Received: 19 April 2018 | Revised: 29 November 2018 | Accepted: 14 January 2019

DOI: 10.1002/pds.4747

ORIGINAL REPORT

WILEY

The accuracy of date of death recording in the Clinical Practice Research Datalink GOLD database in England compared with the Office for National Statistics death registrations

Arlene M. Gallagher[1,2] | Daniel Dedman[1] | Shivani Padmanabhan[1] | Hubert G. M. Leufkens[2] | Frank de Vries[2,3]

[1] Clinical Practice Research Datalink, Medicines and Healthcare products Regulatory Agency, London, UK

[2] Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands

[3] Department of Clinical Pharmacy & Toxicology, Maastricht University Medical Centre+, Maastricht, The Netherlands

Correspondence
A. M. Gallagher, Clinical Practice Research Datalink (CPRD), Medicines and Healthcare products Regulatory Agency, 10 South Colonnade, Canary Wharf, London E14 4PU, UK.
Email: arlene.gallagher@cprd.com

Abstract

Purpose: It is not clear whether all deaths are recorded in the Clinical Practice Research Datalink (CPRD) or how accurate a recorded date of death is. Individual-level linkage with national data from the Office for National Statistics (ONS) and Hospital Episode Statistics (HES) in England offers the opportunity to compare death information across different data sources.

Methods: Age-standardised mortality rates (ASMRs) standardised to the European Standard Population (ESP) 2013 for CPRD were compared with figures published by the ONS, and crude mortality rates were calculated for a sample population with individual linkage between CPRD, ONS, and HES data. Agreement on the fact of death between CPRD and ONS was assessed and presented over time from 1998 to 2013.

Results: There were 33 997 patients with a record of death in the ONS data; 33 389 (98.2%) of these were also identified in CPRD. Exact agreement on the death date between CPRD and the ONS was 69.7% across the whole study period, increasing from 53.4% in 1998 to 78.0% in 2013. By 2013, 98.8% of deaths were in agreement within ±30 days.

Conclusions: For censoring follow-up and calculating mortality rates, CPRD data are likely to be sufficient, as a delay in death recording of up to 1 month is unlikely to impact results significantly. Where the exact date of death or the cause is important, it may be advisable to include the individually linked death registration data from the ONS.

KEYWORDS

CPRD, data linkage, death, HES, mortality, ONS, pharmacoepidemiology, validation, UK

This paper had not previously been published.

1 | INTRODUCTION

Electronic health records (EHRs) are increasingly used in clinical research. The Clinical Practice Research Datalink (CPRD) is an ongoing primary care database of anonymised medical records from general practitioners (GPs) and has been the basis of 886 published papers in the last 5 years.[1] The CPRD GOLD database is based on GP records using the Vision software system, and the quality of primary care data is variable since data are entered by GPs during routine consultations, and not for the purpose of research.

Death is a common outcome measure in EHR studies with multiple articles published each year using CPRD GOLD.[2-7] Clear information on the date of death is important for censoring in observational epidemiology, especially for analyses of mortality and for studies assessing events at the end of life.[8] It is not clear whether all deaths are recorded in CPRD GOLD or how accurate a recorded date of death is.[9] In England and Wales, when a patient dies, it is the statutory duty of the doctor who had attended in the last illness to issue the death certificate.[10] However, GPs may not routinely receive information about their deceased patients for whom they did not complete the death certificate.[11] Primary care software systems allow the recording of death information via multiple methods and changes to the software over time or when converting between systems may impact the completeness of records.[12] CPRD researchers have developed an algorithm to identify multiple potential records of death from across the primary care record and combine this information to identify the best estimate for the date of death; this has not been externally validated (see Appendix S1 for details of the algorithm).

Individual-level linkage with national data from the Office for National Statistics (ONS) and Hospital Episode Statistics (HES) in England offers the opportunity to compare death information across the sources. This external validation will help researchers using EHRs to make decisions over the choice of data source when death is an important study variable. The objective of this study was to compare death date information based on the CPRD GOLD algorithm with national death registration data from the ONS and hospital death records from the HES for England. We examined how the mortality rate for CPRD GOLD was compared with the published figures from the ONS and used individually linked data to investigate how both the fact of death and date of death were compared between data sources.

KEY POINTS

- The majority of deaths are recorded in UK primary care data.
- The mortality rate calculated for CPRD is comparable with the ONS.
- Where the exact date of death or the cause is important in a CPRD study, it may be advisable to include the individually linked national ONS death registration data.
- Researchers should take into consideration the impact on sample size and power of a study when deciding whether to use the linked data as this can restrict the cohort size substantially.

