

Contribution of primary care organisation and specialist care provider to variation in GP referrals for suspected cancer: ecological analysis of national data.

Christopher Burton<sup>1</sup>,

Luke O'Neill<sup>1</sup>,

Phillip Oliver<sup>1</sup>,

Peter Murchie<sup>2</sup>

Author affiliations

1 Academic Unit of Primary Medical Care, University of Sheffield, Sheffield, UK

2 Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK

Correspondence to

Professor Christopher Burton,

Academic Unit of Primary Medical Care,

University of Sheffield,

Sheffield S5 7AU,

UK

[chris.burton@sheffield.ac.uk](mailto:chris.burton@sheffield.ac.uk)

## **Abstract**

**Objectives:** To examine how much of the variation between general practices in referral rates and cancer detection rates is attributable to local health services rather than the practices or their populations.

**Design:** Ecological analysis of national data on fast-track referrals for suspected cancer from general practices. Data was analysed at the levels of general practice, primary care organisation (Clinical Commissioning Group) and secondary care provider (Acute Hospital Trust) level. Analysis of variation in detection rate was by multilevel linear and poisson regression.

**Setting:** 6379 group practices with data relating to more than 50 cancer cases diagnosed over the five years from 2013-2017.

**Outcomes:** Proportion of observed variation attributable to primary and secondary care organisations in standardised fast-track referral rate and in cancer detection rate before and after adjustment for practice characteristics.

**Results:** Primary care organisation accounted for 21% of the variation between general practices in the standardised fast-track referral rate and 42% of the unadjusted variation in cancer detection rate. After adjusting for standardised fast-track referral rate, primary care organisation accounted for 31% of the variation in cancer detection rate (compared to 18% accounted for by practice characteristics). In areas where a hospital trust was the main provider for multiple primary care organisations, hospital trusts accounted for the majority of the variation attributable to local health services (between 63% and 69%).

**Conclusion:** This is the first large-scale finding that a substantial proportion of the variation between GP practices in referrals is attributable to their local healthcare systems. Efforts to reduce variation need to focus not just on individual practices but on local diagnostic service provision and culture at the interface of primary and secondary care.

**Keywords:** Cancer, General Practice, Primary Care, Referral, Variation, Multi-level modelling

## Introduction

Ensuring that patients with symptoms of possible cancer have access to prompt diagnostic tests and treatment is a key concern for healthcare systems <sup>1,2</sup>. Variation between providers in referral or access to such services is thus an important performance issue <sup>3-5</sup>. In the UK and Europe, particular attention has been drawn to variation in referral from primary to secondary care for patients with clinical features possibly indicative of cancer <sup>4,6-8</sup>. This observed variation, particularly the proportion of patients with cancer who have been referred through a fast-track pathway, has been portrayed as unacceptably large <sup>3,9</sup>.

The route from a patient experiencing symptoms to a specialist cancer diagnosis involves several steps before, during, and after primary care <sup>1</sup>. Variation in referral may arise at these different stages and may be either systematic <sup>10</sup> or random <sup>11</sup>. It may arise at patient or population level; within general practices; and at the level of larger organisations such as primary care organisations and secondary care providers of diagnostic services for cancer <sup>12</sup>. Potential sources of systematic variation include patient and population demographics and environmental exposures <sup>13</sup>, accessibility to and quality of primary care <sup>3</sup>, and variation in the capacity and local policies of secondary care services <sup>4</sup>. Common reasons for apparent random variation include the heterogeneity of cancer presentations <sup>14</sup> and small numbers of cases at practice level <sup>11,15</sup>. These are summarised in Table 1.

We have previously demonstrated that the variation in GP urgent referrals for suspected cancer can be considered from the perspective of referral decision accuracy in terms of sensitivity and specificity <sup>16</sup>. From this perspective, sensitivity (referred to in UK national statistics as 'cancer detection rate') equates to the proportion of cancers which are diagnosed via the fast-track pathway and specificity is related to the number of patients referred by the fast-track pathway who did not have cancer. We previously found that there was more variation between practices in

specificity than in sensitivity for a given specificity, and interpret this as indicating that practices vary more in referral threshold (referring more or less patients), than in referral decision accuracy (referring the “right” patients at a given referral threshold).<sup>16</sup>

In this study, we aimed to examine how much variation in the number of fast-track suspected cancer referrals which GP practices made, and in the cancer detection rate from these referrals, could be accounted for by local health services above the level of the general practice (primary care organisations and diagnostic service providers).

## Methods

We conducted an analysis of publicly available data. The data contained no individual patient information and no ethical permissions were needed.

### **Data sources**

We obtained data from the Public Health England Fingertips database<sup>17</sup> which holds extensive data, at GP practice level, about fast-track referrals for suspected cancer. These are widely referred to as two-week wait (2WW) referrals. We extracted and aggregated data for five consecutive years between 2012 and 2017. Data were available for general practices (groups of between 2 and 30 GPs) rather than individual physicians.

### **Selection of variables for analysis**

We extracted the total number of 2WW referrals; total number of cancer diagnoses; and the number of cancers diagnosed after 2WW referral (as opposed to other pathways). For each practice we also extracted the age-standardised 2WW referral rate; the proportion of patients aged over 65; and the practice’s index of multiple deprivation (a measure of socio-economic status of the registered patients).

For this study we considered two approaches to account for different referral thresholds. The first was to use the standardised 2WW referral rate. This represents a value for each practice, based on the crude 2WW referral rate adjusted for practice deprivation and proportion of patients aged over 65 and centred on a value of 100. This standardised rate is approximately normally distributed. The second was to use an estimated specificity for each practice using the definition of specificity from the contingency table of a diagnostic accuracy study as true negatives divided by (true negatives + false positives) . As in this instance true negatives represent patients without cancer but whose clinical presentation may have indicated cancer, we could not measure this directly. Instead we estimated a total population at risk of cancer in the practice by dividing the number of cancer cases by 3% (as this represents the probability of a cancer diagnosis with symptoms which are sufficient to trigger a 2WW referral). This was calculated using our previously published method<sup>16</sup>. Standardised 2WW referral rate and specificity were strongly correlated (Spearman's rho = 0.74). For the primary analysis of detection rate we used the standardised 2WW referral rate because (a) it is simpler to understand (b) it is provided in the data rather than estimated so makes for easier replication (c) it is already adjusted for practice population demographics of age and deprivation.

