

How much do tumor stage and treatment explain socioeconomic inequalities in breast cancer survival? Applying causal mediation analysis to population-based data

Ruoran Li¹ · Rhian Daniel^{2,3} · Bernard Rachet^{1,3} 

Received: 17 November 2015 / Accepted: 28 April 2016 / Published online: 10 May 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Substantial socioeconomic inequalities in breast cancer survival persist in England, possibly due to more advanced cancer at diagnosis and differential access to treatment. We aim to disentangle the contributions of differential stage at diagnosis and differential treatment to the socioeconomic inequalities in cancer survival. Information on 36,793 women diagnosed with breast cancer during 2000–2007 was routinely collected by an English population-based cancer registry. Deprivation was determined for each patient according to her area of residence at the time of diagnosis. A parametric implementation of the mediation formula using Monte Carlo simulation was used to estimate the proportion of the effect of deprivation on survival mediated by stage and by treatment. One-third (35 % [23–48 %]) of the higher mortality experienced by most deprived patients at 6 months after diagnosis, and one tenth (14 % [–3 to 31 %]) at 5 years, was mediated by adverse stage distribution. We initially found no evidence of mediation via differential surgical treatment. However, sensitivity analyses testing some of our study limitations

showed in particular that up to thirty per cent of the higher mortality in most deprived patients could be mediated by differential surgical treatment. This study illustrates the importance of using causal inference methods with routine medical data and the need for testing key assumptions through sensitivity analyses. Our results suggest that, although effort for earlier diagnosis is important, this would reduce the cancer survival inequalities only by a third. Because of data limitations, role of differential surgical treatment may have been under-estimated.

Keywords Breast cancer · Survival · Socioeconomic inequalities · Causal mediation · Population-based data · Tumour stage · Surgical treatment

Introduction

Substantial socioeconomic inequalities in cancer survival have been observed in England for decades [1–3], meaning that many cancer deaths could be avoided [4]. For breast cancer, besides lower screening uptake and differential access to treatment, more advanced stage at diagnosis and severe comorbidity are regularly proposed as the most plausible explanatory factors of these inequalities [5, 6]. However, both factors seem to explain only part of these inequalities, at least for breast and colorectum cancers [7, 8].

Population-based data are crucial to understand the mechanisms affecting all patients and to help define policies. Quantifying the proportion of the effect of deprivation on survival that's attributable to differential stage of diagnosis and treatment is important for better resource allocation to address the gap between the rich and the poor. Methodological issues, however, are inherent to

Electronic supplementary material The online version of this article (doi:10.1007/s10654-016-0155-5) contains supplementary material, which is available to authorized users.

✉ Bernard Rachet
bernard.rachet@lshtm.ac.uk

¹ Cancer Research UK Cancer Survival Group, Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK

² Department of Medical Statistics, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

³ Centre for Statistical Methodology, London School of Hygiene and Tropical Medicine, London, UK

observational data. Most of the previous results were based on conventional analytic approaches (e.g. by describing the deprivation gap after adjusting for or stratifying by stage). However, if stage and treatment are on the causal pathway from deprivation to cancer survival, or if there is an interaction between deprivation and the mediator(s), these conventional approaches may lead to flaws in interpretation [9–12]. Using methods from the causal inference literature, we aim to disentangle the contributions of differential stage at diagnosis and differential treatment to the socioeconomic inequalities in cancer survival. To this end, we use population-based and routinely collected data for all patients diagnosed with a breast cancer within a defined area.

Materials and methods

Data

We included in the analyses all women (aged 15–99 at diagnosis) diagnosed with malignant, invasive breast cancer during 2000–2007, followed up until 31 December 2007, and collected by the Northern and Yorkshire Cancer Registry Information Service (NYCRIS), a population-based cancer registry covering 12 % of the English population. Ascertainment of the vital status was considered to be complete for all patients [13].

Each patient was allocated a socio-economic deprivation score according to her area (Lower Super Output Area) of residence at the time of diagnosis, using the English Indices of Multiple Deprivation (IMD) 2001 (income domain) [14]. These scores were categorised according to the quintiles of their national distribution.

Each patient was allocated one of the four broad tumor TNM stages using a restrictive approach [15].

Information on surgical treatment was retrieved from a routinely collected national hospital dataset (Hospital Episode Statistics or HES). We retained surgical treatment within 1 month before and 6 months after the cancer diagnosis. The treatment (OPCS-4) codes [16] were categorized based on recommendations made by the Site-Specific Clinical Reference Group (SSCRG) for breast cancer [17] (Appendix 1). These categories were then dichotomized into ‘major treatment’ (axillary dissection or other axillary nodal procedures, breast conserving surgery, mastectomy, and plastic surgery) and ‘minor or no surgery’ (other surgical procedures and none).

Analyses

We estimated net survival from breast cancer, for each deprivation group and by stage, using the Pohar-Perme

estimator [18] implemented in the Stata [19] package *stms* [20].

