

Categorical and anti-categorical approaches to US racial/ethnic groupings: revisiting the National 2009 H1N1 Flu Survey (NHFS)

Mulinari, Shai; Wemrell, Maria; Rönnerstrand, Björn; Subramanian, SV; Merlo, Juan

Published in: Critical Public Health

10.1080/09581596.2017.1316831

2018

Document Version: Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):

Mulinari, S., Wemrell, M., Rönnerstrand, B., Subramanian, SV., & Merlo, J. (2018). Categorical and anticategorical approaches to US racial/ethnic groupings: revisiting the National 2009 H1N1 Flu Survey (NHFS). Critical Public Health, 28(2), 177-189. https://doi.org/10.1080/09581596.2017.1316831

Total number of authors:

General rights

Unless other specific re-use rights are stated the following general rights apply: Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights

- Users may download and print one copy of any publication from the public portal for the purpose of private study
- You may not further distribute the material or use it for any profit-making activity or commercial gain
 You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY



Critical Public Health



ISSN: 0958-1596 (Print) 1469-3682 (Online) Journal homepage: http://www.tandfonline.com/loi/ccph20

Categorical and anti-categorical approaches to US racial/ethnic groupings: revisiting the National 2009 H1N1 Flu Survey (NHFS)

Shai Mulinari, Maria Wemrell, Björn Rönnerstrand, S. V. Subramanian & Juan Merlo

To cite this article: Shai Mulinari, Maria Wemrell, Björn Rönnerstrand, S. V. Subramanian & Juan Merlo (2017): Categorical and anti-categorical approaches to US racial/ethnic groupings: revisiting the National 2009 H1N1 Flu Survey (NHFS), Critical Public Health

To link to this article: http://dx.doi.org/10.1080/09581596.2017.1316831

9	© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
	Published online: 26 Apr 2017.
	Submit your article to this journal 🗗
α	View related articles 🗷
CrossMark	View Crossmark data 🗗

Full Terms & Conditions of access and use can be found at http://www.tandfonline.com/action/journalInformation?journalCode=ccph20



RESEARCH PAPER

OPEN ACCESS OPEN ACCESS Check for updates



Categorical and anti-categorical approaches to US racial/ethnic groupings: revisiting the National 2009 H1N1 Flu Survey (NHFS)

Shai Mulinari^{a,b}, Maria Wemrell^b, Biörn Rönnerstrand^c, S. V. Subramanian^d and Juan Merlo^{b,e}

 $^{\mathrm{a}}$ Faculty of Social Sciences, Department of Sociology, Lund University, Lund, Sweden; $^{\mathrm{b}}$ Faculty of Medicine. Department of Clinical Sciences, Unit of Social Epidemiology, Lund University, Malmö, Sweden; Department of Political Science, University of Gothenburg, Göteborg, Sweden; ^dDepartment of Social and Behavioral Sciences, Harvard School of Public Health, Boston, MA, USA; Center for Primary Health Care Research, Region Skåne, Malmö, Sweden

ABSTRACT

Intersectionality theory calls for the understanding of race/ethnicity, sex/ gender and class as interlinked. Intersectional analysis can contribute to public health both through furthering understanding of power dynamics causing health disparities, and by pointing to heterogeneities within, and overlap between, social groups. The latter places the usefulness of social categories in public health under scrutiny. Drawing on McCall we relate the first approach to categorical and the second to anti-categorical intersectionality. Here, we juxtapose the categorical approach with traditional between-group risk calculations (e.g. odds ratios) and the anticategorical approach with the statistical concept of discriminatory accuracy (DA), which is routinely used to evaluate disease markers in epidemiology. To demonstrate the salience of this distinction, we use the example of racial/ethnic identification and its value for predicting influenza vaccine uptake compared to other conceivable ways of organizing attention to social differentiation. We analyzed data on 56,434 adults who responded to the NHFS. We performed logistic regressions to estimate odds ratios and computed the area under the receiver operating characteristic curve (AU-ROC) to measure DA. Above age, the most informative variables were education and household poverty status, with race/ethnicity providing minor additional information. Our results show that the practical value of standard racial/ethnic categories for making inferences about vaccination status is questionable, because of the high degree of outcome variability within, and overlap between, categories. We argue that, reminiscent of potential tension between categorical and anti-categorical perspectives, between-group risk should be placed and understood in relationship to measures of DA, to avoid the lure of misguided individual-level interventions.

ARTICLE HISTORY

Received 14 October 2016 Accepted 27 March 2017

KEYWORDS

Epidemiology; ethnicity; race; racism; intersectionality; vaccination; discriminatory accuracy

Introduction

Over recent decades, intersectionality theory, which calls for understanding of categories like race/ ethnicity, sex/gender and class as interlinked rather than as separate has been advocated and sometimes integrated into studies of population health (Bauer, 2014). McCall (2005) distinguishes between categorical intersectionality research, which aims to analyze how interlocking systems of oppression drive disparities between existing social groupings, and anti-categorical intersectionality, which critiques



categorization *per se*, as use of social categories may in itself contribute to perpetuation, creation or essentialization of difference between groups. In epidemiology, categorical intersectionality can inform the field's traditional mapping of health disparities, through the use of intersectional social categories, in measurement of between-group average risk (Bauer, 2014). In contrast, anti-categorical intersectionality poses a greater challenge to epidemiology since it urges researchers to make explicit the variability within, and overlap between, socially defined groups; and to consider implications of this heterogeneity for the usefulness of social categories and the design of public health policies. However, the important tensions between average risk and heterogeneity, which can be related to potential friction between categorical and anti-categorical perspectives, are seldom teased out in epidemiology, which may result in ambiguous recommendations to researchers and policy-makers regarding the use and value of social categories. For example, Lofters and O'Campo (2012, p. 105) ask epidemiologists to use quantitative intersectional methodologies to 'highlight the most vulnerable subgroups where action is most urgently needed and ensure the best use of resources for ameliorating inequities' *and* to consider heterogeneity within socially defined groups to avoid the lure of misguided individual-level interventions, but without discussing the potential conflict between the two recommendations.

