not afterward comparability will decrease. However, dominance will not change, as a necessary requirement for ω_C to leave the shared path is that she has more conditions than the individuals from Ω_b.

- For individuals from Ω_b who are not on the same path as ω_C before the cure but are on the same path afterward, both comparability and dominance will increase. To move on the same path, ω_C must have the same number of conditions as the members of Ω_b to whom this situation applied before the cure. After the cure they are on the same path and ω_C has one less condition and is therefore dominated. Thus, for each member of $\Omega_b b$ for whom this situation applies the contribution to the CHDI is $\frac{I_a(\omega)+1/n_a}{J_a(\omega)+1/n_a}$, which is bigger than or equal to $\frac{I_a(\omega)}{J_a(\omega)}$ because $I_a(\omega) \leq J_a(\omega)$.

This proves that $\text{CHDI}_{a_C}(\Omega_b) \geq \text{CHDI}_a(\Omega_b)$ holds. $\text{CHDI}_b(\Omega_a) \geq \text{CHDI}_b(\Omega_{a_C})$ is trivial, as only ω_C's contribution to the CHDI is affected and can only decrease. □

Proof of property 3.6. It suffices to show that the CHDI can be decomposed by groups; the decomposition of the PODI follows immediately. The CHDI is given by:

$$\begin{aligned} \text{CHDI}_a(b) &= \frac{1}{n_b} \sum_{\omega \in \Omega_b} \frac{I_a(\omega)}{J_a(\omega)} \\ &= \frac{1}{n_b} \sum_{\omega \in \Omega_b} \Pr(\omega \succ \omega_a | \omega \sim \omega_a), \end{aligned}$$

where $\Pr(\omega \succ \omega_a | \omega \sim \omega_a)$ is the probability that a member ω_a of Ω_a dominates ω from Ω_b, conditional on being comparable. It can be decomposed as

$$\text{CHDI}_a(b) = \sum_{s=1}^{S} \frac{n_a^s n_b^s}{n_a n_b} \frac{1}{n_b^s} \sum_{\omega \in \Omega_b^s} \frac{I_a^s(\omega)}{J_a^s(\omega)} + \frac{1}{n_b} \sum_{\omega \in \Omega_b^s} \frac{I_a^{-s}(\omega)}{J_a^{-s}(\omega)} \frac{1 - n_a^s}{n_a}.$$

Rearranging gives

$$\begin{aligned}\mathrm{CHDI}_a(b) &= \frac{1}{n_b}\sum_{s=1}^{S}\sum_{\omega\in\Omega_b^s}\Pr(\omega \succ \omega_a|\omega \sim \omega_a \wedge \omega_a \in \Omega_a^s)\Pr(\omega_a \in \Omega_a^s) \\ &+ \frac{1}{n_b}\sum_{s=1}^{S}\sum_{\omega\in\Omega_b^s}\Pr(\omega \succ \omega_a|\omega \sim \omega_a \wedge \omega_a \notin \Omega_a^s)\Pr(\omega_a \notin \Omega_a^s),\end{aligned}$$

which equals $\frac{1}{n_b}\sum_{\omega\in\Omega_b}\Pr(\omega \succ \omega_a|\omega \sim \omega_a)$ and completes the proof. □

Proof of property 3.7. We provide a non-transitive example. Suppose we have three groups, Ω_a, Ω_b, and Ω_c, and three conditions, A, B, and C. Assume that in group Ω_a 30% of individuals have no condition and 70% have condition B. In group Ω_b 30% have no condition, 50% have condition A, and 20% have both condition A and B. In group Ω_c 50% have no condition and 50% have both condition A and C. The values of the PODI between the populations described in this example are $P_{ab} = 0.22$, $P_{bc} = 0.05$; thus $P_{ac} = -0.2$, and thus the PODI is not transitive.

□