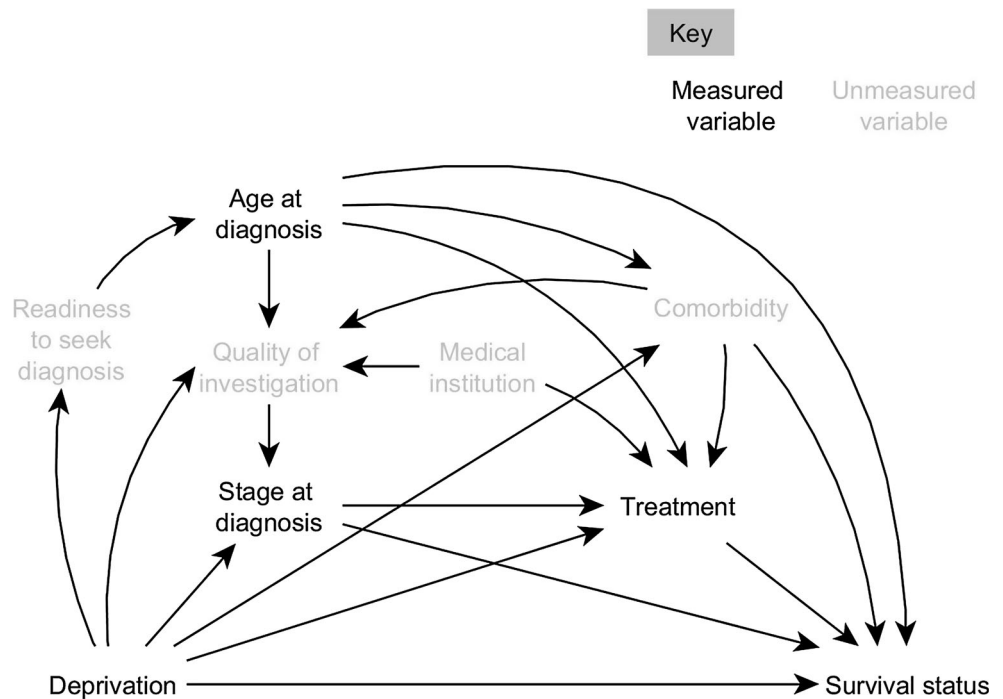
The assumed causal relationships between variables are represented by a Directed Acyclic Graph (DAG) (Fig. 1, Appendix 2). Our main exposure of interest, the patient’s deprivation level, causally influences the age at which a woman was diagnosed with breast cancer, her comorbidity, thoroughness of the disease investigation, stage at diagnosis, the treatment received, and survival status after the cancer diagnosis. Year and regions at diagnosis were considered as baseline confounders. Factors such as the quality of investigation and comorbidity (shown in grey as unmeasured variables) were incorporated in the DAG. The omission of variables and arrows also represents our causal assumptions, e.g. we assume that the quality of the investigation does not affect survival except through its effect on stage at diagnosis.

We examined what proportions of the deprivation gap in survival were explained separately by tumor stage and treatment. Because of our data structure (in particular, the existence of important mediator-outcome confounders affected by exposure, the likely presence of many interactions and the fact that our outcome is binary) we focused on the decomposition of the *total causal effect* (TCE) into what have recently been termed *randomized interventional analogues of natural direct and indirect effects*, henceforth RIANDE and RIANIE [21–23].

The RIANDE and RIANIE can be estimated with an extension of Robins’ g-computation formula [24] implemented using Monte Carlo simulation in the Stata command *gformula* [25]. We chose this method because of flexible modelling that allows interactions and other nonlinearities. Although flexible in terms of parametric modelling assumptions, this method relies on the assumptions of no unaccounted confounding of the exposure–mediator, mediator–outcome or exposure–outcome relationship.

We conducted three analyses to investigate the mediating roles of stage and treatment (Appendix 3, Appendix 4). We first estimated the proportion of the effects of deprivation on survival that was mediated by differences in stage at diagnosis, i.e. we computed the ratio between the effect of deprivation on log odds of death that was mediated by stage (the *RIANIE*) and the total effect of deprivation on log odds of death (the *total causal effect*, *TCE*, which is the sum of the *RIANDE*—the effect not mediated by the mediator stage—and the *RIANIE*). In the second analysis, we estimated the proportion of the effect of deprivation on log odds of death that was mediated by differences in treatment. Stage at diagnosis was here considered to be a confounder of the relationship between treatment and survival, and was allowed to be affected by deprivation. Such a confounder is dealt with using an extension of the g-computation formula [24, 25]. In the

Fig. 1 Direct Acyclic Diagram (DAG) depicting the causal relationships between deprivation and survival status in breast cancer patients. Year of diagnosis and region are considered as baseline confounders, with potentially an arrow to each node in the diagram, and thus are not shown in this DAG



third analysis, we estimated the proportion of the effect of deprivation on treatment that is mediated by differential stage.

