



Harshfield, A., Abel, G. A., Barclay, S., & Payne, R. A. (2018). Do GPs accurately record date of death? A UK observational analysis. *BMJ Supportive and Palliative Care*. <https://doi.org/10.1136/bmjspcare-2018-001514>

Peer reviewed version

License (if available):  
CC BY-NC-ND

Link to published version (if available):  
[10.1136/bmjspcare-2018-001514](https://doi.org/10.1136/bmjspcare-2018-001514)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via BMJ Publishing at <https://spcare.bmj.com/content/early/2018/06/27/bmjspcare-2018-001514> . Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# **How accurately do GPs record date of death? A UK observational analysis of linked primary care and national mortality data.**

Amelia Harshfield<sup>1</sup>

Gary A Abel<sup>2</sup>

Stephen Barclay<sup>1</sup>

Rupert A Payne<sup>3</sup>

1. Primary Care Unit, University of Cambridge, Forvie Site, Robinson Way, Cambridge, CB2 0SR, UK
2. University of Exeter Medical School (Primary Care), Smeall Building, St Luke's Campus, Exeter, EX1 2LU, UK
3. Centre for Academic Primary Care, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK

## **Corresponding author**

Dr Rupert Payne  
Centre for Academic Primary Care  
University of Bristol  
Canynge Hall  
39 Whatley Road  
Bristol, BS8 2PS  
UK

Email: [r.payne@bristol.ac.uk](mailto:r.payne@bristol.ac.uk)

Phone: +44 117 3314545

## **Word count**

3,362

## **Abstract**

### Objective

To examine the concordance between date of death recorded in UK primary care and national mortality records.

### Methods

UK primary care data from the Clinical Practice Research Datalink was linked to Office for National Statistics (ONS) data, for 118,571 patients who died between September 2010 and September 2015. Logistic regression was used to examine factors associated with discrepancy in death dates between datasets.

### Results

Death dates matched in 76.8% of cases with primary care dates preceding ONS date in 2.9%, and following in 20.3% of cases; 92.2% of cases differed by <2 weeks. Primary care date was >4 weeks later than ONS in 1.5% of cases and occurred more frequently with deaths categorised as “external” (15.8% vs. 0.8% for cancer), and in younger patients (15.9% vs. 1% for 18-29 and 80-89 years respectively). General practices with the greatest discrepancies (97.5th percentile) had around 200-times higher odds of recording substantially discordant dates than practices with the lowest discrepancies (2.5th percentile).

### Conclusion

Dates of death in primary care records often disagree with national records and should be treated with caution. There is marked variation between practices, and studies involving young patients, unexplained deaths, and where precise date of death is important, are particularly vulnerable to these issues.

## **Keywords**

Mortality data; Death date; Primary care; Electronic health records

## **Introduction**

The use of large, routinely collected health data for research and policy purposes is increasingly common. In the UK, this includes clinical primary care data, administrative hospital records, disease registries and national mortality statistics. Linkage of the records within such datasets can add considerable value over and above the information held within the separate individual datasets. However, such linkages also present important challenges with respect to data quality(1). One important example is where different sources provide inconsistent results, not due to linkage mismatch, but due to incorrect recording of the same patient attribute in one or both linked datasets.

State-funded general practice provides free-at-point-of-access birth-to-death primary healthcare in the UK to nearly the whole population, and acts as gatekeeper to specialist services. The associated electronic health records are therefore a potentially invaluable source of data for clinical, public health and policy research. However, general practitioners (GPs) seldom record data with research in mind, but rather for the purposes of patient care. This can lead to under-coding or other coding inaccuracies: a fifth of GP record entries use free text only rather than structured coding(2). We have previously found that the use of hospital records as opposed to GP records to identify patients with cardiovascular disease, results in the identification of patient groups with significantly different clinical characteristics(3). A study of suicide and self-harm found that only a quarter of deaths officially recorded by the Office for National Statistics (ONS) were matched by a corresponding diagnostic code in general practice records(4). Others have found delays in the recording of clinical information within the primary care record(5).

Accurately determining the date of death is of importance for many observational research studies, particularly where end-of-life care is being considered; relatively small discrepancies in death date in these circumstances are undesirable. Given the aforementioned concerns about the accuracy of general practice data, the linkage of official mortality data to primary care data has the potential to provide more accurate information about the timing of death. However, linkage is not always available, and it would therefore be of value to have a better understanding of the accuracy of the primary care record alone for determining date of death. The aim of the current study was to examine the degree of, and factors associated with, concordance between date of death recorded in a large UK primary care dataset and that recorded in national death records.