2 | METHODS

The data used for this study were sourced from three routinely collected linked data sources in England. CPRD GOLD primary care data comprise the anonymous longitudinal EHRs of over 14 million patients in the United Kingdom in a dynamic cohort where patients can join/leave a GP practice over the course of follow-up.[13,14] These data are based on the Vision software system and have been shown in numerous validation studies to be generally of high quality.[15,16] Data have been collected since 1987. The ONS death registration data contain the date and coded cause of death for the population of England and Wales from January 1998, based on the death certificate, and are considered the gold standard.[17] The national HES data contain details of all admissions to National Health Service (NHS) hospitals in England from April 1997, including information on deaths in hospital.[18] Linkage between CPRD GOLD data and the ONS and HES data sources is available at an individual level for those patients registered at practices in England that have consented to data linkage. Linkage between data sets is undertaken using a deterministic linkage algorithm, based on a patient's exact NHS identification number, sex, date of birth, and residential postcode.

A sample of one million patients from CPRD GOLD was identified in January 2014 for a previous publication, based on those who had been registered at a contributing practice for at least 1 day before this date.[14] We used the same sample, restricting it further to those alive when linked ONS death registration data became available (1998; 85.7%) and to those in England (81.5%) who were eligible for linkage with both the ONS data and the HES data (76.8%) to create our study cohort (see Figure SA). Follow-up time was assessed based on the time since registering at the GP practice until the patient was deducted from the record (transferred out of the practice for any reason). Analyses were restricted to the overlapping coverage period of active follow-up in all three data sources since the influence of the denominator definition is important in studies of linked data.[19]

In order to compare the overall mortality rate in CPRD GOLD with the ONS data, we calculated age-standardised mortality rates (ASMRs) standardised to the European Standard Population (ESP) 2013 and compared this with figures published by the ONS using the same methodology. Data from the ONS were available from 2001 onwards. The ESP is an artificial population structure that is used in the weighting of mortality data to produce comparable ASMRs.[20] The population with individual linkage between all three data sources was used to calculate crude mortality rates in each data source and compare them. Poisson regression was used to calculate rates per 1000 person-years overall and by calendar year from 1998 onwards. We assessed agreement on the fact of death between the CPRD GOLD and the ONS and compared the difference in dates between the two over time.

3 | RESULTS

Table 1 describes the random sample of patients with data from 1998 to 2014 (n = 857 047) and the study cohort of 536 341 in England who

TABLE 1 Characteristics of a random sample of individuals in CPRD GOLD, active (alive) since 1998, and the subcohort in England eligible for linkage

Total N, %, Unless Stated Otherwise	Random Sample of Patients	Subcohort of the Random Sample: Patients in England, Eligible for ONS and HES Linkage
Total number of patients (N)	857 047	536 341
Time since registration at the GP practice		
Time in years, mean (SD)	11.7 (12.9)	11.2 (12.4)
Time in years, median (IQR)	7.1 (2.3-17.2)	6.7 (2.3-16.5)
Follow-up time during the linkage coverage period		
Follow-up time in years, mean (SD)		7.3 (5.7)
Follow-up time in years, median (IQR)		5.8 (2.0-13.2)
Sex		
Female	442 055 (51.6%)	276 384 (51.5%)
Male	414 992 (48.4%)	259 957 (48.5%)
Age at registration with the GP practice		
Age in years, mean (SD)	28.1 (21.7)	28.4 (21.7)
Age in years, median (IQR)	27.0 (9.0-41.0)	27.0 (10.0-41.0)
Age < 18	271 091 (31.6%)	165 438 (30.8%)
Age 18-29	213 520 (24.9%)	134 027 (25.0%)
Age 30-39	146 978 (17.1%)	94 554 (17.6%)
Age 40-49	84 790 (9.9%)	53 868 (10.0%)
Age 50-64	79 505 (9.3%)	50 021 (9.3%)
Age 65-74	31 534 (3.7%)	19 712 (3.7%)
Age 75+	29 629 (3.5%)	18 721 (3.5%)
Body mass index (BMI), aged 18+		
Mean BMI (SD)	25.9 (5.3)	25.9 (5.2)
Median BMI (IQR)	25.1 (22.3-28.6)	25.0 (22.3-28.5)
BMI unknown or out of range (<10, 70+)	211 792 (30.4%)	127 437 (29.3%)
Practice region		
North-East	13 158 (1.5%)	10 147 (1.9%)
North-West	90 440 (10.6%)	73 149 (13.6%)
Yorkshire and The Humber	29 628 (3.5%)	19 430 (3.6%)
East Midlands	30 160 (3.5%)	16 758 (3.1%)
West Midlands	68 561 (8.0%)	55 029 (10.3%)
East of England	80 799 (9.4%)	61 526 (11.5%)
South-West	71 233 (8.3%)	64 435 (12.0%)
South Central	98 990 (11.6%)	69 935 (13.0%)
London	126 807 (14.8%)	99 927 (18.6%)
South-East Coast	88 714 (10.4%)	66 005 (12.3%)
Northern Ireland	20 677 (2.4%)	0 (0.0%)
Scotland	74 961 (8.7%)	0 (0.0%)
Wales	62 919 (7.3%)	0 (0.0%)

