Preventive Medicine Reports 4 (2016) 199–208

Contents lists available at ScienceDirect

Preventive Medicine Reports

journal homepage: http://ees.elsevier.com/pmedr

Characteristics associated with self-rated health in the CARDIA study: Contextualising health determinants by income group

Shilpa Nayak [a,*], Alan Hubbard [b], Stephen Sidney [c], S. Leonard Syme [b]

[a] Department of Public Health and Policy, The Whelan Building, Quadrangle, The University of Liverpool, Liverpool L69 3GB, UK
[b] School of Public Health, The University of California, Berkeley, Berkeley, USA
[c] Kaiser Permanente Northern California Division of Research, Oakland, USA

ARTICLE INFO

Article history:
Received 30 March 2016
Accepted 5 June 2016
Available online 08 June 2016

Keywords:
Self-rated health
Health inequalities
Socioeconomic factors
Health determinants
Classification trees

ABSTRACT

An understanding of factors influencing health in socioeconomic groups is required to reduce health inequalities. This study investigated combinations of health determinants associated with self-rated health (SRH), and their relative importance, in income-based groups.

Cross-sectional data from year 15 (2000−2001) of the CARDIA study (Coronary Artery Risk Development in Young Adults, USA) - 3648 men and women (mean 40 years) - were split into 5 income-based groups. SRH responses were categorized as 'higher'/'lower'. Health determinants (medical, lifestyle, and social factors, living conditions) associated with SRH in each group were analyzed using classification tree analysis (CTA).

Income and SRH were positively associated ($p < 0.05$). Data suggested an income-based gradient for lifestyle/medical/social factors/living conditions. Profiles, and relative importance ranking, of multi-domain health determinants, in relation to SRH, differed by income group. The highest ranking variable for each income group was chronic burden-personal health problem (<$25,000); physical activity ($25–50,000; $50–75,000; $100,000+); and cigarettes/day ($75–100,000). In lower income groups, more risk factors and chronic burden indicators were associated with SRH. Social support, control over life, optimism, and resources for paying for basics/medical care/health insurance were greater (%) with higher income.

SRH is a multidimensional measure; CTA is useful for contextualizing risk factors in relation to health status. Findings suggest that for lower income groups, addressing contributors to chronic burden is important alongside lifestyle/medical factors. In a proportionate universalism context, in addition to differences in intensity of public health action across the socioeconomic gradient, differences in the type of interventions to improve SRH may also be important.

1. Introduction

The socioeconomic gradient in health is well recognized. Knowledge of differences in characteristics associated with good or poor health in socioeconomic groups is important to inform appropriate interventions, and improve health status across the gradient. Health status is a complex construct. The health implications of a single risk factor or exposure may not be universally identical; that is, health status would depend on interaction with coexisting variables, so that different combinations of risk and protective factors produce different outcomes. A solitary focus on single risk factors overlooks the combined impact of these multi-domain influences on health status (Marmot et al., 1998; Ostlin et al., 2005). The WHO Task Force on Research Priorities for Equity in Health called for research studying the "interrelationships between individual factors and social context that increase or decrease the likelihood of achieving and maintaining good health" (Ostlin et al., 2005).

SRH is a common measure of global health status, and an independent predictor of subsequent morbidity and mortality (CDCa,b,c, 2016; Idler and Benyamini, 1997; Moller et al., 1996; ONS). For high proportions of populations to report good SRH is in itself an important end point. Studies have identified independent determinants of SRH from diverse domains, including demographic, lifestyle, socio-environmental factors, and physical and mental health status; higher education and income are associated with better SRH status (Franks et al., 2003; Kunst et al., 2005; Mackenbach, 2005; Manderbacka et al., 1999; McFadden et al., 2008; Molarius et al., 2007; Shields, 2008; Shields and Shooshtari, 2001; Singh-Manoux et al., 2006). Adult SRH is also influenced by early-life factors (e.g. social circumstances at birth and school qualifications) (Power et al., 1998). The potential modifying effect of socioeconomic status (SES) on the relationship between objective health and SRH has been explored in earlier studies, with inconsistent findings (Delpierre et al., 2009, 2012; Dowd and Zajacova, 2010; Onadja et al., 2013; Singh-Manoux et al.,

* Corresponding author.
E-mail address: shilpan@liv.ac.uk (S. Nayak).

http://dx.doi.org/10.1016/j.pmedr.2016.06.001

2007). There is also evidence suggesting SES does not modify the association between SRH and mortality, and that influence of health-related predictors is similar across socioeconomic groups (Burstrom and Fredlund, 2001; McFadden et al., 2009; Smith et al., 2010). Such inconsistencies may in fact result in an underestimation of health inequalities (Delpierre et al., 2009; Dowd and Zajacova, 2010; McFadden et al., 2009; Singh-Manoux et al., 2007). SES may affect expectations of health and risk, the factors considered in assessing subjective health, or their relative weighting. Socioeconomic circumstances can determine the range of factors pertinent to health; we explore this further, in the context of income, in the present study.

In addition to adverse childhood circumstances, a greater prevalence of adverse material circumstances, unhealthy behaviors and psychosocial factors are important in explaining health inequalities (van Lenthe et al., 2004). Lifestyle choices are rooted in socioeconomic context. In targeting factors such as physical exercise, smoking or alcohol consumption, there is value in understanding the concurrent upstream factors that might influence or restrict these choices (Marmot et al., 1998). Meyer et al., for example, found low SES linked to greater neighborhood safety concerns; these were negatively associated with physical activity, which in turn was negatively linked with mental health and SRH (Meyer et al., 2014). Thus, in low SES groups acting primarily on physical activity levels without addressing contextual factors which influence it, may not impact on health status. Mitigation of cumulative adverse effects requires a multi-level and multi-dimensional approach to intervention (Morello-Frosch et al., 2011; Wen et al., 2006).