This article seeks to further a conceptual and methodological discussion on use of categorical and anti-categorical approaches in studies of population health and US racial/ethnic groupings. We do this by juxtaposing, on the one hand, a categorical approach with traditional between-group risk calculations (e.g. odds ratios, ORs), and, on the other hand, the anti-categorical approach with the statistical concept of discriminatory accuracy (DA), which is routinely used to evaluate the performance of diagnostic, prognostic, or screening markers in epidemiology (Pepe, Janes, Longton, Leisenring, & Newcomb, 2004). The underpinning idea of the concept of DA is that, to be suitable for individual-level inference, most exposure categories, whether social, geographic, or biological, need to be robust in their capacity to discriminate between individuals who do and do not demonstrate the outcome of interest (Merlo, 2014; Merlo & Wagner, 2013). Therefore, measures of DA are highly relevant in public health even if they are still infrequently reported in the literature (Merlo & Mulinari, 2015; Mulinari, Bredström, & Merlo, 2015; Wemrell, Mulinari, & Merlo, 2017a). We demonstrate the salience of this approach using the empirical example of US racial/ethnic identification and its value for predicting non-receipt of seasonal influenza vaccine compared to other conceivable ways of organizing attention to social differentiation in public health.

In the US context, a large number of studies have investigated how seasonal influenza vaccine uptake is linked to socioeconomic and demographic factors such as household income, educational level, age, gender, and race/ethnicity (Ding et al., 2011; Linn, Guralnik, & Patel, 2010; Vlahov, Bond, Jones, & Ompad, 2012). In this literature, some studies focus specifically on racial/ethnic disparities (Lu, Singleton, Euler, Williams, & Bridges, 2013; Lu et al., 2014, 2015). Notably, the US Centers for Disease Control and Prevention (CDC) regularly publishes influenza vaccination rates using a four-level race/ethnicity standard: Hispanic (any race); non-Hispanic white only; non-Hispanic black only; and non-Hispanic, all other races or multiple races (CDC, 2011). Over the last two decades, data have consistently revealed higher influenza vaccination coverage among non-Hispanic White adults than among non-Hispanic Black adults or Hispanic adults (Lu et al., 2013, 2014, 2015), believed to translate into differences in flu-associated morbidity and mortality (Dee et al., 2011). The well-established and persistent racial/ethic disparities found in prior studies, together with the importance of other socioeconomic and demographic factors, provide an appropriate empirical setting for the intersectional approach advanced in this article.

Another reason for selecting seasonal influenza vaccine uptake as an empirical example is the on-going discussions on appropriate policies to reduce racial/ethnic disparities (Fiscella, 2005; Hutchins, Fiscella, Levine, Ompad, & McDonald, 2009). The majority of the suggested policies are broad, including, e.g. increasing vaccine availability; reducing patient 'out of pocket' costs; making the offering of vaccines in health care and other settings as a routine practice; educating about risks and benefit of vaccines; using patient reminder and recall systems; and standing orders for vaccination (Lu et al., 2014, 2015). A shared feature of such policies is that they do not target individuals based on racial or ethnic identification, and may be beneficial across racial/ethnic groups while simultaneously reducing differences

between racial/ethnic groups. For example, offering free or low-cost vaccination may increase vaccination rates in all groups, in particular among low-income individuals, but may also reduce differences because of disproportionately high poverty rates in some racial/ethnic groups.

However, in addition to broad interventions, policies targeting specific racial/ethnic groups have been proposed (Chen, Fox, Cantrell, Stockdale, & Kagawa-Singer, 2007; Phillips, Kumar, Patel, & Arya, 2014; Wooten, Wortley, Singleton, & Euler, 2012). For example, it has been suggested that Black and Hispanic adults should be targeted with a text message campaign prompting them to talk to their doctors about vaccination to help address knowledge gaps and dispel misconceptions (Phillips et al., 2014). Conceptually, racially or ethnically tailored interventions involve the translation of group-level rates to individual-level risk. Yet this translation is questionable at best because of potentially important variability in outcome within groups and overlap between groups (Kaplan, 2014; Merlo, 2014; Mulinari et al., 2015). Leaving concerns about stigmatization aside (Guttman & Salmon, 2004), suggestions to implement racially or ethnically tailored policies raise questions about the value of racial/ethnic identification as a predictor of vaccination status and its predictive value compared to and above other relevant social categorizations, e.g. those based on age, income, education, or gender, or of a combination of social categorizations.

With that in mind, our purpose was threefold. First, we sought to investigate average associations between standard social categorizations and non-receipt of seasonal influenza vaccine, consistent with the conventional mapping of health disparities. Second, we sought to explore the heterogeneity of observational effects within standard racial/ethnic categories by stratifying racial/ethnic groups by gender and education, consistent with a categorical intersectionality perspective. Third, we sought to investigate how well racial/ethnic categories predicted non-receipt of the vaccine compared to and above other relevant social categorizations. Consistent with an anti-categorical intersectionality perspective, the latter analysis of DA may challenge the practical value of standard social categories for individual-level prediction. For all purposes, we used data from 56,434 adults who responded to the National 2009 H1N1 Flu Survey (NHFS) (CDC, 2012).

Methods

The National 2009 H1N1 Flu Survey

The publically available NHFS and survey data have been described elsewhere (Ding et al., 2011). In brief, the NHFS was a one-time telephone survey conducted from October 2009 through June 2010 on behalf of the CDC to monitor and evaluate the 2009–2010 vaccination campaign (CDC, 2012). The survey collected data on the uptake of both the pandemic pH1N1 and usual trivalent seasonal influenza vaccines among adults and children. Among the contacted adults, 56,656 (45.2%) completed the interview. Individual-level and household-level socio-demographic information was requested from interviewees. For some variables (race/ethnicity, gender, age), missing values were imputed. The NHFS used a sequential hot-deck method to assign imputed values, which involves replacing missing values for a non-respondent with observed values from a respondent that is similar to the non-respondent with respect to characteristics observed by both cases (CDC, 2012). There is no information in the NHFS on the amount of imputed values but according to the CDC the amount was 'very small' (personal communication).

Assessment of variables

Outcome variable

The outcome variable was seasonal flu vaccination (yes or no). 'Yes' indicated that the person had received at least one seasonal influenza vaccination since August 2009. Two hundred and two (0.4%) individuals with missing values on this variable were excluded from the analysis.