Abbreviations: CPRD, Clinical Practice Research Datalink; GP, general practitioner; HES, Hospital Episode Statistics; IQR, Interquartile range; ONS, Office for National Statistics; SD, standard deviation.

were eligible for linkage with both the ONS death registration data and the HES data; there were no major differences between the two. In the study cohort, the average follow-up time since registration at the GP practice was 11.2 years (standard deviation [SD] 12.4), of which 6.8 years (SD 5.5) were during the linkage coverage period. There were slightly more females than males (51.5% vs 48.5%). Most registered with their GP in the early years of life with over half registering before their 30th birthday. Body mass index (BMI) data were available for 70.7% of adult patients (aged 18 y or more), where the mean was 25.9 (SD 5.2); BMI information was missing for 97.9% of children. Smoking information was captured for 85.2% of adult patients; 25.4% were non-smokers, 29.6% ex-smokers, and 30.2% smokers. Smoking information was missing for 86.4% of children, and 11.0% were recorded as non-smokers. All patients in the subcohort available for linkage were in England, spread across all regions.

ASMRs per 100 000 population, standardised to the ESP 2013, are presented separately for males and females in CPRD GOLD alongside the published national rates from the ONS, which were available from 2001 onwards (Figure 1). In both data sources, the mortality rate was higher for males than females. The mortality rate published by the ONS was higher than calculated for CPRD GOLD for both males and females.

For our study cohort, the crude mortality rate for 1998 to 2013 using the ONS data was 9.33 per 1000 person-years (95% confidence interval, 9.23-9.43). For the same patients, this was compared with 9.41 (9.31-9.51) for CPRD GOLD data and 4.27 (4.21-4.34) for HES data. Mortality rates by calendar time for all three sources are presented in Figure 2.

There were 33 997 patients with a record of death in the ONS data; 33 389 (98.2%) of these were also in CPRD GOLD (Figure 3). There were 608 patients with a death recorded in the ONS but not in CPRD GOLD; 498 (82%) had been transferred out of the practice for a reason other than due to death. There were 834 additional patients for whom the algorithm used by CPRD identified a date of death; however, there was no record in the ONS; 802 (96.2%) had been transferred out due to death. There were 17 deaths documented in the HES that were not recorded in either the ONS or CPRD GOLD data.

Exact agreement between the death date estimated by the CPRD GOLD algorithm and the date from the ONS was 69.7% across the whole study period, increasing from 53.4% in 1998 to 78.0% in 2013 (Figure 4). An earlier death date was identified in CPRD GOLD for less than 3%, and this did not change over time. A later date was identified in CPRD GOLD for 27.7%, reducing from 44.9% in 1998 to 19.5% in 2013. By 2013, 98.8% of deaths were in agreement within ±30 days.

4 | DISCUSSION

This study showed that 98.2% of deaths in the ONS data are recorded in the CPRD GOLD primary care data. With such a high overall agreement rate, for studies where the fact of death is important, it would be acceptable to use death information from CPRD GOLD. Agreement on the exact date of death increased over time to 78.0% in 2013. Discrepancies were most commonly when the death date recorded at the GP practice was later than the official date on the death certificate; however, both the amount of disagreement and the size of the delay have been reducing over time.

