Social Science & Medicine 188 (2017) 157–165

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Is structural stigma's effect on the mortality of sexual minorities robust? A failure to replicate the results of a published study

Mark Regnerus [a, b, *]

[a] Department of Sociology, University of Texas at Austin, 305 E 23rd St, A1700, Austin, TX 78712-1086, USA
[b] Austin Institute for the Study of Family and Culture, 2021 Guadalupe St., Suite 260, Austin, TX 78705, USA

ARTICLE INFO

Article history:
Received 11 April 2016
Received in revised form
1 November 2016
Accepted 11 November 2016
Available online 14 November 2016

Keywords:
United States
Prejudice
Stigma
Sexual orientation
Mortality
Replication
Data imputation
Scientific transparency

ABSTRACT

Background: The study of stigma's influence on health has surged in recent years. Hatzenbuehler et al.'s (2014) study of structural stigma's effect on mortality revealed an average of 12 years' shorter life expectancy for sexual minorities who resided in communities thought to exhibit high levels of anti-gay prejudice, using data from the 1988–2002 administrations of the US General Social Survey linked to mortality outcome data in the 2008 National Death Index.

Methods: In the original study, the key predictor variable (structural stigma) led to results suggesting the profound negative influence of structural stigma on the mortality of sexual minorities. Attempts to replicate the study, in order to explore alternative hypotheses, repeatedly failed to generate the original study's key finding on structural stigma. Efforts to discern the source of the disparity in results revealed complications in the multiple imputation process for missing values of the components of structural stigma. This prompted efforts at replication using 10 different imputation approaches.

Results: Efforts to replicate Hatzenbuehler et al.'s (2014) key finding on structural stigma's notable influence on the premature mortality of sexual minorities, including a more refined imputation strategy than described in the original study, failed. No data imputation approach yielded parameters that supported the original study's conclusions. Alternative hypotheses, which originally motivated the present study, revealed little new information.

Conclusion: Ten different approaches to multiple imputation of missing data yielded none in which the effect of structural stigma on the mortality of sexual minorities was statistically significant. Minimally, the original study's structural stigma variable (and hence its key result) is so sensitive to subjective measurement decisions as to be rendered unreliable.

1. Introduction

Researchers have successfully documented associations between social stigma toward sexual minorities and the experience of adverse health outcomes among them (Bostwick, 2012; Hatzenbuehler et al., 2009). Stigma, however, is not simple to define or operationalize, prompting measurement challenges that make it difficult to assess just how influential stigma is on health outcomes. Measurement difficulties, moreover, make it harder to develop broad confidence in conclusions across studies. In his widely disseminated and discussed manuscript on the poor validity of most published research findings, Ioannidis (2005: 698) cites the "flexibility in designs, definitions, outcomes, and analytical modes" as well as the relative popularity of a particular research subject as two key factors apt to weaken confidence in published research findings and elevate the risk of scientific missteps. This, together with the rapid expansion of publication outlets and pressure to publish, has contributed to a surge in scientific overstatements, errors, accusations of fabrications, and the issuing of errata or retractions, as well as a renewed call for greater transparency across the research process (Cumming, 2013; Ioannidis, 2008; Simmons et al., 2011).

Together with five co-authors, Mark Hatzenbuehler analyzed data from the 1988–2002 survey administrations of the General Social Survey (GSS), linked to mortality outcome data in the 2008 National Death Index (NDI). That study revealed dramatically shorter life expectancy—approximately 12 years—for sexual

* Department of Sociology, University of Texas at Austin, 305 E 23rd St, A1700, Austin, TX 78712-1086, USA.
E-mail address: regnerus@prc.utexas.edu.

http://dx.doi.org/10.1016/j.socscimed.2016.11.018

minorities who resided in communities believed to exhibit high levels of anti-gay prejudice, even after controlling for a variety of demographic and health-related indicators. Their findings were published in this journal in 2014 in a special volume on structural stigma and health that Hatzenbuehler co-edited.

In the present study, the same GSS-NDI linked data is reanalyzed in order first to replicate—and then to assess alternative explanations for—the findings in the original study of structural stigma and all-cause mortality in sexual minority populations. Given that the GSS and NDI are publicly-accessible datasets, this approach seemed reasonable, feasible, and a scientific value, especially when the original study posed such notable findings. However, after initial attempts to replicate the original study's key result about the influence of social stigma on premature mortality failed—and efforts to obtain more information from the first author about their decisions concerning the imputation of missing data on the stigma measures were unsuccessful—a variety of focused attempts at replication were undertaken, with no success. The results of these efforts are reported herein.