Because the deprivation gap in survival varies by time since diagnosis, the binary survival outcome (dead vs. alive) was stratified according to time since diagnosis: at 6 months, 1 year given (conditioning on) 6-month survival, 3 years given 1-year survival, and 5 years given 3-year survival. The analyses were performed separately on each of these four binary survival outcomes, in order to disentangle early from late mediating effects of stage and treatment on deprivation gap in survival.

We used multinomial regressions to model stage at diagnosis (four categories) and logistic regression for treatment and survival status. Age at diagnosis was modelled using restricted cubic splines [26].

Single stochastic imputation within the g-computation procedures was used to handle missing stage (8 %). All variables in the models (including vital status), exact length of follow-up times and detailed treatment categories were included in the imputation model.

Results

We analyzed 36,793 women diagnosed with breast cancer between 2000 and 2007 in Yorkshire and North East (Table 1). More deprived patients were diagnosed at an older age ($P = 0.001$) and a more advanced stage

($P < 0.001$) than the more affluent. The higher the deprivation level, the more advanced the stage.

Survival from breast cancer

Net survival differed between the most affluent and the most deprived patients by 3 % at 1 year (97 vs. 94 %), and 10 % at 5 years (86 vs. 76 %) after diagnosis (Fig. 2, Appendix 5).

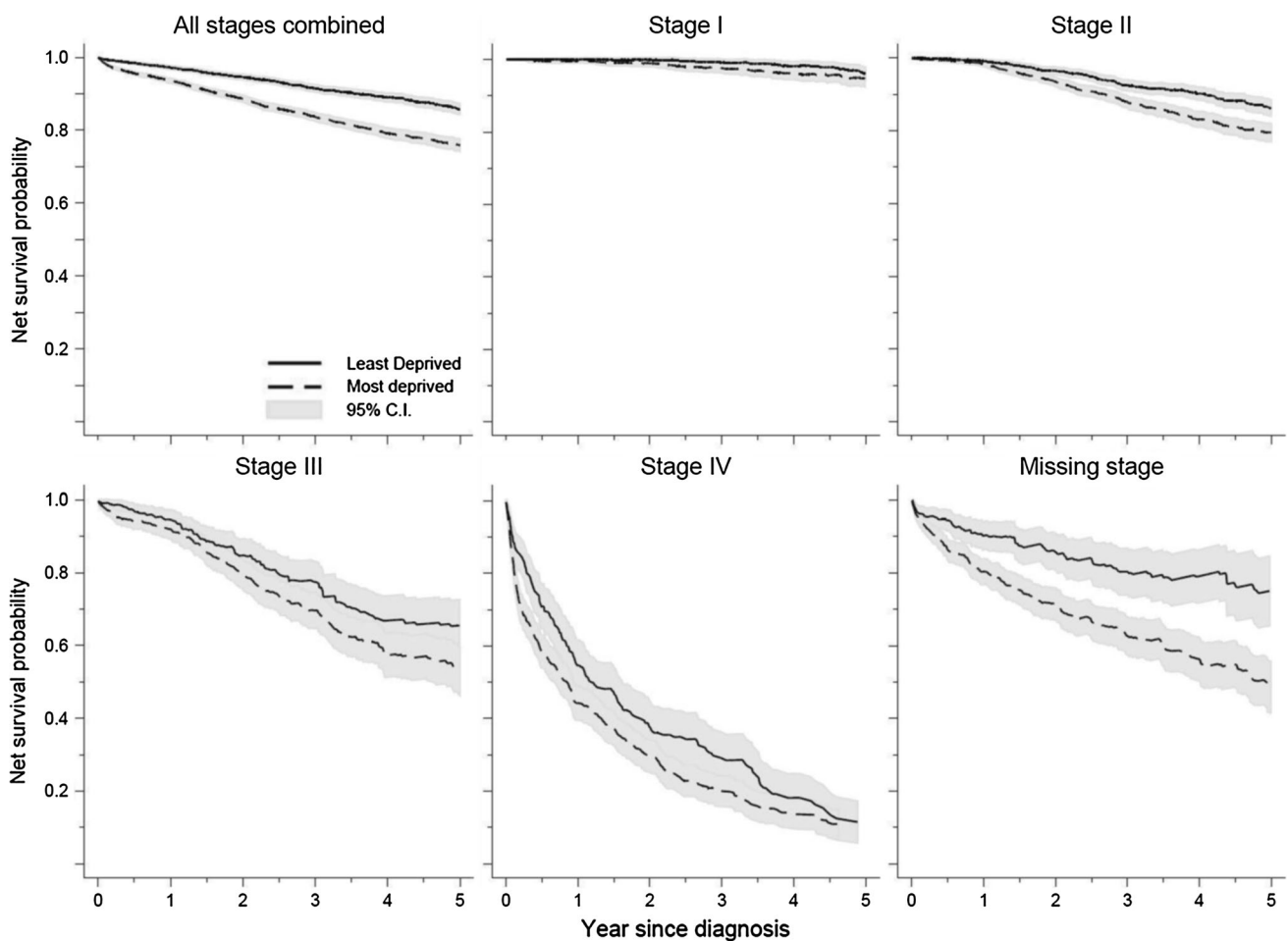
Stage-specific survival estimates were consistently lower in the more deprived patients. Large deprivation gap existed for the short-term survival (at 1 year after diagnosis) in the most advanced stage (IV), and in the long-term survival (at 5 years) in the less advanced stages (II–III). For patients with missing stage information, the more deprived experienced worse survival.