### **Inclusion / exclusion of GP practices' data**

We have previously demonstrated that variation in detection rate is particularly susceptible to error when there are less than 50 cancer cases per GP practice<sup>15</sup>. We thus set a minimum criterion for eligibility for the primary analysis of at least 50 cases of cancer diagnosed over the five-year period. We tested this assumption in a sensitivity analysis using different cut-offs for minimum number of cancers. We excluded a small number of practices which served predominantly student populations (identified by a very low proportion of patients aged over 65) as we found that because of the very low incidence of cancer and two week wait referrals the

algorithm for calculating specificity returned impossible values ( $>1$ ). We also excluded practices from one CCG because data was only available for less than one quarter of practices in this particular CCG.

### **Allocation of practices to primary care organisations and diagnostic services**

The downloaded data linked each practice to a Clinical Commissioning Group (CCG). CCGs are primary care organisations responsible for commissioning both the delivery of primary care services in a geographical area and the commissioning of specialist services for practices in their area. CCGs varied in size between 5 and 108 practices and a population of between 74,000 and 894,000 patients.

We also linked CCGs to Acute Hospital Trusts (AHTs) which act as referral and diagnostic service providers for CCGs. AHTs comprise one or more hospitals within a geographical area with populations which are typically similar to or larger than those of CCGs. There are fewer AHTs than CCGs in England (127 and 207 respectively). We mapped CCGs to AHTs, recognising that CCGs commonly commission services from more than one AHT. Typically, this is because some practices at the CCG boundary are closer to another AHT. In mapping CCG to AHT we used the criteria for mapping listed in Supplementary Table 1. Where the mapping was either single CCG to single AHT or multiple CCG to single AHT we allocated each CCG to the relevant AHT. Where there were multiple possible AHT:CCG mappings, we used information on AHT and CCG websites to identify AHTs which held the largest reported contract from a CCG or where there was other evidence including a particular relationship.

### **Analysis of variation between practices and CCGs**

We examined the variation between practices graphically and by creating multilevel regression models. In both approaches we first examined all included practices in a

two-level model (practices in CCGs) and then examined practices where there were multiple CCGs per AHT in a three-level model (practices in CCG in AHT).

For the plots of standardised 2WW referral rate and detection rate we ordered CCGs (or AHTs) by median value of the relevant variable. For the 2-level analysis we created box plots for each CCG which were ordered by median value. For the 3-level analysis we plotted a single point for each CCG representing the median value of the variable; points were grouped by AHT.

Statistical analysis used multilevel linear regression for standardised 2WW referral rate and multilevel poisson regression for detection rate. For standardised 2WW referral rate, we used a simple model in which the only predictor was the clustering variable. For the primary analysis of detection rate, we used both the simple model and a model including standardised 2WW referral rate as an independent variable. As initial plotting of data suggested a non-linear relationship between detection rate and standardised 2WW referral rate, we also included a quadratic term. For the secondary analysis of detection rate we used specificity, and practice demographics as independent variables. As the distributions of both referral rate and detection appeared unimodal and approximately symmetrical, we compared the appropriateness of linear and non-linear (poisson and logistic) regression models by inspecting standardised residuals. For the final analysis we selected the regression models with residuals best approximating to a normal distribution (linear regression for referral rate and poisson for detection rate).

We analysed clustering at CCG level for all practices and at CCG and AHT for practices where more than one CCG was linked to the same AHT. We conducted two sensitivity analyses: first we repeated the 3-level analysis after excluding CCGs which mapped to multiple AHTs (i.e. those in central London and Newcastle) and

second we examined the effect of different inclusion thresholds of numbers of cases per practice over the 5 years.

For each model we reported two measures of the amount of variance attributable to the clustering variable:  $R^2$  and the intra-class correlation coefficient (ICC).  $R^2$  is widely used as an indicator of the proportion of total variance which can be explained by a statistical model. In the context of a multilevel model, it can be used to partition variance into that explained by fixed effects (of covariates across all units) and random effects (arising from the clusters). We used the formulae for  $R^2$  in linear mixed effects models by Nakagawa<sup>18</sup> modified by Johnson<sup>19</sup>. ICC is broadly similar in purpose, in that it also describes the proportion of variance which can be attributed to clustering in the data. It has the advantage that in a 3-level model it can be split between the different levels. Thus, ICC and  $R^2$  provide similar but complementary information.

Analysis was conducted in R3.6 with the lme4 package for linear mixed effects models using Restricted Maximal Likelihood (REML) fitting. We obtained the ICC using the sjstats package and  $R^2$  using the MuMIn packages within R. We used the Aikake Information Criterion (AIC) in order to compare models.

## Results

The downloaded data included information from 7219 practices with a combined total population in 2015 of 54.7 million patients. We excluded data from 830 practices with less than 50 cancer diagnoses, 4 practices in which the proportion of patients aged 65 or over was <2% and 6 practices from one CCG in which more than two-thirds of practices had missing data. This left 6379 practices (88.4%) with a combined population in 2015 of 51.6 million patients (94.3%). The distribution of small (excluded) practices was uneven across CCGs: in one case this left only 5 practices remaining in the CCG cluster for analysis. Of the 206 CCGs and 127 acute hospital trusts, 71 CCGs and trusts mapped one to one, leaving 135 CCGs and 56 trusts with

more complex mappings. These complex-mapped CCGs included 3558 of the included practices (55.8%) with a combined population in 2015 of 29.0 million patients (56.2%).

