Cancer Epidemiology 52 (2018) 28–42

Contents lists available at ScienceDirect

Cancer Epidemiology

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

Original Research Article

Missing data and chance variation in public reporting of cancer stage at diagnosis: Cross-sectional analysis of population-based data in England

Matthew E. Barclay[a,b], Georgios Lyratzopoulos[a,b,c,*], David C. Greenberg[a,b], Gary A. Abel[d]

[a] Cambridge Centre for Health Services Research, Department of Public Health and Primary Care, Forvie Site, Robinson Way, Cambridge, CB2 0SR, United Kingdom
[b] National Cancer Registration and Analysis Service, Public Health England, Victoria House, Capital Park, Fulbourn, Cambridge, CB21 5XA, United Kingdom
[c] Epidemiology of Cancer Healthcare and Outcomes (ECHO) Research Group, Department of Behavioural Science and Health, University College London, WC1E 7HB, United Kingdom
[d] University of Exeter Medical School (Primary Care), Smeall Building, St Luke's Campus, Exeter, EX1 2LU, United Kingdom

ABSTRACT

Background: The percentage of cancer patients diagnosed at an early stage is reported publicly for geographically-defined populations corresponding to healthcare commissioning organisations in England, and linked to pay-for-performance targets. Given that stage is incompletely recorded, we investigated the extent to which this indicator reflects underlying organisational differences rather than differences in stage completeness and chance variation.

Methods: We used population-based data on patients diagnosed with one of ten cancer sites in 2013 (bladder, breast, colorectal, endometrial, lung, ovarian, prostate, renal, NHL, and melanoma). We assessed the degree of bias in CCG (Clinical Commissioning Group) indicators introduced by missing-is-late and complete-case specifications compared with an imputed 'gold standard'. We estimated the Spearman-Brown (organisation-level) reliability of the complete-case specification. We assessed probable misclassification rates against current pay-for-performance targets.

Results: Under the missing-is-late approach, bias in estimated CCG percentage of tumours diagnosed at an early stage ranged from −2 to −30 percentage points, while bias under the complete-case approach ranged from −2 to +7 percentage points. Using an annual reporting period, indicators based on the least biased complete-case approach would have poor reliability, misclassifying 27/209 (13%) CCGs against a pay-for-performance target in current use; only half (53%) of CCGs apparently exceeding the target would be correctly classified in terms of their underlying performance.

Conclusions: Current public reporting schemes for cancer stage at diagnosis in England should use a complete-case specification (i.e. the number of staged cases forming the denominator) and be based on three-year reporting periods. Early stage indicators for the studied geographies should not be used in pay-for-performance schemes.

1. Introduction

The percentage of cancer patients diagnosed at an 'early stage' (i.e. TNM stages 1–2) has been routinely reported for National Health Service commissioning organisations (Clinical Commissioning Groups, CCGs) since 2014 [1], following recommendations in the 2011 national cancer strategy for England [2]. Recently, this indicator has been adopted into a pay-for-performance scheme for CCGs [3]. Typical CCGs meeting the relevant targets in a given year would receive a financial incentive of £250,000. The aim of these public reporting and pay-for-performance schemes is to promote diagnosis of cancer at an earlier stage and thereby improve outcomes for patients across England. We further summarise this policy context and the technical aspects of the indicator in Box 1.

Indicators used for comparing the performance of healthcare organisations should, among other considerations, be both valid and reliable. Valid indicators truly measure the intended construct of interest, while reliability indicates the precision by which the construct is measured. The validity of performance indicators based on routinely-collected healthcare data may be undermined by missing information [4,5]. Low reliability, where measures are not precise enough to distinguish organisational performance, is a prevailing concern when person-level measures are aggregated into organisation-level scores [6–9]. Frequently, indicators are published and used in pay-for-

* Corresponding author at: 1-19 Torrington Place, London, WC1E 7HB, United Kingdom.
E-mail address: y.lyratzopoulos@ucl.ac.uk (G. Lyratzopoulos).

https://doi.org/10.1016/j.canep.2017.11.005
Received 10 August 2017; Received in revised form 9 November 2017; Accepted 11 November 2017

Box 1
Early stage at diagnosis indicator

In the English National Health Service (NHS), the planning, funding and monitoring of healthcare delivery is the responsibility of 'healthcare commissioning' organisations currently known as Clinical Commissioning Groups. These are responsible for geographically-defined populations. There are about 200 Clinical Commissioning Groups across England, covering an average general population of about 250,000 residents. To support and promote their planning, funding and monitoring function, high level performance indicators for Clinical Commissioning Groups are published annually, across different disease areas, including cancer. In England, a nationwide population-based cancer registration system has been in existence since 1971. In recent years, the modernisation of cancer registration systems has enabled the capturing of information on stage at diagnosis for a high proportion of patients. This has allowed for the introduction of the 'early diagnosis' indicator for Clinical Commissioning Groups studied in our paper. This indicator relates to the stage at diagnosis of 10 different solid tumour sites, and can be met by a Clinical Commissioning Group if either of the following criteria apply: a) 60% or greater proportion of all registered cases with relevant tumours are known to have been diagnosed in TNM stages 1 or 2; or b) there has been a 4% or greater absolute increase within a year in the proportion of all registered cases with relevant tumours known to have been diagnosed in TNM stages 1 or 2.

performance schemes without these concerns being examined or addressed.