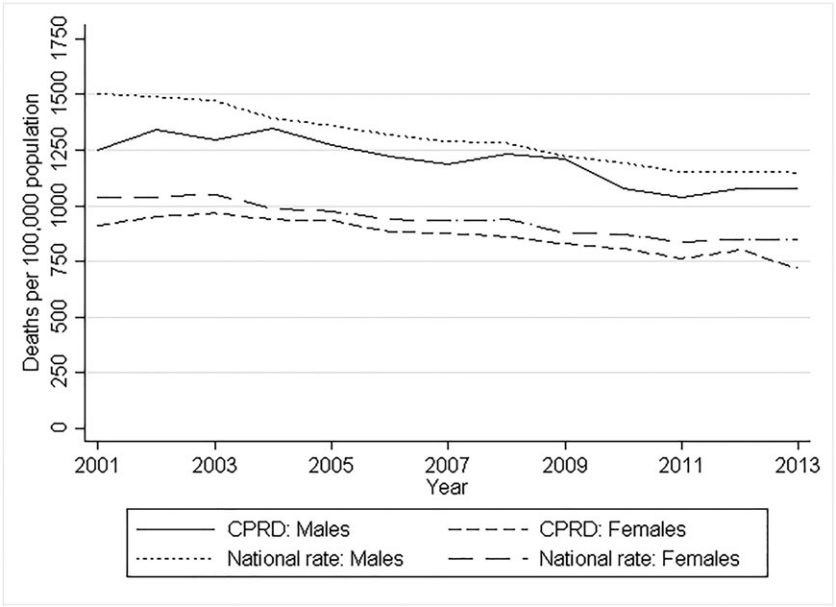

FIGURE 1 Age-standardised mortality rates (ASMRs), 2001 to 2013 for England. CPRD, Clinical Practice Research Datalink

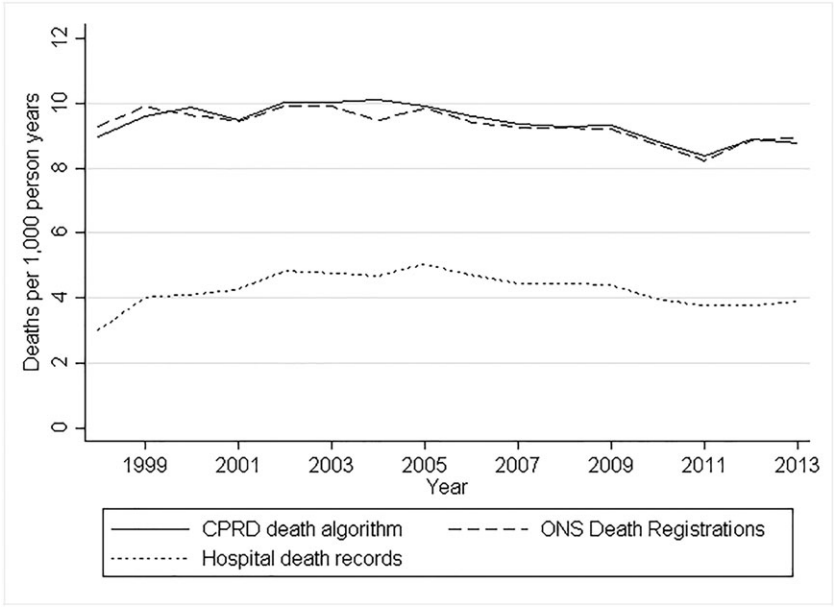

FIGURE 2 Mortality rates, 1998 to 2013 for the same patients in England, as recorded in three data sources. CPRD, Clinical Practice Research Datalink; ONS, Office for National Statistics

The objective of this study was to compare death date information based on the CPRD GOLD algorithm with national information from the ONS and hospital records, since GPs may not routinely receive information about their deceased patients and there is more than one way to record a death using the primary care software system. We were able to compare the information at the GP practice with the death certificate data by comparing individual records in both data sources.

The CPRD GOLD algorithm has not been externally validated, and questions have been posed as to whether all deaths are recorded at the practice. The high level of agreement on the fact of death found in this study is encouraging and suggests that the majority of deaths are recorded in CPRD GOLD. Of the 1.8% that were not recorded in CPRD GOLD, a large proportion (82%) had been transferred out of the practice for another reason (not due to death), suggesting that their record had been closed before the GP had been informed of their death. The HES data were limited to patients admitted to hospital, where as many as half of all deaths may occur.[21] For 2015/2016, 46.4% of deaths occurred in hospital in England[22] and we found a similar rate in this study. For this reason, we would not recommend using the HES as the sole source of data for identifying death as an outcome.

This study found that even in recent years, one-fifth of patients have a date of death recorded at the GP practice that is later than the date recorded in the ONS, which is based on the official death certificate. There are several reasons why there may be delayed recording of death information at the GP practice. Firstly, the practice may not be informed within a reasonable time frame. For expected deaths, the registered GP is likely to be informed as soon as it is practicable; however, it is not essential for them to attend to verify a death[23] and the procedure for ensuring the registered GP is notified is not clear. They may be informed by the attending GP, the hospital, and relatives, or information may first arrive from a central request to deduct the patient's record. Secondly, there may be limited incentive at the practice to close a patient's record immediately since their income relates to the list size.[24,25] Thirdly, there may be little reason to ensure that