## **Methods**

### **Data sources**

The Clinical Practice Research Datalink (CPRD) (6) is a large database of routinely collected UK primary care health records, widely used for epidemiological, public health and health service research, and covers approximately 8% of the UK population. It contains a range of clinical and administrative primary care data pertaining to UK general practice. This includes registration details: when a patient first registered with a practice, when they left a practice, and the reason for leaving a practice (e.g. moving away from the area, death). The primary care death date is derived using a CPRD algorithm to determine the most likely date from several sources within the GP record, including transfer-out date, death administration data (usually filled in from the death certificate), or explicit coded record of death(7). A subset of around three-quarters of the English practices represented in CPRD consent to having their patients' data linked to other national datasets, including ONS mortality data.

Mortality data published by ONS (8) are derived from official death registrations, based on medical certificates of the cause of death, which are completed by a doctor. It is a legal requirement in the UK that all deaths are registered within 5 days; in cases of unexpected or suspicious deaths, certificates are usually completed by a pathologist after post-mortem examination, which may lead to delays in registration. Only the date and underlying cause of death are available in the ONS data that is linked to CPRD.

The study was approved by the CPRD Independent Scientific Advisory Committee (Protocol 15\_239MnAR).

## Population

Our population was defined based on the June 2016 release of CPRD primary care data. Inclusion criteria were adults with a specified gender of male or female, aged between 18 and 105 at death, with a death date in primary care records occurring between September 2010 and September 2015, and flagged by CPRD as having acceptable data quality. We did not restrict inclusion to patients with a minimum period of registration prior to the index date, to avoid excluding patients who had moved recently prior to death. Linkage of primary care and ONS data was conducted by a trusted third party (Health and Social Care Information Centre) using a deterministic algorithm, with only the subset of English practices participating in the linkage scheme included. Stepwise CPRD matching criteria were applied, with records with the same NHS number, as well as postcode or sex and date of birth, considered a definitive match (corresponding to a CPRD *match\_rank* value of  $\leq 5$ )(9). Complete data were available for age, gender, and cause of death.

## Data analysis

Age at death was calculated by subtracting patients' year of death using primary care death date from year of birth, and categorised as 18-29 years, seven 10-year age bands from 30 to 99 years, and 100 years or over. Cause of death was based on ICD-10 coded "Underlying Cause of Death" as recorded in ONS data, and grouped into ten categories reflecting the reporting approach taken by ONS (see Table A.1 for classification).

For each patient, the difference in the two dates of death was calculated by subtracting the primary care death date from the ONS death date. The discrepancies in death date were grouped into 11 categories: exact date match, plus 5 categories each for the primary care date being before or after the ONS date (<1 week, 1 to <2 weeks, 2 to <4 weeks, 4 weeks to 6 months, and >6 months).

We hypothesised that the recording of clinical activity would decrease following death, with the potential to provide insight into whether the primary care or ONS date reflected the true date of death. We therefore calculated clinical activity recorded in the 2 weeks prior to and including the date of death, and in the 2 weeks after the date of death, for those patients where primary care and ONS dates differed; this was carried out for both primary care and ONS dates. Two measures were considered as proxies for clinical activity: firstly, recording of any prescription in the clinical record (irrespective of whether it had been previously prescribed); and secondly, recording of a consultation with either a doctor or nurse (the criteria used to determine such consultations are provided in Table A.2).

A descriptive analysis was undertaken, examining how the distribution of discrepancy in death date varied with age, gender and cause of death, with logistic regression used to model associations. We also explored the variation in death date discrepancy across all practices using a mixed-effects model. The random effect for practice utilised in this model represents the unmeasured variation between practices. Wilcoxon-rank tests were used to

examine whether there was evidence of a difference in the change in clinical activity before and after death between primary care and ONS dates.

All analyses were carried out in Stata 13.1 (10).

## **Results**

There were 118,571 patients in the primary care dataset matching the inclusion criteria, who died between September 2010 and September 2015 and had a linked complete ONS death date (Figure A.1). Linked ONS death date was not available for a further 1798 patients (Table A.3). The median age at death was 82 (interquartile range 72 to 89) years, and 47% of patients were male. This is similar to national death statistics over the same time period (median age 81 (IQR 70 to 87) years, 48% male) (8). The distribution of cause of death also matched that reported nationally(11), with cancer accounting for 29% of deaths (Table A.4).