NHFS explanatory variables

We used socio-demographic variables defined in the NHFS. 'Race and ethnicity' were based on self-reported information. It included the following groups: Hispanic (any race), non-Hispanic White, non-Hispanic Black, and non-Hispanic, other races or multiple races. This four-level race and ethnicity variable was derived from answers to two questions in the NHFS. Consistent with the revised Office of Management and Budget (OMB, 1997) standards for classification of race and ethnicity, the first question was 'Are you of Hispanic or Latino origin?' The interviewer was instructed to offer the following alternatives: 'Mexican/Mexicano, Mexican-American, Central American, South American, Puerto Rican, Cuban/Cuban American, or other Spanish-Caribbean'. This was followed by a second question: '[In addition to being Hispanic or Latino,] Are you White, Black or African-American, American Indian, Alaska Native, Asian, Native Hawaiian or other Pacific Islander?'The race/ethnicity variable in the NHFS, however, contains only four race/ethnicity categories; the NHFS other races or multiple races' category includes Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander, and other races, as well as any non-Hispanic respondent selecting more than one race.

'Gender' was either man or woman. While from an intersectionality perspective, binary classification of gender is a limitation; an 'other' category was not permitted by the survey data. 'Age' was divided into five groups (18–34; 35–44; 45–54; 55–64; and 65 or more years). We assessed socioeconomic position using two variables: the 'poverty status' of the person's household and the participant's self-reported 'level of education' (college graduate; some college; 12 years; <a href="https://documents.org/level-nc-number-new-n

Intersectional explanatory variables

Recent public health studies have stressed the importance of considering social categories not only distinctly but also intersectionally (i.e. simultaneously in individuals) (Lofters & O'Campo, 2012). For instance, it is possible that the average risk of non-receipt of the vaccine is similar in intersectional subgroups defined by different 'race/ethnicity' (e.g. Black women vs. White men) but divergences within the same racial/ethnic group (e.g. White men vs. White women). If this was true, it would point to important heterogeneity of effects within and between standard racial/ethnic categories. Therefore, in addition to existing variables in the NHFS, we created two novel intersectional variables by stratifying the 'race and ethnicity' categories by, first, 'gender' and, second, 'gender' and 'education'. We used education rather than household poverty as a proxy for socioeconomic position in this combined variable because fewer values were missing for the former (5% vs. 17%).

Statistical analysis

Measures of association

We used logistic regression to examine the association between the potentially explanatory variables and non-receipt of seasonal influenza vaccine. We developed a series of analyses that modeled one variable at a time followed by more elaborate models that adjusted for age, household poverty, and level of education. In addition, we conducted separate analyses using the two intersectional variables mentioned above, created to investigate heterogeneity of effects within and between racial/ethnic groups. In all analyses, we used the provided survey weights that are calculated using a number of socioeconomic and demographic variables including age, gender, race/ethnicity, and state of residence (CDC, 2012). We expressed associations by means of ORs and 95% confidence intervals (CIs). The reference groups in the analyses were those presenting the highest vaccination rates.

Analysis of discriminatory accuracy

DA measures the ability of a diagnostic tool, marker or category to correctly discriminate between people with or without an outcome of interest (Merlo, 2014; Pepe et al., 2004). In principle, diagnostic

tools, markers, or categories, often included as covariates in statistical models, need to have high DA to be deemed valid for diagnostic or prognostic assessment. It is well known that measures of association alone are inappropriate for gauging the DA of statistical models (Pepe et al., 2004). In fact, what we normally consider a strong association between an exposure and an outcome (e.g. an OR of 10) may be related to a rather low capacity of the exposure to discriminate cases and non-cases. For linear regression models, DA corresponds with the concept of variance explained (r^2) used to evaluate the general strength of findings in research fields including epidemiology (Merlo & Wagner, 2013). For logistic regression models, DA is assessed by means of receiver operating characteristic (ROC) curve analysis. The ROC curves were created by plotting sensitivity, or the true positive fraction (TPF), vs. 1-specificity, or the false positive fraction (FPF), at various threshold settings of predicted risk obtained from the logistic regression models. The TPF expresses the probability that given some covariates an unvaccinated individual belongs to the class coded as 1 (the individual is predicted to be unvaccinated) at a specific threshold setting of predicted risk. The FPF expresses the probability that, using the same threshold, a vaccinated individual belongs to the class coded as 1, i.e. the individual is misclassified as unvaccinated. We calculated the area under the ROC curve (AU-ROC), or C statistic, as a measure of DA. AU-ROC assumes a value from 0.5 to 1 where 1 is perfect discrimination and 0.5 is as informative as flipping an unbiased coin (i.e. the covariates have no predictive power) (Pepe et al., 2004). Here, the AU-ROC can be interpreted as the probability that a randomly selected non-vaccinated individual will have a higher predicted risk of non-receipt than a randomly selected vaccinated individual. For example, an AU-ROC = 0.6 means that if we randomly select one unvaccinated and one vaccinated individual, the probability of having a higher predicted risk of non-receipt for the unvaccinated individual is 60%. If the AU-ROC = 1, every unvaccinated individual would have higher predicted risk of non-receipt than every vaccinated individual.

In an initial series of simple logistic regression models, we calculated the AU-ROCs with 95% CIs of models including age alone or age plus one or more other variables. We assessed the incremental discriminatory value of a model by calculating the increase in AU-ROC. We used the AU-ROC of age as the baseline from which to assess the incremental discriminatory value of other models because age is a major determinant of influenza vaccine receipt and also a confounder of the association between race/ethnicity and influenza vaccination receipt (Lu et al., 2013, 2014, 2015). In a second series of logistic regression models, we calculated the AU-ROCs with 95% CIs of models including age and the variable 'race and ethnicity' together with 'gender' or with 'gender', 'household poverty status', and 'educational level'. This second series of modeling was done to assess the incremental discriminatory value of more elaborate models. Finally, we calculated the AU-ROCs with 95% CIs of models including age and the two intersectional variables to test whether the use of intersectional sub-groupings lead to improvement of DA compared to models that include 'race/ethnicity', 'gender' and 'education' as separate terms.

We performed the statistical analyses using SPSS Version 22.0 (SPSS Inc., Chicago, Illinois, USA) and STATA (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP).