Total effect of deprivation on cancer survival status

We first estimated the *total causal effect* of deprivation on survival status, which is the sum of all effects shown in Fig. 1, adjusted for the confounding effect of region and year of diagnosis. The odds of dying within the first 6 months since diagnosis roughly increased linearly with increasing deprivation (odds ratio [OR] comparing most deprived to most affluent patients: 2.77, 95 % confidence interval [CI] 2.17, 3.53) (Fig. 3a, Appendix 6). This deprivation effect decreased slightly as follow-up time increased. However, the effect remained fairly strong at 5 years

Table 1 Characteristics of women diagnosed with breast cancer, Yorkshire and North East (England), 2000–2007

	All patients	Deprivation				
		Least deprived	2	3	4	Most deprived
Number of patients	36,793	6411	7549	6982	7642	8209
% Alive at end of follow-up	78.9	85.1	81.2	79.2	76.7	73.6
Mean age at diagnosis	62.9	61.0	61.3	62.4	63.1	63.7
(interquartile range)	51.9–73.9	51.2–70.3	51.8–72.7	52.1–74.3	52.1–75.7	52.2–75.2
Stage at diagnosis (%)						
I	37.6	40.9	38.8	38.0	36.3	34.8
II	43.0	43.7	43.9	41.7	42.5	43.1
III	6.8	6.2	6.2	7.0	7.0	7.4
IV	4.6	3.5	3.7	4.6	5.0	5.9
Missing	8.0	5.7	7.4	8.7	9.2	8.8
% Receiving major treatment	26.9	25.9	25.7	27.6	27.1	28.0

**Fig. 2** Net survival by deprivation and stage at diagnosis, women diagnosed with breast cancer, Yorkshire and North East (England), 2000–2007

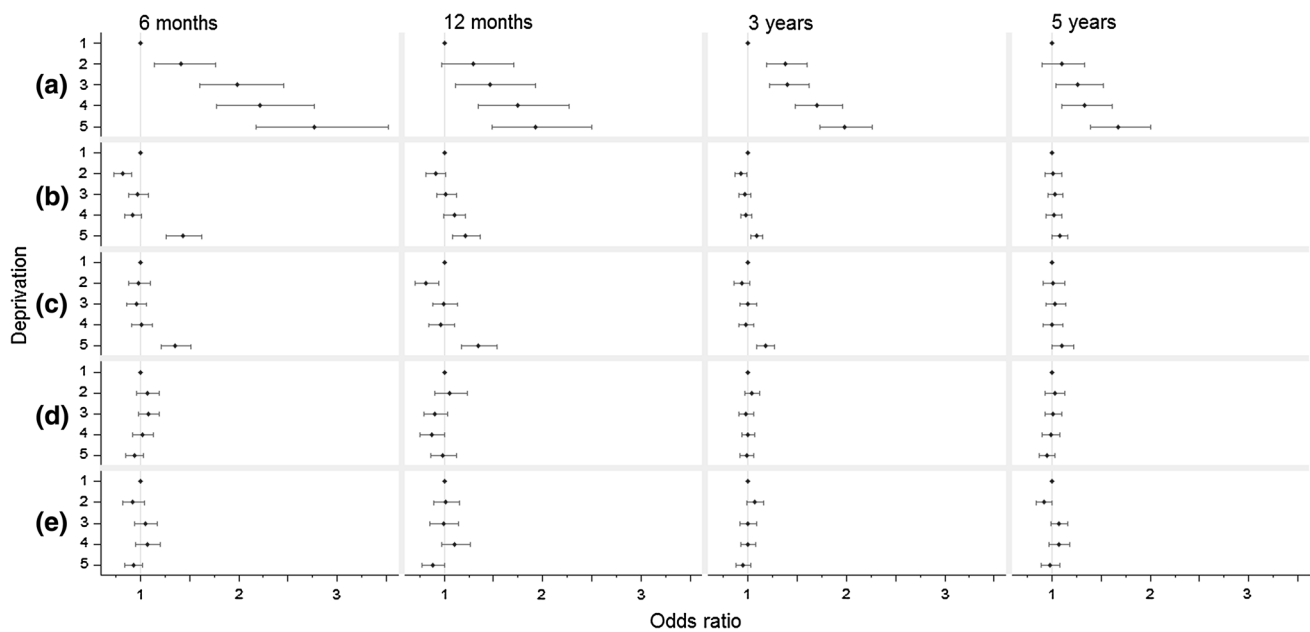


Fig. 3 Effect of deprivation on odds of deaths at different times since breast cancer diagnosis, results from the G-computation Analyses. (a) Shows the total effect of increasing deprivation on odds of death. (b) shows the effect that were mediated via stage at diagnosis, and (c) shows the effect mediated via surgical treatment. Odds of death at 12 months since diagnosis are conditional on surviving the first

