

The challenges of choosing and explaining a phenomenon in epidemiological research on the “Hispanic Paradox”

Sean A. Valles¹

Published online: 11 January 2016
© Springer Science+Business Media Dordrecht 2016

Abstract According to public health data, the US Hispanic population is far healthier than would be expected for a population with low socioeconomic status. Ever since Kyriakos Markides and Jeannine Coreil highlighted this in a seminal 1986 article, public health researchers have sought to explain the so-called “Hispanic paradox.” Several candidate explanations have been offered over the years, but the debate goes on. This article offers a philosophical analysis that clarifies how two sets of obstacles make it particularly difficult to explain the Hispanic paradox. First, different research projects define the Hispanic paradox phenomenon in substantially different ways. Moreover, using Bas van Fraassen’s pragmatic theory of explanation and Sean Valles’s extension of it with the concept of “phenomenon choice,” it also becomes clear that there are also multiple ways of explaining each individual definition of the phenomenon. A second set of philosophical and methodological challenges arises during any attempt to study “Hispanic” phenomena, with one key challenge being that the “Hispanic” panethnic concept was intentionally made vague as it was developed and popularized during the 1960s–1970s. After comparing this case with similar cases in the philosophical literature, the article concludes with observations on what makes this problem unique, particularly its ethical features.

Keywords Explanation · Health disparities · Hispanic paradox · Philosophy of epidemiology · Philosophy of medicine · Public health ethics

Despite methodological limitations of much of the research, it can be concluded with some certainty that the health status of Hispanics in the Southwest is much more similar to the health status of other whites than that of

✉ Sean A. Valles
valles@msu.edu

¹ Lyman Briggs College and Department of Philosophy, Michigan State University, Holmes Hall, 919 E. Shaw Lane, Room E35, East Lansing, MI 48825, USA

blacks although socioeconomically, the status of Hispanics is closer to that of blacks. This observation is supported by evidence on such key health indicators as infant mortality, life expectancy, mortality from cardiovascular diseases, mortality from major types of cancer, and measures of functional health. On other health indicators, such as diabetes and infectious and parasitic diseases, Hispanics appear to be clearly disadvantaged relative to other whites.
~Kyriakos S. Markides and Jeannine Coreil [1, p. 253]

Introduction

The passage above appears in Kyriakos Markides and Jeannine Coreil's 1986 article, "The Health of Hispanics in the Southwestern United States: An Epidemiologic Paradox." The article informally reviews existing evidence from the preceding two decades and concludes that Hispanic health overall is inexplicably good in "key health indicators." Three decades later, the epidemiological community still has not explained the paradox, despite continuing attempts to do so.

The "Hispanic paradox" is taken to be paradoxical by epidemiologists because the data seem to contradict epidemiology's understanding of the "social determinants of health," the network of social factors affecting populations' health (malnutrition due to poverty, dangerous working conditions, limited access to healthcare services, etc.) [2]. These determinants of health would presumably manifest in health outcomes such as shorter lifespans in the relatively socioeconomically disadvantaged US Hispanic population. Yet, data continue to show that Hispanics fare better than expected in a variety of health measures [3]. This failure to make sense of basic health patterns in a population of over 50 million Americans is surprising [4]. The existence of such a mystery also presents a serious equity problem since ignorance about a minority population's health needs is an obstacle to the creation of ethically sound health policies. This article will examine two sets of philosophical and methodological impediments that have made the Hispanic paradox particularly difficult to explain.

First, the article will demonstrate that attempts to explain the Hispanic paradox are hindered by the widely varying definitions of "Hispanic paradox" used in the literature.¹ Furthermore, using work by Bas van Fraassen and Sean Valles, it will be demonstrated that individual definitions of the Hispanic paradox can each be interpreted as requiring either of two different types of explanations. The article will then analyze some sample articles in the Hispanic paradox literature to show how they are hampered by these explanatory challenges. Next, independently of the first set of obstacles, the article will argue that research on the Hispanic paradox is impeded by the heterogeneity and fuzziness of the "Hispanic" concept. These

¹ Sometimes the terms "epidemiologic paradox" and "epidemiological paradox" are used in place of "Hispanic paradox," which is in keeping with the 1986 Markides and Coreil article and an even earlier article on Mexican American mental health [5]. However, these terms are ambiguous since they are sometimes used to refer to a "healthy migrant effect" in any population [6]. Other Hispanic paradox terminological variants include "Latino Paradox" and "Latino Epidemiological Paradox" [7].

features, traced to the history of the demographic concept's development in the 1960s and 1970s, continue to affect contemporary data collection practices. Two other examples from philosophy of science will be offered as illuminating analogous cases. Finally, the article will argue that the Hispanic paradox case is made unique from the other philosophy of science cases due to the ethical (including social and disciplinary/professional) factors involved in public health research about an ethnic minority population.

Variations in definition of the Hispanic Paradox

“Hispanic paradox” means many different things to many different people. The ambiguity in the Hispanic paradox literature is not just a matter of an imprecise definition or incautious use of the term. The Hispanic paradox literature has a more fundamental scientific and philosophical challenge: Hispanic paradox researchers have each legitimately chosen from a wide range of possible phenomena to investigate, and hence have attempted to explain substantially different things. Table 1 includes sample definitions of the Hispanic paradox phenomenon, all appearing in publications from 2012–2014.

The discrepancies between the definitions of “Hispanic paradox” have drawn attention and critique. A 2001 article by Alberto Palloni and Jeffrey Morenoff judges that the Hispanic paradox is “a remarkably slippery idea, a moving target of sorts that refers to a number of very different things [13, p. 149].” Palloni and Morenoff analyze this “moving target” as consisting of combinations of three different dimensions: four versions of the health outcome (“(1) infant and child mortality, (2) adult mortality, (3) birthweight, or (4) adult health status” [13, p. 149]), three versions of the target population (US residents born in Mexico, US residents born in any Latin American country, or all Hispanics as indicated by having a typically Hispanic surname), and two versions of the comparison population (non-Hispanic Whites or non-Hispanic Blacks).² These three dimensions, and the possible states identified for each, leave Palloni and Morenoff frustrated by a theoretical total of 24 unique versions of the Hispanic paradox created by the mixing and matching of variables [13, p. 149]. As can be seen in Table 1's quotes from citations [3] and [11], Palloni and Morenoff's three dimensions do not take into account an additional dimension of how versions of the Hispanic paradox diverge in the current literature: different versions of the Hispanic paradox state that Hispanics have, with respect to some given metric(s), (1) better health, (2) similar health, or (3) similar or better health.

² As Bennett Holman has helpfully suggested, the Hispanic paradox could perhaps be described without reference to a comparison population if the phenomenon can be reframed as Hispanic populations failing to manifest the typical gradient of health and wealth, in which low socioeconomic status is associated with poor health. There is indeed a body of literature on this subject, particularly noting that education and health are weakly related in US Hispanics (unlike non-Hispanic Whites) [14]. However, this alternative approach to examining Hispanics' mysteriously good average health still requires comparing Hispanics to other population(s) since there must always be some set of reference data against which the Hispanic health-wealth data is considered aberrant.