1.1. Background

Anti-gay stigma, as the original study's authors and others have pointed out using diverse data sources, is often found to be corrosive to the mental and physical health of sexual minorities (Hatzenbuehler, 2009; Herek and Garnets, 2007; Meyer, 2003). In the original study under scrutiny here, Hatzenbuehler and his co-authors note that while researchers have believed *structural* stigma to be harmful to individuals' health, few have been able to adequately construct and test a contextual measure of such stigma. Indeed, they note "little or no variation to study" in previous attempts, given "the pervasiveness of structural stigma" in American communities (Hatzenbuehler et al., 2014: 34).

The original study cites scholarly support for the observation that sexual minorities live in social environments that vary widely in their support for gays and lesbians, and notes evidence suggesting that higher rates of contextual stigma, such as state-level amendments prohibiting same-sex marriage, are associated with elevated experience of adverse psychological disorders and attempted suicide (Hatzenbuehler, 2011; Hatzenbuehler et al., 2009). The key research question they pose in their 2014 publication is whether such stigma contributes to premature death among sexual minorities. The matched GSS-NDI data allow for a unique test of the hypothesis.

The authors found that, after controlling for individual and community-level risk factors, structural stigma was still strongly associated with premature mortality among sexual minorities, displaying a hazard ratio of 3.03 (95% CI: 1.50, 6.13), which translates into a life expectancy difference of 12 years, on average (with a range of 4–20 years). This would indicate that sexual minorities living in communities displaying "high" stigma against homosexuality are apt to die notably sooner than sexual minorities living in communities with lower average stigma. For purposes of comparison, 12 years of reduced life span is greater than that found by the Centers for Disease Control and Prevention (CDC) among regular smokers, among whom life spans are documented to be, on average, 10 years shorter than among nonsmokers (Sakata et al., 2012). The magnitude of this finding—that personal and political attitudes among one's co-residents could be more harmful than the damage self-inflicted by smoking—prompted concern about possible alternative explanations and pathways of influence.

While no research effort is flawless, the original study seemed to overlook several possible confounding variables, including a primary sampling unit (PSU) measure of proportion Black. Given that African Americans are historically both politically liberal and yet cool toward LGBT rights, and communities comprising a higher share of them are more apt to suffer from higher (and earlier) mortality rates, questions about possible omitted variable bias arose. Additionally, the failure to include a measure of personal religiosity in the model seems unusual as well, given the proliferation of a religion-and-health literature in the 1990s that culminated in documenting a seven-year average difference in life expectancy between religious attenders and non-attenders using data from the same source as the Hatzenbuehler et al. study—the NDI (Hummer et al., 1999). It is the original study's process of imputing missing data for its four key social stigma items, however, that appears to bar the way to the successful recreation of the original key predictor variable—structural stigma—and hence hamper the ability to replicate the study's key findings and test alternative pathways of influence.

2. Methods

2.1. The original study's stigma measures

The merged GSS-NDI dataset is publicly available and was prepared for replication, to be followed by the test for possible confounds. The original reported sample of 914 sexual minority respondents out of 21,045 total respondents (4.34 percent) was successfully replicated, as was the 14 percent of respondents who had died by 2008. The individual-level control variables and PSU-level control measures were also replicated, with only tiny differences in a small number of measures.

The effort to replicate the original study was successful in everything except the creation of the PSU-level structural stigma variable. The study's authors constructed this PSU-level structural stigma variable from the following four GSS-NDI items:

1. "If some people in your community suggested that a book in favor of homosexuality should be taken out of your public library, would you favor removing this book, or not?" (GSS variable name: libhomo)
2. "Should a man who admits that he is a homosexual be allowed to teach in a college or university, or not?" (GSS variable name: colhomo)
3. "Suppose a man who admits that he is a homosexual wanted to make a speech in your community. Should he be allowed to speak, or not?" (GSS variable name: spkhomo)
4. "Do you think that sexual relations between two adults of the same sex is always wrong, almost always wrong, wrong only sometimes, or not wrong at all?" (GSS variable name: homosex)

To construct the PSU-level structural stigma variable, the researchers dichotomized and summed the responses to these four survey items for each case, averaged this value for each PSU, and then constructed a dichotomous (i.e., threshold) measure of "high structural stigma" based on this PSU-level average. That top-quartile cut point, the authors note, was at 1.77, indicating that respondents were considered as living in a PSU with high structural anti-gay stigma if PSU residents responded with an anti-gay answer to (slightly) fewer than two of the four questions.

2.2. Analysis of missing data

The authors noted that, "given the structure of the GSS, not all questions were asked among all respondents each year," and that "(e)ach of these measures had greater than five percent missing due to this planned missing design, meaning that not all respondents were given the chance to respond to all questions." Fig. 1 displays the proportion of missing data for each of the four