Validation of suicide and self-harm records in the Clinical Practice Research Datalink

Kyla H. Thomas, Neil Davies, Chris Metcalfe, Frank Windmeijer, Richard M. Martin & David Gunnell

School of Social and Community Medicine, University of Bristol, Bristol, UK

Correspondence

Dr Kyla Thomas MBBS, MSc, MSc (Research), MFPH, School of Social and Community Medicine, University of Bristol, 39 Whatley Road, Bristol BS8 2PS, UK.
Tel.: +44 1179 287200
Fax: +44 0117 928 7325
E-mail: kyla.thomas@bristol.ac.uk

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT

- The Clinical Practice Research Datalink (CPRD), formerly known as the General Practice Research Database (GPRD), is being increasingly used to investigate suicide-related adverse drug reactions, although no previous studies have comprehensively validated the recording of suicide and nonfatal self-harm in the CPRD. Recent linkages of CPRD General Practices with Office for National Statistics mortality data and Hospital Episode Statistics data provide new opportunities for validation.

WHAT THIS STUDY ADDS

- Use of diagnostic codes ('Read codes') recorded by general practitioners for suicide identification has low sensitivity and positive predictive value compared with Office for National Statistics-confirmed suicides obtained by record linkage (the gold standard).
- Approximately 31.6% of hospital-admitted cases of self-harm are not recorded in the CPRD. Compared with estimates derived from registers of hospital attendance for self-harm, CPRD underestimates the incidence of self-harm by approximately 54.5%. The lack of a gold standard for nonfatal self-harm makes it difficult to validate this outcome fully.

AIMS

The UK Clinical Practice Research Datalink (CPRD) is increasingly being used to investigate suicide-related adverse drug reactions. No studies have comprehensively validated the recording of suicide and nonfatal self-harm in the CPRD. We validated general practitioners' recording of these outcomes using linked Office for National Statistics (ONS) mortality and Hospital Episode Statistics (HES) admission data.

METHODS

We identified cases of suicide and self-harm recorded using appropriate Read codes in the CPRD between 1998 and 2010 in patients aged ≥15 years. Suicides were defined as patients with Read codes for suicide recorded within 95 days of their death. International Classification of Diseases codes were used to identify suicides/hospital admissions for self-harm in the linked ONS and HES data sets. We compared CPRD-derived cases/incidence of suicide and self-harm with those identified from linked ONS mortality and HES data, national suicide incidence rates and published self-harm incidence data.

RESULTS

Only 26.1% ($n = 590$) of the 'true' (ONS-confirmed) suicides were identified using Read codes. Furthermore, only 55.5% of Read code-identified suicides were confirmed as suicide by the ONS data. Of the HES-identified cases of self-harm, 68.4% were identified in the CPRD using Read codes. The CPRD self-harm rates based on Read codes had similar age and sex distributions to rates observed in self-harm hospital registers, although rates were underestimated in all age groups.

CONCLUSIONS

The CPRD recording of suicide using Read codes is unreliable, with significant inaccuracy (over- and under-reporting). Future CPRD suicide studies should use linked ONS mortality data. The under-reporting of self-harm appears to be less marked.

Introduction

Large primary care databases, such as the Clinical Practice Research Datalink [CPRD (http://www.cprd.com); formerly known as the General Practice Research Database (GPRD)] and The Health Improvement Network (THIN) database (http://www.thin-uk.com), are being increasingly used in epidemiological research [1]. These databases provide excellent opportunities for investigating the incidence and prevalence of diseases and for conducting pharmacoepidemiology, pharmacovigilance and health services research, because over 98% of people in the UK are registered with a general practitioner (GP), and almost all GPs use computerized records [2].

The value of primary care databases comes from their size (power) and ability to investigate drug use in the wider population outside of the often tightly controlled clinical environments of randomized controlled trials [3]. Whilst randomized controlled trials and meta-analyses are the preferred methods for evaluating drug therapies, they are usually only large enough to detect primary outcomes and common side effects. Rare end-points, such as suicide, are difficult to investigate using experimental studies, because the sample size requirements for such outcomes are prohibitively large [4]. In the last 15 years, several observational studies have used primary care databases to investigate drug safety issues for rare disease outcomes [5–10].

There are well-known limitations of using primary care databases for research. Given that the data are primarily collected for use in routine clinical practice, they may not be sufficiently accurate for research purposes. Lawrenson et al. [2], in their review of the use of general practice databases, described data accuracy in terms of validity (i.e. what is the likelihood that a patient with a diagnosis recorded on the database genuinely has that particular condition) and completeness or sensitivity (i.e. if a patient has been diagnosed as having diabetes, what is the likelihood that it would be recorded on the database). The recording of chronic conditions, such as diabetes, has a high sensitivity (>90%), because chronic illnesses require regular follow-up [2]. With the introduction of the Quality and Outcomes Framework in 2004, it is likely that certain chronic conditions would be even better recorded; for instance, the recording of HbA_{1c} for diabetes was 13% higher in 2005 than 2004 [11, 12]. However, recording is likely to be lower for self-harm (which is not included in the Quality and Outcomes Framework) and major one-off events, such as suicide.

Self-harm is one of the commonest reasons for attendance at hospital emergency departments [13]. Approximately 50% of patients who self-harm consult their GPs in the month after the episode [14]. However, one study found that in almost half of cases involving a mental health specialist, there was no communication with the individual's GP [15]. This may result in significant under-recording of self-harm on primary care databases if such incidents are not also routinely captured by GP information systems. Furthermore, cause of death may not be confirmed for many months after a death has occurred, because the average time taken to complete a coroner's inquest in England and Wales is over 6 months [16]. If GPs are not notified of the outcome of the inquest, cause of death may not be entered onto the electronic medical records. If under-reporting leads to differential misclassification of outcomes in the CPRD, then biased estimates of associations of suicide and self-harm and exposures may be obtained.

To identify health outcomes of interest, CPRD researchers use algorithms based on combinations of specific codes used to record diagnoses, medications, referrals and hospital attendance. Although such approaches for detecting suicides have been validated using the THIN database, the studies were based on subsets of patients [17, 18]. The last review of the quality of suicide reporting in the CPRD was carried out over 15 years ago [19, 20]; nonfatal self-harm has not been previously assessed in CPRD or THIN.

The electronic medical records of some English CPRD practices have recently been linked with Office for National Statistics (ONS) mortality data and Hospital Episode Statistics (HES). The ONS data (the gold standard) record suicides, while the HES data provide an opportunity to validate the recording of nonfatal self-harm in the CPRD, albeit limited to hospital-admitted cases.

The aim of this paper is to validate the recording of self-harm and suicide using Read code algorithms in the CPRD. Specifically, we assess the reliability of these algorithms for the identification of cases of suicides and self-harm and compare our findings with the following: (i) national incidence rates; (ii) information obtained from recent linkages to ONS mortality data and HES; and (iii) free text searches of GP consultation records.

Methods

Ethics approval

Ethics approval for this study was obtained from the MHRA's Independent Scientific Advisory Committee.

Source data and population

The CPRD contains electronic medical records from approximately 5.1 million active patients, representing about 8.3% of the UK population (based on the April 2011 release). Approximately 50% of CPRD general practices in England were recently linked with ONS mortality and HES data.