Death dates matched perfectly in 76.8% of cases. ONS date of death was later than primary care date of death in only 2.9% of cases. Primary care date of death was between 1 day and 4 weeks later than ONS in 19.0% of cases, with 1.5% of primary care dates more than 4 weeks later than that recorded in ONS (Figure 1). The perfect agreement between dates was slightly worse in men than women (76.1% vs. 77.4%;  $p < 0.001$ ).

Differences in the frequency of discordant death dates are seen across different causes of death. Discrepancies in death dates were far more common with deaths categorised as “external causes”, where 15.8% of primary care death dates were 4 weeks or more later than the ONS death date (Figure 2A and Table A.5). In comparison, this degree of discrepancy was only observed in between 0.5% and 1.6% of other causes of death. Similar differences were seen in the degree to which there was exact agreement in date of death; 63.6% of cases with death recorded as “external causes” had exact agreement, compared with between 73.5% and 80.5% for other causes (Figure 2A and Table A.5). In comparison to deaths due to cancer (the commonest cause of death), deaths due to “external causes” had 25-times the odds of being recorded 4 weeks or more later in primary care than ONS (Table 1; unadjusted odds ratio 25.3, 95% confidence interval 21.8 to 29.4).

**Table 1. Association between discrepancy in death date and gender, age at death, and cause of death**

	N	Odds ratios (95% confidence interval)			
		Primary care death date >4 weeks later than ONS date		Any discrepancy in death date	
		Unadjusted	Adjusted*	Unadjusted	Adjusted*
<b>Gender</b>					
Male	56,181	Ref	Ref	Ref	Ref
Female	62,390	0.51 (0.47, 0.57)	0.70 (0.63, 0.78)	0.93 (0.90, 0.95)	0.96 (0.93, 0.99)
<b>Age at death</b>					
18-29	635	18.7 (14.8, 23.6)	4.47 (3.36, 5.95)	2.01 (1.71, 2.36)	1.69 (1.40, 2.04)
30-39	1,061	12.7 (10.4, 15.9)	4.29 (3.31, 5.55)	1.79 (1.57, 2.03)	1.62 (1.40, 1.88)
40-49	3,027	7.72 (6.53, 9.13)	3.75 (3.08, 4.58)	1.55 (1.43, 1.68)	1.45 (1.33, 1.59)
50-59	6,132	3.46 (2.92, 4.09)	2.53 (2.10, 3.05)	1.21 (1.14, 1.28)	1.21 (1.13, 1.30)
60-69	13,634	1.84 (1.57, 2.16)	1.70 (1.44, 2.01)	1.07 (1.02, 1.12)	1.09 (1.04, 1.15)
70-79	24,168	1.12 (0.96, 1.30)	1.12 (0.95, 1.31)	1.02 (0.99, 1.06)	1.02 (0.98, 1.07)
80-89	42,386	Ref	Ref	Ref	Ref
90-99	25,815	0.75 (0.64, 0.89)	0.76 (0.64, 0.91)	0.96 (0.92, 1.00)	0.95 (0.91, 0.99)
100+	1,713	0.64 (0.35, 1.17)	0.63 (0.34, 1.16)	0.73 (0.64, 0.83)	0.69 (0.60, 0.79)
<b>Cause of death</b>					
Ischaemic heart disease	14,513	2.04 (1.70, 2.45)	2.28 (1.89, 2.75)	1.37 (1.31, 1.43)	1.54 (1.46, 1.62)
Cancer	34,398	Ref	Ref	Ref	Ref
Dementia	11,186	0.70 (0.53, 0.93)	1.11 (0.82, 1.49)	0.92 (0.87, 0.97)	1.05 (0.98, 1.11)
Chronic lower resp. disease	6,616	1.67 (1.30, 2.14)	1.90 (1.47, 2.46)	1.17 (1.10, 1.25)	1.24 (1.16, 1.33)
Cerebrovascular disease	8,866	0.95 (0.72, 1.25)	1.28 (0.97, 1.71)	1.10 (1.04, 1.16)	1.23 (1.15, 1.31)
Influenza/pneumonia	6,743	0.88 (0.64, 1.22)	1.26 (0.91, 1.75)	1.24 (1.17, 1.32)	1.43 (1.34, 1.54)
Digestive disorders	5,713	1.81 (1.40, 2.35)	1.69 (1.30, 2.21)	1.31 (1.23, 1.40)	1.38 (1.29, 1.49)
Other circulatory	10,182	1.68 (1.36, 2.09)	2.16 (1.73, 2.69)	1.21 (1.15, 1.27)	1.38 (1.30, 1.47)
External causes	3,866	25.3 (21.8, 29.5)	21.3 (18.0, 25.3)	2.16 (2.02, 2.32)	2.28 (2.10, 2.48)
Other causes	16,488	2.21 (1.86, 2.63)	2.63 (2.20, 3.14)	1.10 (1.05, 1.15)	1.22 (1.16, 1.28)
<b>Practice</b>					
95% mid-ranget†		130 (79, 225)	189 (112, 340)	72 (53, 100)	74 (54, 104)