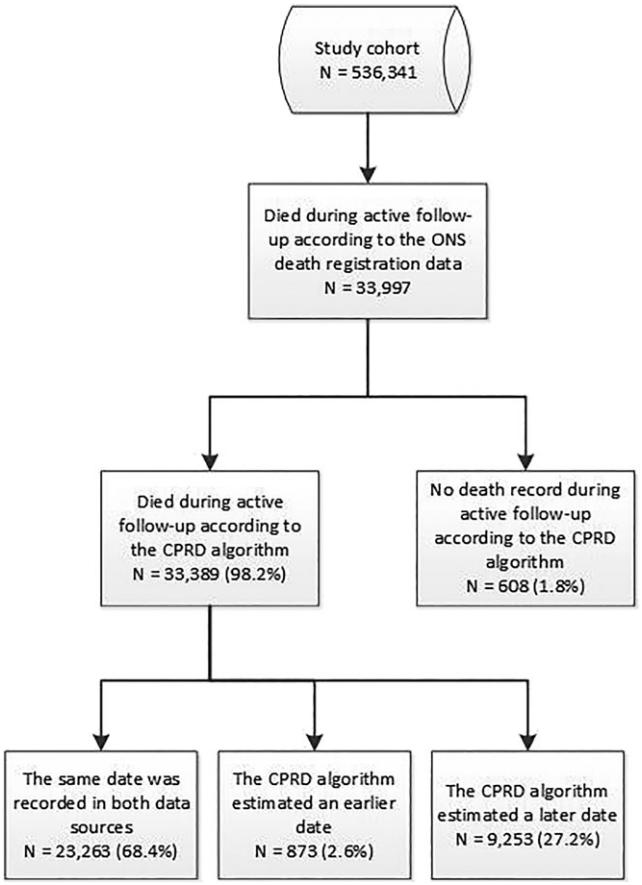

FIGURE 3 Concordance between the Clinical Practice Research Datalink (CPRD) algorithm and the gold standard Office for National Statistics (ONS) data using individually linked data

the exact date of death is recorded since further clinical care is no longer required and this type of administrative task is unlikely to be high priority for the practice. The lag time may be affected by the cause of death, which we did not assess in this study. For some conditions, patients are highly likely to be in hospital at the time of death, as was found for heart failure,[26] and certain types of death, such as suicide, may be recorded differently.[27] Further studies could investigate the differences in lag time by cause of death. Delayed recording at the GP practice may contribute to an immeasurable time bias where CPRD GOLD is used for a study without linking to the ONS death registration data.[28] This type of bias is likely to have more impact on case-control study designs that use mortality as an outcome. In this study, we found the median delay to be 16 days. For studies where it is important to be exact on the date of death, it may be advisable to use the linked ONS data in order to minimise including immeasurable person time; however, the impact is likely to be less in more recent years as both the proportion of patients and the median delay have reduced over time. This is in line with the improvement in the completeness of recording of many types of record at the GP practice in recent years.[14]

This study used the same random sample of patients from a previous publication, which provided a detailed overview of the data available in CPRD GOLD.[14] The sample was limited to patients eligible for linkage with other data sources and to those with follow-up since 1998. This immediately reduced the original sample of one million patients to 53% in a similar way to previously published studies.[29] Researchers need to balance in the loss in patient numbers alongside the gain in information from the linked data. Several CPRD GOLD studies have used the linked ONS data in order to provide additional information on the cause of death, which would otherwise be less complete in primary care.[30-32] However, some studies have specifically chosen not to include linked ONS data even when death was an important study outcome, in order to maximise patient numbers and increase external validity by including all four countries of the United Kingdom.[7] In this study, no major differences were found between the original random sample of patients and the subcohort eligible for linkage with both the ONS and the HES data, who were based in England only. In addition, previous studies have shown no difference between mortality rates in a broad CPRD GOLD cohort compared with a subset eligible for linkage.[33]