Results

Mapping of disparities through measurement of between-group average risk

As shown in Table 1, the overall non-receipt of seasonal influenza vaccine in the sample was 53.3%. According to the raw data, coverage was higher for individuals identified as non-Hispanic White compared to each of the other racial/ethnic groups, as well as in men compared to women. Vaccination coverage also generally increased with increasing age, household income, and educational level.

Our analyses revealed that, compared to the non-Hispanic White group, rates of non-vaccination receipt were significantly higher among non-Hispanic Blacks (OR = 1.72, CI 95% 1.52-1.94), Hispanics (OR = 1.88, Cl 95% 1.63-2.17), and people identified as being of other or multiple races (OR = 1.19, Cl 95% 1.63-2.17)95% 1.04–1.37) (Table 2). The associations remained conclusive for non-Hispanic Blacks and Hispanics after adjustment for age, but the strength of the associations diminished for both groups and especially



Table 1. Characteristics of sample.

	Total (n)	Total (%)	Non-receipt of seasonal influenza vaccine (%)
All	56,434	100	53.3
Racial/ethnic category			
Non-Hispanic White	44,909	79.6	51.0
Non-Hispanic Black	4553	8.1	63.5
Hispanic, any race	3651	6.5	66.7
Non-Hispanic, other or multiple races	3321	5.9	57.0
Gender			
Female	33,458	59.3	50.5
Male	22,976	40.7	57.5
Age			
18–34	11,022	19.5	71.0
35–44	8244	14.6	63.6
45–54	11,077	19.6	60.3
55–64	11,699	20.7	48.8
>=65	14,392	25.5	32.3
Education			
College graduate	21,390	37.9	48.6
Some college	14,882	26.4	54.9
12 years	12,164	21.6	54.7
<12 years	5020	8.9	60.3
Missing or unknown	2978	5.3	62.2
Poverty status of household			
>=\$75,000	14,398	25.5	49.9
<\$75,000, above poverty threshold	26,994	47.8	52.1
Below poverty threshold	5587	9.9	64.5
Missing or unknown	9455	16.8	55.5

for Hispanics (OR = 1.35, CI 95% 1.18-1.56). Additional adjustment for educational level and household poverty status further weakened associations but they remained statistically conclusive (Table 2). Moreover, men had a higher rate of non-receipt of seasonal influenza vaccine than women, and there were conclusive differences across age groups, as well as across household poverty and educational level categories (Table 2).

Heterogeneity of effects between and within racial and ethnic categories

The combination of the race/ethnicity and gender variables that created 8 different intersectional subgroups revealed that in comparison to non-Hispanic White women, all other subgroups except women identified as being of 'other or multiple races' had higher rates of non-vaccination receipt (Table 3). However, ORs were similar for non-Hispanic White men (OR = 1.20, CI 95% 1.11-1.30) and Hispanic women (OR = 1.41, CI 95% 1.19–1.67), showing that the risk of non-vaccination receipt is heterogeneously distributed within and between racial/ethnic categories. Combining race/ethnicity, gender, and education variables to create 40 different intersectional subgroups resulted in an even more complex picture: we observed substantial heterogeneity of effects within and between groups defined by race/ ethnicity (Table 3).

Measuring the discriminatory accuracy of social categorizations

Despite these statistically significant associations, the DA of the categories studied was very low. Table 4 shows the AU-ROCs of models that included age alone or age together with one or more of the explanatory variables. The AU-ROC for age alone was 0.658 (Model 1) and it increased only slightly (+0.005) when information on race/ethnicity was included (Model 2). That is, if we randomly select one unvaccinated and one vaccinated individual from the NHFS, the probability of having a higher predicted risk of non-receipt for the unvaccinated individual in the two models is 65.8 and 66.3%, respectively. Similarly, information on gender did little to improve the DA above the model that included age (+0.006) (Model



Table 2. Measures of association between social categories and non-receipt of seasonal influenza vaccine.

	Una	djusted	Age-a	adjusted	,	age, educational ld poverty status
_	ORa	CI ^b 95%	OR	CI 95%	OR	CI 95%
Racial/ethnic category						
Non-Hispanic White	1		1		1	
Non-Hispanic Black	1.72	1.52-1.94	1.57	1.38-1.78	1.40	1.23-1.60
Hispanic, any race	1.88	1.63-2.17	1.35	1.18-1.56	1.18	1.02-1.36
Non-Hispanic, other or multiple races	1.19	1.04-1.37	0.97	0.84-1.13	0.93	0.80-1.08
Gender						
Female	1		1		1	
Male	1.27	1.19-1.35	1.19	1.11-1.27	1.23	1.14-1.32
Age						
18–34	4.98	4.48-5.54				
35-44	3.24	3.45-4.04				
45-54	3.14	2.84-3.47				
55-64	1.92	1.74-2.16				
>=65	1					
Education						
College graduate	1		1			
Some college	1.24	1.14-1.34	1.25	1.15-1.37		
12 years	1.29	1.18-1.41	1.48	1.35-1.63		
<12 years	1.60	1.41-1.80	1.92	1.69-2.18		
Missing or unknown	1.81	1.56-2.10	1.60	1.37-1.87		
Poverty status of household						
>=\$75,000	1		1			
<\$75,000, above poverty threshold	1.12	1.04-1.22	1.43	1.32-1.56		
Below poverty threshold	1.73	1.53-1.96	1.86	1.64-2.10		
Missing or unknown	1.31	1.18-1.45	1.76	1.58-1.96		

^aOdds ratio.

3) or age and race (\pm 0.004) (Model 4; compare to Model 2). Household poverty status and educational level were the most informative variables beyond age (each \pm 0.014, not shown), but the model including age, household poverty status, and educational level still reached only an AU-ROC = 0.678 (\pm 0.020) (Model 5). Notably, including race/ethnicity only added \pm 0.001 (Model 6), which is consistent with a strong relationship between class and race/ethnicity. We observed the highest DA (AU-ROC = 0.681) for the model that included all explanatory variables (Model 7). However, this higher DA compared to the model including age only (\pm 0.022) was mainly due to the socioeconomic variables. In the final analysis, we tested whether the composite intersectional variables improved the DA compared with the models where the 'race and ethnicity', 'gender' and 'educational level' variables were kept separate; we found that use of intersectional sub-groupings did little to further improve DA (Models 4 vs. 8 and 7 vs. 9).