The characteristics of the total population, the sample analysed for CCG clustering, and the sample with clustering by CCG and acute hospital trusts are shown in Table 2. The distributions of standardised 2WW referral rate, detection rate and specificity unimodal and only mildly skewed. For practices with a minimum of 50 cancers, the minimum and maximum values for detection rate were 0.12 and 0.73 respectively. Only 2/6379 practices had values for detection rate of less than 0.2 (both of these had between 50 and 100 cancers). Mean detection rate was 0.48 (standard deviation 0.073). For the practices with a minimum of 50 cancers the mean standardised two-week referral rate which was 99.3 (originally centred at 100) with a standard deviation of 28.9.

Figure 1 shows the distribution within CCGs of practice values for (a) Standardised 2WW Referral Rate and (b) Detection Rate. Box plots each represent one CCG and they are ordered by median value; outlying points have been excluded for clarity.

Table 3 contains the results of the 2-level regression models of practices in CCGs. The columns relating to R-squared include the total  $R^2$ , the amount attributable to fixed effects (practice level characteristics) and the amount attributable to random effects (clustering within CCGs). Adding covariates to the primary model for detection rate increased the overall  $R^2$ , while reducing the random effects  $R^2$  and ICC, however these remained very substantial at 0.31 and 0.37 respectively. This indicates that around one third of the variation between practices in detection rate (after adjusting for standardised 2WW referral rate) is attributable to clustering within their CCG.

Figure 2 shows the relationship between CCG level results and AHTs in the subset of data from practices linked to AHTs which have multiple CCGs. The plots show the median values for each CCGs of (a) practice Standardised 2 Week Wait Referral

Rate and (b) practice Detection Rate. In these plots, AHTs are ordered by median value.

Table 4 contains the results of the 3-level regression models of practices in CCGs in AHTs. The columns relating to subsets of  $R^2$  shown in table 3 have been omitted for clarity and the columns for ICC have been expanded to include the total ICC, its partition between CCG and AHT and the proportion of total ICC attributable to AHT. The total R-squared and ICC values are slightly lower than in table 3, reflecting that table 4 comes from only a subset of the data in table 3. However, they remain substantial, with total ICC between 0.34 and 0.39 in the primary analysis. The right-hand columns of table 4 confirm the visual impression from figure 2 that AHTs account for a substantial proportion of the observed clustering effects. For the primary analysis model, AHTs account for between 66% and 69% of the total ICC.

The sensitivity analysis which excluded CCGs mapped to multiple AHTs showed similar results to the main analysis (Supplementary Table 2). Supplementary table 3 shows the effect of setting different thresholds for the number of included practices on the clustering for both standardised two week referral rate and detection rate (with and without adjustment). As the threshold increases, the ICC and the  $R^2$  attributable to the CCG increases indicating that the effect of clustering by CCG increases as “noise” from small practices is removed.

## Discussion

### **Summary of main findings**

This study shows that variation between GP practices in cancer referral within the same national health system is substantially associated with their local commissioning and specialist diagnostic services. Diagnostic service providers

account for a greater proportion of this variation than primary care commissioning organisations.

### **Strengths and limitations**

This study extends published research investigating the accuracy of cancer referrals<sup>11 14-16</sup> in terms of detection rate adjusted for referral rate. We have not identified any previous studies which have used multilevel regression to examine clustering of referral practice for possible cancer. A large current national dataset was used with high standards of data quality. Our selection of a five-year data period, a threshold of 50 cancers and the use of process measures rather than clinical outcomes reflects our previous published work and that of others<sup>11 15 16</sup>.

There are a number of limitations to our analysis. First, the lowest level of analysis was GP practice: this meant that meaningful patient level data was absent (e.g. cancer type<sup>14</sup>) or included in summary form only (e.g. proportion of population aged 65 or over). Thus, we can only comment on the variation between practices, not on the factors which affect individual patients. While this is not ideal, this is the data on which practice performance is monitored and so it is appropriate to examine it. Second, we did not include other covariates relating to practice characteristics which have been reported recently and relate both to the population served<sup>7</sup> and the way the practice meets the population needs<sup>8</sup>. However, these had modest effects on the two outcome measures we included in our analysis and none of these seems particularly likely to explain the clustering which we observed in the data. Third, we had to aggregate data from several consecutive years rather than using a single set of multi-year data, this may have introduced some error due to incomplete coding<sup>20</sup>. Our analysis may also have obscured temporal trends, however as changes to guidelines were introduced nationally during this time there is no strong reason why these should influence local clustering effects<sup>2</sup>. Fourth, we were only able to examine aggregate performance of practices, not of individual practitioners. This was

neither possible from the data, nor feasible because of the small numbers of cases per practice. Variation has been shown to occur at both the level of the clinician and the practice<sup>21</sup>.

While mapping of practices to CCGs was robust, mapping of CCGs to AHTs is more challenging in some areas and may have introduced error. We used multiple steps to optimise these mappings (Supplementary table 1) and carried out a sensitivity analysis excluding CCGs where there was uncertainty, which showed little effect on the results. While we demonstrate local clustering effects at the level of AHT where there are multiple CCGs linked to one AHT, this ecological study cannot demonstrate the directions of causal influence (whether AHTs influence CCGs or vice versa) nor can it exclude additional clustering levels (such as practice neighbourhoods or networks).

### **Relationship to other research**

A recent UK study of variation in process and outcomes of care between specialists clustered within hospitals showed that doctor variation exceeded hospital variation<sup>22</sup>, but clustering effects were small (most ICCs between 0.005 and 0.05) in contrast to those seen in our study (0.21 to 0.37). Two studies from Italy found that post-discharge care for cardiac<sup>23</sup> and respiratory<sup>24</sup> disease showed greater variation due to hospital than primary care practitioner, but this may not be surprising as one would expect treatment to be dictated by the hospital and followed by the GPs. A further study examining primary care antibiotic prescribing demonstrated variation attributable to region<sup>25</sup>.