Table 1 Contrasting Hispanic Paradox definitions

What the Hispanic Paradox Is (emphases added)	Source and notes
Despite Hispanics' lower socioeconomic position (i.e., lower education and income) and lower rate of health insurance coverage, they exhibit lower all-cause mortality rates than do non-Hispanics . This "Hispanic paradox" for adult mortality outcomes has been researched over the past 2 decades.	[8, p. 836]
Despite the higher prevalence of risk factors and a lower socio-economic status, Hispanics have been observed to have a lower risk of coronary death (CD) compared to non-Hispanic whites (NHW), leading several investigators to refer to a "Hispanic paradox."	[9, p. 303]
...the Hispanic paradox, a term used to describe the finding that Hispanics in the United States have similar mortality rates to those of [European Americans] in spite of worse disease profiles and lower SES...	[10, p. 1522]
Studies consistently document a Hispanic paradox in U.S. adult mortality, whereby Hispanics have similar or lower mortality rates than non-Hispanic whites despite lower socioeconomic status. This study extends this line of inquiry to disability, especially among foreign-born Hispanics, since their advantaged mortality seemingly should be paired with health advantages more generally.	[11, p. 81]
Despite a significantly more disadvantaged risk factor profile, Hispanics in the United States often experience similar or better health outcomes across a range of health and disease contexts compared with non-Hispanic Whites (NHWs), an epidemiological phenomenon commonly referred to as the "Hispanic paradox."	[3, p. e52] "Systematic Review and Meta-Analysis"
Despite the increased prevalence of [cardiovascular] risk factors, several studies have shown a paradoxically lower rate of [cardiovascular disease] among Hispanics as compared to NHW, including lower [cardiovascular] mortality. This observation has been referred to as the Hispanic paradox.	[12, p. 791] "A Systematic Review and Meta-Analysis of the Hispanic Paradox"

Since 2001, the number of variations within the original three dimensions has grown, at least in the dimension of health outcomes: Palloni and Morenoff identified four total, to which coronary death and cardiovascular disease can now be added. This expansion and the additional fourth dimension increase the total theoretical number of combinations from Palloni and Morenoff's 24 to a new total of 108 combinations (six versions of health outcomes; three versions of target population; two versions of comparison population; three versions of relative difference between populations). Admittedly, this includes some combinations that would be unlikely to get labeled as paradoxical in practice (e.g., if Hispanic adult mortality were found to be similar to Black adult mortality). Nevertheless, even a non-systemic review of the literature reveals that Hispanic paradox researchers mix and match components from different dimensions, yielding a range of different target phenomena, with each different phenomenon treated as *the* Hispanic paradox.

Candidate explanations of the Hispanic paradox

Four types of Hispanic paradox explanation have been offered over the years, with no one (or combination) gaining consensus support from the research community:

Healthy Migrant Effect Hispanic migrants entering the US (often in order to work) are more likely to be relatively young and able-bodied compared to their peers in their home countries; migrants enter the US because they are already atypically healthy (see, e.g., [15]).

Return Migration Effect (a.k.a. “*Salmon Bias Effect*”) Migrants are more likely to return to their countries of origin if they begin suffering from a long-term illness; the US Hispanic population is healthier only by having its sick members disappear from the US and hence from US data (see, e.g., [16]).

Ethnic Enclave Advantage (a.k.a. “*Barrio Advantage*”) Hispanics living in certain ethnic enclaves (neighborhoods with high concentrations of other Hispanics with similar cultural backgrounds) experience a variety of social support structures that benefit health. While poverty may be high in many such neighborhoods, dynamics such as strong social support for the elderly can buffer the negative effects of low socioeconomic status (see, e.g., [17]).

Systematic Data Error(s) One or more systematic biases in data collection, such as the misreporting of ethnicity on death certificates, is simply giving the illusion of better Hispanic health [18].

Each type of explanation is supported by empirical evidence, and each has been proposed as a partial or complete explanation of the paradox. This article will not attempt to assess the credibility of these hypotheses, individually or in conjunction. Rather, this article is concerned with making sense of why decades of data collection and debate within the community have not settled the dispute about which explanation(s) account for the Hispanic paradox.

Phenomenon choice

Hispanic paradox researchers must not only contend with different definitions of the Hispanic paradox but also with differing underlying understandings of what sort of explanation is required. Valles proposes the concept of “phenomenon choice” as an underappreciated early step in scientific model construction, using it to identify what has prevented researchers from reaching a consensus on how to explain the evolution of common single-gene genetic diseases [19]. The article explores how different researchers’ choices of target phenomenon (e.g., the total rate of cystic fibrosis cases vs. the rate of one cystic fibrosis mutation/variant only) can drastically impact later model evaluation (e.g., whether high rates are judged to be the consequence of natural selection or simply of random fluctuations in gene frequencies). Adopting philosophical work by Sylvain Bromberger and Bas van Fraassen, the article describes explanations as responses to “why questions,” in the sense of asking “Why *A* is the case, rather than some other possible state of affairs

from a set of possibilities [S],” the so-called “contrast class” [19]. Contextual considerations (goals, interests, parties involved, etc.) establish which particular set [S] of contrast class alternatives is called for in any given situation [19; 20, p. 129]. In other words, first, I choose a phenomenon to explain. Second, based on the exact delineation of that phenomenon, I tacitly or explicitly formulate a “why question” and seek to answer it. Formulating that “why question” requires that I choose a set of alternative states of the world for comparison: “why do I see the phenomenon behaving this way when it could behave in any of these other ways?” My interests, goals, and overall context will determine how I interpret what the why question needs to accomplish, and that interpretation determines which “other ways” I use when formulating the contrast class.

During the initial phenomenon choice stage of explanation, Hispanic paradox researchers have made quite different choices about what exactly the phenomenon is (as illustrated in Table 1). After that stage, Hispanic paradox researchers face an additional stage during which they must make non-obvious choices about contrast classes in order to proceed in the explanation. Surveying the Hispanic paradox literature, it seems that there are two different contrast classes of alternatives used by Hispanic paradox researchers, $[S_1]$ and $[S_2]$, that could each be invoked when expressing the Hispanic paradox in the format, “Why is the world this way, as opposed to some other way from the set of alternatives [S]?” That is, a single potential version of a Hispanic paradox question, “Why is Hispanic adult mortality similar to non-Hispanic White adult mortality?” has two possible interpretations and two different corresponding types of explanations. The first seeks an account of which causal factors make two populations similar with regard to a single outcome (as opposed to different), while the second interpretation seeks an account of why the two populations are similar with respect to some measure (but not in other measures). Both are legitimate interpretations.

Interpretation 1: If one reads the why question as asking why the *Hispanic population is similar to the non-Hispanic White population with respect to this outcome/measure, instead of being higher or lower* (Contrast Class 1), then it calls for an account of which specific causal factors are able to generate the similar effects in that outcome despite the different social determinants of health present in each population. This type of contrast class interpretation appears, for example, in [8, 21].