Discussion

Eliminating health disparities along lines of race/ethnicity is an important goal of public health policy. Our results confirm findings that adult seasonal influenza vaccination coverage is higher among non-Hispanic White adults than among non-Hispanic Black adults or Hispanic adults (Lu et al., 2013, 2014, 2015; CDC, 2011). The group defined as 'non-Hispanic, other races or multiple races' also had lower vaccination coverage than the White majority group, but the difference disappeared when we controlled for age. When faced with no evidence of a difference between broadly defined racial/ethnic groups, researchers have sometimes sought to disaggregate groups since aggregating data can conceal inequities between sub-groups. For example, a study found no differences in vaccination coverage between the non-Hispanic White group and the broad Asian/Pacific Islander group, but found differences between the non-Hispanic White group and the Filipino American sub-group (Chen et al., 2007).

^bConfidence interval.

 Table 3. Age-adjusted measures of association between subgroups and non-receipt of seasonal influenza vaccine.

								Educati	educational level				
				College	College Graduate	Some	Some College	12 }	12 years	<12	<12 years	Missing c	Aissing or unknown
Racial/ethnic category Gender	Gender	ORª	95% Cl ^b	OR	12 %56	OR	95% CI	OR	12% CI	OR	95% CI	OR	95% CI
Non-Hispanic White	Female	-		-		1.22	1.08-1.37	1.36	1.19–1.56	1.81	1.47-2.22	1.99	1.57–2.53
	Male	1.20	1.11-1.30	1.14	1.01-1.29	1.62	1.38-1.88	1.87	1.60-2.19	2.07	1.60-2.69	1.58	1.23-2.02
Non-Hispanic Black	Female	1.64	1.41–1.91	1.43	1.07-1.94	2.00	1.55-2.57	2.14	1.58-2.90	3.34	2.38-4.69	1.58	0.76 - 3.32
	Male	1.79	1.43-2.26	2.70	1.88-3.87	1.26	0.78-2.04	2.65	1.87-3.76	3.58	2.24-5.72	1.84	0.89 - 3.81
Hispanic, any race	Female	1.41	1.19–1.67	1.57	1.11-2.22	1.38	0.97-1.86	2.47	1.69–3.62	1.68	1.23-2.29	1.75	0.99 - 3.10
	Male	1.53	1.21-1.93	0.89	0.54 - 1.49	1.68	1.01-2.78	2.92	2.00-4.26	2.50	1.77-3.54	2.82	1.56-5.10
Non-Hispanic, other or	Female	0.91	0.74-1.12	1.14	0.83-1.55	0.98	0.66-1.45	1.12	0.67-1.90	1.55	0.87-2.77	0.81	0.33-1.94
multiple races	Male	1.26	1.03-1.55	1.31	0.95-1.82	1.52	1.00–2.32	1.45	0.96–2.18	3.32	1.71–6.44	1.65	0.91–3.02

^aOdds ratio. ^bConfidence interval.

Table 4. AU-ROC analysis to evaluate the DA of different models for non-receipt of seasonal influenza vaccine.

Variables in model	Model no.								
NHFS original variables	1	2	3	4	5	6	7	8	9
Age									
Racial/ethnic category									
Gender									
Education									
Poverty status of household									
Intersectional variables									
Racial/ethnic category stratified by gender									
Racial/ethnic category stratifed by gender and education									
AU-ROC ^a	0.658	0.663	0.664	0.668	0.678	0.679	0.681	0.669	0.682
ΔAU-ROC	Ref.	0.005	0.006	0.010	0.020	0.021	0.023	0.011	0.024

 $^{a}95\%$ confidence intervals are \pm 0.005 or 0.004.

The gray shading indicates which variables are included in Models 1-9. For example, Model 1 only included the variable age.

A recognized problem with sub-group analyses is that conclusive findings may represent spurious associations (Sun, loannidis, Agoritsas, Alba, & Guyatt, 2014). However, our study highlights another issue of major importance to public health practice and research: while aggregate data may conceal differences *between* groups (Pande & Yazbeck, 2003), aggregating data can also conceal substantial outcome variability (and thus inequality) *within* groups and overlap *between* groups (Bleich, Thorpe, Sharif-Harris, Fesahazion, & LaVeist, 2010). If this heterogeneity is considerable, references to between-group differences in mean values, without simultaneous reference to within-group variation and between-group overlap, risk overemphasizing the value of racial/ethnic categories as a means of predicting the health-related or health care-seeking behavior of individuals (Mulinari, Juárez, Wagner, & Merlo, 2015; Mulinari et al., 2015). Reminiscent of potential tension between categorical and anti-categorical approaches (McCall, 2005), then, between-group average risk should be placed and understood in relationship to measures of DA to avoid the lure of misguided individual-level interventions.

Assertion of the limited value of racial/ethnic categories for individual-level prediction is not new (Kaplan, 2014; Kaplan & Bennett, 2003), and its relevance extends beyond medicine and public health, e.g. to profiling by law enforcement and security personnel (Engel, 2008). In medicine, a meta-analysis of racial differences in response to antihypertensive drugs found that despite differences between US Whites and Blacks at the aggregate level, race has little value in predicting response to antihypertensive drugs, because Whites and Blacks overlap greatly in their response to all categories of drugs (Sehgal, 2004). Similarly, the use of human racial/ethnic categories in genetics has been heavily criticized because of the large genetic diversity within groups and continuous overlap between groups despite average differences in allele frequencies (Lewontin, 1972; Holsinger & Weir, 2009). The novelty of our study is the introduction of ROC curves as a measure of DA to gauge the overlap between US racial/ethnic categories. ROC curve analysis, or similar approaches like the multilevel analysis of individual heterogeneity (Merlo, 2003, 2014; Wemrell, Mulinari, & Merlo, 2017b), can be used to identify when biological (Juarez, Wagner & Merlo, 2014), geographical (Merlo, Viciana-Fernández, Ramiro-Fariñas, & Research Group of the Longitudinal Database of the Andalusian Population (LDAP), 2012), socioeconomic or ethnic (Beckman et al., 2004; Mulinari et al., 2015) categorizations are valid as instruments for individual-level predictions. In the present case, the large overlaps in vaccination coverage are reflected in the low DA of the racial/ ethnic categories used. A low DA effectively refutes the argument that although not every individual within a racial/ethnic group possesses a particular trait, racial/ethnic categories function well enough in predicting which individuals possess it. Because standard racial/ethnic categories do not function well enough for individual-level prediction, the reliance on racial/ethnic identification as a proxy in medical decision-making may lead to inappropriate treatment based on stereotyping (Kaplan, 2014). This does not preclude the possibility of other racial/ethnic categorizations having a higher DA, or that existing categorizations are more relevant for predicting other outcomes, but to our knowledge such a case awaits empirical confirmation.