Analyses at practice level have indicated that lower fast-track referral rates were associated with poorer one-year cancer survival<sup>4</sup> and that lower endoscopy referral rates are associated with lower rates of curative surgery for upper gastrointestinal surgery<sup>26</sup>. Meanwhile analysis at regional level has recently demonstrated

substantial variation in aggressive treatment for lung cancer is associated with differences in survival <sup>27</sup>.

### **Meaning of the study**

Our data indicate a pressing need to extend the focus on variation beyond the GP consulting room and into the wider health services. While the majority of variation between practices remains unexplained, the extent of variation at the level of primary care organisation and secondary care provider suggests that GPs in different areas are effectively referring into different systems which are performing differently. Until this is better understood, there needs to be caution about publishing national GP benchmarking figures which may be highly dependent on wider services and circumstances.

### **Unanswered questions**

This higher level variation requires further exploration. From the current data, we are not able to establish how much the observed differences are due to capacity within diagnostic and specialist services, to ways in which referral criteria are applied locally, or to other cultural factors. The finding that clustering of detection rate is greater than clustering of referral rate suggests that capacity in secondary care is not the only factor. While adjusting for referral rate reduced the observed clustering of detection rate (from ICC of 0.42 to 0.37) this remained very substantial. Further research should consider combining analysis of routine data with in-depth approaches such as ethnography.

### **Conclusion**

Sources of primary care practice variation appear to extend beyond the practice boundaries. Efforts to reduce variation in suspected cancer referrals need to focus not just on individual practices but on local diagnostic and specialist service provision and culture.



### **Funding and sponsorship**

This study was conducted with no external funding.

### **Competing Interests**

All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

### **Data sharing statement**

The data used for this analysis are publicly available at <https://fingertips.phe.org.uk/profile/cancerservices>

Table 1: Systematic and random sources of variation in GP referrals for suspected cancer

	Source of variation		
Type of variation	Patient / practice population	GP practice	Health care organisation
<b>Systematic</b>	Age Socio-economic status Environmental exposures	Accessibility Quality of care	Capacity Implementation of guidelines/pathways Culture
<b>Random</b>	Heterogeneity of clinical presentation	Sampling (small number of cases)	Competing demands due to exceptional events

**Table 2 Characteristics of included data**

	All Practices <sup>1</sup>	Practices in 2-level analysis <sup>2</sup>	Practices in 3-level analysis <sup>3</sup>
Number of Practices	7219	6379	3435
Number of CCGs	206	206	133
Number of AHTs	126	126	54
Total patients	54660743	51561572	28142558
Index of Multiple Deprivation Median, (IQR)	21.6 (13.8 – 31.6)	20.3 (13.3 to 29.7)	20.8 (13.5 – 29.9)
Percentage of patients aged over 65: median, (IQR)	17.2 (12.3 – 21.1)	17.9 (13.8 – 21.7)	17.6 (13.4 to 21.3)
Two week wait referral detection rate: mean (SD)	0.474 (0.081)	0.480 (0.073)	0.483 (0.071)

1 Excluding those with very few patients aged over 65 and from the single CCG with substantial missing data

2 Excluding those with less than 50 cancer diagnoses over the five years

3 More than one CCG per AHT.

CCG: Clinical Commissioning Group; AHT: Acute Hospital Trust; IQR: Interquartile range; SD: standard deviation

**Table 3 Attributed variation within multilevel models of practices nested within CCGs**

Outcome variable	Covariates	R-squared			ICC	AIC
		Total	Fixed	Random		
Multilevel Linear Regression						
Standardised 2WWRR	None	0.209	0	0.209	0.209	58954
Multilevel Poisson Regression (primary analysis)						
Detection Rate	None	0.435	0	0.435	0.421	46077
Detection Rate	a	0.477	0.140	0.338	0.379	45276
Detection Rate	a+b	0.492	0.178	0.314	0.369	44848
Multilevel Poisson Regression (secondary analysis)						
Detection Rate	c	0.496	0.122	0.375	0.413	45437
Detection Rate	c+d	0.492	0.129	0.36	0.403	45277
Detection Rate	c+d+e+f	0.477	0.154	0.323	0.368	44972

Covariates: a = standardised 2WWRR; b= (standardised 2WWRR)<sup>2</sup>; c= specificity; d = (specificity)<sup>2</sup>;e = proportion patients aged 65+ and index of multiple deprivation.

CCG: Clinical Commissioning Group; ICC: Intraclass correlation coefficient; 2WWRR: two week wait referral rate AIC: Aikake Information Criterion

Table 4 Attributed variation within multilevel models of practices nested within CCGs within Acute Hospital Trusts

Outcome variable	Covariates	Total R2	Intraclass correlation coefficient				AIC
			Total	CCG	Acute	% acute	
Multilevel Linear Regression Standardised 2WWRR	None	0.212	0.209	0.081	0.131	63%	31993
Multilevel Poisson Regression (primary analysis)							
Detection Rate	None	0.405	0.391	0.135	0.256	66%	24780
Detection Rate	a	0.460	0.346	0.111	0.235	68%	24181
Detection Rate	a+b	0.474	0.335	0.104	0.231	69%	24082
Multilevel Poisson Regression (secondary analysis)							
Detection Rate	c	0.479	0.380	0.119	0.261	69%	24414
Detection Rate	c+d	0.473	0.369	0.115	0.253	69%	24327
Detection Rate	c+d+e+f	0.458	0.337	0.095	0.243	72%	24177

