Spatial and Spatio-temporal Epidemiology 19 (2016) 1–9

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

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

Small-area based smoothing method for cancer risk mapping

Toni Patama[a,*], Eero Pukkala[a,b]

[a] Finnish Cancer Registry, Institute for Statistical and Epidemiological Cancer Research, Helsinki, Finland
[b] School of Health Sciences, University of Tampere, Tampere, Finland

ARTICLE INFO

Article history:
Received 15 November 2015
Revised 7 April 2016
Accepted 19 May 2016
Available online 27 May 2016

Keywords:
Cancer
Spatial
Mapping
Epidemiology
Incidence

ABSTRACT

Lay people and health professionals are increasingly interested in health issues in their region. The high-quality cancer and population registries in the Nordic countries allow fine spatial and temporal visualization of mapped data. This article describes – with real-data examples – a mapping method developed in Finland for such visualization but also utilized for data from numerous other countries and non-cancer outcomes. The Finnish smoothing is based on weighting small-area specific observations with population sizes and distance without losing the interpretability of the values. The approach has essentially improved the readability of spatio-temporal trends of incidence and mortality rates of cancers and has been applied in many studies. With help of the tool, decision makers can understand the key message of existing data and use it more effectively in their efforts to provide better health care services.

1. Introduction

High-quality cancer and population registries in several countries allow fine spatial and temporal visualization. Cancer registration started in Finland in 1953 and maps of cancer incidence or mortality have been published since the 1960 for geographical units such as health care districts. However, sometimes marked variations exist within such larger units because the patterns of aetiological factors do not necessarily follow the administrative borders (Pukkala et al., 2001). For example, the way of life of people living near the administrative centre of an area is often different from that of people living in the periphery. To overcome the issue of having too large and heterogeneous areal units, maps have been produced for small administrative units such as municipalities. Historically, the average population of municipalities in Finland was about 10,000, but currently it is increasing due to a merging process.

There are two major problems with using municipality level for mapping purposes: (1) due to the small numbers of cases per municipality, the random variation is disturbingly high; and (2) the geographical area of municipalities with highest populations (cities) may be so small that the most important areas are hardly visible when mapped. Further, with traditional mapping methods, it was rather difficult (or impossible) to gain any visual impression of geographical trends (Pukkala, 1989).

The first attempt at using the spatial interpolation method of smoothing small area based rates was in Norway in 1980 (Glattre et al., 1985) and the method was developed further by the Finnish Cancer Registry for better illustration of spatial patterns of incidence rates. The first Finnish smoothed maps were published in 1987 in an atlas of cancer incidence in Finland in the period 1953–1982 (Pukkala et al., 1987). The experience of adopting the smoothing technique for cancer incidence and mortality maps for Northern Europe as a whole (Pukkala et al., 2001) indicated the need to produce visualizations of cancer-related measures that allow easy comparisons between countries and between calendar periods. The Finnish Cancer Registry started to develop computer software for

* Corresponding author.
E-mail address: toni.patama@gmail.com (T. Patama).

http://dx.doi.org/10.1016/j.sste.2016.05.003

Table 1
Examples of regions and phenomena that have been visualized with the smoothing method described in this article.

