Preventive Medicine 56 (2013) 365–371

Contents lists available at SciVerse ScienceDirect

Preventive Medicine

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

Review

A scoping review of statistical approaches to the analysis of multiple health-related behaviours

Kareena McAloney [a,*], Hilary Graham [b], Catherine Law [c], Lucinda Platt [d]

[a] Department of Health Sciences, Seebohm Rowntree Building, University of York, Heslington, York, YO10 5DD, UK
[b] Department of Health Sciences, University of York, UK
[c] Centre for Paediatric Epidemiology and Biostatistics, UCL Institute of Child Health, UK
[d] Institute of Education, University of London, UK

ARTICLE INFO

Available online 18 March 2013

Keywords:
Health behaviours
Clustering
Co-occurrence
Cluster analysis
Latent class analysis

ABSTRACT

Background. Smoking, diet, exercise, and alcohol are leading causes of chronic disease and premature death, many engage in two or more of these behaviours concurrently. The paper identified statistical approaches used to investigate multiple behavioural risk factors.

Method. A scoping review of papers published in English from 2000 to 2011 was conducted; papers are related to concurrent participation in at least two of the behaviours. Statistical approaches were recorded and categorised.

Results. Across 50 papers, two distinct approaches were identified. Co-occurrence analyses focused on concurrent but independent behaviours, represented by prevalence of behavioural combinations and/or by the summing behaviours into risk indexes. Clustering analyses investigated underlying associations between the concurrent behaviours, with clustering identified by divergences in observed and expected prevalence of combinations or through identification of latent or unobservable clusters. Co-occurrence was more frequently reported, but the use of clustering techniques and, in particular, cluster analytic and latent variable techniques increased across the study period.

Discussion. The two approaches investigate concurrent participation in multiple health behaviours but differ in conceptualisation and analysis. Despite differences, inconsistency in the terminology describing the study of multiple health behaviours was apparent, with potential to influence understandings of concurrent health behaviours in policy and practice.

Contents

Background

Behavioural factors like smoking, physical inactivity, diet and alcohol consumption are major proximal causes of chronic disease and premature mortality (WHO, 2008) and have a high economic cost (Scarborough et al., 2011).

To date, research and policy have tended to focus on single behaviours; for example, alcohol research and policy have developed in relative isolation from tobacco control research and policy (Department of Health, 2010; Home Office, 2012). However, lifestyles are made up of multiple behaviours, and adults can fail to meet government

* Corresponding author. Fax: +44 1904 321320.
E-mail address: kareena.mcaloney@york.ac.uk (K. McAloney).

0091-7435/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ypmed.2013.03.002

recommendations for more than one behaviour. In England, for example, 64% of those aged 16 and over take less than the recommended level of physical activity and 52% of adults do not consume the recommended daily portions of fruit and vegetables; 35% consume more than the recommended quantity of alcohol on their heaviest drinking day and 20% are regular smokers (NHS Information Centre, 2011). As this suggests, many young people and adults engage in more than one behaviour associated with an elevated risk of ill-health (Coups et al., 2004; Mistry et al., 2009; Poortinga, 2007) and there is evidence that doing so can have a multiplicative rather than additive effect on health risk (Conry et al., 2011).

The need for a broader lifestyle perspective is widely appreciated but its development presents a number of methodological challenges. A fundamental challenge concerns on how to represent and analyse multiple risk behaviours, a challenge addressed through the statistical techniques employed to capture and investigate sets of health-related behaviours (Abel, 1991). Scoping reviews are often undertaken to provide a rapid overview of emergent research areas and to assist in identifying questions to be addressed by a subsequent systematic review (Arksey and O'Malley, 2005). We therefore undertook a scoping review to review and assess the statistical approaches used to characterise and investigate multiple behavioural risk factors.

Methods

Scoping reviews provide a broad and rapid overview of a research area, and are regarded as particularly useful for rapidly-developing fields marked by a diversity of approaches (Arksey and O'Malley, 2005; Davis et al., 2009; Grant and Booth, 2009; Levac et al., 2010). A scoping review was conducted in autumn 2011 of studies of multiple risk behaviours. We applied Arksey and O'Malley's widely-used framework, which parallels to that used for systematic reviews (Arksey and O'Malley, 2005) but recommends broad search terms and inclusion criteria without quality appraisal filters (Davis et al., 2009; Grant and Booth, 2009). This framework provides a structured methodological approach that, in contrast to the tightly-defined questions that underpin systematic reviews, is designed to answer broad and general research questions. The scoping review is an iterative process, with search terms subject to refinement in the light of the studies identified (Arksey and O'Malley, 2005).

We searched three major electronic bibliographic databases (Medline, PsychInfo, and ScienceDirect) for papers published in English from 2000 to 2011. We used search terms for smoking, drinking, physical activity and diet, with separate searches performed for pairs of these terms (smoking + drinking terms; smoking + physical activity terms; smoking + diet terms etc.). Additional searches used terms specifically related to concurrent behaviours (e.g. 'multiple health behaviours' 'cluster', 'co-occurrence'). No other restrictions were applied.