Interpretation 2: If one reads the why question as asking why Hispanic and non-Hispanic White populations are *similar in their adult mortality rates, instead of being similar in other health outcomes/metrics* (Contrast Class 2), then it calls for an account of why Hispanic adult mortality rates are similar, while diabetes rates are higher, infant mortality rates are lower, etc. This type of contrast class interpretation appears, for example, in [9, 11].

There is no single “right” way to read the question, no single “right” contrast class to use in answering the question, and hence no single “right” way to answer. All depends on the research interests of the researchers involved and the overall context of the work.

The contrast class ambiguities of how to go about explaining the Hispanic paradox add an additional layer of choice for researchers. Not only are there 108 theoretical variations of the Hispanic paradox, there are also two distinct ways to interpret a request to explain a Hispanic paradox phenomenon. There are then 216 potential ways of explaining 108 epidemiological phenomena. These 216 potential explanations are not meant to be an exhaustive list, nor do they all appear in the literature. Rather, the objective has been to show how distinct Hispanic paradox phenomena can each be explained in multiple distinct ways. This presents a serious obstacle to those who would attempt to solve the Hispanic paradox to the satisfaction of their various colleagues, many of whom will reasonably disagree about what exactly needs to be explained.

Definition and phenomenon choice as obstacles to explaining the Hispanic paradox

Contradictory ideas about the meaning of “Hispanic paradox” directly impede the Hispanic paradox community’s efforts to make sense of the various Hispanic paradox explanations. For example, a 2006 article by David Smith and Benjamin Bradshaw reanalyzes existing US mortality data to correct for demographic classification changes, concluding after the readjustments, “there is no ‘Hispanic paradox’” [22, p. 1686]. They take this position because, after applying their adjusted estimations, “life expectancy [in 2000] was lower for both Hispanic males and Hispanic females than for the non-Hispanic White population, by approximately 0.5 year for females and 1.1 years for males” [22, p. 1689]. Since the article defines the Hispanic paradox as Hispanics having “lower mortality than the non-Hispanic White population,” the authors judge that the purported Hispanic paradox phenomenon was simply a data artifact that has now been revealed to be an illusion, and hence explained away [22, p. 1686]. This explanation is not so successful at addressing the original formulation of the Hispanic paradox in Markides and Coreil’s 1986 article, which posited that Hispanic health is more similar to non-Hispanic White health than to Black health [1]. Smith and Bradshaw use their own definition of the Hispanic paradox (comparing only Hispanics and non-Hispanic Whites) and so have no reason to mention that in 2000, the life expectancy at birth for Blacks was far lower than for Whites: 6.5 years lower for males and 4.8 years lower for females [23]. That is, Smith and Bradshaw believe that they have explained away the Hispanic paradox by showing that the adjusted Hispanic rates are slightly worse than the non-Hispanic White rates (instead of better). Meanwhile the Black rates are far worse than both—this is the very pattern that Markides and Coreil found to be paradoxical in the first place [1]. The year after Smith and Bradshaw’s article was published, Robert Hummer et al. adopted the Markides and Coreil definition and found previous explanations, including Smith and Bradshaw’s, lacking [24].

Hummer et al.’s 2007 article, “Paradox Found (Again),” cites the Smith and Bradshaw article, but insists that the Hispanic paradox has not been explained since they adopt the original 1986 definition of the Hispanic paradox (quoted in full at the

beginning of this article): “Markides and Coreil ... did not define the paradox as *better* health or mortality for Hispanics compared with non-Hispanic whites” [24]. Under that definition, the 2006 Smith and Bradshaw data correction simply reconfirms that there is an unexplained Hispanic paradox. There is still no explanation for why Hispanic life expectancy is just beneath the non-Hispanic White rate while the Black rate is far beneath both; the paradox remains intact. In sum, Smith and Bradshaw reasonably believe they have explained the Hispanic paradox with data showing Hispanic mortality is somewhat worse than non-Hispanic White mortality. Hummer et al., however, reasonably believe that such data cannot explain the paradox since the readjustment still leaves Hispanic mortality closer to non-Hispanic White mortality than to Black mortality (the very phenomenon they believe needs to be explained).

An additional factor in the dispute is that Smith and Bradshaw and Hummer et al. appear to be pursuing explanations of different contrast classes [22, 24]. Smith and Bradshaw seek to explain if and why Hispanic mortality rates are lower than non-Hispanic White rates, as opposed to being similar or higher (contrast class 1, above). After they produce data suggesting the rates are in fact different, they consider the explanatory task accomplished. Hummer et al., however, are interested in Hispanic health broadly, citing a wide array of data but paying particular attention to infant mortality rates, partly because it is unlikely that many newborns are skewing the data by returning to Mexico immediately after birth [24]. Citing the wide array of data is directly relevant (and not just incidental background information) because their explanatory task is to account for why infant mortality rates appear to be slightly lower than those of non-Hispanic Whites, while other health metrics vary in how they compare to non-Hispanic White rates (contrast class 2, above). Hummer et al. remind readers:

... among Mexican American (U.S.-born) women, the statistical parity in infant mortality with non-Hispanic whites observed in the first week after birth disappeared in the later periods of infancy, when the Mexican American women exhibited a moderate disadvantage compared with U.S.-born, non-Hispanic white women. Such patterns of less-favorable Mexican American health over time and across generations in comparison with non-Hispanic whites are consistent with a negative acculturation interpretation of Hispanic health or longer exposure to a less-healthy social environment for the Mexican-origin population compared with non-Hispanic whites. Thus, the most important issue in moving forward in this area of research is not whether an epidemiologic paradox of Mexican-origin infant mortality exists in the United States; it does, at least for now. Rather, the more important issue is whether Mexican-origin health and mortality outcomes will continue to be characterized by parity or near-parity with non-Hispanic whites in a context of continuing social disadvantage.... [24, p. 455]

In sum, the two sets of authors are not only divided by their definitions of the Hispanic paradox, but also their phenomenon choice decisions about contrast classes.

The added layer of phenomenon choice means that explanations have a further way of diverging. Smith and Bradshaw found a curiously low Hispanic mortality rate and sought to explain why it was lower than the non-Hispanic White rate (instead of being higher). They adjusted the mortality data, found the Hispanic rate to be higher than the non-Hispanic White rate, and considered the explanation complete. Hummer et al. looked at the mortality data and not only disagreed with the Hispanic paradox definition (the adjusted mortality data was still paradoxical under their definition) but also disagreed about what it means to explain the Hispanic paradox—they wondered what makes some health outcomes similar in Hispanics and non-Hispanic Whites (e.g., mortality in early infancy), while other outcomes are poorer (e.g., mortality in later infancy) [24].