The CPRD contains consultation, prescribing, referral and health outcome information for individual patients; this is entered by GPs and their staff onto computer systems [21]. These records are the primary administrative medical records used by the practices. Participating general practices use a computerized system called

Vision, which includes built-in data-collection software. This software extracts data from practice computers, excludes personal data, such as names and addresses, and assigns anonymized patient and prescriber identification numbers. During a consultation, GPs can enter descriptions of patients' symptoms and diagnoses using either Read codes or written text. General practitioners select Read codes to indicate diagnoses or symptoms using automated drop-down lists of codes. Each Read code is linked to a specific phrase of text, which indicates a single diagnosis or symptom.

Study population

Patients were eligible for inclusion in this study if they were aged 15 years and over and were enrolled in a practice at any time from 1 January 1998 to 31 December 2010. This time period was chosen to coincide with the availability of HES and ONS linkage for English practices within the CPRD. To maximize quality and completeness, we restricted our analyses to data from the following sources.

1 Patient records that had been classified as 'acceptable' by the CPRD; this means that certain quality criteria had been met, for example there were no breaks in the patient records and year of birth, and the patient's sex and first registration date were recorded.
2 Practices that had been designated as 'up to standard' by the CPRD, i.e. the practice provided continuity in data recording, with exclusion of practices which did not sufficiently record whether patients had died or been transferred out of the practice.

Other data sources

Three other sources of data were used in this study, as follows.

1 ONS data. The ONS produces annual reports on specific causes of death, including suicides in the UK [22]. The ONS mortality data, including date of death and causes of death, have been linked to approximately half of all English practices in the CPRD.
2 HES. This is a secure data warehouse that contains details of all admissions to NHS hospitals in England only (http://www.hesonline.nhs.uk). Integrated HES data for 50% of CPRD practices are provided free of cost to CPRD users.
3 Multicentre Study of Self-Harm. This study records all episodes of self-harm (including those not admitted to a hospital bed) presenting to hospital emergency departments in three centres, namely Oxford, Manchester and Derby [23].

Case identification

Cases of suicide and self-harm (the 'events') were identified by extracting all records with Read codes for suicide,

attempted suicide and self-harm (see Appendices 1 and 2). Given that suicide-related Read codes may refer to both fatal and nonfatal suicide attempts, completed suicides within the CPRD were identified using the conventional CPRD approach of linking patient deaths to Read codes for suicide that were recorded within 95 days of the CPRD-derived death dates (CPRD personal communication and Appendix 2).

Validation of cases

We assessed the reliability of our Read code algorithms (see Appendix 1) to identify cases of suicide and nonfatal self-harm using two approaches:

1 We estimated age- and sex-specific incidence of suicide and nonfatal self-harm in the CPRD using events defined by Read codes from 1998 to 2010. We defined incidence as the patient's first ever event. We then compared these incidence rates to rates given in the ONS mortality statistics for the UK (for suicide) and self-harm registry data described in the Multicentre Study of Self-Harm (for nonfatal self-harm) [23, 24].
2 We compared the number of suicides and episodes of self-harm identified using algorithms based on Read codes with those retrieved using linked ONS mortality and HES data in the approximately 50% of general practices with ONS–HES linkage (England only). Cases of suicide (source, ONS mortality) and nonfatal self-harm (source, HES data) were identified by the following International Classification of Disease (ICD) codes; ICD 10: X60–X84 (intentional self-harm, which includes purposely self-inflicted poisoning or injury and attempted suicide) and Y10–Y34 (event of undetermined intent), excluding Y33.9 where verdict was still pending [25]; ICD 9: E950–E959 (recorded suicides and self-inflicted injury) and E980–E989 (injury undetermined, whether accidentally or purposely inflicted), excluding E988.8. International Classification of Disease codes for undetermined deaths were included, because most of these deaths are probable suicides [26].

We calculated the sensitivity and positive predictive values (PPV) of our Read code algorithms for detecting suicide. We performed a sensitivity analysis for the CPRD-detected suicides using death dates within 30, 180 and 360 days of the suicide/self-harm Read code date. We could not use self-harm (as defined in HES) as the gold standard for self-harm, because only half of self-harm episodes that present to hospital are admitted [27, 28]; therefore, we identified the percentage of those patients with self-harm records in HES who also had records indicating self-harm in the CPRD within 6 months of hospital admission.

Free text searches

To assess whether free text searches of GP consultation records might improve the identification of cases of

suicide and self-harm that were not Read coded as such, we carried out free text searches of CPRD patient records for missed cases of HES-identified self-harm in 2010 and all missed cases of suicide from January 1998 to December 2010. We used the following search terms: suicid* (including suicide, suicidal, suicidality), overdose, depress* (including depression, depressed), self harm*(including self harming, self harmed), self injur* (including self injury, self injurious) and self poison* (including self poisoning).

Statistical analyses

We performed all analyses in Stata version 12 (StataCorp LP, College Station, TX, USA). Rates were age standardized using the European Standard Population [29]. Incidence rates were calculated using mid-year population denominators from the ONS (http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk–england-and-wales–scotland-and-northern-ireland/population-estimates-timeseries-1971-to-current-year/index.html, accessed 7 February 2012).

Results

Suicide

UK We identified 1214 male suicides and 553 female suicides using Read code algorithms in all eligible CPRD practices. The age-standardized suicide rates were 5.5 per 100 000 for males and 2.2 per 100 000 for females, whereas ONS suicide rates over the same time period were 18.5 per 100 000 for men and 5.7 per 100 000 for women. Figure 1 shows the trends in age-standardized, sex-specific suicide rates in the UK from 1998 to 2010 in those aged 15 years and over based on: (i) national ONS data; and (ii) using CPRD Read code algorithms. Suicide rates were underestimated using Read code algorithms for both men and women. Figure 2 shows the age- and sex-specific suicide rates based on CPRD Read code-defined suicides in 1998–2010 compared with national rates. Suicide rates were consistently higher in men than women using ONS and CPRD data; however, the age distribution of suicide rates using CPRD Read code algorithms was different from that observed using ONS data. Rates were highest in males and females aged over 75 years in the CPRD, whereas in the ONS men aged 15–44 years and women aged 45–74 years had the highest suicide rates.

England (ONS-linked CPRD data) There were 2260 ONS-confirmed suicides between 1998 and 2010 in the 50% of English practices with linked mortality data; 1728 (76.5%) in males and 532 (23.5%) in females. We identified 1063 suicides using Read code algorithms in these linked practices; 590 of these [55.5%, 95% confidence interval (CI): 52.5–58.5%] were true, ONS-confirmed suicides (PPV). These 590 cases represented 26.1% (95% CI: 24.3–28.0%) of all the 2260 ONS-confirmed suicides (sensitivity). In a sensitivity analysis, when we increased the time period between the clinical event date and the CPRD-derived death date from 30 to 360 days, the sensitivity increased from 25.0 (95% CI: 23.2–26.8%) to 35.5 (95% CI: 33.5–37.5%), but the PPV decreased from 63.2 (95% CI: 59.9–66.3%) to 46.0 (95% CI: 43.6–48.3%).