6 months after cancer diagnosis. Similarly, odds of death at 3 years are conditional on survival the first 12 months; odds of death at 5 years are conditional on survival the first 3 years. We used conditional odds of death in order to disentangle early from late mediating effects of stage and treatment on deprivation gap in survival

conditioning on 3-year survival (for most deprived compared with least deprived, OR: 1.67, 95 % CI 1.39, 2.00).

Role of stage on the socio-economic differences in cancer survival status

The effect of socioeconomic status on survival mediated through stage (the *RIANIE*, Fig. 3b) was only apparent when comparing the most deprived with the least deprived. This indirect effect through stage decreased as time since diagnosis increased (OR for 6-month mortality: 1.43, CI: 1.27, 1.62; OR for 5-year conditional mortality: 1.08, CI 1.00, 1.16) (Fig. 3b). On the log odds scale, stage only accounted for about one-third of the total effect of deprivation at 6-month and 1 year (proportion mediated [PM]: 35 %, CI 23, 48 %; 30 %, CI 5, 54 %, respectively), a proportion which decreased to just over a tenth at 3 and 5 years since diagnosis (PM: 12 %, CI 4, 21 %; 14 %, CI –3, 31 %, respectively) (Appendix 6).

We also treated both age and stage as mediators (in place of just stage). We assumed here linear associations between the logarithm of age and treatment or mortality. The overall pattern hardly changed although adding age tended to slightly increase the long-term PM. This might reflect the long-term effect of age on all-cause mortality (Appendix 6).

Role of treatment on the socio-economic differences in cancer survival status

The higher the stage, the less likely a patient would receive major surgical treatment (Fig. 4). For patients under 70 years when diagnosed with early stages (stages I and II), more deprived patients received more treatment. By contrast, for patients aged 70 and over, more deprived patients received less treatment for all stages.

We did not find any evidence for total effect of deprivation and effect mediated by stage on treatment (Appendix 8). Although treatment patterns differ between different socioeconomic classes, the main mediation analysis found no evidence for the effect of deprivation on mortality mediated through differential treatment (Fig. 3c, Appendix 7).

Discussion

Our results suggest that, for the most deprived patients only, earlier breast cancer diagnosis would reduce the deficit in short-term survival by up to a third and in longer-term survival by up to a tenth. The available crude information on treatment seems to show that differential surgical treatment between deprivation groups played a minor



Fig. 4 Probability of receiving major surgery for women diagnosed with breast cancer, Yorkshire and North East (England), 2000–2007

role in socioeconomic inequalities in breast cancer survival.

For the younger (15–69) patients diagnosed at stage I or II, the finding that more deprived patients received more treatment contradicts the *a priori* hypothesis by some oncologists: more deprived patients may have more comorbidity, and thus less aggressive diagnostic investigation and treatment. Prevalence of both obesity and tobacco smoking widely varies in the general population between deprivation groups [27, 28], but we did not have reliable information about comorbidity of the cancer patients. However, the surgical differences observed between the socioeconomic groups may reflect that, within a given stage, more deprived patients were diagnosed with more advanced disease. To investigate this hypothesis, we will need more detailed information on tumor stage and diagnostic investigation. In addition, more affluent patients may have received treatment within private facilities, information not available to us.

In the absence of individual measure of socioeconomic status for population-based studies in England, we used an ecological measure of deprivation [14]. Because LSOAs (the geographical level of the deprivation measure) are relatively small (1500 inhabitants on average) and have been made as socially homogenous as possible, the ecological bias is probably small. An ecological measure reflects both the individual and contextual dimensions of deprivation. We are not able to disentangle individual and contextual dimensions of deprivation and this affects conceptualizing hypothesized interventions. The English healthcare system is strongly territorialised, and any perceived intervention should primarily target these territories in which individual-level actions (via the general practices) are also possible. Such interventions correspond to our conceptual framework, i.e. we asked: what would be the outcome of women in the deprived group, had they lived in the same area as those lived in the most affluent areas, with similar background risk factors and access to primary and secondary healthcare for their cancer diagnosis and treatment.

We identified three main plausible reasons that could bias our results: misclassification of the stage at diagnosis, misclassification of the treatment and unmeasured confounders between the mediator(s) and the outcome(s).