The definitional disagreements and the diverging phenomenon/contrast class choices constitute one set of challenges for explaining the Hispanic paradox. There is an additional and separate obstacle facing attempts to explain the Hispanic paradox, which is that the very meaning of “Hispanic” is, by design, both ambiguous and heterogeneous as a demographic category. I have shown above that definitions of “Hispanic paradox” vary from one research project to another. But it turns out that the concept “Hispanic” is also ambiguous, for a different set of reasons.

Defining “Hispanic”

Ever since the original 1986 Markides and Coriel article, Hispanic paradox research has struggled with the meaning of “Hispanic” and the measurement of Hispanic public health. In that article, predominantly Mexican American health data was used to make inferences regarding all Hispanics in the US southwest, and that data on southwestern Hispanics, in turn, kindled research into a US-wide Hispanic paradox [1].

“Hispanic origin” has a unique legal status in the United States, as it is the only ethnicity to be a required category in federal demographic data collection guidelines. Its inclusion has had a number of positive impacts on the population it describes. A unified Hispanic community is capable of wielding greater social power than a series of disjointed communities (Mexican immigrants who speak only Spanish, US citizens of Cuban descent who speak only English, etc.). Self-identified Hispanic ethnicity has proven to be a valuable tool for promoting social justice. For example, a recent review of rates of high school non-completion between racial and ethnic groups found that Hispanics had by far the worst rate: 37.7% non-completion by age 25 [25, p. 10–11]. Such information can provide a starting point for investigations of subgroup patterns (e.g., how non-completion rates differ for those educated in the US vs. in other countries), differences between regions of the US, and so on. Those follow-up data, in turn, make it possible to promote more equitable education policies. “Hispanic”-level data collection is a necessary step in this process. “Hispanic” is a valuable concept in social policy. However, its history points to some of its weaknesses as a variable in epidemiological research.

G. Cristina Mora provides a history of the Hispanic concept in her 2014 book, *Making Hispanics: How Activists, Bureaucrats, and Media Constructed a New American*. The book lays out the way the curious term evolved. One notable feature is that “Hispanic” is not quite an ethnic concept. It is a panethnic concept, akin to “Asian American,” since it is an aggregate concept that lumps together a diverse collection of distinct ethnic identities that are rooted in a wide range of places, cultures, and populations (Spain, the Caribbean, South America, etc.). During its development in the 1960s and 1970s, the Hispanic panethnic concept was conceived so broadly, so inclusively, that it left supporters struggling to articulate its conceptual coherence or even its basic definition [26].

In effect, Hispanic panethnicity became institutionalized over time as activists, bureaucrats, and media executives forged a new field centered around the new category... Ambiguity was a critical element of this new Hispanic field. Activists, media executives, and census officials never really defined who Hispanics were, nor did they argue definitively that characteristics like language, place of birth, or surname made Hispanics Hispanic. Instead, they reiterated that, above all, Hispanics were Hispanic because they shared a common set of values and a common culture. The stakeholders used descriptors like *hardworking*, *religious*, and *family-oriented*—adjectives that could be applied to any group—to describe the unique characteristics uniting Hispanics. [26, p. 156]

“Hispanic” was federally codified in 1976 [26, p. 98] and continues to appear in places such as the 2010 census, according to which, “‘Hispanic or Latino’ refers to a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race” [4]. As a result of such a broad characterization of the panethnicity, “Hispanic” lumps together foreign-born and US-born, South Americans and Spaniards, descendants of indigenous populations, descendants of African populations, and descendants of European populations, etc.

Once established, socially constructed demographic categories can themselves cause health effects. Most directly, being identified as a member of a certain demographic group can lead to implicit biases and discrimination in medical care. Such biases seem to be one of the factors responsible for the widespread under-prescription of opioid pain medication (morphine, oxycodone, etc.) to Hispanic patients—even those who report severe pain [27].

The history and application of the Hispanic concept as an obstacle to explaining the Hispanic paradox

Even if one sets aside concerns about the fuzzy definition and heterogeneous population of “Hispanic,” there are additional challenges stemming from the measurement of Hispanic populations. An Institute of Medicine report notes that federal guidelines advocate self-reporting as the best method for race and ethnicity data collection partly because it “respects individual dignity,” and claims that researchers consider self-reports to be the “gold standard” [28, p. 134; 29].

However, other ethnicity researchers disagree, either denying that any gold standard exists [30] or pointing out that a person's unprompted way of self-identifying is a more plausible gold standard than how one responds to pre-determined questions on an official form [31]. This issue came to the forefront in a heated 2002 debate between Kenneth Smith [32] and Modood et al. [33] regarding (among other issues) whether or not one's ethnicity is inherently a matter of self-identification, or whether self-identification is just a good method for measuring ethnicity.

These are not exclusively theoretical questions, as they affect the reliability of any "Hispanic" data. As M. Anne Visser notes, since many people with Latin American backgrounds prefer to self-identify by country of origin (e.g., Mexican American) and reject the "Hispanic" label, the Hispanic concept has, in a sense, undercounted the population it was intended to count [34]. This undercounting, in turn, leads to inequitable distribution of resources [34]. On the other hand, if one is Hispanic by virtue of self-identifying as such, then someone who identifies as Mexican American and denies being Hispanic cannot be truly undercounted by the Hispanic concept.

Research on Hispanics necessitates taking stances on difficult philosophical questions. If a longitudinal study of Americans of Mexican descent finds that some alternate between embracing and denying the terms "Mexican American" origin or "Hispanic" ethnicity (which indeed happens regularly), then what does that mean for Hispanic paradox data [35]? If, one day, a person stops identifying as Hispanic, is such a person then automatically a non-Hispanic [35]? Under which ethnicity column should experts record that person's data when he or she dies? No matter how researchers answer such questions, it affects crucial Hispanic paradox data such as mortality counts for Hispanics. How one understands and operationalizes "Hispanic" impacts the explanatory power of many of the data error explanations of the Hispanic paradox (one of the four types of Hispanic paradox explanation) [36]. The Smith and Bradshaw article, discussed above, uses its re-estimation techniques to cope with methodological gaps in the Hispanic data record, which includes datasets of Hispanics as identified by self-reporting, datasets of Hispanics as identified by an observer (e.g., "funeral home personnel"), and datasets of Hispanics as identified by having a surname that is associated with Hispanics [22]. The need to make non-obvious choices about how to operationalize "Hispanic" and the fact that Hispanic is an especially heterogeneous panethnic category that includes a wide range of health profiles, together make it unclear whether it is accurate to call it a *Hispanic* paradox in the first place.

Is the Hispanic paradox really a Hispanic phenomenon? The name certainly denotes this, as do each of the definitions in Table 1. However, as reviewed above, the Hispanic panethnic label includes a diverse collection of populations. Most notably, US Hispanics include both US-born and foreign-born people, and it is not clear whether both groups experience the Hispanic paradox, or to what extent if they do. Or, if not, then the Hispanic paradox term is misleading (see a similar suggestion in [37]). Moreover, two of the primary types of explanations for the Hispanic paradox (the healthy migrant effect and the return migration effect) are about who migrates to/from the US Hispanic population, which means that some explanations' credibility will entirely hinge on whether foreign-born, US-born, or both

subpopulations experience the effect. Unfortunately, the data are decidedly mixed on the topic of whether the Hispanic paradox effect extends to US-born Hispanics, in which health measures, and to what degree (e.g., see contrasting data in [8, 24, 37, 38]).