The validity and reliability of the early stage indicator for CCGs as currently specified have not been evaluated. Currently, patients with cancer with no recorded stage are treated as though they had late stage cancer, but an alternate specification excluding such patients may be more appropriate. Furthermore, the annual reporting period may be either unnecessarily long or too short to allow for reliable estimation of performance. In this article, we demonstrate how appropriate statistical techniques may be used to examine the properties of this indicator, and identify specific improvements to reduce bias and improve its reliability.

2. Materials and methods

2.1. Data sources

We used population-based data (Public Health England National Cancer Registration and Analysis Service) on TNM stage at diagnosis and other patient and tumour characteristics of patients diagnosed during 2013 with 10 common cancers: bladder (ICD10 C67); female breast (C50); colorectal (C18–C20); endometrial (C54); lung (C33–C34); ovarian (C56–C574); prostate (C61); and renal (C64) cancers; melanoma (C43); and non-Hodgkin lymphoma (C82–C85). The choice of cancer sites and definition of early stage (TNM stages 1–2) reflected those included in the Public Health Outcomes Framework and the CCG Quality Premium; for both, data relating to patients diagnosed in 2013 was reported in 2014 [1,3,10,11].

2.2. Analysis

2.2.1. Examining bias arising from missing data in indicators of early stage at diagnosis

In the study year (2013) stage completeness across all 10 cancer sites was 82%, ranging from 71% to 91% for renal and endometrial cancer, respectively. We used multiple imputation by chained equations (MI) to produce a 'best estimate' early stage indicator, which we treated as the gold standard. Separately by cancer site, a binary early stage indicator for each patient was imputed with logistic regression [12], using auxiliary information on important patient and tumour characteristics associated with stage at diagnosis including patient age, sex, tumour grade (partially missing), CCG, and survival time from diagnosis [13–16]. The MI indicator for each CCG was estimated as the mean percentage of tumours diagnosed at early stage over ten imputed datasets [17]. Appendix A contains further details of the imputation model.

We judged *a priori* that indicators based on the MI approach were not suitable for routine use in public reporting, primarily due to the need for follow-up periods to have elapsed to obtain survival information for use in imputation models, as well as the computational complexity and lack of end-user familiarity with the underlying statistical methods. Instead simpler approaches would be preferable if they are not associated with a substantial degree of bias. We therefore investigated the degree of bias in CCG scores using two simpler approaches for producing early stage indicators. First, the 'missing-is-late' indicator, where the percentage of all tumours with recorded early stage is estimated assuming that those without recorded stage information are advanced stage tumours. The missing-is-late approach is currently used to produce early stage indicators [1,3,10]. Second, the 'complete-case' indicator, where the percentage of staged tumours diagnosed at early stage is estimated based only on tumours with observed stage. We described the degree of bias in either missing-is-late or complete-case indicators by comparing organisational estimates against the 'best estimate' MI indicator.

2.2.2. Examining the reliability of early stage indicators

The statistical reliability of a measure indicates its reproducibility (consistency) in repeated measurement and its robustness to random measurement error. Here we are concerned with organisation-level (or Spearman-Brown) reliability which represents the extent to which organisational measures (in our case the measured percentages of cancer patients diagnosed in early stage) reflect true differences between organisations, as opposed to random (i.e. chance) variation [7,18–20]. For further details of the calculation of reliability for binary indicators, see Appendix B.

Mixed effects logistic regression models were used to model variation in the percentage of tumours diagnosed at early stage estimated using the complete-case indicator. Our main focus was the composite (all 10 cancers) indicator for CCGs, but we performed similar analyses for each individual cancer site (see Appendix B) and for local government organisations (local authorities) and general practices. These models produced an estimate of the organisation-level variance on the log-odds scale. The estimated variance was used to calculate odds ratios for diagnosis at early rather than late stage comparing the 75th/25th and 95th/5th percentiles of the distribution to illustrate the variation between organisations. Importantly, this was the underlying (true) variation which can be thought of as that which would be seen with very large sample sizes in each organisation, such that the influence of sampling variation would be minimal. This underlying (true) variation will be less than the variation in observed stage metrics as the latter will also include a contribution from chance/sampling [19]. The organisation-level variance on the log-odds scale was also used to calculate the reliability for each indicator based on the number of cases in the study year.

In addition to estimating the reliability of the observed data, model outputs were used to estimate the number of tumours required for each