Misclassification of stage at diagnosis

More deprived cancer patients may more likely be managed by non-specialized centres and low-workload surgeons [29]. Evaluating the spread of their cancers (i.e. staging) may not be thorough enough (Fig. 1) and, as a result, they might be more often under-staged and receive non-optimal treatment [8]. We tested this hypothesis by assuming different proportions of the most deprived patients were under-staged. We randomly up-staged 10, 30 and 50 % of the most deprived patients by one level (stage I to II, etc.) ten times and reran the analyses to estimate the PM distributions. The proportion of survival inequalities mediated by stage hardly changed with 10 % of under-staged most deprived patients, but increased substantially with 30 and 50 % of under-staged, more particularly for conditional survival at 1 year and over (data on request). For example, more than half of the lower conditional 1-year survival among the most deprived patients would be mediated by stage if above 30 % of them were under-staged (vs. 30 % mediated if stage was not misclassified). Changing our main conclusion about the role of stage on survival inequalities would require above 30 % of the most deprived patients were systematically under-staged, compared to none in the most affluent group, a rather extreme assumption that is not supported by the literature.

Misclassification of treatment

Surgery, often in conjunction with other treatments, remains the main curative treatment of breast cancer. Information on radiotherapy and chemo/hormono-therapy was too poor to be used here. The quality, completeness and intention (whether curative or not) of the surgical procedure were not known. It was reported that 3.6 % of surgical treatment for breast cancer were made in private hospital in NYCRIIS [30]. Such under-estimation of the surgery proportion is likely to affect primarily the more affluent patients. We conducted a sensitivity analysis to investigate how such misclassification would influence the mediating effect of treatment on the socioeconomic differences in breast cancer survival. We randomly changed the treatment status for 3.6 % of the patients from no/minor surgical procedures to major surgery, according to the stage and age distribution of those who had records of receiving major treatment. We assumed that those patients were entirely from the most affluent group. We generated 100 new datasets on which we carried out g-computation analyses, estimating the proportion of effect of deprivation on survival mediated by treatment. We confirmed the absence of indirect effect through differential treatment on cancer survival status for deprivation groups 2–4. However, treatment did mediate around 30–40 % of the differential mortality between the most deprived patients and the most affluent, regardless time since diagnosis (Fig. 5), under the assumption that only the most affluent patients had surgeries in private hospitals. In addition, around 10 % of the cancer registry cases could not be matched with HES

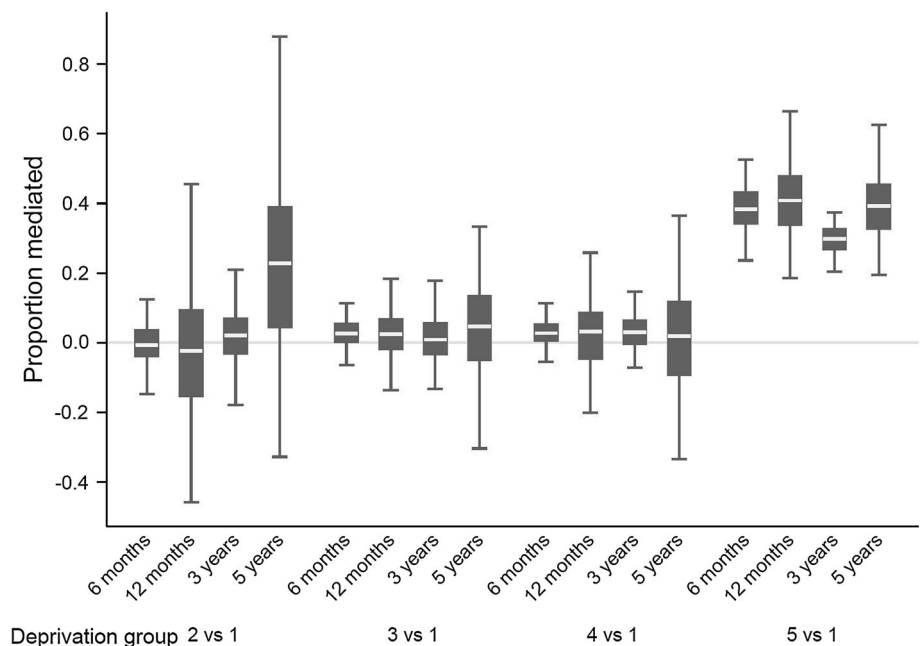
(inpatient data from the National Health Services) [30]. Surgical information is likely to be missing completely at random for such patients and we do not expect this to bias our results.

Our analyses crudely dichotomized treatment into ‘major’ and ‘minor or no’ surgery categories. When ‘major’ surgery was further categorized into ‘breast conserving surgery’ and ‘mastectomy’, the results remained unchanged (Appendix 7). Nevertheless, the crude treatment information may explain why the mediating effect of treatment on deprivation gap in survival remained moderate and was not affected in the sensitivity analysis on the misclassification of tumor stage.