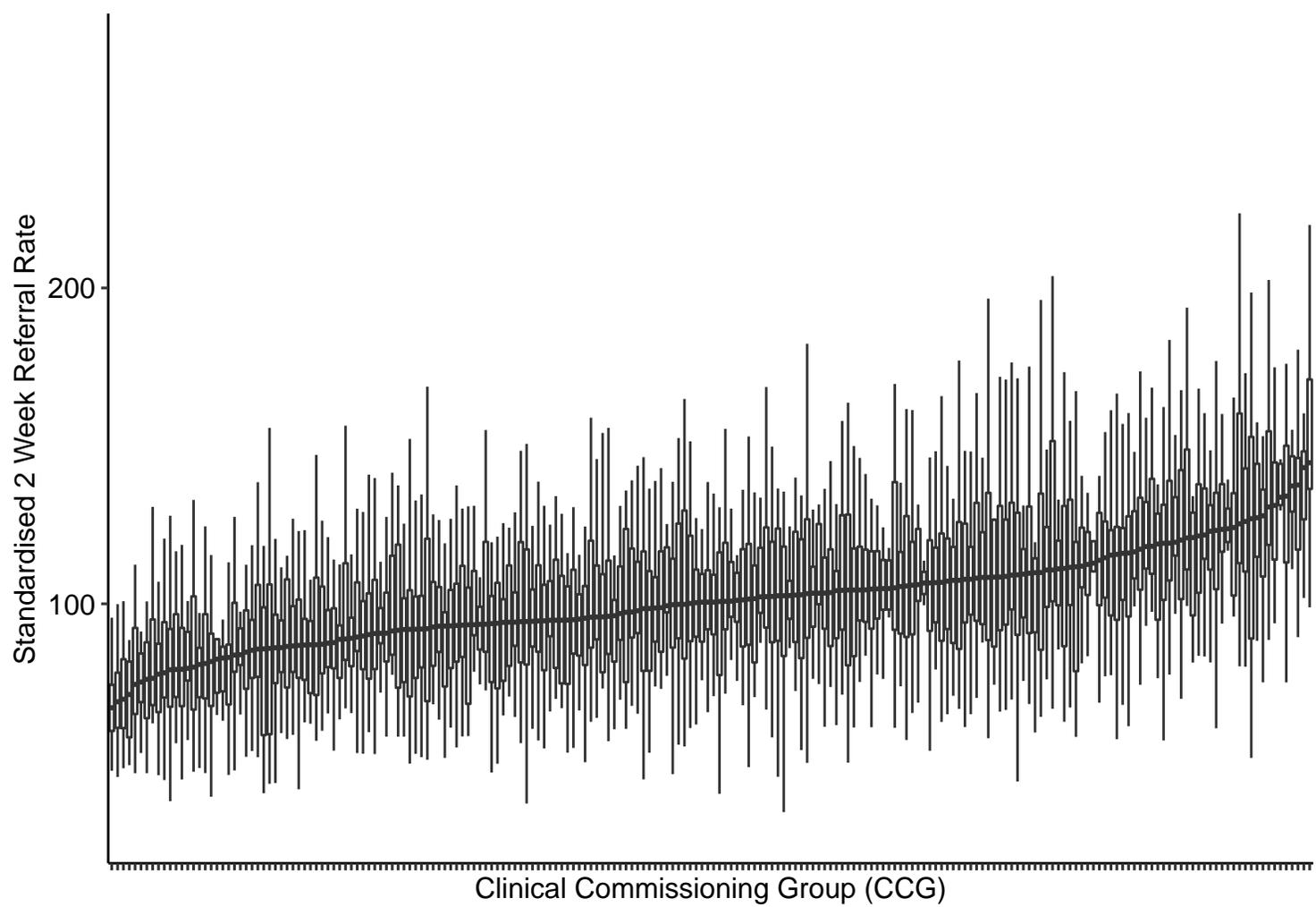
Covariates: a = standardised 2WWRR; b= (standardised 2WWRR)<sup>2</sup>; c= specificity; d = (specificity)<sup>2</sup>; e = proportion patients aged 65+ and index of multiple deprivation.

CCG: Clinical Commissioning Group; ICC: Intraclass correlation coefficient; 2WWRR: two week wait referral rate AIC: Aikake Information Criterion