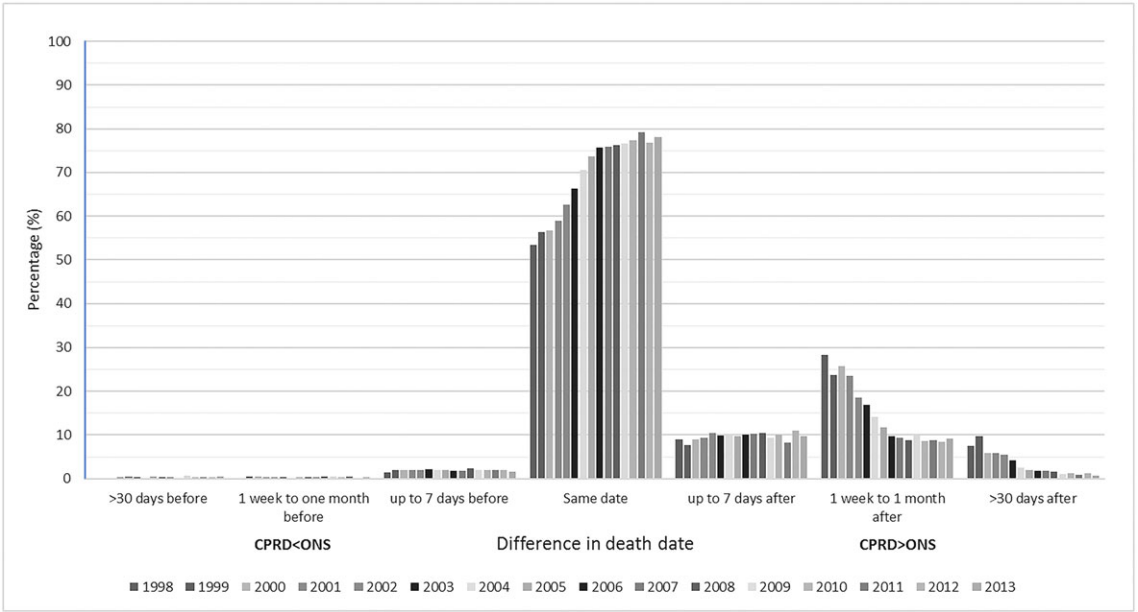

FIGURE 4 Difference between the Clinical Practice Research Datalink (CPRD) death date and the Office for National Statistics (ONS) death date [Colour figure can be viewed at wileyonlinelibrary.com]

The ONS death registration data are based on official information from the Medical Certificate of Cause of Death (MCCD) and are seen as the gold standard, as all deaths in the United Kingdom should be registered. Comparing this with the information available in CPRD GOLD provided the opportunity to see at the individual level how accurate the data at the GP practice are. However, there are some limitations to the ONS data. The ONS data include only deaths that occur in England and Wales and not those that occur in other parts of the United Kingdom or abroad. Our patient sample may have included patients who died elsewhere, which may explain the small number of deaths that did not appear in the ONS records. Delays in the registration of deaths can have an impact on the ONS mortality statistics.[34] Firstly, there can be delays in the issuing of the MCCD. In order to complete the MCCD, the certifying doctor must have seen the deceased in the previous 2 weeks of life. Delays can occur when there is a need for a post-mortem and/or inquest or when the death has been referred to a coroner, such as deaths with unknown cause; violent, unnatural, or suspicious deaths; those as a result of an accident, neglect, or suicide; or those that were employment related, occurred during an operation or when in custody.[35] Around 10 000 deaths a year in England and Wales are not registered for 6 months.[36] A very small number (0.3%) of deaths remain legally uncertified.[35] In addition, there is also a lag time between the issue of the MCCD and the record appearing in the ONS data. There is a 5-day legal limit on registering a death in England and Wales, which is usually done by the next of kin or a family member at a register office. However, there has been an increase in the number missing this deadline in recent years due to difficulties in getting an appointment.[37] Future studies could investigate the difference between the date of death and the date of registration of the death.

To our knowledge, this is the first study to directly compare the recording of the date of death in primary records with the gold standard from the ONS. Previous papers have relied upon comparing overall mortality rates,[38] probabilistic linkage,[39] or contacting GP practices for more information.[40]

This study showed that the CPRD GOLD algorithm provides comparable results to using the ONS death registration data for the date of death. For censoring follow-up and calculating mortality rates, the primary care data are likely to be sufficient as a delay in death recording of up to 1 month is unlikely to impact results significantly. Where the exact date of death or the cause is important, it may be advisable to use the individually linked national ONS death registration data. Researchers should take into consideration the impact on sample size and power of a study when deciding whether to use the linked data as this can restrict the cohort size substantially.

ETHICS STATEMENT

This study was approved by the Independent Scientific Advisory Committee (ISAC) for the Medicines and Healthcare products Regulatory Agency (MHRA) database research under 1_090V2Mn.

ACKNOWLEDGEMENTS

This paper was also reviewed by Puja Myles and William T. Hamilton. The authors received no specific funding for this work. The CPRD is owned by the UK Department of Health and operates within the Medicines and Healthcare Products Regulatory Agency (MHRA). Bert Leufkens receives no direct funding or donations from private parties, including the pharmaceutical industry. Research funding from public-private partnerships, ie, IMI and TI Pharma (www.tipharma.nl), has been accepted under the condition that no company-specific product or company-related study is conducted. He has received unrestricted research funding from public sources, ie, the Netherlands Organisation for Health Research and Development (ZonMW), the EU 7th Framework Program (FP7), the Dutch Medicines Evaluation Board (MEB), the National Health Care Institute (ZIN), and the Dutch Ministry of Health. Frank de Vries supervises a PhD student who is employed at F. Hoffmann-La Roche Ltd, Basel, Switzerland. He has not received any fees.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY

All data are available from the Clinical Practice Research Datalink for researchers who meet the criteria for access via the Independent Scientific Advisory Committee. URL for the Clinical Practice Research Datalink (CPRD): http://www.cprd.com/home/. Contact details for CPRD ISAC: ISAC Secretariat, Medicines & Healthcare products Regulatory Agency, CPRD Division, 151 Buckingham Palace Road, London, SW1W 9SZ, UK England. Electronic submission of protocols and emails to isac@cprd.com.