Another argument professed in favor of using racial/ethnic identification to predict vaccination behavior is based on reports of unique barriers to adult influenza vaccination in different racial/ethnic groups (Chen et al., 2007). Yet on closer inspection, most of those barriers are not unique to any particular group. For example, Chen et al. (2007) found that 32% of African-American influenza vaccination absentees cited concerns over the vaccine causing influenza or serious side effects, while 18% of Whites, 13% of Latinos, 11% of Japanese Americans, and 22% of Filipino Americans cited the same reason. Nonetheless, the authors called for ethnic specific strategies to address the issues of mistrust by African-American expressed in sentiments such as their concern that the influenza vaccine causes influenza' (Chen et al., 2007). While there may be issues of mistrust among African-Americans related to racism and social exclusion, mistrust is not a racially unique phenomenon (Boulware, Cooper, Ratner, LaVeist, & Powe, 2003), nor is it a racially unique reason for not being vaccinated (Chen et al., 2007). Social inequity in vaccination coverage and social patterning of trust are unlikely to be effectively addressed by racially tailored interventions. On the contrary, experiences with tailored social programs suggest they tend to undermine social trust (Kumlin & Rothstein, 2005). Interventions may be particularly misguided when targeted at altering the behavior of selected individuals, as opposed to changing macro- or mesolevel factors that enable and constrain behaviors because targeting individuals carries a higher risk of stigmatization (Guttman & Salmon, 2004). To be clear, we are not questioning the importance of race/ ethnicity as an identity, or the lived experience of people in a racialized society. Rather, our concern is with the use of racial/ethnic categories for individual-level prediction and profiling. We believe this use would be dramatically reduced, if measures of DA be routinely reported alongside measures of associations when gauging group-level differences.

Our study also raises questions about the value of racial/ethnic identification for predicting vaccination status compared to other conceivable ways of organizing attention to social differentiation in public health. That the CDC routinely releases vaccination coverage data by race/ethnicity is consistent with federal mandates requiring agencies under the Department of Health and Human Services to collect and report race/ethnicity-based statistics to monitor and combat inequalities (Epstein, 2008). A major argument for collecting race/ethnicity-based statistics is that race/ethnicity is a primary axis of social distinction and is therefore associated with a broad array of factors with important modifying effects on health and health care delivery (Kaplan & Bennett, 2003). However, as pointed out by Epstein (2008), the federal endorsement of a specific set of racial/ethnic categories has resulted in the proliferation of studies that treat these taxonomic categories as the standardized formal units of analysis; in the process, other ways of classifying health risks, such as behavioral practices, and other ways of classifying populations, such as by social class, receive far less attention.

The CDC does not consistently report influenza vaccination coverage by socioeconomic status indicators such as income or education. The CDC acknowledges that racial/ethnic disparities in influenza vaccination coverage have been studied more extensively compared to other potentially relevant disparity domains, such as gender and socioeconomic position (Setse et al., 2011), suggesting that disparities along these lines are considered of lesser concern. Yet information on variables relevant to other disparity domains is readily available, and our analysis shows conclusive differences between women and men irrespective of age (i.e. not fully explained by pregnancy) and across socioeconomic groups, consistent with the results reported by others (Setse et al., 2011). These differences appear to be as large as or larger than those observed between individuals identified as Black or White. In fact, the ROC curve analysis showed that above age, the most informative variables were education and household poverty status (+0.020), with race/ethnicity providing very little additional information (+0.001). It is important to note that race/ethnicity and socioeconomic position are not independent, as the disadvantage that members of some minority groups suffer will translate into, on average, lower income and educational levels. Polices that effectively address socioeconomic inequities are therefore predicted to diminish, albeit not eliminate, racial/ethnic gaps. Ignoring socioeconomic inequalities risks diverting attention away from policies that could have major impact on vaccination rates among minority group members while simultaneously benefitting the large group of deprived Whites.

Intersectionality theory posits that social differentiation takes place along multiple, non-independent, and possibly interacting axes (McCall, 2005). In the case of vaccination coverage, one consequence of this social complexity is that most individuals can be construed as belonging to one or more major social groups with lower vaccination coverage than one or more comparison groups. It also means that, through application of a categorical intersectionality perspective, groups can be split into a number of smaller taxonomic units through the combination of more than one major axis of social differentiation, as we have done in this paper. Yet the ROC curve analysis showed that the composite intersectional variables did little to improve the DA compared with the models where the 'race and ethnicity', 'gender' and 'educational level' variables were kept separate. This highlights the fact that splitting the population into increasingly smaller taxonomic units to 'hone in on ... the most vulnerable subgroups' (Lofters & O'Campo, 2012, p. 105) may not ensure the best use of resources for ameliorating inequalities because of the high degree of outcome variability within, and overlap between, social categories. The problem, therefore, is how to justify focusing on one particular axis of social differentiation rather than any other. Decisions to focus on one particular set of social positions or intersection of positions will be guided by political, theoretical, and pragmatic choices and constraints. This point is underlined by the fact that routine stratification by race/ethnicity is primarily a US practice bolstered by federal mandates and standards (Epstein, 2008). While measures of DA provide no escape from this situation, at least they underscore the important points that social structures, such as racism, generate persistent patterns of inequality but not law-like regularities (Muntaner, 2013), and that there is a great deal of variance in health and health care seeking behavior that is not readily mapped onto social position (Dunn, 2012).