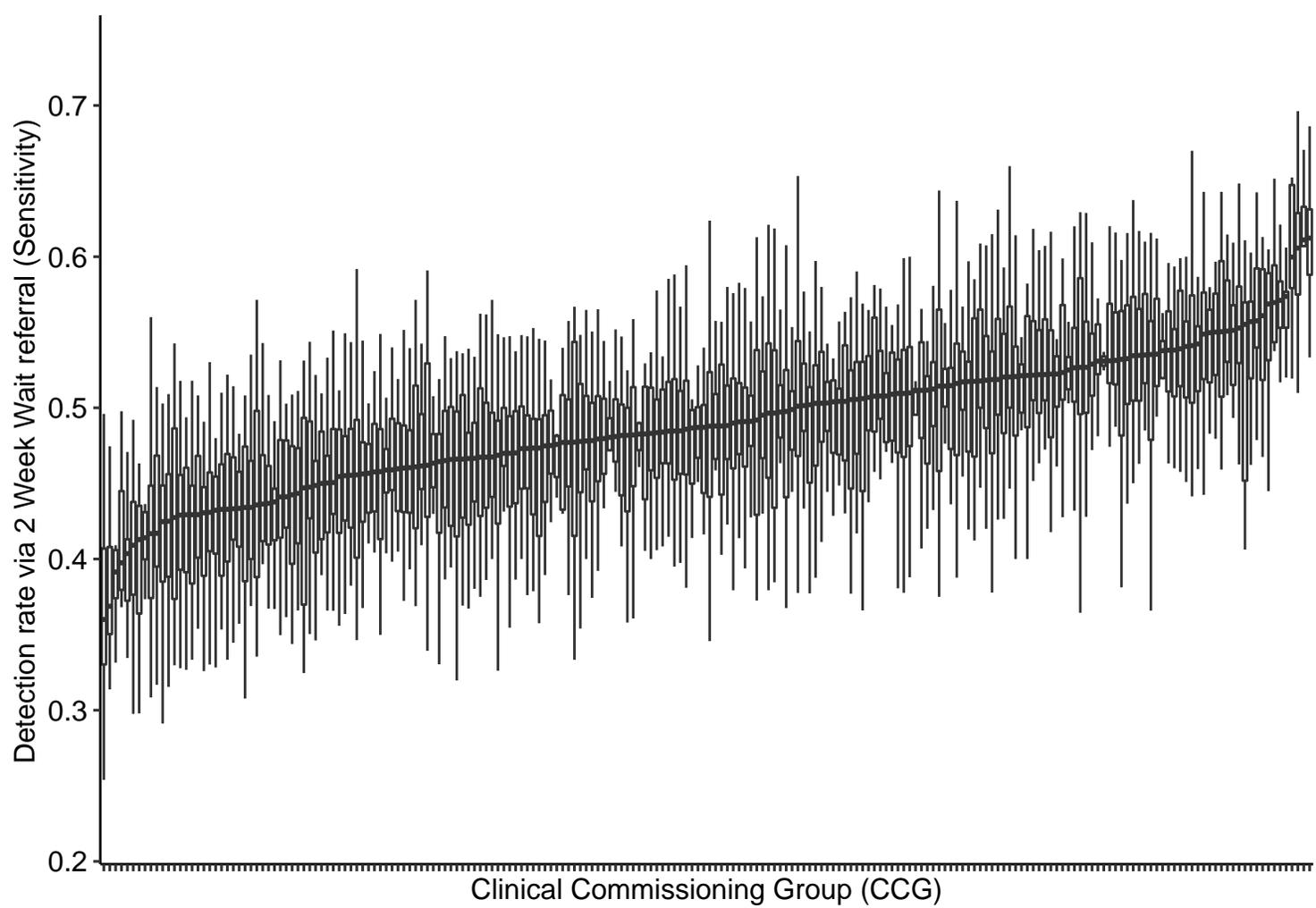
**Figure 1 Distribution of indicators of GP practice fast-track referrals for suspected cancer grouped by CCG**

CCG: Clinical Commissioning Group;