Region	Phenomenon
Finland	Cancers incidence and mortality (Pukkala and Patama, 2008, Pukkala and Patama, 2010)
	Human papilloma virus infections (Lehtinen et al., 2006)
	Leucocyte antigens and cervical cancer incidence (Castro et al., 2007)
	Kaposi sarcoma (Kaasinen et al., 2013)
	Cancer incidence forecast (Vanhatalo et al., 2010)
North Europe (13 countries)	Cancer incidence, mortality, mortality/incidence ratios (Pukkala et al., 2001)
Nordic countries (Denmark, Finland, Iceland, Norway and Sweden)	Cancer incidence (Pukkala et al., 2007; Patama et al., 2013)
	Cancer mortality (Patama et al., 2008)
Sweden, Stockholm and Gotland	Cancer incidence (Adolfsson et al., 2008)
Poland	Cancer incidence and mortality (Zatonski et al., 1993a, 1993b)
Belarus and Ukraine	Incidence of breast cancer after Chernobyl accident (Pukkala et al., 2006)
The Netherlands	Cancer sites incidence and mortality (Siesling et al., 2008; Elferink et al., 2012; van der Aa et al., 2008)
Belgium	Cancer incidence (Belgian Cancer Registry, 2011)
United Kingdom	Cancer incidence in South-West England (SWPHO, 2008)
	Cancer survival in England (Queresma et al., 2014)
Slovenia	Cancer incidence (Žagar et al., 2011)
USA (Ohio)	Cancer incidence, mortality, mortality/incidence ratios (Tyczynski et al., 2006)
Philippines (Metro Manila and Rizal province)	Cancer incidence (Laudico et al., 2008; Redaniel et al., 2008; Medina et al., 2011)

building animated smoothed cancer maps based on the smoothing method. This application has been used to produce smoothed maps of cancer incidence and other phenomena for several countries worldwide (Table 1).

In this article we describe the current version of the Finnish Cancer Registry's mapping method based on smoothing the small area specific rates of any frequency measure, taking account distance to neighbourhood and population. Options and challenges of the approach are illustrated with several real-data examples. This article is based on the materials used in a chapter of the book *Geography and Health – A Nordic Outlook* (Patama et al., 2014), with some updates.

2. Materials, methods and results

2.1. Visualization of map data

Visualization of spatial disease frequency data has traditionally been done using maps on which the rates are presented based on administrative boundaries (i.e. thematic maps). An example of this type of visualization is shown in Fig. 1, which shows the main example phenomenon selected for this presentation, the incidence of lung cancer among Finnish women in the period 1993–2002. A part (a) shows the rates for the health care districts under the responsibility of university hospitals in Finland, which have 0.7–1.7 million inhabitants. The full geographical variation of lung cancer incidence among women in Finland is colour coded. Clearly, only part of the variation in lung cancer incidence among women in Finland is illustrated in the large-area presentation.

Fig. 1b shows the actual rates for each of the c. 400 municipalities in Finland (in which the female population ranges from less than 200 to 300,000). The incidence rates in sparsely populated municipalities vary from 0 to 40 per 10^5, mainly due to random variation. The municipalities with a large area but low population density located in the north can have dramatic effect on map, whereas usually the highly populated cities have small geographical areas and therefore they can be almost invisible on the maps.

Many of the problems related to issues of too large areal units and too small areal units can be overcome by using the smoothing technique described below.

2.2. The smoothing technique

A computerized mapping method for smoothing was originally developed for smoothed averages of real measurements of the proportion of minerals in soils (Björklund and Gustavsson, 1987). Subsequently, the technique was adopted for mapping incidence of cancer (Pukkala, 1989). The main purpose in the method is to model trends in the spatial variation of risk of developing cancer.

The rate for each square (in our example maps of Finland 2 km × 2 km) of grid on the map was defined as a weighted average of the age-adjusted incidence rates in the municipalities, with population centres within 250 km (R) from the centre of the grid (Fig. 2). The weights (w_i) were inversely associated with distance; the weight being halved where the distance d_0 was 25 km (see weighting function using a bell-shaped function known as the Butterworth's function in Fig. 3). In addition, frequencies, the populations (directly proportional) of the rates I_i are considered as weights too. The formula of the weighted averages I_j within moving window with radius (R) can be deduced as follows.

At first we need to calculate weights (w_i) for the distances (d_i) of the municipalities using formulas 1 for every geographical unit point ($i = 1, 2, 3, \ldots$) (municipalities) in the given floating window:

$$w_i = \frac{1}{1 + (\mathrm{d}_i/d_0)^2} \quad (1)$$

The weighting factor W_i is adjusted by population P_i of the rate, we get

$$W_i = w_i P_i \quad (2)$$