Table 1 shows age-, era- (before/after 2004) and sex-specific sensitivities and PPVs for Read code-identified suicides in the CPRD. There was little difference between the sensitivity and PPV by sex or era, although suicides amongst males had a lower sensitivity and higher PPV than females (the difference in sensitivity for males and females was consistent with chance $p = 0.210$). The PPV and sensitivity were somewhat better in more recent years. The Read code algorithms were most sensitive for 45–74 years old,
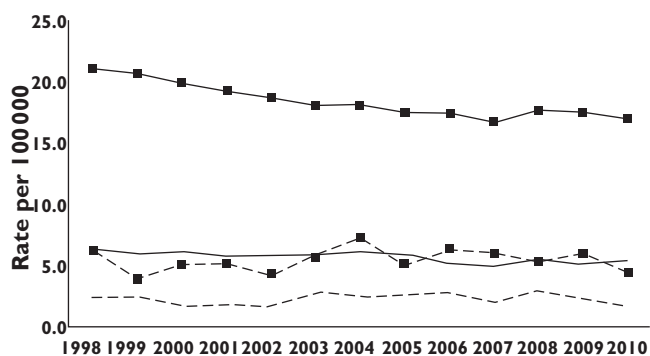

Figure 1

Trends in sex-specific age-standardized suicide rates per 100 000 in the UK from 1998 to 2010 in ages 15 years and over using data derived from: official Office for National Statistics (ONS) mortality statistics for the UK (A) and Clinical Practice Research Datalink (CPRD) Read code algorithms (B). –■–, ONS males; ——, ONS females; –■-, CPRD males; – – –, CPRD females
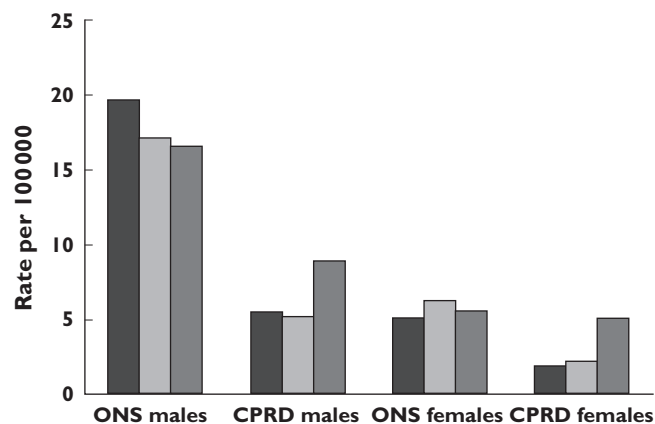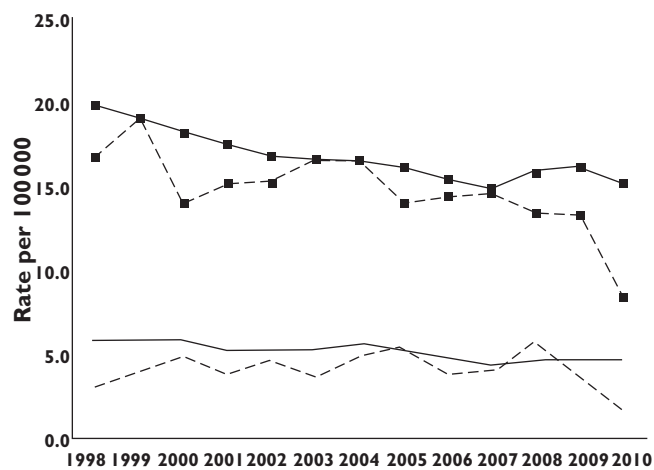

Figure 2

Comparison of UK ONS suicide incidence rates with Read code-identified CPRD suicide rates per 1 000 000 between 1998 and 2010. ■, 15–44 years; ▦, 45–74 years; ■, 75+ years

Table 1

Sensitivity and positive predictive values (95% confidence intervals) for CPRD Read code algorithms to detect suicides compared with ONS-confirmed suicides in the CPRD–ONS linked practices

	Sensitivity (%)	95% Confidence interval	PPV (%)	95% Confidence interval
Sex				
Male	25.5	23.4–27.6	59.7	56.1–63.3
Female	28.2	24.4–32.2	46.0	40.5–51.6
Age (years)				
15–44	25.1	22.5–27.7	61.7	57.1–66.3
45–74	28.7	25.8–31.7	61.6	56.9–66.2
75+	20.4	15.4–26.3	25.4	19.3–32.4
Years				
1998–2003	22.9	20.3–25.6	52.7	47.8–57.5
2004–2010	28.6	26.1–31.1	57.4	53.4–61.2

Abbreviations are as follows: CPRD, Clinical Practice Research Datalink; and ONS, Office for National Statistics.
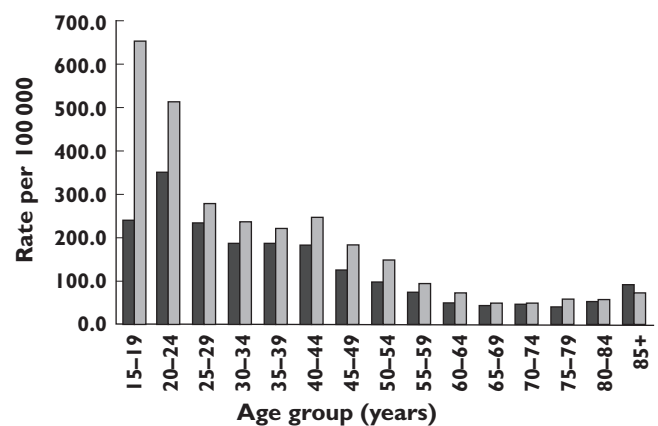

Figure 3

Trends in sex-specific age-standardized suicide rates per 1 000 000 in England from 1998 to 2010 in ages 15 years and over using data derived from: official ONS mortality statistics for England (A) and data from linked CPRD–ONS practices in England (B). –■–, ONS males; ——, ONS females; – ■ –, ONS linked males; – – –, ONS linked females

but a considerably lower PPV was obtained for those aged over 75 years (25.4%, 95% CI: 19.3–32.4%).

Figure 3 shows the trends in age-standardized suicide rates in England as a whole, as well as for ONS-linked English practices in the CPRD, based on all suicides identified through the ONS linkage. Excluding 2010, suicide rates in ONS-linked practices followed a similar pattern to that for England as a whole, although rates in ONS-linked practices tended to be slightly lower.

Self-harm

UK There were 30 449 episodes of male self-harm and 43 787 episodes of female self-harm identified using CPRD Read codes between 1998 and 2010. Figure 4 shows the self-harm rates per 100 000 by sex for the year 2007 based

Figure 4

Self-harm rates per 1 000 000 by age and sex for the year 2007 in UK CPRD practices using Read codes to identify self-harm cases. ■, males; □, females

on Read code algorithms. Female self-harm rates were consistently higher than male rates for all age groups, with the exception of those aged 85 years and older. The highest self-harm rates were seen in females aged 15–19 years and males aged 20–24 years. The age and sex distributions were similar to those seen in the HES data and the combined data from the three hospital registers in the Multicentre Study of Self Harm (Table 2) [24].