**A** Standardised 2 week wait referral rate



**B** Cancer detection rate

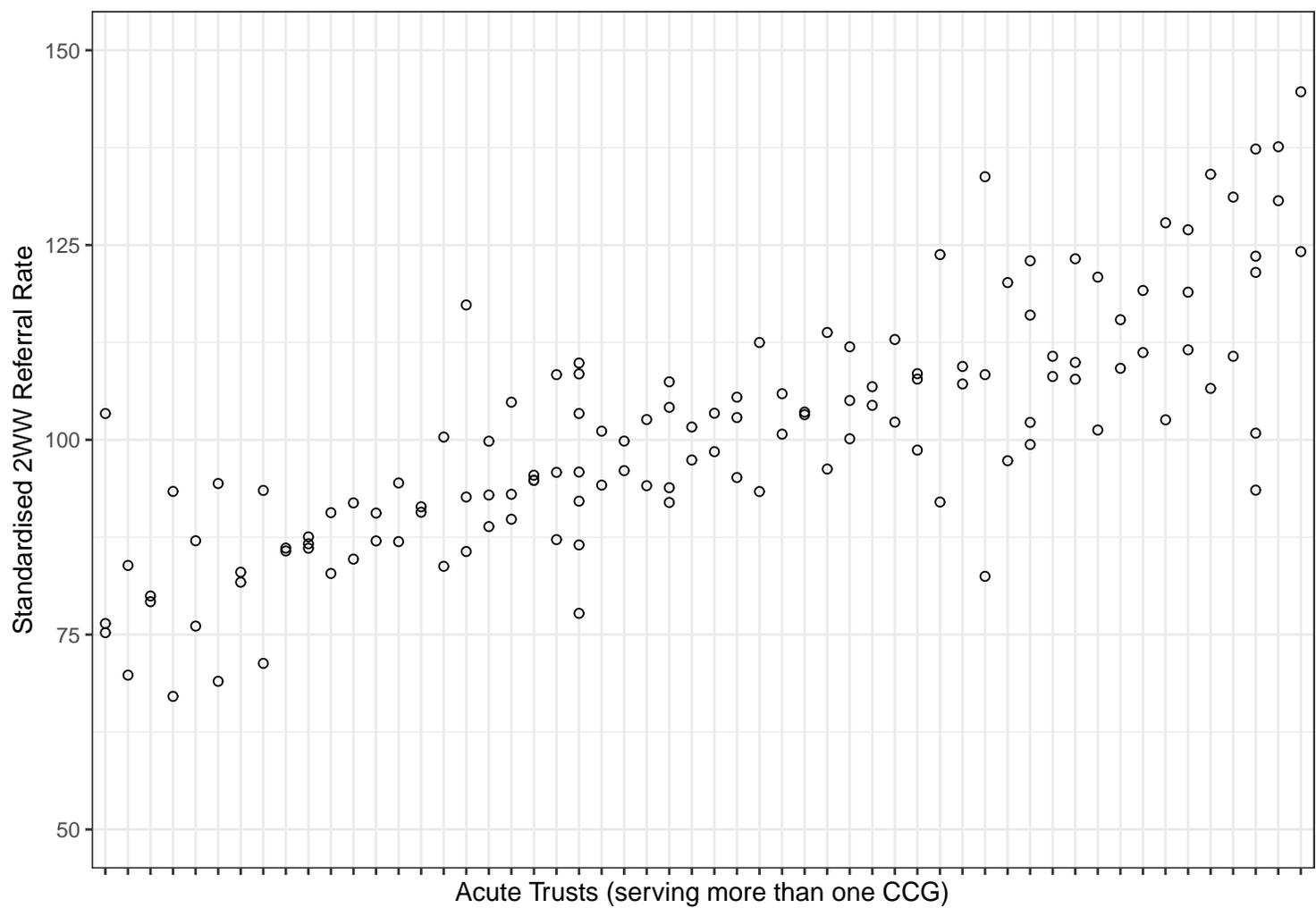


**Figure 2 Distribution of median indicators of GP fast track referrals per CCG grouped by Acute Hospital Trust**

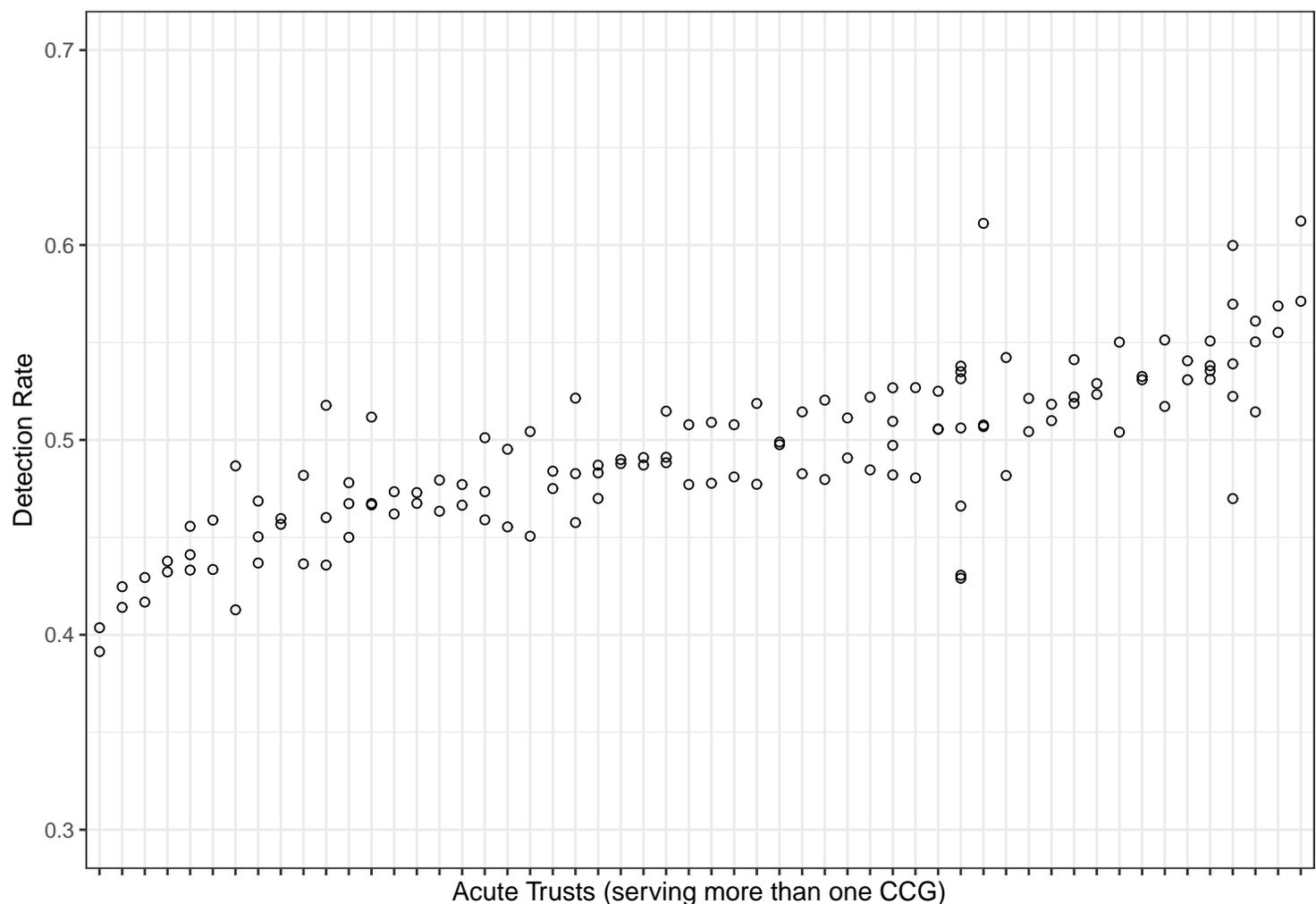


CCG: Clinical Commissioning Group; 2WW: Two week wait.