\* adjusted for gender, age at death, cause of death, and practice

† calculated from the variance of the random effect ( $\sigma^2$ ) and is given by  $e^{2 \times 1.96 \times \sigma}$  and represents the odds ratio comparing a practice at the 2.5th percentile of the distribution of practices to one at the 97.5th percentile

The discrepant death dates tended to be more common in those of younger age (Table 1, Figure 2B and Table A.5). Primary care and ONS date of death had exact agreement for 77.4% of individuals dying in their 80s (the commonest age category), with primary care 4 weeks or more later than ONS in only 1% of cases. In comparison those in their 40s had over 7-times the odds of having a primary care date of death 4 weeks or more later than ONS (7.2%; unadjusted OR 7.7, 95% CI 6.5 to 9.1); the difference in those aged 18 to 29 years was even more marked (15.9%; OR 18.7, 95% CI 14.8 to 23.6) (Table 1, Figure 2B and Table A.5). The association with age also persisted after adjustment for cause of death and gender but was substantially attenuated (Table 1), with those dying aged 18 to 29 years substantially more likely than those dying in their 80s to have a primary care date of death 4 weeks or more later than ONS (adjusted OR 3.8, 95% CI 2.9 to 4.9).

There was considerable variation in discrepancy in death dates across practices (Figure 3). Using the estimated variance of the random intercept in our adjusted mixed model we estimate that practices at the top of the 95% mid-range of practices (i.e. 97.5th percentile compared to 2.5th percentile) had almost 200 times the odds of recording death dates discordant with ONS (OR 189, 95% CI 112 to 340; Table 1).

Associations between all the factors of interest (including practice) and discrepancy in death date were weaker when we used a very stringent definition of discrepancy (i.e. any difference  $\geq 1$  days). The odds ratios in a model considering any discrepancy in death dates are considerably smaller than a model considering discrepancies  $>4$  weeks (Table 1).

Rates of clinical activity before and after date of death are illustrated in Figure 4. Prescribing activity was generally higher in the fortnight before ONS date of death than it was in the fortnight after (4.2 vs. 0.5 prescriptions), with a similar pattern observed for the primary care date of death (2.6 vs. 0.4). The magnitude of decrease was significantly greater for the ONS date of death ( $p < 0.001$ ). Rates of face-to-face consultations also decreased following ONS date of death (1.1 vs. 0.2 consultations). Although relatively less marked than for prescribing, this change was again greater ( $p < 0.001$ ) than that observed following primary care date of death (0.8 vs. 0.1 consultations).

## **Discussion**

This study found that there was a discrepancy in recorded death dates between primary care records (from CPRD data) and ONS national data in almost a quarter of cases, although there is very considerable variation between GP practices in the degree of discrepancy. In the majority of cases of discrepancy, the date of death recorded by the GP comes after that recorded by ONS. When broken down by cause of death and age, it is apparent that those individuals with an external cause of death, and younger individuals, are more likely to have a substantial delay between the date of death recorded by ONS and that recorded in the GP record. However, unmeasured practice factors have by far the greatest impact on discrepancy in death date.

### **Interpretation of findings**

There is a legal obligation to register deaths in a timely fashion in the UK: this formal process might thus be expected to result in relatively accurate national records. Indeed, the ONS date of death should probably still be regarded as the “gold standard”. Recording the date of death in primary care records is not covered by such legislation, but can be important in clinical practice, to avoid causing distress for relatives by mistakenly attempting to contact a deceased patient and for audit purposes(12). Since GP records are predominantly kept to support the provision of direct clinical care, the accurate recording of death information may

not be prioritised. GP records may therefore reflect the date the practice is notified of a death (if the death has not been certified by the GP, this notification is usually received from family, carers or a hospital), rather than the actual date of death, resulting in the general tendency observed for GP-recorded dates to be later than ONS dates. It is also possible that the data provided to GPs is inaccurate, although this is unlikely to be the main reason for the observed discrepancies as these are predominantly in one direction.