However, Read code algorithm-defined self-harm incidence rates were lower than those derived from self-harm hospital registry data, e.g. male rates at all ages were 148.6 per 100 000 *vs.* 317 per 100 000 based on the Multicentre Study register data. The lower rates derived from HES (see Table 2) reflect the fact that these are based on hospital admissions, whereas the Multicentre Study register data record all hospital-presenting cases of self-harm, regardless of whether or not they led to admission. Figure 5 compares the age- and sex-specific incidence of self-harm

Table 2

Comparisons of CPRD Read code algorithm-defined self-harm rates per 100 000 for the year 2007 with rates derived from HES admission data and emergency department attendances from the Multicentre Study of Self-Harm [23]

2007 Age group (years)	CPRD	HES	Register
Males			
15–24	292.2	167.8	391.7
25–34	208.4	154.3	356.8
35–44	183.5	151.5	417.9
45–54	113.1	94.9	367.4
55+	53.0	42.9	96.8
All ages	148.6	108.7	317.0
Females			
15–24	587.5	381.6	851.6
25–34	256.4	186.9	524.1
35–44	236.9	199.2	585.7
45–54	164.4	158.3	423.5
55+	65.6	44.7	90.6
All ages	213.1	158.3	481.1

Abbreviations are as follows: CPRD, Clinical Practice Research Datalink; and HES, Hospital Episode Statistics.
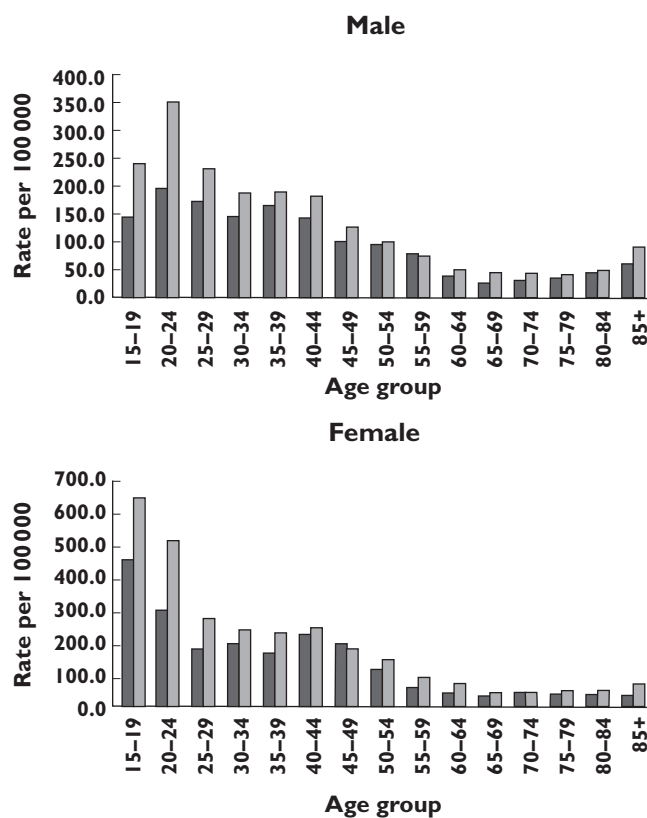

Figure 5

Comparison of Hospital Episode Statistics (HES) self-harm incidence rates with Read code-identified CPRD self-harm incidence rates per 100 000 for 2007 in the HES-linked English practices in the CPRD. ■, HES; ▨, CPRD

derived from HES with that based on Read code algorithms in 2007. The rates based on Read code algorithms were approximately twice the HES rates for men aged 15–19 and 20–24 years.

England (HES-linked CPRD data) Approximately 68.4% of patients had a self-harm Read code in the CPRD within 6 months of a relevant HES self-harm ICD code, indicating that around one-third of hospital-admitted cases of self-harm are not recorded by their GP using these Read codes on the CPRD. Patients recorded on both HES and CPRD were of a similar age to those who did not have a CPRD Read code after hospital admission, although they were more likely to be female (62.2% female compared with 52.7% male).

Free text searches

Read code records failed to identify 1670 out of the 2260 ONS-confirmed suicides in 1998–2010. Free text searches of these records identified an additional 179 cases of suicides, amounting to 10.7% of the missed cases.

There were 622 patients who were admitted to hospital with self-harm in 2010 who did not have a Read-coded CPRD record of self-harm within 6 months of hospital admission. One hundred and one (16.2%) of these patients would have been identified as cases of nonfatal self-harm from searching the free text records.

Discussion

Main findings

We validated the reporting of suicide and self-harm based on Read code algorithms using various data sources. We found fewer suicides than expected using Read code algorithms for all UK CPRD practices. In the 50% of English practices linked to the ONS data, Read code algorithms had low sensitivity (26.1%; 95% CI: 24.3–28.0%) and PPV (55.5%; 95% CI: 52.5–58.5%) compared with the gold standard ONS data, and underestimated suicide rates in both sexes for all age groups. Sensitivity and PPV were lowest in those aged 75 years and over. This is likely to be a direct result of the methodology (i.e. suicide-related Read codes associated with death in a certain time period) used to identify suicide prior to the availability of linked ICD-coded mortality data, because older people are less likely to self-harm and more likely to die from other (nonsuicide) causes of death compared with other age groups. This reduces sensitivity (because fewer suicides are preceded by self-harm in the elderly) and increases the number of false-positive 'suicides' (reduces the PPV). The PPV for Read code-identified suicides was higher in males than females; this is most probably because the risk of suicide amongst those people who self-harm is considerable higher in males than females [30]. Sensitivity increased with increasing number of days from the clinical event date to the

CPRD-derived death date; this may be related to the time required for the completion of coroners' inquests for suicides and hence delays in GPs being notified of patient deaths.

Suicide rates in English ONS-linked practices were comparable to ONS rates for England for most years except 2010. This discrepancy in 2010 may be due to delays in receiving coroners' reports, because our CPRD data set was constructed in May 2011, and 2010 suicide rates were published by the ONS only in early 2012 (http://www.ons.gov.uk/ons/rel/subnational-health4/suicides-in-the-united-kingdom/2010/stb-statistical-bulletin.html, accessed 7 February 2012). Approximately two-thirds of hospital admissions for self-harm had a self-harm Read code recorded within 6 months of their admission. Although self-harm rates based on Read codes followed a similar age and sex distribution to those recorded in hospital registries of patients presenting with self-harm, the Read code-defined events underestimated the rates in the majority of age groups when self-harm registry data were used for comparison. The CPRD self-harm rates were approximately twice the HES self-harm rates for men aged 15–19 and 20–24 years. This may be due to the better recording of self-harm in the CPRD for these age groups or because when they attend hospital for self-harm they are less likely to be admitted, because self-harm in these groups may be considered a weaker predictor of suicide risk. The risk of suicide increases with an individual's age at the time of self-harm [30]. Searches of free text records in the CPRD identified 10.7% of the suicides missed by Read code searches and 16.1% of the missed cases of nonfatal self-harm.

Strengths and limitations

This is the first study to validate self-harm recording in the entire CPRD and the largest and most comprehensive appraisal of suicide validation to date. Although suicide recording has been validated before in the Value Added Medical Products (VAMP) health resource, the precursor to the CPRD, it was based on a cohort of patients prescribed antidepressants and was done over 15 years ago [20].

Although best practice for identification of fatal outcomes in the CPRD involves the follow-up of all deaths (for example, the examination of free text records for all deaths as well as death certificate data for all deaths where cause of death was unknown), such an approach is extremely expensive. Therefore, we took a pragmatic approach and validated the methods that have been used in practice in recent research literature to identify suicides and self-harm outcomes in the CPRD [7, 31, 32].

We had no gold standard for 'true' cases of nonfatal self-harm, so we could not directly compare the self-harm rates based on the Read code algorithms with the HES data; the latter cover admitted patients who comprise only 50% of hospital-presenting cases of self-harm [28]. The similarities of the age and sex distributions of the CPRD-identified nonfatal self-harm cases to those seen in HES and the Multicentre Study of Self-Harm register data provide some reassurance that while CPRD under-records self-harm, there may not be any age or sex bias in the under-recording. Although data were available from hospital registries of patients presenting with self-harm in Oxford, Manchester and Derby, these cities may not be representative of the entire UK.