Given the complexity of the socioeconomic gradient in health, Adler et al. discussed conceptual and methodological issues constraining earlier research on SES and health; one such issue is the limited ability of parametric multivariate regression to capture a large number of multi-domain interrelated variables, and fully unravel the mechanisms that might contribute to the gradient (Adler et al., 1994). Classification tree analysis (CTA), a form of recursive partitioning, provides an alternative approach with several advantages: this non-parametric technique is valuable for studying a complex set of predictor variables, and large sample size; it is data-adaptive, handles high dimensionality, a mixture of data types, and non-standard data structure, while providing insight into the predictive structure of the data (Breiman et al., 1984). Tree-based methods have been used to partition individuals and establish high risk groups by clinical signs and symptoms (Kershaw et al., 2007); they may also uncover interactions potentially overlooked in logistic regression, unless modeled a priori (Forthofer and Bryant, 2000; Lemon et al., 2003; Nelson et al., 1998).

The aim of this study is to apply CTA to investigate combinations of multi-domain health determinants associated with self-rated health (SRH), and conduct an exploratory analysis of their combinations and relative importance in income-based groups. The factors considered represent multiple influences from the social-ecological model of health; a fundamental aim of the study was to contextualize these multi-domain factors, and study their potential joint impact and interactions. It is unclear whether the relative importance of risk factors associated with health status remains the same across income-based groups. We propose these would vary based on interaction of lifestyle choices, psychosocial factors, and living and working conditions, influenced by socioeconomic context. An understanding of these differences is important in planning interventions to improve health status, and reduce income-based health inequalities.

2. Methods

This study utilized cross-sectional data collected by the CARDIA longitudinal study (Coronary Artery Risk Development in Young Adults), started in 1985 with a cohort of 5115 men and women aged 18–30 years (1.1% of participants were 17–35 years), recruited in Birmingham, Alabama; Chicago, Illinois; Minneapolis, Minnesota; and Oakland, California. For this study, data were taken from the year 15 follow-up, conducted 2000–2001 (except race/ethnicity - 1985–1986, family history - 1995), through interviewer and self-administered questionnaires, to examine associations between SRH and many health determinants assessed for adults (mean age 40 years) in that year. From 5115 participants at baseline, 3672 were followed-up in year 15; participants with a response for SRH, coded as male or female, were included in the study sample of 3648 participants.

2.1. Study variables

Outcome, SRH, was assessed on a 5-point scale: "In general would you say your health is excellent, very good, good, fair or poor?" Responses of poor/fair/good were grouped as 'lower' SRH. Responses of very good/excellent were grouped as 'higher' SRH, as they were more definite positive statements of better health; respondents may have regarded a response of good, being the centre of a 5-point scale, as a neutral or 'average' value. This grouping also resulted in more equal group sizes. In a previous study, fair and good self-ratings of health were associated with higher mortality, so that risk was not associated solely with the poor group, but a gradient was observed. (Idler et al., 1990).

Predictor variables used in the analysis (appendix, Table 1A) represented multiple domains and a range of health determinants based on the social-ecological model of health: age, sex, and hereditary factors; individual lifestyle factors and medical history; social and community influences; living and working conditions (Gebbie et al., 2003; Dahlgren and Whitehead, 1991).

2.2. Statistical analyses

The study sample was split into 5 groups based on respondents' total family income. Mantel-Haenszel chi-square test for trend was used to assess the relationship between predictor variables and ordinal income categories. Continuous variables were analysed using the Kruskall-Wallis test.

For each income-based group, CTA was run using all predictor variables, excluding total family income, to segment the group into smaller mutually exclusive subgroups of individuals, and identify predictor variables associated with the outcome measure, SRH. At every node of the tree model formed in the analysis, the sample of individuals was split based on the predictor variable that maximised the goodness of split function, i.e. resulted in the largest decrease in impurity of the prior 'parent' node (a node that is split further into subgroups), in terms of distribution of SRH status. A ranking of predictor variables was based upon normalized importance, ranging from 0 to 100, with the variable with the greatest relative measure of importance scored at 100, and other variables scored in the range 0 to 100 (Breiman et al., 1984). Tree growing criteria were set to a minimum 'parent node' size of 20 (individuals) and 'child node' size of 10 (a child node is a subgroup formed from splitting of a parent node). Data were analysed using IBM SPSS Statistics (SPSS v21).

3. Results

Distribution of the study sample by income-based group: under $25,000: 16% ($n$ = 578); $25,000–$50,000: 25% (n = 911); $50,000–$75,000: 22% (n = 791); $75,000–$100,000: 15% (n = 527); and $100,000 and over: 22% ($n$ = 797) (N = 3604; income data missing for 44 participants).

Distribution of study variables by income-based group (Table 1):

SRH: Proportion of 'higher' SRH increased with income ($p < 0.05$).

Sex, race/ethnicity and hereditary factors: There was an increasing proportion of males and whites, with higher income, and an inverse income gradient for proportion of respondents with family history of diabetes, stroke, maternal high blood pressure, maternal angina, and