In sum, our study shows that the practical value of standard racial/ethnic categories, and other relevant social categorizations, for making inferences about individuals' vaccination status is questionable despite seemingly large and conclusive differences between groups. More generally, our study highlights the tension between average, between-group, risk and measures of DA, related to and understood by means of categorical and anti-categorical intersectionality. While quantitative intersectionality research has often been of the categorical type, anti-categorical approaches have usually been furthered through qualitative research, often encompassing philosophical critique of social categorization as potentially leading to demarcation, exclusion and furthered inequality. Operationalized through measurement of DA, anti-categorical approaches can also be investigated, expressed and developed within a quantitative framework.

Limitations

Because it is based on a cross-sectional telephone survey, our study has several weaknesses. Among these, it should be stressed that the response rate was relatively low (45.2%), which increases the risk of non-response bias, and that information was self-reported and may be subject to recall error. According to the CDC (2011), the survey overestimates seasonal influenza vaccination coverage; in part this may because of misclassification of pandemic pH1N1 vaccine for seasonal influenza vaccine. To test if the low DA of racial/ethnic categories was limited to seasonal influenza vaccination, we ran the analyses with 2009 pandemic pH1N1 vaccination status as the outcome, but conclusions were the same (available upon request). Finally, our analysis does not consider the fact that vaccination levels changed over the duration of survey administration which could a have slight effect on vaccination coverage estimates.

There is a substantial body of literature discussing the strength and weakness of different methods for assignment to racial/ethnic categories including self-report, investigator-assigned, based on administrative records, and using genetic markers; and study results can differ substantially depending on the method used (reviewed in Kaplan, 2014). In epidemiology, the 'gold standard' for racial/ethnic assignment is self-report, consistent with the principle that people are who they say they are. Yet the complexity and fluidity of individual identity make it impossible to divide the population into non-overlapping racial/ethnic groups, or to validly and reliably allocate people to any given set of categories. Accordingly, research studies have found inconsistencies in the way that race and ethnicity are self-reported and recoded by investigators (Kaplan, 2014). However, because our purpose was to



evaluate standard racial/ethnic categories used regularly by public health researchers and authorities, any limitations of race/ethnicity data, although important to acknowledge, do not undermine our finding that standard racial/ethnic categories have low DA for the studied outcome.

Acknowledgement

We thank Judith B. Kaplan for important comments on an earlier draft.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Swedish Research Council (VR) [grant number 2013-1268 to S.M], [grant number 2013-1695 to S.M] and [grant number 2013-2484 to J.M] and the Public Health Unit at Scania County Council (Region Skåne) [grant number 2013 to J.M].

References

- Bauer, G. R. (2014). Incorporating intersectionality theory into population health research methodology: Challenges and the potential to advance health equity. Social Science & Medicine, 110, 10–17.
- Beckman, A., Merlo, J., Lynch, J. W., Gerdtham, U. G., Lindström, M., & Lithman, T. (2004). Country of birth, socioeconomic position, and healthcare expenditure: a multilevel analysis of Malmö, Sweden. Journal of Epidemiology and Community Health, 58, 145-149.
- Bleich, S. N., Thorpe, R. J., Sharif-Harris, H., Fesahazion, R., & LaVeist, T. A. (2010). Social context explains race disparities in obesity among women. Journal of Epidemiology and Community Health, 64, 465-469.
- Boulware, L. E., Cooper, L. A., Ratner, L. E., La Veist, T. A., & Powe, N. R. (2003). Race and trust in the health care system. Public Health Reports, 118, 358.
- Centers for Disease Control and Prevention (CDC) (2012, August 20). Datasets and Related Documentation for the National 2009 H1N1 Flu Survey (NHFS). Retrieved May 26, 2016, from http://www.cdc.gov/nchs/nis/data_files_h1n1.htm
- Centers for Disease Control and Prevention (CDC). (2011, May 13). Final estimates for 2009-2010 seasonal influenza and influenza A (H1N1) 2009 monovalent vaccination coverage—United States, August 2009 through May, 2010. Retrieved May 26, 2016, from http://www.cdc.gov/flu/fluvaxview/coverage_0910estimates.htm
- Chen, J. Y., Fox, S. A., Cantrell, C. H., Stockdale, S. E., & Kagawa-Singer, M. (2007). Health disparities and prevention: Racial/ ethnic barriers to flu vaccinations. Journal of Community Health, 32, 5-20.
- Dee, D. L., Bensyl, D. M., Gindler, J., Truman, B. I., Allen, B. G., D'Mello, T., ... Fowlkes, A. (2011). Racial and ethnic disparities in hospitalizations and deaths associated with 2009 pandemic influenza A (H1N1) virus infections in the United States. Annals of Epidemiology, 21, 623–630.
- Ding, H., Santibanez, T. A., Jamieson, D. J., Weinbaum, C. M., Euler, G. L., Grohskopf, L. A., ... Singleton, J.A. (2011). Influenza vaccination coverage among pregnant women-National 2009 H1N1 Flu Survey (NHFS). American Journal of Obstetrics and Gynecology, 204, S96-S106.
- Dunn, J. R. (2012). "Explanation", philosophy and theory in health inequalities research: Towards a critical realist approach. In P. O'Campo & J. R. Dunn (Eds.), Rethinking social epidemiology (pp. 23–42). New York, NY: Springer.
- Engel, R. S. (2008). A critique of the "outcome test" in racial profiling research. Justice Quarterly, 25(1), 1-36.
- Epstein, S. (2008). Inclusion: The politics of difference in medical research. Chicago, IL: University of Chicago Press.
- Fiscella, K. (2005). Commentary-anatomy of racial disparity in influenza vaccination. Health Services Research, 40, 539-550. Guttman, N., & Salmon, C. T. (2004). Guilt, fear, stigma and knowledge gaps: Ethical issues in public health communication interventions. Bioethics, 18, 531-552.
- Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: Defining, estimating and interpreting FST. Nature Reviews Genetics, 10, 639-650.
- Hutchins, S. S., Fiscella, K., Levine, R. S., Ompad, D. C., & McDonald, M. (2009). Protection of racial/ethnic minority populations during an influenza pandemic. American Journal of Public Health, 99, S261-S270.
- Juárez, S. P., Wagner, P., & Merlo, J. (2014). Applying measures of discriminatory accuracy to revisit traditional risk factors for being small for gestational age in Sweden: A national cross-sectional study. BMJ Open, 4, e005388.
- Kaplan, J. B. (2014). The quality of data on "race" and "ethnicity": Implications for health researchers, policy makers, and practitioners. Race and Social Problems, 6, 214–236.