Comparison with previous studies

A variety of methods have been previously used to validate diagnoses within the CPRD [33]. These methods were either internal, e.g. using free text (where GPs record uncoded comments related to the consultation) or sensitivity analyses using different diagnostic algorithms, or external, such as GP questionnaires, provision of copies of anonymized clinical paper records and comparison of incidence rates. However, most validations are compromised by the low practice participation rates, which limit the generalizability of the findings. The quality of reporting of the validations also differed among studies, and many papers did not provide lists of the Read code algorithms that were used.

We were surprised by the extent of under-reporting for suicide and self-harm in the CPRD, which was in contrast to the almost 100% accuracy reported in the VAMP database [20], but more in keeping with recent estimates described by Hall [17], where only one case out of seven confirmed suicides was recorded as such in THIN. Another validation of suicide recording in THIN in a cohort of patients with epilepsy found more promising results, with a high PPV of 88% [18]; however, the authors compared Read codes with GP questionnaires; findings were not reported for comparisons with death certificates, which had also been requested. Also, many intentional self-harm Read codes were omitted from their algorithms; no clear rationale was provided for their omission.

Conclusions

We found that the use of conventional Read code algorithms in the CPRD to detect cases of suicide misses approximately three-quarters of suicides and generates a high number of false positives (PPV 55.5%). Free text searching detects only around 10% of missed cases. Whilst linked data are available for only around half of all CPRD practices, more suicides are detected in this subset of CPRD, because the sensitivity is 100% instead of the 26.1% obtained when Read codes are used to identify cases. This more than compensates for the reduction in the total patient population available from the limited number of practices that consented to linkage. Therefore, future studies of suicide in the CPRD should use linked ONS mortality data. The under-reporting of self-harm in the CPRD appears to be less marked than that of suicide (about 50% when compared with hospital registry data of hospital attendances for self-harm and 32% for hospital-admitted

cases; see Table 2); however, our assessment of self-harm was less comprehensive, because there was no appropriate gold standard with a record of all incident episodes of self-harm. Free text searching detected only 16% of the missed cases.

It may be useful to study HES self-harm outcomes as well as Read code algorithm-identified self-harm. Advantages of using HES self-harm data include greater accuracy and the fact that hospital admissions are likely to be more severe cases of self-harm. The HES self-harm may therefore be a more relevant outcome for researchers who are interested in studying nonfatal self-harm with high suicidal intent, because such cases may be more likely to require hospital admission. However, only 50% of English CPRD practices have consented to linkage, so studies which use HES self-harm as their primary outcome may have reduced power compared with those based on Read code-identified cases. Furthermore, hospital-admitted cases of self-harm may be different in relation to methods used, degree of suicidal intent and subsequent management.

The creation of the CPRD from the GPRD in March 2012 will be hugely beneficial to pharmacoepidemiological research, because there will be greater coverage of the UK population and an excellent opportunity for increasing linkage to other databases, thus improving the accuracy of determining certain clinical outcomes. This study highlights the potential benefits of further data linkage. Persuading more practices to consent to linkage may provide considerable rewards, although these data will also require validation.

Competing Interests

Appendix 1 List of Read codes used to identify suicides and nonfatal self-harm in the CPRD

Read code	Description
SL..14	Overdose of biological substance
SL..15	Overdose of drug
SLHz.00	Drug and medicament poisoning not otherwise specified
TK..00	Suicide and self-inflicted injury
TK..11	Cause of overdose – deliberate
TK..12	Injury – self-inflicted
TK..13	Poisoning – self-inflicted
TK..14	Suicide and self-harm
TK..15	Attempted suicide
TK..17	Para-suicide
TK0.00	Suicide + self-inflicted poisoning by solid/liquid substances
TK00.00	Suicide + self-inflicted poisoning by analgesic/antipyretic
TK01.00	Suicide + self-inflicted poisoning by barbiturates
TK01000	Suicide and self-inflicted injury by amylobarbitone

Appendix 1
Continued

Read code	Description
TK01100	Suicide and self-inflicted injury by barbitone
TK01400	Suicide and self-inflicted injury by phenobarbitone
TK02.00	Suicide + self-inflicted poisoning by other sedatives/hypnotics
TK03.00	Suicide + self-inflicted poisoning tranquillizer/psychotropic
TK04.00	Suicide + self-inflicted poisoning by other drugs/medicines
TK05.00	Suicide + self-inflicted poisoning by drug or medicine not otherwise specified
TK06.00	Suicide + self-inflicted poisoning by agricultural chemical
TK07.00	Suicide + self-inflicted poisoning by corrosive/caustic substance
TK0z.00	Suicide + self-inflicted poisoning by solid/liquid substance not otherwise specified
TK1.00	Suicide + self-inflicted poisoning by gases in domestic use
TK10.00	Suicide + self-inflicted poisoning by gas via pipeline
TK11.00	Suicide + self-inflicted poisoning by liquified petrol gas
TK1y.00	Suicide and self-inflicted poisoning by other utility gas
TK1z.00	Suicide + self-inflicted poisoning by domestic gases not otherwise specified
TK2.00	Suicide + self-inflicted poisoning by other gases and vapours
TK20.00	Suicide + self-inflicted poisoning by motor vehicle exhaust gas
TK21.00	Suicide and self-inflicted poisoning by other carbon monoxide
TK2z.00	Suicide + self-inflicted poisoning by gases and vapours not otherwise specified
TK3.00	Suicide + self-inflicted injury by hang/strangulate/suffocate
TK30.00	Suicide and self-inflicted injury by hanging
TK31.00	Suicide + self-inflicted injury by suffocation by plastic bag
TK3y.00	Suicide + self-inflicted injury by other means than hang/strangle/suffocate
TK3z.00	Suicide + self-inflicted injury by hang/strangle/suffocate not otherwise specified
TK4.00	Suicide and self-inflicted injury by drowning
TK5.00	Suicide and self-inflicted injury by firearms and explosives
TK51.00	Suicide and self-inflicted injury by shotgun
TK52.00	Suicide and self-inflicted injury by hunting rifle
TK54.00	Suicide and self-inflicted injury by other firearm
TK5z.00	Suicide and self-inflicted injury by firearms/explosives not otherwise specified
TK6.00	Suicide and self-inflicted injury by cutting and stabbing
TK60.00	Suicide and self-inflicted injury by cutting
TK60100	Self-inflicted lacerations to wrist
TK60111	Slashed wrists self-inflicted
TK61.00	Suicide and self-inflicted injury by stabbing
TK6z.00	Suicide and self-inflicted injury by cutting and stabbing not otherwise specified
TK7.00	Suicide and self-inflicted injury by jumping from high place
TK70.00	Suicide + self-inflicted injury – jump from residential premises
TK71.00	Suicide + self-inflicted injury – jump from other manmade structure
TK72.00	Suicide + self-inflicted injury – jump from natural sites
TK7z.00	Suicide + self-inflicted injury – jump from high place not otherwise specified
TKx.00	Suicide and self-inflicted injury by other means
TKx0.00	Suicide + self-inflicted injury – jump/lie before moving object
TKx0000	Suicide + self-inflicted injury – jumping before moving object
TKx1.00	Suicide and self-inflicted injury by burns or fire
TKx2.00	Suicide and self-inflicted injury by scald
TKx3.00	Suicide and self-inflicted injury by extremes of cold
TKx4.00	Suicide and self-inflicted injury by electrocution