ORCID

Arlene M. Gallagher https://orcid.org/0000-0002-7507-4923

REFERENCES

1. Kontopantelis E, Stevens RJ, Helms PJ, Edwards D, Doran T, Ashcroft DM. Spatial distribution of clinical computer systems in primary care in England in 2016 and implications for primary care electronic medical record databases: a cross-sectional population study. *BMJ Open*. 2018 Feb 28;8(2):e020738.
2. Alatorre CI, Hoogwerf BJ, Deeg MA, et al. Factors associated with stroke, myocardial infarction, ischemic heart disease, unstable angina, or mortality in patients from real world clinical practice with newly-diagnosed type 2 diabetes and early glycemic control. *Curr Med Res Opin*. 2017 Oct;24:1-19.
3. Lane DA, Skjøth F, Lip GYH, Larsen TB, Kotecha D. Temporal trends in incidence, prevalence, and mortality of atrial fibrillation in primary care. *J Am Heart Assoc*. 2017 Apr 28;6(5):e005155.
4. Parisi R, Webb RT, Carr MJ, et al. Alcohol-related mortality in patients with psoriasis: a population-based cohort study. *JAMA Dermatol*. 2017 Sep 15;153(12):1256-1262.
5. Stewart D, Han L, Doran T, McCambridge J. Alcohol consumption and all-cause mortality: an analysis of general practice database records for patients with long-term conditions. *J Epidemiol Community Health*. 2017 Aug;71(8):729-735. https://doi.org/10.1136/jech-2017-209241. Epub 2017 Jul 5
6. Strongman H, Korhonen P, Williams R, et al. Pioglitazone and risk of mortality in patients with type 2 diabetes: results from a European multidatabase cohort study. *BMJ Open Diabetes Res Care*. 2017 May 29;5(1):e000364.
7. Zghebi SS, Steinke DT, Carr MJ, Rutter MK, Emsley RA, Ashcroft DM. Examining trends in type 2 diabetes incidence, prevalence and

mortality in the UK between 2004 and 2014. *Diabetes Obes Metab*. 2017 Apr 7;19(11):1537-1545.