Using broad ethnic or racial identifiers risks concealing the heterogeneity of the people labeled by those identifiers. For example, Raj Bhopal has critiqued the use of “Asian” in public health due to the enormous heterogeneity between the different populations included therein [39]. Valles has similarly critiqued the use of “African American” and “White” for their uses in contexts where within-population heterogeneity is already known—hypertension in the former case and cystic fibrosis in the latter [40]. For example, “African American” was used in a set of federal dietary salt guidelines even though “US-born African American” would have easily bracketed off the drastically healthier subpopulation of foreign-born African descendants. The decision to define the observed epidemiological paradox as a *Hispanic* paradox raises the possibility of similar problems.

Within both the foreign-born and US-born Hispanic populations there appears to be a great deal of variation between populations of different national origins, though these data are limited since those of Mexican origin, and sometimes Cuban and Puerto Rican origin, tend to get the most subgroup attention while other Hispanics are often lumped into regional or “other” categories [8, 24]. This is consistent with 2010 census practices, in which these same groups get separate tick-boxes for self-identification in the Hispanic ethnicity section, while other Hispanics are instructed to provide write-in responses regarding their specific Hispanic origin [4].

Until the precise boundaries of the Hispanic paradox effect are well established, debates over the explanation will have a weak empirical basis. It is not enough to know that there is an average benefit of one or more kinds to the Hispanic population. If it is simply the case that Hispanics are healthier because foreign-born people are healthier, and Hispanics include a large number of foreign-born people, then that suggests one set of follow-up research and candidate explanations (including investigating migration and acculturation behaviors). If it is the case that only Hispanics with ancestry from certain countries benefit, then it suggests another set of follow-up research and candidate explanations (including investigating cultural practices in the relevant countries). A recent *Lancet* editorial, summarizing a CDC report on Hispanic health data, concludes that in future Hispanic healthcare research and practice, “interventions should be personalised to account for original ethnic origin and birth location.... The USA cannot ignore the health of Hispanic people and, as the spotlight falls on patient-centred medicine, it is important to recognise the health differences and needs of subpopulations in all societies worldwide” [41].

Multivalent concepts and “explanatory divides” in biomedicine

The aforementioned sorts of challenges share characteristics with other cases previously discussed in the philosophical literature. The similarities and differences between these cases and the Hispanic paradox case help to identify what is most challenging about Hispanic paradox research and why.

In a case analogous to the Hispanic paradox case, Jacqueline Sullivan has explored how the “Morris Water Maze” apparatus has been used to investigate a number of distinct phenomena, with researchers fundamentally disagreeing about what precisely is being investigated in water maze experiments [42]. When a rat is placed into the apparatus—a pool of water surrounded by assorted objects and escape platforms that are adjusted to test how the rat navigates the space—it is not clear exactly what phenomenon is under investigation: “Morris water maze performance,” “spatial memory,” “spatial navigation,” etc. [42, p. 262]. This multiplicity of target phenomena is an obstacle to creating accounts of the neurological mechanisms that generate the (ambiguous) phenomenon [42]. It is interesting that such divergences can appear even when interpreting a simple laboratory device and setup, making it seem less persuasive to attribute the Hispanic paradox divergences to the limitations of researching human populations outside of controlled lab settings.

The obstacles in the murky Hispanic paradox literature become clearer by examining the similarities and differences between the Hispanic paradox case and an analogous behavioral genetics case explored in James Tabery’s 2014 book on the nature-nurture debates, *Beyond Versus*. While meta-analyses are typically used to settle empirical disputes, Tabery shows in great detail how medical meta-analyses are limited by their foundational assumptions regarding the variables investigated. He demonstrates this in the case of research on the relationship between the 5-HTTLPR gene, stress in one’s environment, and psychological traits related to anxiety. Two camps have emerged since the early 2000s, each touting meta-analysis results that suggest opposite evaluations of how the variables are related. As Tabery shows, the resulting “dueling meta-analyses” have quite different criteria for which studies to include during meta-analysis, with one camp restricting its meta-analysis to “direct replications” of the original study on the topic, while the other camp includes a much larger set of studies on the subject, including “indirect replications” studying the same subject but with different research designs [43, pp. 87–91].

Tabery ultimately traces these differing study inclusion criteria to two longstanding theoretical traditions in biology. One tradition aims to understand *how much* population variation in some trait (e.g., depression) is attributable to a factor (e.g., a gene), and another tradition aims to understand *how* biological mechanisms give rise to a trait [43]. Since each tradition asks different questions (“how much” vs. “how”), the result is an “explanatory divide” between the camps. The divergent definitions of “Hispanic paradox” can also be understood as a type of “explanatory divide,” though with illuminating dissimilarities.

In his case study, Tabery finds “dueling meta-analyses” reaching different conclusions about the same literature. By contrast, the Hispanic paradox literature’s two 2013 meta-analyses (see Table 1) were released nearly simultaneously—neither cites the other—and they choose different phenomena to investigate [3, 12]. The two meta-analyses do not duel; both conclude that the Hispanic paradox does exist, but they ultimately affirm substantially different hypotheses using almost entirely different datasets. One dataset contains studies of Hispanic morality [3], and the other contains studies of Hispanic cardiovascular morbidity and mortality [12]. The two meta-analyses agree on comparing Hispanics to non-Hispanic Whites, but disagree whether Hispanic health is better or whether it is “similar or better.” They

also disagree about the scope of the phenomenon. One meta-analysis takes it to be a broad phenomenon involving many health outcomes (but only tests mortality effects) and the other restricts it to only cardiovascular phenomena.

Comparing the two Hispanic paradox meta-analyses to Tabery's, his meta-analyses are separated by the strictness of their inclusion criteria (one has 54 studies included, while another has 14 since it only allows exact replications of the study design). Analogously, the Hispanic paradox meta-analyses are also split in their specificity—one is on cardiovascular mortality [12] and the other is on mortality in any context [3]. Hence, they also have different sizes: the former has only 18 articles included and the latter has 58. But, the Hispanic paradox meta-analyses overlap by only a single study [44]. They come close at other points—e.g., each cites a different 1996 Wei et al. publication based on the San Antonio Heart Study—but it is striking to see that two systematic reviews of Hispanic paradox mortality, one narrow and one broad, only share a single point of overlap. Intuitively, one would expect the broad Ruiz et al. article to include all of the studies included in the Cortes-Bergoderi et al. article, in addition to many more studies.