Appendix 1
Continued

Read code	Description
TKx5.00	Suicide and self-inflicted injury by crashing motor vehicle
TKx6.00	Suicide and self-inflicted injury by crashing of aircraft
TKx7.00	Suicide and self-inflicted injury caustic substance, excluding poison
TKxy.00	Suicide and self-inflicted injury by other specified means
TKxz.00	Suicide and self-inflicted injury by other means not otherwise specified
TKy.00	Late effects of self-inflicted injury
TKz.00	Suicide and self-inflicted injury not otherwise specified
U2..00	[X]Intentional self-harm
U2..11	[X]Self-inflicted injury
U2..12	[X]Injury – self-inflicted
U2..13	[X]Suicide
U2..14	[X]Attempted suicide
U2..15	[X]Para-suicide
U20.00	[X]Intentional self-poisoning/exposure to noxious substances
U20.11	[X]Deliberate drug overdose/other poisoning
U200.00	[X]Intentional self-poisoning/exposure to non-opioid analgesic
U200.11	[X]Overdose – paracetamol
U200.12	[X]Overdose – ibuprofen
U200.13	[X]Overdose – aspirin
U200000	[X]Intentional self-poisoning/exposure to non-opioid analgesic at home
U200100	[X]Intentional self-poisoning non-opioid analgesic at residential institution
U200400	[X]Intentional self-poisoning non-opioid analgesic in street/highway
U200500	[X]Intentional self-poisoning non-opioid analgesic trade/service area
U200y00	[X]Intentional self-poisoning non-opioid analgesic other specified place
U200z00	[X]Intentional self-poisoning non-opioid analgesic unspecified place
U201.00	[X]Intentional self-poisoning/exposure to antiepileptic
U201000	[X]Intentional self-poisoning/exposure to antiepileptic at home
U201z00	[X]Intentional self-poisoning antiepileptic unspecified place
U202.00	[X]Intentional self-poisoning/exposure to sedative hypnotic
U202.11	[X]Overdose – sleeping tablets
U202.12	[X]Overdose – diazepam
U202.13	[X]Overdose – temazepam
U202.15	[X]Overdose – nitrazepam
U202.16	[X]Overdose – benzodiazepine
U202.17	[X]Overdose – barbiturate
U202.18	[X]Overdose – amobarbital
U202000	[X]Intentional self-poisoning /exposure to sedative hypnotic at home
U202400	[X]Intentional self-poisoning sedative hypnotic in street/highway
U202y00	[X]Intentional self-poisoning sedative hypnotic other specified place
U202z00	[X]Intentional self-poisoning sedative hypnotic unspecified place
U204.00	[X]Intentional self-poisoning/exposure to psychotropic drug
U204.11	[X]Overdose – antidepressant
U204.12	[X]Overdose – amitriptyline
U204.13	[X]Overdose – SSRI
U204000	[X]Intentional self-poisoning /exposure to psychotropic drug at home
U204100	[X]Intentional self-poisoning psychotropic drug at residential institution
U204y00	[X]Intentional self-poisoning psychotropic drug other specified place

Appendix 1

Continued

Read code	Description
U204z00	[X]Intentional self-poisoning psychotropic drug unspecified place
U205000	[X]Intentional self-poisoning/exposure to narcotic drug at home
U205y00	[X]Intentional self-poisoning narcotic drug other specified place
U205z00	[X]Intentional self-poisoning narcotic drug unspecified place
U206.00	[X]Intentional self-poisoning/exposure to hallucinogen
U206400	[X]Intentional self-poisoning hallucinogen in street/highway
U207.00	[X]Intentional self-poisoning/exposure to other autonomic drug
U207000	[X]Intentional self-poisoning/exposure to other autonomic drug at home
U207z00	[X]Intentional self-poisoning other autonomic drug unspecified place
U208.00	[X]Intentional self-poisoning/exposure to other/unspecified drug/medicament
U208400	[X]Intentional self-poisoning other/unspecified drug/medication in street/highway
U208y00	[X]Intentional self-poisoning other/unspecified drug/medication other specified place
U208z00	[X]Intentional self-poisoning other/unspecified drug/medication unspecified place
U20A.00	[X]Intentional self-poisoning organic solvent, halogen hydrocarbon
U20A.11	[X]Self-poisoning from glue solvent
U20A000	[X]Intentional self-poisoning organic solvent, halogen hydrocarbon, home
U20A400	[X]Intentional self-poisoning organic solvent, halogen hydrocarbon, in highway
U20Az00	[X]Intentional self-poisoning organic solvent, halogen hydrocarbon, unspecified place
U20B.00	[X]Intentional self-poisoning/exposure to other gas/vapour
U20B.11	[X]Self carbon monoxide poisoning
U20B000	[X]Intentional self-poisoning/exposure to other gas/vapour at home
U20B200	[X]Intentional self-poisoning other gas/vapour school/public admin area
U20By00	[X]Intentional self-poisoning other gas/vapour other specified place
U20Bz00	[X]Intentional self-poisoning other gas/vapour unspecified place
U20C.00	[X]Intentional self-poisoning/exposure to pesticide
U20C.11	[X]Self-poisoning with weedkiller
U20C.12	[X]Self-poisoning with paraquat
U20C000	[X]Intentional self-poisoning/exposure to pesticide at home
U20Cy00	[X]Intentional self-poisoning pesticide other specified place
U20y.00	[X]Intentional self-poisoning/exposure to unspecified chemical
U20y000	[X]Intentional self-poisoning/exposure to unspecified chemical at home
U20y200	[X]Intentional self-poisoning unspecified chemical school/public admin area
U20yz00	[X]Intentional self-poisoning unspecified chemical unspecified place
U21.00	[X]Intentional self-harm by hanging/strangulation/suffocation
U210.00	[X]Intentional self-harm by hanging/strangulation/suffocation at home
U211.00	[X]Intentional self-harm by hanging/strangulation/suffocation occurrence at residential institution
U21y.00	[X]Intentional self-harm by hanging/strangulation/suffocation other specified place
U21z.00	[X]Intentional self-harm by hanging/strangulation/suffocation unspecified place
U22.00	[X]Intentional self-harm by drowning and submersion
U221.00	[X]Intentional self-harm by drowning/submersion occurrence at residential institution

