



Control of data quality for population-based cancer survival analysis[☆]



Ruoran Li^{*}, Louise Abela, Jonathan Moore, Laura M. Woods, Ula Nur, Bernard Rachet, Claudia Allemani, Michel P. Coleman

Cancer Research UK Cancer Survival Group, London School of Hygiene and Tropical Medicine, London, UK

ARTICLE INFO

Article history:

Received 14 February 2014
Accepted 23 February 2014
Available online 29 March 2014

Keywords:

Cancer
Survival
Research design
Population register
Epidemiology

ABSTRACT

Background: Population-based cancer survival is an important measure of the overall effectiveness of cancer care in a population. Population-based cancer registries collect data that enable the estimation of cancer survival. To ensure accurate, consistent and comparable survival estimates, strict control of data quality is required before the survival analyses are carried out. In this paper, we present a basis for data quality control for cancer survival.

Methods: We propose three distinct phases for the quality control. Firstly, each individual variable within a given record is examined to identify departures from the study protocol; secondly, each record is checked and excluded if it is ineligible or logically incoherent for analysis; lastly, the distributions of key characteristics in the whole dataset are examined for their plausibility.

Results: Data for patients diagnosed with bladder cancer in England between 1991 and 2010 are used as an example to aid the interpretation of the differences in data quality. The effect of different aspects of data quality on survival estimates is discussed.

Conclusions: We recommend that the results of data quality procedures should be reported together with the findings from survival analysis, to facilitate their interpretation.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Population-based cancer survival is one of the most important measures of the overall effectiveness of cancer care and control in a population, alongside incidence and mortality. Trends in cancer survival provide an indication of improvements in diagnosis and treatment [1].

Standard checks required for cancer incidence data have been described [2–4] and are embodied in the widely used IARC Check program [5]. However, additional quality checks are required for survival analysis, because the completeness and validity of data on vital status (alive, dead or lost to follow up) and follow-up time of the patients become crucial.

The interpretation of survival comparisons between countries or populations (defined by calendar period, socio-economic status, race or ethnicity) relies on the thoroughness of quality control

procedures, which ensure that incomplete, ineligible or incoherent tumour records are flagged and excluded. We describe a set of quality control procedures that have been applied to population-based data for several recent national and international studies of cancer survival [6–9]. This set of procedures can form a basis for data quality control in cancer survival analysis.

2. Materials and methods

2.1. Cancer registry data

Cancer registries collate data from sources such as hospitals, general practitioners, pathology departments, cancer referral units and screening programmes, and obtain one record for each tumour including data on the patient (date of birth, sex, residence or postcode, ethnicity, identifier), the tumour (date of diagnosis, topography, morphology, behaviour, microscopic confirmation, stage at diagnosis), the treatment (surgical procedure, chemotherapy, radiotherapy) and the outcome (date and place of death) [10,11]. This process may not be completed for six to nine months, until a patient's course of treatment has finished.

Information on the patient's vital status is later added from sources such as regional or national death indexes, social security, health insurance, death certificates, physician or hospital contacts

[☆] This work was funded by Office for National Statistics and a Cancer Research UK programme grant (C1336/A11700).

^{*} Corresponding author at: Cancer Research UK Cancer Survival Group, Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. Tel.: +44 020 7927 2855.

E-mail address: ruoran.li@lshtm.ac.uk (R. Li).

and/or home visits. The key concern is that the eventual death of all registered cancer patients is recorded. The quality and completeness of this information is essential for accurate estimation of survival.

2.2. Defining the cancers

Cancers are defined by their anatomic location (site) and microscopic appearance (morphology), and whether they are benign, in situ, malignant or of uncertain behaviour (behaviour), under the International Classification of Diseases [12] or the International Classification of Diseases for Oncology [13]. Various utilities exist to convert ICD codes between the various revisions [14,15].

In what follows, we write from the perspective of a general cancer registry, with data on all cancers.

2.3. Quality control

Quality control procedures are designed to ensure that survival analyses include only patients resident in the defined territory who were diagnosed with a primary, invasive, malignant neoplasm during a defined calendar period, and whose tumour record is valid and logically coherent [16].

We propose three distinct phases for the quality control of cancer data for survival analysis (Fig. 1). In the following sections, we will describe the rationale and process for each of these phases with accompanying examples. As in any data quality control process, feedback is provided to the data sources, i.e. the registries, which may lead to modifications. In a study involving several registries, quality control would entail discussion between the analytic centre and each registry concerned.