The result of the large and small differences between the two Hispanic paradox systematic reviews/meta-analyses is that they ask different questions, compile almost entirely different surveys of the published literature, and end up “agreeing” in the hollow sense of both concluding the Hispanic paradox exists, but meaning different things in each case. Tabery's case illustrates how an “explanatory divide” can separate different lines of research on the basis of distinct theoretical traditions, which then manifests methodological differences. But there does not appear to be any analogous theoretical or ideological division in the Hispanic paradox case. The Hispanic paradox case shows that the accumulation of small discrepancies between research terms' research designs can create similar explanatory divides even without any theoretical dispute actively pushing those discrepancies apart.³ The different research teams exploring diverging definitions of the Hispanic paradox seem to have simply drifted apart, rather than intentionally sailing toward different destinations.

The Morris water maze and 5-HTTLPR cases serve as an illustration, and warning to Hispanic paradox researchers, of how implicit disagreements about explanatory tasks do not automatically resolve themselves. Productive dialogue is undermined by the opaque and unspoken reasons for the disagreements. The creation of the Morris water maze, as a shared technical apparatus, has served to organize observations of experimental rat behavior, but also left subsequent researchers to struggle with lingering ambiguities about which precise rat mental representations and cognitive processes are actually tested with the apparatus. And, like in the Hispanic paradox case, the ambiguity has continued even as the line of research has developed over the decades—decades of time and close attention have not been enough to resolve the ambiguity. The 5-HTTLPR case offers a similar cautionary tale to Hispanic paradox researchers, since it shows that even meta-analyses' structure, precision, and wide scopes are not necessarily sufficient to resolve disagreements about how to explain disputed phenomena (i.e., “explanatory divides,” in Tabery's phrasing). Together, the cases show that the explanatory

³ See [45] and [46] for additional discussion of the evidential limitations of meta-analyses.

challenges faced by Hispanic paradox researchers fit into a recent philosophical literature analyzing biomedical cases with similar challenges, with all three cases showing that subtle and implicit explanatory disputes can nevertheless make very large impacts on the literature. Despite the Hispanic paradox case's similarities to these other cases, the Hispanic paradox case also has some ethical features that make it unique and worthy of special attention.

What makes the Hispanic paradox case unique

The Morris Water Maze, 5-HTTLPR behavioral genetics, and the Hispanic paradox have some structural similarities in their debates, but the debates each have different social constraints and consequences. All three are cases where subtle disagreements about the nature of the phenomenon being explained have led to confusion among researchers. Ethically, though, the Hispanic paradox case is distinguished by its status as a dispute over the health of a large minority population. There are thus equity and social justice issues at stake in Hispanic paradox research. To understand how these ethical issues affect and distinguish the Hispanic paradox case, one can turn to the principles of the 2013 Leeds Consensus. The Consensus provides ten principles for appropriate ethnicity and health research, generated during a consensus building exercise conducted with an international group of leaders in the field of ethnicity and health research. The exercise revealed that experts in the field share a set of commitments to certain ethical principles.

The first Leeds Consensus principle states, "the purpose of research on ethnicity and health should be for the well-being and betterment of populations being studied" [47, p. 507]. In a 2005 update on the Hispanic paradox, Markides (co-author of the seminal 1986 article) et al. warn that research on the Hispanic paradox creates risks for Hispanics, since public awareness of the Hispanic paradox could be used as a justification for taking away healthcare resources from Hispanics, who suffer from "remediable disparities in health care access and the burden of infectious diseases, diabetes, and disability," even if there are some other Hispanic paradox benefits [48, p. 74]. These particular risks to vulnerable Hispanic populations make Hispanic paradox research unusual. Research on minority populations' health typically involves evidence of poor health, which introduces risks of health-related stigma. Markides et al.'s warning is a vital reminder that Hispanic paradox research creates its own set of risks through its potential to bluntly misrepresent Hispanics as being unworthy of health resources.

Markides et al.'s warning [48] also helps to guide the application of a second Leeds Consensus principle, which states: "Equity should be the guiding ethical principle for ethnic health research; researchers must be alert to the dangers of discriminatory thinking and behaviour and guard against actual and potential harm" [47, p. 507]. Ethnic health disparities research and social determinants of health research are driven by a commitment to equity. A variety of scholars have pointed to the importance of particular historical factors in establishing the foundations of social epidemiology and of contemporary understandings of social determinants of health. These factors include the 1840s' social justice movements (Marxism, labor reform, etc.) and their theoretical

leaders (Rudolph Virchow, Edwin Chadwick, etc.) [49, 50], the leadership of the World Health Organization since its founding in the 1940s [51], and the work of epidemiologists in the last few decades, such as that of John Cassel and Mervyn Susser [49], and more. Whatever the origins of epidemiology's focus on investigating social determinants of health, it is now deeply embedded within epidemiology, including the WHO and member state commitments to the issue as a component of promoting health equity [52]. Hispanic paradox researchers are divided regarding how to manage the methods and concepts in their research, but at the very least the Hispanic paradox research community is tightly united in its ultimate goal of promoting health equity and addressing the social determinants of health. A shared set of broad goals and priorities creates a sort of unity in the Hispanic paradox community even as it struggles with the explanatory divides generated by differing definitions and interpretations of explanatory tasks.

A third Leeds principle helps to make sense of the Hispanic paradox case's unique stakes and constraints. It states: "Ethnicity is significantly correlated with disadvantage and ill-health and researchers in the field of health inequalities have both a professional and ethical responsibility to incorporate evidence on ethnicity into their work and recommendations" [47, p. 507]. In many life cases of biology research, the simplest response to a problematic population concept is to simply abandon the population concept. That is not an acceptable option in the case of the Hispanic concept. In the absence of any other routine mechanisms for monitoring the health of US population now called "Hispanic," public health researchers are stuck with the Hispanic concept. It is no contradiction that, as stated above, it would also be desirable to supplement Hispanic panethnicity data with Colombian ethnicity data, etc. The need for additional fine-grained data does not change the fact that there is a massive volume of Hispanic data due to the US legal requirement to collect such demographic data alongside racial data. The available data reveals a number of health disparities, including troublingly high rates of infectious diseases and diabetes [48]. In the face of that data, the Leeds Consensus demands that researchers address the data and build upon it. It would be unethical to just discard Hispanic data because of its frustrating fuzziness and heterogeneity.

Choices regarding how to define, research, and explain the Hispanic paradox have very real consequences for public health. As Katikireddi and Valles argue, even something as simple as the choice of how to define a variable in epidemiological research is an act of great ethical and epistemic significance for the communities described by that variable [53]. Similar concerns lead Michael Root to reject race as a proxy variable for genetic variations in patients [54]. The Hispanic paradox disputes are in many ways very esoteric disputes about epidemiological and social scientific methodology, yet those nuances have enormous direct and indirect consequences for Hispanic social justice and public health ethics.

Conclusion

Researchers have spent several decades unable to explain the Hispanic paradox to one another's satisfaction. Articles offer arguments, counter-arguments, alternative framings, and so on, never settling what ought to be explained or how it ought to be

explained. Meanwhile, the Hispanic population's health needs cannot be properly met because it remains quite unclear what those needs are.