- Kaplan, J. B., & Bennett, T. (2003). Use of race and ethnicity in biomedical publication. *Journal of the American Medical Association*, 289, 2709–2716.
- Kumlin, S., & Rothstein, B. (2005). Making and breaking social capital the impact of welfare-state institutions. *Comparative Political Studies*, 38, 339–365.
- Lewontin, R. C. (1972). The apportionment of human diversity. Evolutionary Biology, 6, 381–398.
- Linn, S. T., Guralnik, J. M., & Patel, K. V. (2010). Disparities in influenza vaccine coverage in the United States, 2008. *Journal of the American Geriatrics Society, 58*, 1333–1340.
- Lofters, A., & O'Campo, P. (2012). Differences that matter. In P. O'Campo & J. R. Dunn (Eds.), *Rethinking social epidemiology* (pp. 93–109). New York, NY: Springer.
- Lu, P. J., Singleton, J. A., Euler, G. L., Williams, W. W., & Bridges, C. B. (2013). Seasonal influenza vaccination coverage among adult populations in the United States, 2005–2011. *American Journal of Epidemiology, 178*, 1478–1487.
- Lu, P. J., O'Halloran, A., Bryan, L., Kennedy, E. D., Ding, H., Graitcer, S. B., ... Singleton, J. A. (2014). Trends in racial/ethnic disparities in influenza vaccination coverage among adults during the 2007–2008 through 2011–2012 seasons. *American Journal of Infection Control*, 42(7), 763–769.
- Lu, P. J., O'Halloran, A., Williams, W. W., Lindley, M. C., Farrall, S., & Bridges, C. B. (2015). Racial and ethnic disparities in vaccination coverage among adult populations in the US. *American Journal of Preventive Medicine*, 49, S412–S425.
- McCall, L. (2005). The complexity of intersectionality. Signs: Journal of Women in Culture and Society, 30, 1771–1800.
- Merlo, J. (2003). Multilevel analytical approaches in social epidemiology: Measures of health variation compared with traditional measures of association. *Journal of Epidemiology and Community Health*, *57*, 550–552.
- Merlo, J. (2014). Invited commentary: multilevel analysis of individual heterogeneity A fundamental critique of the current probabilistic risk factor epidemiology. *American Journal of Epidemiology*, 180, 208–212.
- Merlo, J., & Mulinari, S. (2015). Measures of discriminatory accuracy and categorizations in public health: A response to Allan Krasnik's editorial. *The European Journal of Public Health*, *25*, 910–910.
- Merlo, J., Viciana-Fernández, F. J., Ramiro-Fariñas, D., & Research Group of the Longitudinal Database of the Andalusian Population (LDAP) (2012). Bringing the individual back to small-area variation studies: A multilevel analysis of all-cause mortality in Andalusia. Spain. Social Science & Medicine, 75, 1477–1487.
- Merlo, J., & Wagner, P. (2013). The tyranny of the averages and the indiscriminate use of risk factors in public health: A call for revolution. *European Journal of Epidemiology*, 28(Suppl. 1), 148.
- Mulinari, S., Bredström, A., & Merlo, J. (2015). Questioning the discriminatory accuracy of broad migrant categories in public health: Self-rated health in Sweden. *The European Journal of Public Health*, 25, 911–917.
- Mulinari, S., Juárez, S. P., Wagner, P., & Merlo, J. (2015). Does maternal country of birth matter for understanding offspring's birthweight? A multilevel analysis of individual heterogeneity in Sweden. *PLoS One, 10*, e0129362.
- Muntaner, C. (2013). Invited commentary: On the future of social epidemiology A case for scientific realism. *American Journal of Epidemiology*, 178, 852–857.
- Office of Management and Budget (OMB) (1997, October 30). Revisions to the standards for the classification of federal data on race and ethnicity. Retrieved January 11, 2017, from http://www.whitehouse.gov/omb/fedreg_1997standards
- Pande, R. P., & Yazbeck, A. S. (2003). What's in a country average? Wealth, gender, and regional inequalities in immunization in India. *Social Science & Medicine*, *57*, 2075–2088.
- Pepe, M. S., Janes, H., Longton, G., Leisenring, W., & Newcomb, P. (2004). Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American Journal of Epidemiology, 159*, 882–890.
- Phillips, A. L., Kumar, D., Patel, S., & Arya, M. (2014). Using text messages to improve patient–doctor communication among racial and ethnic minority adults: An innovative solution to increase influenza vaccinations. *Preventive Medicine*, 69, 117–119.
- Sehgal, A. R. (2004). Overlap between whites and blacks in response to antihypertensive drugs. *Hypertension*, *43*, 566–572. Setse, R. W., Euler, G. L., Gonzalez-Feliciano, A. G., Bryan, L. N., Furlow, C., Weinbaum, C. M., & Singleton, J. A. (2011). Influenza vaccination coverage United States, 2000–2010. *Morbidity and Mortality Weekly Report (MMWR) Surveillance Summaries*, *60*, 38–41.
- Sun, X., Ioannidis, J. P., Agoritsas, T., Alba, A. C., & Guyatt, G. (2014). How to use a subgroup analysis: Users' guide to the medical literature. *Journal of the American Medical Association*, 311, 405–411.
- Vlahov, D., Bond, K. T., Jones, K. C., & Ompad, D. C. (2012). Factors associated with differential uptake of seasonal influenza immunizations among underserved communities during the 2009–2010 influenza season. *Journal of Community Health*, 37, 282–287.
- Wemrell, M., Mulinari, S., & Merlo, J. (2017a). Intersectionality and risk for ischemic heart disease in Sweden: Categorical and anti-categorical approaches. *Social Science & Medicine*, 177, 213–222.
- Wemrell, M., Mulinari, S., & Merlo, J. (2017b). An intersectional approach to multilevel analysis of individual heterogeneity (MAIH) and discriminatory accuracy. Social Science & Medicine, 178, 217–219.
- Wooten, K. G., Wortley, P. M., Singleton, J. A., & Euler, G. L. (2012). Perceptions matter: Beliefs about influenza vaccine and vaccination behavior among elderly white, black and Hispanic Americans. *Vaccine*, *30*, 6927–6934.