Read code	Description
U22y.00	[X]Intentional self-harm by drowning/submersion occurrence at other specified place
U22z.00	[X]Intentional self-harm by drowning/submersion occurrence at unspecified place
U24.00	[X]Intentional self-harm by rifle shotgun/larger firearm discharge
U241.00	[X]Intentional self-harm by rifle shotgun/larger firearm discharge occurrence at residential institution
U242.00	[X]Intentional self-harm by rifle shotgun/larger firearm discharge in school/public admin area
U25.00	[X]Intentional self-harm by other/unspecified firearm discharge
U250.00	[X]Intentional self-harm other/unspecif firearm discharge occurrence at home
U26.00	[X]Intentional self-harm by explosive material
U27.00	[X]Intentional self-harm by smoke, fire and flames
U270.00	[X]Intentional self-harm by smoke fire/flames occurrence at home
U274.00	[X]Intentional self-harm by smoke fire/flame occurrence in street/highway
U27z.00	[X]Intentional self-harm by smoke fire/flames occurrence in unspecified place
U28.00	[X]Intentional self-harm by steam hot vapours/hot objects
U280.00	[X]Intentional self-harm by steam hot vapours/hot objects occurrence at home
U28z.00	[X]Intentional self-harm by steam hot vapours/hot objects occurrence in unspecified place
U29.00	[X]Intentional self-harm by sharp object
U290.00	[X]Intentional self-harm by sharp object occurrence at home
U291.00	[X]Intentional self-harm by sharp object occurrence at residential institution
U294.00	[X]Intentional self-harm by sharp object occurrence in street/highway
U29y.00	[X]Intentional self-harm by sharp object occurrence at other specified place
U29z.00	[X]Intentional self-harm by sharp object occurrence at unspecified place
U2A.00	[X]Intentional self-harm by blunt object
U2A0.00	[X]Intentional self -arm by blunt object occurrence at home
U2A1.00	[X]Intentional self -arm by blunt object occurrence at residential institution
U2A3.00	[X]Intentional self -arm by blunt object occurrence at sports/athletic area
U2B.00	[X]Intentional self-harm by jumping from a high place
U2B0.00	[X]Intentional self-harm by jumping from high place occurrence at home
U2B4.00	[X]Intentional self-harm by jumping from high place occurring in street/highway
U2B6.00	[X]Intentional self-harm by jumping from high place industrial/construction area
U2By.00	[X]Intentional self-harm by jumping from high place occurrence other specified place
U2Bz.00	[X]Intentional self-harm by jumping from high place occurrence unspecified place
U2C.00	[X]Intentional self-harm by jumping/lying before moving object
U2C1.00	[X]Intentional self-harm by jumping/lying before moving object occurrence at residential institution
U2C4.00	[X]Intentional self-harm by jumping/lying before moving object occurrence in street/highway
U2Cy.00	[X]Intentional self-harm by jumping/lying before moving object occurrence other specified place
U2D.00	[X]Intentional self-harm by crashing of motor vehicle
U2D0.00	[X]Intentional self-harm by crashing of motor vehicle occurrence at home

Appendix 1
Continued

Read code	Description
U2D4.00	[X]Intentional self-harm by crashing of motor vehicle occurrence in street/highway
U2D6.00	[X]Intentional self-harm by crashing of motor vehicle occurrence industrial/construction area
U2E.00	[X]Self-mutilation
U2y.00	[X]Intentional self-harm by other specified means
U2y0.00	[X]Intentional self-harm by other specified means occurrence at home
U2y1.00	[X]Intentional self-harm by other specified means occurrence at residential institution
U2yz.00	[X]Intentional self-harm by other specif means occurrence at unspecified place
U2z.00	[X]Intentional self-harm by unspecified means
U2z0.00	[X]Intentional self-harm by unspecified means occurrence at home
U2z2.00	[X]Intentional self-harm by unspecified means occurrence school/institution/public administrative area
U2zy.00	[X]Intentional self-harm by unspecified means occurrence other specified place
U2zz.00	[X]Intentional self-harm by unspecified means occurrence at unspecified place
U30.11	[X]Deliberate drug poisoning
U41.00	[X]Hanging strangulation + suffocation undetermined intent
U44.00	[X]Rifle shotgun + larger firearm discharge undetermined intent
U45.00	[X]Other + unspecified firearm discharge undetermined intent
U4B.00	[X]Falling jumping/pushed from high place undetermine intent
U4Bz.00	[X]Fall jump/push from high place undetermine intent occurring at unspecified place
U72.00	[X]Sequelae of intentional self-harm assault + event of undetermined intent
U720.00	[X]Sequelae of intentional self-harm
ZRLfC12	Health of the Nation Outcome Scales item 2 – nonaccidental self-injury
ZX..00	Self-harm
ZX..11	Self-damage
ZX1.00	Self-injurious behaviour
ZX1.12	SIB – self-injurious behaviour
ZX1.13	Deliberate self-harm
ZX11.00	Biting self
ZX11.11	Bites self
ZX12.00	Burning self
ZX13.00	Cutting self
ZX13.11	Cuts self
ZX15.00	Drowning self
ZX18.00	Hanging self
ZX19.00	Hitting self
ZX19100	Punching self
ZX19200	Slapping self
ZX1B.00	Jumping from height
ZX1B100	Jumping from building
ZX1B200	Jumping from bridge
ZX1B300	Jumping from cliff
ZX1C.00	Nipping self
ZX1E.00	Pinching self
ZX1G.00	Scratches self
ZX1H.00	Self-asphyxiation
ZX1H100	Self-strangulation
ZX1H200	Self-suffocation
ZX1I.00	Self-scalding
ZX1J.00	Self-electrocution
ZX1K.00	Self-incineration

Appendix 1
Continued

Read code	Description
ZX1K.11	Setting fire to self
ZX1K.12	Setting self alight
ZX1L.00	Self-mutilation
ZX1L100	Self-mutilation of hands
ZX1L200	Self-mutilation of genitalia
ZX1L300	Self-mutilation of penis
ZX1L600	Self-mutilation of ears
ZX1LD00	[X]Self mutilation
ZX1M.00	Shooting self
ZX1N.00	Stabbing self
ZX1Q.00	Throwing self in front of train
ZX1Q.11	Jumping under train
ZX1R.00	Throwing self in front of vehicle
ZX1S.00	Throwing self onto floor

Appendix 2 Further information on methodology

Code identification

Two authors, K.H.T. and D.G., identified a list of potential Read codes for self-harm. Where there was disagreement, R.M.M. was asked to provide a third opinion, after which a consensus opinion was reached.

Three categories of self-harm were identified, as follows.

1 Definite self-harm. These included Read codes where intent was more explicitly implied, such as TK...17 Para-suicide, TK01.00 Suicide + self-inflicted poisoning by barbiturates, ZX1.13 Deliberate self-harm, U20.00 [X] Intentional self-poisoning/exposure to noxious substances, U30.11 [X] Deliberate drug poisoning. Read codes for overdose which specified drugs commonly implicated in suicide, such as antidepressants and analgesics, were also categorized as definite self-harm [34].

2 Possible self-harm. These included codes such as SLD6.00 Emetic drug poisoning, SLG.12 eye drug poisoning, SLC.00 cardiovascular drug poisoning and U205.11 [X] Overdose – heroin.

3 Accidental injury. These Read codes specifically included 'accidental' in their definition, such as T8...11 Cause of overdose – accidental, T840.00 Accidental poisoning by antidepressants.

For the validation study, definite self-harm was the outcome of interest so we used all the Read codes that we included in this category. Initially, the code SL...15 Overdose of Drug was classified as possible self-harm. However, the age and sex distribution for people with this code was identical to that for the other codes for definite self-harm and 210 confirmed ONS suicides were recorded in the

CPRD using this Read code; it was second only to the Read code U2...13 [X] Suicide, which identified 251 ONS-confirmed suicides. For this reason we included the Read code SL...15 Overdose of drug in the definite self-harm category. Further information can be obtained from the authors on request.

Identification of deaths within the CPRD

Deaths may be identified by three ways in the CPRD (Shivani Padmanabhan, Medicines and Healthcare Regulatory Agency, personal communication).

1 A transferred out patient (i.e. a patient who is no longer registered with that CPRD practice) with a transfer out reason that has been specified as death.
2 A clinical or referral event with a Read code indicating a death category, including statement of death.
3 A record in the death administration structured data area in the additional clinical details file.

Although using a transfer out reason of death is the most reliable way of identifying deaths, the transfer out date may not be the date of death. Death dates are included in the CPRD patient files. These death dates are derived using an algorithm which is not publicly shared. We used the CPRD-derived death dates, because they were identical to the ONS dates of death in ~100% of cases (in the sample of patients with ONS-confirmed suicides). Transfer out dates were less accurate than CPRD-derived death dates when compared with ONS dates of death.