Unmeasured confounders

We assumed that, conditional on deprivation, age, stage and year at diagnosis, and government regions, there were no further (unmeasured) common causes of treatment and survival status. However, in addition to staging thoroughness, comorbidity could be an important confounder for treatment and mortality, which we did not account for due to lack of reliable individual information. Ignoring the confounding effect of comorbidity would potentially lead to over-estimation of the beneficial effect of major surgery on mortality: patients with high levels of comorbidity experience high mortality, and may have lower rate of major surgery. Since we found little evidence for treatment to mediate the effect between deprivation and mortality in the original dataset, inclusion of comorbidity would not change this overall interpretation, but only if stage and

Fig. 5 Proportion of the effect of deprivation on odds of death mediated by treatment in a sensitivity analysis taking into account of the potentially missing surgical information due to treatment in private hospitals



treatment were not misclassified. If reliable information on comorbidity becomes available, we could potentially treat it as a mediator between deprivation and mortality, and estimate how much contribution it has to the deprivation gap in survival.

Concluding remarks

Our results are based on population-based data, i.e. on virtually all patients diagnosed with a breast cancer in a given region, including those who were diagnosed with advanced stage and those who were not optimally managed. Since our main focus is to better understand the causal relationships between deprivation and breast cancer survival, and to divide it into path-specific components, applying methods from the growing literature on causal mediation is highly appropriate.

To our knowledge, very few studies attempted to disentangle the effects of deprivation on breast cancer survival. Two studies used data from an earlier periods (late 1990s) of the same region as our study [31, 32]. A complete-case analysis found adverse stage distribution and less surgical treatment (even after adjustment for stage) among more deprived patients [31]. No stage-specific results were provided on treatment. Lower overall 5-year survival was associated with deprivation after adjustment for age and stage, but underlying pathways could not be deduced from these results. A second analysis using latent class modelling [32] clearly identified two groups of patients according to their prognosis: more advanced stage seemed to play a role in the deprivation gap in 5-year survival only in one group. The conclusions were weakened by the fact that overall survival was analysed, while mortality from causes other than breast cancer varies greatly by deprivation within 5 years since diagnosis. Our study is also based on overall mortality. Not adjusting for competing risks of death will dilute the mediating effect of stage. However, this effect would be minimal for short-term survival, as mortality from causes not related to breast cancer does not play a significant role in short-term survival status, especially at 6 months after diagnosis. Using conditional survival also reduced this bias.

Contrasting with our results, a study in another English region found that, in 2006–2010, adverse stage distribution explained half of the deficit in 5-year breast cancer relative survival observed among the most deprived patients, but all of it in other deprivation groups [33]. However, stage-standardisation, used in order to eliminate differences in stage distribution by deprivation, cannot fully identify the effect of deprivation mediated by tumor stage on such observational data.

Applying another causal inference approach, Valeri et al. [34] found that the contribution of stage to the disparities in survival from colorectal cancer between Blacks and Whites in the US was similar to our results for the socio-economic disparities in breast cancer survival in England. They however concluded that the mediation effect of stage represented a “substantial reduction” while we talked about a small reduction, which reflects differences in the study context. Contrasting with the US (at least until recently), the healthcare system in England is universal with free access to diagnosis and treatment. In theory, most disparities in cancer survival should be therefore due to patient and tumour factors, more specifically tumour stage at diagnosis and comorbidity, and not to healthcare system factors. Contrasting this belief, our results add to the growing evidence that one of the strongest prognostic factors, stage, plays a relatively small role in the socio-economic inequalities in cancer survival. Comorbidity (or health performance status) is likely to contribute to inequalities, but will reduce the stage contribution estimated further. It means that, in the context of a supposedly equitable healthcare system, a large proportion of these inequalities remain unexplained; inequalities within the healthcare system are likely to play a key role.

Despite data limitations, we were able to estimate the proportions of the deprivation gap in cancer survival mediated via tumor stage and treatment separately. It informs us about their respective roles, and ultimately, what may be done to most effectively reduce the deprivation gap in cancer survival. In particular, effort for earlier diagnosis would reduce the cancer survival inequalities only by a third. Our conclusions may, however, be altered by unmeasured confounders such as comorbidity, staging thoroughness and detailed treatment information, of which quality and completeness are improving dramatically in the population-based cancer registry data in England. The changes in results after sensitivity analyses demonstrate the vital importance of using reliable and correctly classified surgical treatment data in similar studies.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Rachet B, Woods LM, Mity E, et al. Cancer survival in England and Wales at the end of the 20th century. *Br J Cancer*. 2008; 99(Suppl. 1):2–10.