**A** Standardised 2 week wait referral rate



**B** Cancer detection rate



## References

1. Weller D, Vedsted P, Rubin G, et al. The Aarhus statement: improving design and reporting of studies on early cancer diagnosis. *Br J Cancer* 2012;106(7):1262-7. doi: 10.1038/bjc.2012.68 [published Online First: 2012/03/15]
2. Excellence NfHaC. Suspected cancer: recognition and referral (NG12). London: NICE, 2015.
3. Meechan D, Gildea C, Hollingworth L, et al. Variation in use of the 2-week referral pathway for suspected cancer: a cross-sectional analysis. *The British journal of general practice : the journal of the Royal College of General Practitioners* 2012;62(602):e590-7. doi: 10.3399/bjgp12X654551; 10.3399/bjgp12X654551
4. Møller H, Gildea C, Meechan D, et al. Use of the English urgent referral pathway for suspected cancer and mortality in patients with cancer: cohort study. *BMJ* 2015;351 doi: 10.1136/bmj.h5102
5. Rose PW, Rubin G, Perera-Salazar R, et al. Explaining variation in cancer survival between 11 jurisdictions in the International Cancer Benchmarking Partnership: a primary care vignette survey. *BMJ open* 2015;5(5):e007212-2014-12. doi: 10.1136/bmjopen-2014-007212 [doi]
6. Rubin G, Gildea C, Wild S, et al. Assessing the impact of an English national initiative for early cancer diagnosis in primary care. *Br J Cancer* 2015;112 Suppl 1:S57-64. doi: 10.1038/bjc.2015.43 [published Online First: 2015/03/04]
7. Mendonca SC, Abel GA, Gildea C, et al. Associations between general practice characteristics with use of urgent referrals for suspected cancer and endoscopies: a cross-sectional ecological study. *Family Practice* 2018 doi: 10.1093/fampra/cmz118
8. Lyratzopoulos G, Mendonca SC, Gildea C, et al. Associations between diagnostic activity and measures of patient experience in primary care: a cross-sectional ecological study of English general practices. *Br J Gen Pract* 2018;68(666):e9-e17. doi: 10.3399/bjgp17X694097 [published Online First: 2017/12/20]
9. Baughan P, Keatings J, O'Neill B. Urgent suspected cancer referrals from general practice: audit of compliance with guidelines and referral outcomes. *The British journal of general practice : the journal of the Royal College of General Practitioners* 2011;61(592):e700-6. doi: 10.3399/bjgp11X606591; 10.3399/bjgp11X606591
10. Rodriguez HP, Scoggins JF, von Glahn T, et al. Attributing sources of variation in patients' experiences of ambulatory care. *Med Care* 2009;47(8):835-41. doi: 10.1097/MLR.0b013e318197b1e1 [published Online First: 2009/07/09]
11. Abel G, Saunders CL, Mendonca SC, et al. Variation and statistical reliability of publicly reported primary care diagnostic activity indicators for cancer: a cross-sectional ecological study of routine data. *BMJ Qual Saf* 2018;27(1):21-30. doi: 10.1136/bmjqs-2017-006607 [published Online First: 2017/08/30]
12. Fung V, Schmittiel JA, Fireman B, et al. Meaningful variation in performance: a systematic literature review. *Medical care* 2010;48(2):140-48. doi: 10.1097/MLR.0b013e3181bd4dc3 [doi]
13. Sullivan CO, Omar RZ, Ambler G, et al. Case-mix and variation in specialist referrals in general practice. *The British journal of general practice : the journal of the Royal College of General Practitioners* 2005;55(516):529-33.
14. Zhou Y, Mendonca SC, Abel GA, et al. Variation in 'fast-track' referrals for suspected cancer by patient characteristic and cancer diagnosis: evidence from

- 670 000 patients with cancers of 35 different sites. *Br J Cancer* 2018;118(1):24-31. doi: 10.1038/bjc.2017.381 [published Online First: 2017/11/29]
15. Murchie P, Chowdhury A, Smith S, et al. General practice performance in referral for suspected cancer: influence of number of cases and case-mix on publicly reported data. *Br J Cancer* 2015;112(11):1791-8. doi: 10.1038/bjc.2015.110
  16. Burton CD, McLernon DJ, Lee AJ, et al. Distinguishing variation in referral accuracy from referral threshold: analysis of a national dataset of referrals for suspected cancer. *BMJ Open* 2017;7(8):e016439. doi: 10.1136/bmjopen-2017-016439 [published Online First: 2017/08/23]
  17. England PH. Cancer Services: Public Health England, 2018. <https://fingertips.phe.org.uk/profile/cancerservices>
  18. Nakagawa SS, H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* 2013;4(2):133-42. doi: doi:10.1111/j.2041-210x.2012.00261.x
  19. Johnson PCD. Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution* 2014;5(9):944-46. doi: doi:10.1111/2041-210X.12225
  20. Barclay ME, Lyratzopoulos G, Greenberg DC, et al. Missing data and chance variation in public reporting of cancer stage at diagnosis: Cross-sectional analysis of population-based data in England. *Cancer epidemiology* 2018;52:28-42. doi: 10.1016/j.canep.2017.11.005 [published Online First: 2017/11/28]
  21. Guthrie B, Donnan PT, Murphy DJ, et al. Bad apples or spoiled barrels? Multilevel modelling analysis of variation in high-risk prescribing in Scotland between general practitioners and between the practices they work in. *BMJ Open* 2015;5(11):e008270. doi: 10.1136/bmjopen-2015-008270 [published Online First: 2015/11/08]
  22. Gutacker N, Bloor K, Bojke C, et al. Should interventions to reduce variation in care quality target doctors or hospitals? *Health Policy* 2018;122(6):660-66. doi: 10.1016/j.healthpol.2018.04.004 [published Online First: 2018/04/29]
  23. Di Martino M, Alagna M, Cappai G, et al. Adherence to evidence-based drug therapies after myocardial infarction: is geographic variation related to hospital of discharge or primary care providers? A cross-classified multilevel design. *BMJ Open* 2016;6(4):e010926. doi: 10.1136/bmjopen-2015-010926 [published Online First: 2016/04/06]
  24. Di Martino M, Ventura M, Cappai G, et al. Adherence to Long-Acting Bronchodilators After Discharge for COPD: How Much of the Geographic Variation is Attributable to the Hospital of Discharge and How Much to the Primary Care Providers? *Copd* 2017;14(1):86-94. doi: 10.1080/15412555.2016.1202225 [published Online First: 2016/07/16]
  25. Di Martino M, Lallo A, Kirchmayer U, et al. Prevalence of antibiotic prescription in pediatric outpatients in Italy: the role of local health districts and primary care physicians in determining variation. A multilevel design for healthcare decision support. *BMC Public Health* 2017;17(1):886. doi: 10.1186/s12889-017-4905-4 [published Online First: 2017/11/19]
  26. Shawihdi M, Thompson E, Kapoor N, et al. Variation in gastroscopy rate in English general practice and outcome for oesophagogastric cancer: retrospective analysis of Hospital Episode Statistics. *Gut* 2014;63(2):250-61. doi: 10.1136/gutjnl-2012-304202 [published Online First: 2013/02/22]

27. Moller H, Coupland VH, Tataru D, et al. Geographical variations in the use of cancer treatments are associated with survival of lung cancer patients. *Thorax* 2018;73(6):530-37. doi: 10.1136/thoraxjnl-2017-210710 [published Online First: 2018/03/08]