Identification of suicides within the CPRD

Suicide Read codes may refer to completed suicides, attempted suicide or suicide in a family member. The CPRD recommends (Shivani Padmanabhan, Medicines and Healthcare Regulatory Agency, personal communication) that dates of suicide Read codes are valid as dates of death only if there is a transfer out date with a transfer out reason of death in the patient's record within 95 days of the event date, or the patient has a record in the death administration area of the additional clinical details file. Transfer out dates were in most cases identical or within a few days of the CPRD-derived death dates; therefore, we opted to use the CPRD-derived death dates. Owing to the problem of delays in coroners' reporting [16], which could result in delayed notification of suicides to general practices, we carried out a sensitivity analysis using varying time periods (30, 180 and 360 days) between the event date of the Read code record and the CPRD-derived date of death.

REFERENCES

1 Rodriguez LAG, Gutthann SP. Use of the UK General Practice Research Database for pharmacoepidemiology. Br J Clin Pharmacol 1998; 45: 419–25.

2 Lawrenson R, Williams T, Farmer R. Clinical information for research; the use of general practice databases. J Public Health Med 1999; 21: 299–304.

3 Harpaz R, Dumouchel W, Shah NH, Madigan D, Ryan P, Friedman C. Novel data-mining methodologies for adverse drug event discovery and analysis. Clin Pharmacol Ther 2012; 91: 1010–21.

4 Gunnell D, Saperia J, Ashby D. Selective serotonin reuptake inhibitors (SSRIs) and suicide in adults: meta-analysis of drug company data from placebo controlled, randomised controlled trials submitted to the MHRA's safety review. BMJ 2005; 330: 385–88.

5 Martinez C, Rietbrock S, Wise L, Ashby D, Chick J, Moseley J, Evans S, Gunnell D. Antidepressant treatment and the risk of fatal and non-fatal self harm in first episode depression: nested case-control study. BMJ 2005; 330: 389–93.

6 Rubino S, Roskell N, Tennis P, Mines D, Weich S, Andrews E. Risk of suicide during treatment with venlafaxine, citalopram, fluoxetine, and dothiepin: retrospective cohort study. BMJ 2007; 334: 242.

7 Gunnell D, Irvine D, Wise L, Davies C, Martin RM. Varenicline and suicidal behaviour: a cohort study based on data from the General Practice Research Database. BMJ 2009; 339: b3805.

8 Arana A, Wentworth CE, Ayuso-Mateos JL, Arellano FM. Suicide-related events in patients treated with antiepileptic drugs. N Engl J Med 2010; 363: 542–51.

9 Black C, Jick H. Etiology and frequency of rhabdomyolysis. Pharmacotherapy 2002; 22: 1524–26.

10 Schade R, Andersohn F, Suissa S, Haverkamp W, Garbe E. Dopamine agonists and the risk of cardiac-valve regurgitation. N Engl J Med 2007; 356: 29–38.

11 Department of Health. Investing in General Practice: The New General Medical Services Contreact. London: Department of Health, 2003.

12 Parkinson J, Davies S, Van Staa T. The General Practice Research database: now and future. In: Pharmacovigilance, 2nd edn, eds Mann R, Andrews E. Chichester: Wiley, 2007; 341–8.

13 Hawton K, Fagg J, Simkin S, Bale E, Bond A. Trends in deliberate self-harm in Oxford, 1985–1995. Implications for clinical services and the prevention of suicide. [see comments]. Br J Psychiatry 1997; 171: 556–60.

14 Gunnell D, Bennewith O, Peters TJ, Stocks N, Sharp DJ. Do patients who self-harm consult their general practitioner soon after hospital discharge? A cohort study. Soc Psychiatry Psychiatr Epidemiol 2002; 37: 599–602.

15 Cooper J, Murphy E, Jordan R, Mackway-Jones K. Communication between secondary and primary care following self-harm: are National Institute of Clinical Excellence (NICE) guidelines being met? Ann Gen Psychiatry 2008; 7: 21. doi:10.1186/1744-859X-7-21.

16 Coroners Statistics 2010 England and Wales. In: Justice Mo, editor: Crown copyright, 2011.

17 Hall GC. Validation of death and suicide recording on the THIN UK primary care database. Pharmacoepidemiol Drug Saf 2009; 18: 120–31.

18 Arana A, Wentworth CW, Arellano FM. Validity of the codes of suicidality in the THIN database. Pharmacoepidemiol Drug Saf 2010; 19: 1316–17.

19 Jick SS, Kaye JA, Vasilakis-Scaramozza C, Rodriguez LAG, Ruigomez A, Meier CR, Schlienger RG, Black C, Jick H. Validity of the General Practice Research Database. Pharmacotherapy 2003; 23: 686–89.

20 Jick SS, Dean AD, Jick H. Antidepressants and suicide. BMJ 1995; 310: 215–18.

21 Garcia-Rodriguez LA, Perez-Gutthann S, Jick S. The UK General Practice Research Database. In: Pharmacoepidemiology, ed. Strom BL. Chichester: John Wiley & Sons Ltd, 2002; 375–85.

22 Office for National Statistics. Suicide rates in the United Kingdom 2000–2009. In: Statistical Bulletin, Newport: Crown Copyright, 2011: 1–10.

23 Hawton K, Bergen H, Casey D, Simkin S, Palmer B, Cooper J, Kapur N, Horrocks J, House A, Lilley R, Noble R, Owens D. Self-harm in England: a tale of three cities – multicentre study of self-harm. Soc Psychiatry Psychiatr Epidemiol 2007; 42: 513–21.

24 Bergen H, Hawton K, Waters K, Cooper J, Kapur N. Epidemiology and trends in non-fatal self-harm in three centres in England: 2000–2007. Br J Psychiatry 2010; 197: 493–98.

25 Brock A, Baker A, Griffiths C, Jackson G, Fegan G, Marshall D. Suicide trends and geographical variations in the United Kingdom, 1991–2004. Health Stat Q 2006; 31: 6–22.

26 Linsley KR, Schapira K, Kelly TP. Open verdict v. suicide – importance to research. Br J Psychiatry 2001; 178: 465–68.

27 Kapur N, House A, Creed F, Feldman E, Friedman T, Guthrie E. Management of deliberate self poisoning in adults in four teaching hospitals: descriptive study. BMJ 1998; 316: 831–32.

28 Bennewith O, Gunnell D, Peters TJ, Hawton K, House A. Variations in the hospital management of self-harm in adults in England: an observational study. BMJ 2004; 328: 1108–09.

29 Doll R. Comparison between registries. Age-standardized rates. In: Cancer Incidence in Five Continents, Vol. III, eds Waterhouse J, Muir C, Correa P, Powell J. Lyon: International Agency for Research on Cancer, IARC Scientifc Publications No. 15, 1976: 453–59.

30 Hawton K, Zahl D, Weatherall R. Suicide following deliberate self-harm: long-term follow-up of patients who presented to a general hospital. Br J Psychiatry 2003; 182: 537–42.

31 Webb RT, Kontopantelis E, Doran T, Qin P, Creed F, Kapur N. Risk of self-harm in physically ill patients in UK primary care. J Psychosom Res 2012; 73: 92–7.

32 Hesdorffer DC, Ishihara L, Mynepalli L, Webb DJ, Weil J, Hauser WA. Epilepsy, suicidality, and psychiatric disorders: a bidirectional association. Ann Neurol 2012; 72: 184–91.

33 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; 69: 4–14.

34 Thomas KH, Beech E, Gunnell DJ. Changes in commonly used methods of suicide in England and Wales from 1901–1907 to 2001–2007. J Affect Disord 2013; 144: 235–9.