Results

One hundred and twelve papers were identified as potentially relevant (Fig. 1). Titles, abstracts and method sections were independently screened by two reviewers. Of these, only fifty papers included analyses related to concurrent participation in at least two of the four health behaviours. For each paper, data were extracted on year of publication, study population, health-related behaviours and other risk factors, and analytic approach (Table 1; papers, numbered and prefixed by P). Study's results were reported in two papers (P5 and P17), and another study in three papers (P2, P21, P35). For each paper, data were extracted on year of publication, study population, health-related behaviours and other risk factors, and analytic approach.

The 50 papers were skewed towards the later years of the review period. Thus, there were no papers in 2000–2002 and only one paper in 2003; in 2004–2007 and 2008–2011, there were respectively 11 and 38 papers. Behaviours related to smoking and activity levels (including sedentary behaviour and physical inactivity) were most frequently-reported (45 of the 50 papers for each behaviour) and over half of the papers (n = 27) included all four lifestyle factors. Additionally, 20 papers included anthropometric (e.g. weight, BMI) and clinical (e.g. blood pressure, cholesterol) measures, and four papers included other health-related behaviours such as sexual risk taking, condom use and illegal drug use.

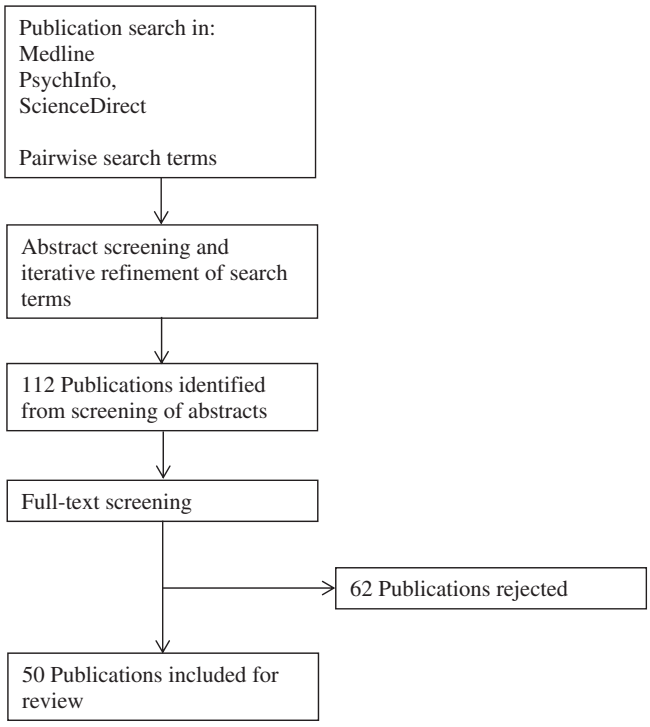

Fig. 1. Flow diagram of publication selection procedure.

A wide range of terms was used to describe the patterning of these behaviour-related risk factors, with titles of papers referring to prevalence, association, co-occurrence and clustering. The latter was the most frequently used; 22 papers referred to clustering or cluster(s) in their titles (P1, P3–4, P7, P10, P22, P26, P29–31, P33, P36–37, P39, P41, P43–48, P50).

With respect to analytical techniques, two major approaches were identified (Table 1). In line with distinctions used in epidemiological research (Ebrahim et al., 2004), we label these *co-occurrence* and *clustering*. Analyses of co-occurrence are directed at the concurrent, but independent, engagement in two or more health-related behaviours (e.g. being a smoker, being physically inactive and binge drinking); analyses of clustering investigate underlying associations between co-occurring health-related behaviours. Of the 50 papers, 37 (74%) included analyses of co-occurrence (P1–37) and 25 (50%) of clustering (P26–50). The two approaches were not mutually exclusive; analyses of co-occurrence were included as a preliminary step in 12 of the 25 papers that analysed clustering (P26–37). Interestingly, six of the papers referring to clustering/cluster(s) in their titles did not include an analysis of clustering in their results, instead they reported co-occurrence among the behaviours studied (P1, P3–4, P7, P10, P22).

Co-occurrence of risk behaviours (37 papers; P1–37): There were two broad approaches to presenting information on co-occurrence. These approaches tended to dichotomise the behaviours of interest into 'risky' (e.g. smoking) or 'not risky' (not smoking), but differ in how they analyse the concurrent nature of the behaviours.

The first approach is to report the prevalence of the different behaviour combinations, an approach adopted in sixteen papers (P5, P14, P18, P20, P 25, P26–36). While simple in principle, the co-occurring patterns of behaviour can quickly generate a large number of behavioural combinations ($2^{\text{number of behaviours}}$). P14 examined physical activity and diet in teenagers, using three lifestyle factors (physical activity, fruit and vegetable consumption, and breakfast consumption) and reported the prevalence of each of the eight potential