2. Lyratzopoulos G, Barbiere JM, Rachet B, et al. Changes over time in socioeconomic inequalities in breast and rectal cancer survival in England and Wales over a 32-year period (1973–2004): the potential role of health care. *Ann Oncol*. 2011;22:1661–6.
3. Rachet B, Ellis L, Maringe C, et al. Socioeconomic inequalities in cancer survival in England after the NHS Cancer Plan. *Br J Cancer*. 2010;103:446–53.
4. Ellis L, Coleman MP, Rachet B. How many deaths would be avoidable if socioeconomic inequalities in cancer survival in England were eliminated? A national population-based study, 1996–2006. *Eur J Cancer*. 2012;48:270–8.
5. Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer survival: a review. *Ann Oncol*. 2006;17:5–19.
6. Leary A, Smith IE. Survival from breast cancer in England and Wales up to 2001. *Br J Cancer*. 2008;99(Suppl 1):S56–8.
7. Woods LM, Rachet B, O'Connell DL, et al. Large differences in patterns of breast cancer survival between Australia and England: a comparative study using cancer registry data. *Int J Cancer*. 2009;124:2391–9.
8. Lejeune C, Sassi F, Ellis L, et al. Socioeconomic disparities in access to treatment and their impact on colorectal cancer survival. *Int J Epidemiol*. 2010;39:710–7.
9. Cole SR, Hernán MA. Fallibility in estimating direct effects. *Int J Epidemiol*. 2002;31:163–5.
10. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992;3:143–55.
11. Pearl J. Direct and indirect effects. In: 17th Conference on Uncertainty in Artificial Intelligence. San Francisco, CA: Morgan Kaufmann; 2001. pp. 411–420.
12. Vanderweele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. *Am J Epidemiol*. 2010;172:1339–48.
13. Office for National Statistics. Cancer statistics: registrations of cancer diagnosed in 2007, England. In: Series MB1 No. 38. Newport: Office for National Statistics; 2010. pp. 1–80.
14. Department of the Environment Transport and the Regions. Measuring multiple deprivation at the small area level: the indices of deprivation 2000. London: DETR; 2000.
15. Walters S, Maringe C, Butler J, et al. Comparability of stage data in cancer registries in six countries: lessons from the International Cancer Benchmarking Partnership. *Int J Cancer*. 2013;132:676–85.
16. Health & Social Care Information Centre. OPCS-4 Classification. 2014. <http://systems.hscic.gov.uk/data/clinicalcoding/codingstandards/opcs4>. Accessed 1 Sept 2014.
17. National Cancer Intelligence Network. Site Specific Clinical Reference Groups (SSCRG) for breast cancer. 2013. http://www.ncin.org.uk/cancer_type_and_topic_specific_work/cancer_type_specific_work/breast_cancer/. Accessed 1 Sept 2014.
18. Perme MP, Stare J, Estève J. On estimation in relative survival. *Biometrics*. 2012;68:113–20.
19. StataCorp. STATA statistical software. 13th ed. College Station, TX: Stata Corporation; 2013.
20. Grzebyk M, Urmès I, Hédélin G. Net survival estimation with stns. *Stata J*. 2014;14:87–102.
21. VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology*. 2014;25:300–6.
22. Didelez V, Dawid P, Geneletti S. Direct and indirect effects of sequential treatments. In: Dechter R, Richardson TS, editors. Proceedings of the 22nd Annual Conference on Uncertainty in Artificial Intelligence. Arlington, VA: AUAI Press; 2006. pp. 138–146.
23. Geneletti S. Identifying direct and indirect effects in a non-counterfactual framework. *J R Stat Soc Ser B Stat Methodol*. 2007;69:199–215.
24. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Model*. 1986;7:1393–512.
25. Daniel RM, Stavola BLD, Cousens SN. gformula: estimating causal effects in the presence of time-varying confounding or mediation using the g-computation formula. *Stata J*. 2011;11:479–517.
26. Royston P, Sauerbrei W. Multivariable modeling with cubic regression splines: a principled approach. *Stata J*. 2007;7:45–70.
27. Sutton R. Adult anthropometric measures, overweight and obesity. In: Craig R, Mindell J, editors. Health Survey for England 2011. Health and Social Care Information Centre; 2012. p. 37.
28. Office for National Statistics. Smoking and drinking among adults, 2008. In: General Lifestyle Survey 2008. Office for National Statistics; 2010. p. 74.
29. Morris EJA, Quirke P, Thomas JD, et al. Unacceptable variation in abdominoperineal excision rates for rectal cancer: Time to intervene? *Gut*. 2008;57:1690–7.
30. Lawrence G. Further analysis of ICBP treatment data (version 1.2). 2013. p. 2 (unpublished report).
31. Downing A, Prakash K, Gilthorpe MS, et al. Socioeconomic background in relation to stage at diagnosis, treatment and survival in women with breast cancer. *Br J Cancer*. 2007;96:836–40.
32. Downing A, Harrison WJ, West RM, et al. Latent class modelling of the association between socioeconomic background and breast cancer survival status at 5 years incorporating stage of disease. *J Epidemiol Community Health*. 2010;64:772–6.
33. Rutherford MJ, Hinchliffe SR, Abel GA, et al. How much of the deprivation gap in cancer survival can be explained by variation in stage at diagnosis: an example from breast cancer in the East of England. *Int J Cancer*. 2013;133:2192–200.
34. Valeri L, Chen JT, Garcia-Albeniz X, et al. The role of stage at diagnosis in colorectal cancer black–white survival disparities: a counterfactual causal inference approach. *Cancer Epidemiol Biomarkers Prev*. 2016;25:83–9.