This article has used philosophical tools to elucidate two sets of conceptual and methodological obstacles that have been impeding efforts to explain the Hispanic paradox. One set of problems arises from the fact that "Hispanic paradox" means different and sometimes contradictory things to different researchers. Meanwhile, in a second set of obstacles, the Hispanic paradox field continues to struggle with data on a demographic category that was designed to be heterogeneous and fuzzy. It is essential that public health researchers resolve the Hispanic paradox because the existence (or not) of the paradox and its boundaries are vital pieces of information for promoting equitable healthcare policies for Hispanics.

The Leeds Consensus ethical principles serve as valuable guides for doing future research [47]. One, Hispanic paradox research should be done first and foremost as a means of benefitting the Hispanic population; the Hispanic paradox offers a lens for examining health disparities research more broadly, but Hispanics' interests must remain paramount. Two, discrimination and inequity are a constant danger in research on the Hispanic paradox. Markides et al. warn that a simplistic belief that Hispanics are healthy could be wielded against them by using it to justify further decreases in their healthcare services [48]. Politicians already trade in innuendo about foreign-born Hispanics being diseased, using it to justify harsher immigration policies [55]. There is particular risk that any Hispanic paradox findings can be wielded by xenophobes or racists to justify harming Hispanics. Three, "Hispanic" is a vexingly broad and imprecise panethnic concept, and the Hispanic paradox phenomenon may only apply to some subpopulations within it. But there is an ethical obligation to investigate known health disparities in ethnicity data, and it would be unethical to ignore the wealth of existing Hispanic data due to the Hispanic concept's frustrating qualities.

Three decades of data have made it abundantly clear that Hispanic health data yield inexplicable patterns on a number of different health measures. These patterns must be explained. The public health of the US's largest racial or ethnic minority group must not remain a mystery. This article has sought to make a small contribution to resolving that mystery by shining a light on some of the factors that render it so hard to explain the Hispanic paradox.

Acknowledgements I am grateful to Raj Bhopal and Jamie Nelson for their comments on this article. I am also grateful for input on the project from: Matthew Haber, James Tabery, the audience members of the University of Utah Department of Philosophy colloquium series, the attendees of the 2013 Philosophy of Medicine Roundtable, the 2014 meeting of the American Philosophical Association-Eastern, and the 2014 meeting of the Philosophy of Science Association.

References

1. Markides, Kyriakos S., and Jeannine Coreil. 1986. The health of hispanics in the Southwestern United States: An epidemiologic paradox. *Public Health Reports* 101(3): 253–265.
2. Waldstein, Anna. 2010. Popular medicine and self-care in a Mexican migrant community: Toward an explanation of an epidemiological paradox. *Medical Anthropology* 29(1): 71–107.

3. Ruiz, John M., Patrick Steffen, and Timothy B. Smith. 2013. Hispanic mortality paradox: A systematic review and meta-analysis of the longitudinal literature. *American Journal of Public Health* 103(3): e52–e60.
4. Humes, Karen R., Nicholas A. Jones, and Roberto R. Ramirez. 2011. *Overview of race and Hispanic origin: 2010*. Washington, DC: United States Census Bureau.
5. Karno, Marvin, and Robert B. Edgerton. 1969. Perception of mental illness in a Mexican-American community. *Archives of General Psychiatry* 20(2): 233–238.
6. Auger, N., Z.-C. Luo, R.W. Platt, and M. Daniel. 2008. Do mother's education and foreign born status interact to influence birth outcomes? Clarifying the epidemiological paradox and the healthy migrant effect. *Journal of Epidemiology and Community Health* 62(5): 402–409.
7. Hayes-Bautista, David E., Paul Hsu, Maria Hayes-Bautista, Delmy Iñiguez, Cynthia L. Chamberlin, Christian Rico, and Rosa Solorio. 2002. An anomaly within the Latino epidemiological paradox: The Latino adolescent male mortality peak. *Archives of Pediatrics and Adolescent Medicine* 156(5): 480–484.
8. Borrell, Luisa N., and Elizabeth A. Lancet. 2012. Race/ethnicity and all-cause mortality in US adults: Revisiting the Hispanic paradox. *American Journal of Public Health* 102(5): 836–843.
9. Willey, Joshua Z., Carlos J. Rodriguez, Yeseon Park Moon, et al. 2012. Coronary death and myocardial infarction among Hispanics in the Northern Manhattan Study: Exploring the Hispanic paradox. *Annals of Epidemiology* 22(5): 303–309.
10. Espinoza, Sara E., Inkyung Jung, and Helen Hazuda. 2013. The Hispanic paradox and predictors of mortality in an aging biethnic cohort of Mexican Americans and European Americans: The San Antonio Longitudinal Study of Aging. *Journal of the American Geriatrics Society* 61(9): 1522–1529.
11. Hayward, Mark D., Robert A. Hummer, Chi-Tsun Chiu, César González-González, and Rebeca Wong. 2014. Does the Hispanic paradox in US adult mortality extend to disability? *Population Research and Policy Review* 33(1): 81–96.
12. Cortes-Bergoderi, Mery, Kashish Goel, Mohammad Hassan Murad, et al. 2013. Cardiovascular mortality in Hispanics compared to non-Hispanic whites: A systematic review and meta-analysis of the Hispanic paradox. *European Journal of Internal Medicine* 24(8): 791–799.
13. Palloni, Alberto, and Jeffrey D. Morenoff. 2001. Interpreting the paradoxical in the Hispanic paradox. *Annals of the New York Academy of Sciences* 954(1): 140–174.
14. Turra, Cassio M., and Noreen Goldman. 2007. Socioeconomic differences in mortality among US adults: Insights into the Hispanic paradox. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 62(3): S184–S192.
15. Abraido-Lanza, Ana F., Bruce P. Dohrenwend, Daisy S. Ng-Mak, and J. Blake Turner. 1999. The Latino mortality paradox: A test of the “salmon bias” and healthy migrant hypotheses. *American Journal of Public Health* 89(10): 1543–1548.
16. Turra, Cassio M., and Irma T. Elo. 2008. The impact of salmon bias on the Hispanic mortality advantage: New evidence from social security data. *Population Research and Policy Review* 27(5): 515–530.
17. Eschbach, Karl, Glenn V. Ostir, Kushang V. Patel, Kyriakos S. Markides, and James S. Goodwin. 2004. Neighborhood context and mortality among older Mexican Americans: Is there a barrio advantage? *American Journal of Public Health* 94(10): 1807–1812.
18. Arias, Elizabeth, Karl Eschbach, William S. Schauman, Eric L. Backlund, and Paul D. Sorlie. 2010. The Hispanic mortality advantage and ethnic misclassification on US death certificates. *American Journal of Public Health* 100(S1): S171–S177.
19. Valles, Sean A. 2010. The mystery of the mystery of common genetic diseases. *Biology and Philosophy* 25(2): 183–201.
20. van Fraassen, Bas C. 1990. *The scientific image. Clarendon library of logic and philosophy*. New York: Oxford University Press.
21. Fenelon, Andrew. 2013. Revisiting the Hispanic mortality advantage in the United States: The role of smoking. *Social Science and Medicine* 82: 1–9.
22. Smith, David P., and Benjamin S. Bradshaw. 2006. Rethinking the Hispanic paradox: Death rates and life expectancy for US non-Hispanic White and Hispanic populations. *American Journal of Public Health* 96(9): 1686–1692.
23. Arias, Elizabeth. 2015. United States life tables, 2011. *National Vital Statistics Reports* 64(11): 1–62.
24. Hummer, Robert A., Daniel A. Powers, Starling G. Pullum, Ginger L. Gossman, and W. Parker Frisbie. 2007. Paradox found (again): Infant mortality among the Mexican-origin population in the United States. *Demography* 44(3): 441–457.