8. Charlton J, Ravindrarajah R, Hamada S, Jackson SH, Gulliford MC. Trajectory of total cholesterol in the last years of life over age 80 years: cohort study of 99,758 participants. *J Gerontol A Biol Sci Med Sci*. 2018 Jul 9;73(8):1083-1089.
9. Harshfield A, Abel G, Barclay S, Payne R. Do GPs accurately record date of death? A UK observational analysis. *BMJ Support Palliat Care*. 2018 Jun 27;0:1-8.
10. Death Certification Advisory Group. Guidance for doctors completing Medical Certificates of Cause of Death in England and Wales. Office for National Statistics' Death Certification Advisory Group, Revised July 2010. Available: https://www.gro.gov.uk/Images/medcert_July_2010.pdf [Accessed 18 July 2017]
11. Wagstaff R, Berlin A, Stacy R, Spencer J, Bhopal RA. Information about patients' deaths: general practitioners' current practice and views on receiving a death register. *Br J Gen Pract*. 1994 Jul;44(384):315-316.
12. Maguire A, Blak BT, Thompson M. The importance of defining periods of complete mortality reporting for research using automated data from primary care. *Pharmacoepidemiol Drug Saf*. 2009 Jan;18(1):76-83.
13. Williams T, van Staa T, Puri S, Eaton S. Recent advances in the utility and use of the General Practice Research Database as an example of a UK primary care data resource. *Ther Adv Drug Saf*. 2012;3(2):89-99.
14. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. 2015 Jun 6;44(3):827-836.
15. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol*. 2010 Jan;69(1):4-14.
16. Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. *Br J Gen Pract*. 2010;60(572):e128-e136.
17. NHS Digital. Mortality data from the Office for National Statistics. Available: http://content.digital.nhs.uk/onsmortality [Accessed 23 December 2016].
18. NHS Digital. Hospital Episode Statistics. Available: http://content.digital.nhs.uk/hes [Accessed 23 December 2016].
19. Gallagher AM, Williams T, Leufkens HGM, de Vries F. The impact of the choice of data source in record linkage studies estimating mortality in venous thromboembolism. *PLOS ONE*. 2016;11(2):e0148349.
20. Eurostat. Revision of the European Standard Population: Report of Eurostat's task force, 2013 Available: http://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/KS-RA-13-028 [Accessed 15 February 2018]
21. Public Health England. Recent trends in place of death in England. Available: http://www.endoflifecare-intelligence.org.uk/view?rid=893 [Accessed 7th March 2017]
22. National End of Life Care Intelligence Network. Number and proportion of deaths by place of occurrence. Available: http://www.endoflifecare-intelligence.org.uk/data_sources/place_of_death [Accessed 20 February 2018]
23. BMA. Confirmation and certification of death. Available: https://www.bma.org.uk/advice/employment/gp-practices/service-provision/confirmation-and-certification-of-death. [Accessed 20 February 2018]
24. NHS Digital. NHS payments to general practice: England, 2015/16.21 September 2016. Available: http://digital.nhs.uk/media/29211/NHS-Payments-to-General-Practice-England-2015-16-Report/Any/nhspaymentsgp-15-16-rep [Accessed 20 February 2018]
25. NHS Digital. General practice trends in the UK to 2016.19 September 2017. Available: http://content.digital.nhs.uk/media/25201/General-Practice-Trends-in-the-UK-to-2016/pdf/General_Practice_Trends_in_the_UK_to_2016.pdf [Accessed 20 February 2018]
26. Hollingworth W, Biswas M, Maishman RL, et al. The healthcare costs of heart failure during the last five years of life: a retrospective cohort study. *Int J Cardiol*. 2016 Dec 1;224:132-138.
27. Thomas KH, Davies N, Metcalfe C, Windmeijer F, Martin RM, Gunnell D. Validation of suicide and self-harm records in the Clinical Practice Research Datalink. *Br J Clin Pharmacol*. 2013 Jul;76(1):145-157.
28. Suissa S. Immeasurable time bias in observational studies of drug effects on mortality. *Am J Epidemiol*. 2008 Aug 1;168(3):329-335.
29. Spaetgens B, de Vries F, Driessen JHM, et al. Risk of infections in patients with gout: a population-based cohort study. *Sci Rep*. 2017 May 3;7(1):1429.
30. Wing K, Bhaskaran K, Smeeth L, et al. Optimising case detection within UK electronic health records: use of multiple linked databases for detecting liver injury. *BMJ Open*. 2016 Sep 2;6(9):e012102.
31. Glover G, Williams R, Heslop P, Oyinlola J, Grey J. Mortality in people with intellectual disabilities in England. *J Intellect Disabil Res*. 2017 Jan;61(1):62-74.
32. Ratib S, Fleming KM, Crooks CJ, Walker AJ, West J. Causes of death in people with liver cirrhosis in England compared with the general population: a population-based cohort study. *Am J Gastroenterol*. 2015 Aug;110(8):1149-1158.
33. Springate DA, Ashcroft DM, Kontopantelis E, Doran T, Ryan R, Reeves D. Can analyses of electronic patient records be independently and externally validated? Study 2—the effect of β-adrenoceptor blocker therapy on cancer survival: retrospective cohort study. *BMJ Open*. 2015 Apr 13;5(4):e007299.
34. Devis T, Rooney C. The time taken to register a death. *Population Trends*. 1997 Summer;88:48-55.
35. Patel V. User guide to mortality statistics, July 2017. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/userguidetomortalitystatisticsjuly2017 [Accessed 15 February 2018]
36. Hawkes N. Why the delays in counting the dead? *BMJ*. 2014 Jul 9;349:g4305.
37. BBC. Death-registering delays 'rise by 70%'. Available: http://www.bbc.co.uk/news/uk-england-41080306. [Accessed 9 February 2018]
38. Ohlmeier C, Langner I, Garbe E, Riedel O. Validating mortality in the German Pharmacoepidemiological Research Database (GePaRD) against a mortality registry. *Pharmacoepidemiol Drug Saf*. 2016 Jul;25(7):778-784.
39. Sieswerda E, Font-Gonzalez A, Dijkgraaf MG, Geskus RB, Heinen RC, van der Pal HJ, van Leeuwen FE, Caron HN, Kremer LC, Reitsma JB. Studying hospitalizations and mortality in the Netherlands: feasible and valid using two-step medical record linkage with nationwide registers. *PLoS One*. 2015 Jul 6;10(7):e0132444;1-13.
40. Hall GC. Validation of death and suicide recording on the THIN UK primary care database. *Pharmacoepidemiol Drug Saf*. 2009 Feb;18(2):120-131.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Gallagher AM, Dedman D, Padmanabhan S, Leufkens HGM, de Vries F. The accuracy of date of death recording in the Clinical Practice Research Datalink GOLD database in England compared with the Office for National Statistics death registrations. *Pharmacoepidemiol Drug Saf*. 2019;28:563–569. https://doi.org/10.1002/pds.4747