In a smaller study using the similar THIN GP dataset, 78% of 584 deaths had perfect agreement of death date between GP record and external paper records; this is very similar to our own findings(13). Of note, that study also found that precise agreement was poor (5%) for those cases where the GP record only comprised data about the transfer of the patient out of the care of the GP (i.e. with no formal recording of death); in that situation, the GP record of death was 21 days later than that recorded in the external data.

The issue of delayed recording can be expected to be particularly marked in situations where a coroner is involved and post-mortem examinations undertaken. Such investigations are common in the case of unanticipated or unnatural deaths, such as those classified as being due to “external causes” and those in younger age groups. This is reflected in the notable delays we observed in these groups of individuals. For example, the average delay nationally in registration for deaths due to drug misuse is over 5 months(14). Discrepancies in diagnostic coding between GP records and ONS data has previously been noted in the majority of cases of suicide and self-harm(4), although that study did not examine the accuracy of death dates. Of note, the tendency of younger individuals to have more substantial delays is, to a degree, independent of cause of death, reflecting the likelihood that further investigation of the circumstances surrounding a death in a younger person will be undertaken by a coroner as an unexpected death, regardless of cause.

However, even though age and cause of death impact upon the dates recorded in GP records, our analysis shows that unmeasured practice factors have a greater impact on death date discrepancy. Our data do not allow us to examine the underlying reasons, which are likely to be multifactorial including variations in local administrative processes and practice staffing. Understanding these factors may provide insights into how the accuracy of death recording can be improved. Our observation that associations between discrepancy in death dates and age, gender, cause of death and practice all become weaker when a looser definition of concordant date is used, is consistent with small discrepancies being subject to a relatively random element irrespective of circumstances.

Our findings that prescribing tends to decrease more noticeably following ONS date of death than date of death in the primary care record is consistent with the ONS date being more likely to represent the true date. This pattern was not obvious when we examined consultations (although we do note consultations decreased substantially following date of death, however defined). This is probably because consultation type is poorly recorded in GP systems, with a “face-to-face” consultation often recorded as the default even if the consultation is simply an administrative event. For both prescribing and consultations, genuine clinical activity can still occur after death; in the former case, an order for further medication may be placed prior to death occurring, and in the latter case consultations may take place with a relative or other third party, or may represent recording of post-mortem administrative information.

### **Strengths and limitations**

To our knowledge, this is the first study that has documented the quality of recording of date of death within GP records. Key strengths are the use of a large, contemporary dataset which is representative of the UK population as a whole, and makes comparisons against a

robust, mandatory national dataset. However, there are also important limitations to consider. First, is the representativeness of the data. The availability of linked data was limited to a subset of English CPRD practices. However, discrepancies in death date recording are unlikely to be limited to geographical boundaries, and we expect similar patterns will exist across the wider UK. In addition, it is possible that practices enrolled with CPRD record death more reliably than other practices. We have no data to support this supposition, although given recording of death is so important for practice administration purposes, we have no particularly strong reason to believe that there would be a significant difference between CPRD and non-CPRD practices in this regard. Second, we cannot be sure that ONS date of death is accurate. Nevertheless, our analysis of prescribing activity provides some reassurance that this date is more likely to represent the true date of death than that recorded in the GP record. Third, it is difficult to identify genuine face-to-face consultations within CPRD. Our approach to identifying such events was based on work commissioned by the UK Department of Health(15). However, misclassification of consultations remains likely, and it is possible that a spike in recording of administrative data in the immediate post-mortem period may have blunted the apparent decrease in clinical activity that we anticipated. Fourth, we cannot account for the small proportion of patients for whom no ONS record of death existed. It is worth noting that this group was disproportionately young and male (Table A.3), so missing records may reflect delays in ONS data being reported for unanticipated deaths. Finally, we were unable to examine to what degree the algorithm used by CPRD to calculate primary care date of death may have contributed to discrepancies. It may be possible to refine this algorithm based on knowledge of how age and cause of death impact upon recording of death in the GP record.