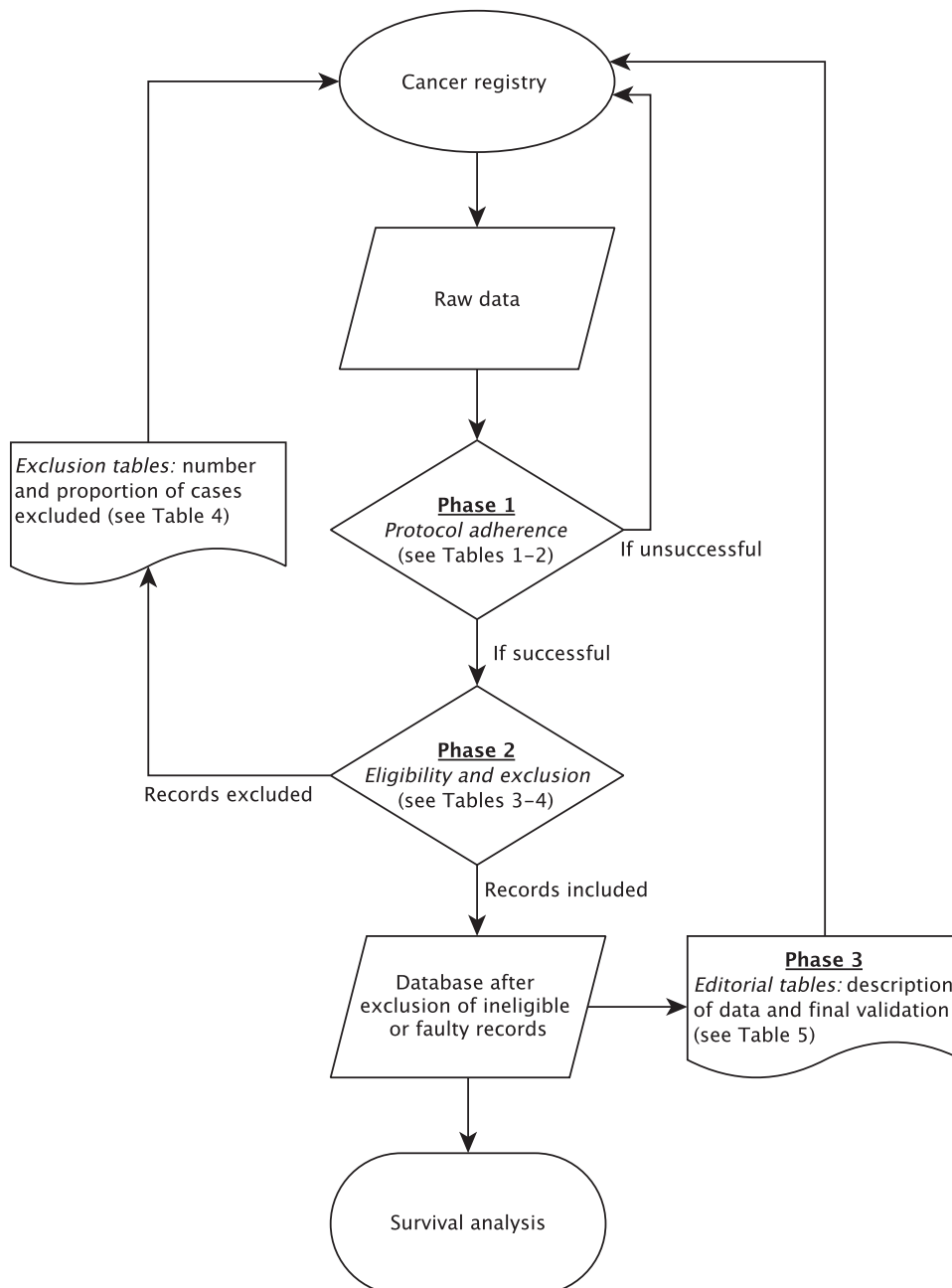


Fig. 1. Schematic of the cleaning process for population-based cancer survival data.

Download English Version:

<https://daneshyari.com/en/article/2108993>

Download Persian Version:

<https://daneshyari.com/article/2108993>

[Daneshyari.com](https://daneshyari.com)