25. Beckles, Gloria L., and Benedict I. Truman. 2013. Education and income—United States, 2009 and 2011. *Morbidity and Mortality Weekly Report* 62(3, suppl.): 9–19.
26. Mora, G.Cristina. 2014. *Making Hispanics: How activists, bureaucrats, and media constructed a new American*. Chicago: University of Chicago Press.
27. Pletcher, Mark J., Stefan G. Kertesz, Michael A. Kohn, and Ralph Gonzales. 2008. Trends in opioid prescribing by race/ethnicity for patients seeking care in US emergency departments. *Journal of the American Medical Association* 299(1): 70–78.
28. Ulmer, Cheryl, Bernadette McFadden, and David R. Nerenz. 2009. *Race, ethnicity, and language data: Standardization for health care quality improvement*. Washington, DC: The National Academies Press.
29. Office of Management and Budget. 1997. Revisions to the standards for the classification of federal data on race and ethnicity. *Federal Register* 62(210): 58782–58790.
30. Callister, Paul, Robert Didham, Deborah Potter, and Tony Blakely. 2007. Measuring ethnicity in New Zealand: Developing tools for health outcomes analysis. *Ethnicity and Health* 12(4): 299–320.
31. Aspinall, P.J. 2011. The utility and validity for public health of ethnicity categorization in the 1991, 2001 and 2011 British Censuses. *Public Health* 125(10): 680–687.
32. Smith, Kenneth. 2002. Some critical observations on the use of the concept of ‘ethnicity’ in Modood et al., ethnic minorities in Britain. *Sociology* 36(2): 399–417.
33. Modood, Tariq, Richard Berthoud, and James Nazroo. 2002. ‘Race’, racism and ethnicity: A response to Ken Smith. *Sociology* 36(2): 419–427.
34. Visser, M.Anne. 2012. Two plus two equals three: Classification error and the Hispanic undercount in United States census surveys. *American Review of Public Administration* 44(2): 233–251.
35. Alba, Richard, and Tariqul Islam. 2009. The case of the disappearing Mexican Americans: An ethnicity mystery. *Population Research and Policy Review* 28(2): 109–121.
36. Lariscy, Joseph T. 2011. Differential record linkage by Hispanic ethnicity and age in linked mortality studies: Implications for the epidemiologic paradox. *Journal of Aging and Health* 23(8): 1263–1284.
37. Lee, Min-Ah, and Kenneth F. Ferraro. 2007. Neighborhood residential segregation and physical health among Hispanic Americans: Good, bad, or benign? *Journal of Health and Social Behavior* 48(2): 131–148.
38. Palloni, Alberto, and Elizabeth Arias. 2004. Paradox lost: Explaining the Hispanic adult mortality advantage. *Demography* 41(3): 385–415.
39. Bhopal, R.S. 2006. Race and ethnicity: Responsible use from epidemiological and public health perspectives. *Journal of Law, Medicine and Ethics* 34(3): 500–507.
40. Valles, Sean A. 2012. Heterogeneity of risk within racial groups, a challenge for public health programs. *Preventive Medicine* 55(5): 405–408.
41. The Hispanic paradox. 2015. *The Lancet* 385(9981): 1918.
42. Sullivan, Jacqueline A. 2010. Reconsidering ‘spatial memory’ and the Morris water maze. *Synthese* 177(2): 261–283.
43. Tabery, James. 2014. *Beyond versus: The struggle to understand the interaction of nature and nurture*. Cambridge, MA: MIT Press.
44. Swenson, Carolyn J., Mary Jo Trepka, Marian J. Rewers, Sharon Scarbro, William R. Hiatt, and Richard F. Hamman. 2002. Cardiovascular disease mortality in Hispanics and non-Hispanic whites. *American Journal of Epidemiology* 156(10): 919–928.
45. Borgerson, Kirstin. 2009. Why reading the title isn’t good enough: An evaluation of the 4S approach to evidence-based medicine. *International Journal of Feminist Approaches to Bioethics* 2(2): 152–175.
46. Stegenga, Jacob. 2011. Is meta-analysis the platinum standard of evidence? *Studies in History and Philosophy of Biological and Biomedical Sciences* 42(4): 497–507.
47. Mir, Ghazala, Sarah Salway, Joe Kai, Saffron Karlsen, Raj Bhopal, George T.H. Ellison, and Aziz Sheikh. 2013. Principles for research on ethnicity and health: The Leeds Consensus Statement. *The European Journal of Public Health* 23(3): 504–510.
48. Markides, Kyriakos S., and Karl Eschbach. 2005. Aging, migration, and mortality: Current status of research on the Hispanic paradox. *Journals of Gerontology: Series B, Psychological sciences and social sciences* 60(spec. 2): 68–75.
49. Berkman, Lisa F., and Ichiro Kawachi. 2000. A historical framework for social epidemiology. In *Social epidemiology*, ed. Lisa F. Berkman, and Ichiro Kawachi. New York: Oxford University Press.

50. Krieger, Nancy, and Anne-Emanuelle Birn. 1998. A vision of social justice as the foundation of public health: Commemorating 150 years of the spirit of 1848. *American Journal of Public Health* 88(11): 1603–1606.
51. Irwin, Alec, and Emile Scali. 2007. Action on the social determinants of health: A historical perspective. *Global Public Health* 2(3): 235–256.
52. World Health Organization. 2013. What are social determinants of health? http://www.who.int/social_determinants/sdh_definition/en/. Accessed June 27, 2014.
53. Katikireddi, S. Vittal., and Sean A. Valles. 2015. Coupled ethical–epistemic analysis of public health research and practice: Categorizing variables to improve population health and equity. *American Journal of Public Health* 105(1): e36–e42.
54. Root, Michael. 2003. The use of race in medicine as a proxy for genetic differences. *Philosophy of Science* 70(5): 1173–1183.
55. Kessler, Glenn. 2015. Is there a link between the measles outbreak and illegal immigration? *The Washington Post*. <http://www.washingtonpost.com/blogs/fact-checker/wp/2015/02/06/is-there-a-link-between-the-measles-outbreak-and-illegal-immigration/>. Accessed Oct 11, 2015.