## **Conclusions**

We have found that date of death recorded in GP data disagrees with that in national records in around a quarter of cases. ONS date of death should be favoured where available. However, over 90% of GP dates agree within 2 weeks of the corresponding ONS record, and it is likely that the GP date is adequate for the majority of research and policy needs. Nevertheless, in circumstances where a more accurate date of death is required, such as studying care in the last few weeks or days of life, or in situations where deaths may be unexpected or occur in younger individuals, the accuracy of the GP date of death should be treated with caution. The accuracy of data from practices with historically large discrepancies should also be carefully considered; further research into the cause of such wide inter-practice variation is needed. Future work should examine the accuracy with which cause of death can be determined from the GP record alone, as well as examining the integrity of the CPRD algorithm used to calculate primary care date of death.

## **Contributors**

AH designed the study protocol, extracted, organised, analysed and interpreted the data, drafted and revised the article, and gave final approval of the version to be published. GAA gave statistical input, interpreted the data, reviewed and edited draft the article and gave final approval of the version to be published. SB interpreted the data, reviewed and edited draft the article and gave final approval of the version to be published. RP designed the study protocol, interpreted the data, reviewed and edited draft the article and gave final approval of the version to be published.

## **Funding**

This paper presents independent research funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR) (grant reference: FR9/251). The views expressed are those of the authors and not necessarily those of the NIHR, the NHS, or the Department of Health.

## **Disclaimer**

The study is based on data form the Clinical Practice Research Datalink obtained under license from UK Medicines and Healthcare Products Agency. The interpretation and conclusions contained in this study are those of the authors alone.

## **Licence for Publication**

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ Supportive and Palliative Care and any other BMJGL products and sublicences such use and exploit all subsidiary rights, as set out in our licence (<http://group.bmj.com/products/journals/instructions-for-authors/licence-forms>).

## **Competing interests**

None declared.

## **Ethics approval**

The protocol (15\_239MnAR) for this study was approved on 2 August 2017 by the Independent Scientific Advisory Committee (ISAC), the independent body that approved use of CPRD data.

## **Figure legends**

Figure 1. Differences between ONS and primary care death dates

Figure 2A. Discrepancies in death dates by cause of death

Figure 2B. Discrepancies in death dates by age at death

Figure 3. Difference in death dates categories by different practices

Figure 4. Clinical activity in 2 weeks before and after ONS and primary care death dates

## **References**

1. Harron K, Wade A, Gilbert R, et al. Evaluating bias due to data linkage error in electronic healthcare records. *BMC Medical Research Methodology*. 2014;14(1):36.
2. Shah AD, Martinez C, Hemingway H. The freetext matching algorithm: a computer program to extract diagnoses and causes of death from unstructured text in electronic health records. *BMC medical informatics and decision making*. 2012;12:88.
3. Payne RA, Abel GA, Simpson CR. A retrospective cohort study assessing patient characteristics and the incidence of cardiovascular disease using linked routine primary and secondary care data. *BMJ open*. 2012;2(2):e000723.
4. Thomas KH, Davies N, Metcalfe C, et al. Validation of suicide and self-harm records in the Clinical Practice Research Datalink. *British Journal of Clinical Pharmacology*. 2013;76(1):145-57.
5. Leite A, Andrews NJ, Thomas SL. Assessing recording delays in general practice records to inform near real-time vaccine safety surveillance using the Clinical Practice Research Datalink (CPRD). *Pharmacoepidemiology and drug safety*. 2017;26(4):437-45.
6. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology*. 2015;44(3):827-36.
7. Clinical Practice Research Datalink. Identifying dates of death in CPRD. London: Medicines and Healthcare products Regulatory Agency.
8. Office for National Statistics. Quality and Methodology Information (QMI): Mortality statistics in England and Wales QMI. Newport, UK; 2016.
9. Clinical Practice Research Datalink. ONS Death Registration Data: Data Specification v1.2. London: Medicines and Healthcare regulatory Agency; 2015.
10. StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
11. Office for National Statistics. Deaths Registered in England and Wales: 2014. 15 July 2015 [Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregistrationsummarytables/2015-07-15>].
12. Singh S. How not to contact the dead. *The British Journal of General Practice*. 2013;63(611):309.
13. GC Hill. Validation of death and suicide recording on the THIN UK primary care database. *Pharmacoepidemiology and drug safety*. 18:120-31.
14. Office for National Statistics. Deaths related to drug poisonings in England and Wales: 2015 registrations. Newport, UK; 2016.
15. Hobbs FDR, Bankhead C, Mukhtar T, et al. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007-2013;14. *The Lancet*.387(10035):2323-30.