FiMT forces in mind trust

COMMUNITY FUND

KING'S College LONDON

KCMHR KING'S CENTRE FOR MILITARY HEALTH RESEARCH

Exploring veterans physical and mental health outcomes: *the use of the Clinical Practice Research Datalink*

Pamela Almeida-Meza, Bethany Croak, Alexandru Dregan, Shehan Hettiaratchy, Nicola Fear, Sharon Stevelink

December 2025

Foreword

Over recent years, significant research has correctly been undertaken regarding the mental health of military veterans in the United Kingdom, however there has been relatively little research on differences or similarities in the physical health of veterans compared with their civilian counterparts. Using large-scale routine NHS data, primarily from England, this study highlights these similarities and differences.

Key findings are the higher recorded prevalence of musculoskeletal problems, especially lower back pain, as well as higher rates of deafness, chronic obstructive pulmonary disease and coronary heart disease. This study also highlights higher rates of prostate and breast cancer, Alzheimer's disease and myocardial infarction (heart attack) when compared with non-veterans, although absolute prevalence remained low. Importantly, the study also shows that more military veterans are living with more than one long term condition (also known as multimorbidity).

From a methodological perspective, this study has shown the ability and benefit of being able to access routine NHS data to identify veterans who are coded within general practitioner computer systems and compare their profiles with the wider population. Given this success, the Clinical Practice Research Datalink (CPRD) can be used to research other conditions which may be of specific interest to the military community.

This study is very important for two main reasons. First, it identified areas where physical health conditions are more commonly recorded among veterans accessing primary care, helping inform clinicians and service planning. Second, it showcases the methods developed by King's Centre for Military Health Research (KCMHR), which offer potential for advancing research into veterans' health. As such, the study is of value not only to the military community, but also to policymakers, healthcare professionals, and other with an interest of improving population health.

Dr (Colonel Retired) Jonathan Leach OBE
NHS England Associate Medical Director for Armed Forces and Veterans Health

Contents

King's Centre for Military Health Research

The King's Centre for Military Health Research (KCMHR), known previously as the Gulf War Illness Research Unit, was launched in 2004 as a joint initiative between the Institute of Psychiatry, Psychology and Neuroscience and the Department of War Studies, King's College London. KCMHR is led by Professor Sir Simon Wessely and Professor Nicola Fear. The centre draws upon the experience of a multi-disciplinary team to undertake research studying military life by using quantitative and qualitative methods. Its flagship study is an ongoing epidemiological multiphase investigation of the health and well-being of approximately 20,000 UK Armed Forces personnel. The study, funded by the UK Ministry of Defence (MoD) and most recently the Office for Veterans' Affairs (OVA), has been running since 2003. Data from the studies conducted at KCMHR have been used to analyse various military issues, and papers have been published in peer reviewed scientific journals. The findings are regularly reported in the press and have been used to inform military policies.

Authors

Pamela Almeida-Meza, Bethany Croak, Alexandru Dregan, Shehan Hettiaratchy, Nicola Fear, Sharon Stevelink

Acknowledgements

We are grateful to Forces in Mind Trust for funding this research and for their support and continued engagement throughout the project, in particular Kirsteen Waller and Tom McBarnet for their thoughtful guidance.

We would also like to thank the members of our Project Advisory Board, whose expertise and insights greatly enriched the study. Members included Dr Phil Moore and Dr Frank O'Kelly, both experienced NHS GPs with a long-standing interest in veterans' health, with Dr O'Kelly's perspective shaped in part by his service in the Armed Forces; Ann Griffiths from the Royal British Legion, who brought expertise in policy and service provision; Professor Catherine Kinane, formerly based at Combat Stress and a current trustee for Veterans Outreach Support, who contributed knowledge of specialist mental health care for veterans; Professor Karen Walker-Bone, an expert in occupational medicine; and Ms Susie Schofield, a medical statistician who provided advice and input into the project's analyses. We are grateful for their time, commitment, and the depth of expertise they brought to the project.

Our discussions with the King's Centre for Military Health Research Veterans Research Advisory Group were insightful and imperative to this project and we thank them for their input and time.

We thank the Clinical Practice Research Datalink (CPRD) for granting access to the data that made this study possible as well as the GPs who participated in the validation of veteran records.

Terminology

COPD	Chronic obstructive pulmonary disease.
CPRD	The Clinical Practice Research Datalink is a large, anonymised database of UK primary care records used for research.
EHRs	Electronic Health Records.
GP	General Practitioner, a doctor who works in general practice and is usually the first point of contact for any health concerns.
HES APC	Hospital Episode Statistics Admitted Patient Care. Dataset containing detailed records of all inpatient admissions to NHS hospitals in England.
ICD-10	The International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10). A globally recognised system developed by the World Health Organization (WHO) for the classification of diseases and other health conditions. It is used worldwide to record and code diagnoses, symptoms, and causes of death, supporting both statistical analysis and research.
IMD	Index of Multiple Deprivation. A widely used measure of relative deprivation in the UK. It assesses the level of deprivation in small areas across England, Wales, Scotland, and Northern Ireland. The IMD combines various indicators across different domains, such as income, employment, education, health, crime, and housing.
MICE	Multiple Imputation by Chained Equations is a method for handling missing data by creating several complete datasets. The estimated values are generated through a series of regression models. The results are then combined to produce estimates that account for uncertainty due to missingness
National Service	National Service was compulsory military service in the UK, introduced in January 1949. It required all physically fit men aged 18 to 21 years to serve in the Armed Forces for at least 18 months (later extended to two years in 1950), followed by four years on the reserve list. The policy formally ended in 1960, with the last National Service personnel discharged in 1963.
NHS	National Health Service.
PAB	Project Advisory Board.
PTSD	Post-Traumatic Stress Disorder.
PR	A prevalence ratio compares how common a condition or outcome is in one group relative to another. A PR of 1 indicates that the outcome is equally common in both groups. A PR greater than 1 indicates that the outcome is more common in the exposed or comparison group, while a PR less than 1 indicates that it is less common.
SNOMED	The Systematised Nomenclature of Medicine is a comprehensive and internationally recognised clinical vocabulary designed for recording and exchanging patient information in electronic health records.
Veteran	In the UK, a veteran is officially defined as anyone who has served for at least one day in His Majesty's Armed Forces, either as a regular or reservist.

Executive Summary

Why is this research important?

Military veterans in the UK have unique and complex health needs, yet national evidence on how they access and use primary care is limited. UK research has largely relied on surveys or hospital records, which capture only part of the picture. Primary care records hold valuable, detailed, and timely information on a broad range of physical and mental health conditions, but to date, veterans have not been identified in these systems at a national level.

What was this research about?

This study explored the feasibility and value of using the Clinical Practice Research Datalink (CPRD), a large database of anonymised UK primary care records, to identify veterans, compare their health with non-veterans, and assess the benefits of linking CPRD to other datasets (e.g., socioeconomic deprivation and hospital admissions).

How did this work?

We identified veterans using veteran codes in CPRD and validated the veteran records we found through a confirmation process involving GP practices. A matched comparison group of non-veterans was selected based on age, gender, practice, and registration date. Health risk factors, physical and mental health conditions, and service use were analysed. Additionally, linked data from hospital records (admitted patients) and deprivation indices were used to improve data completeness and accuracy.

What did we find?

We identified 138,457 veterans in CPRD (122,484 in Aurum and 15,973 in Gold). Compared to non-veterans, veterans had more GP consultations, and their health records were more complete, with fewer missing entries on key health risk factors such as alcohol use, smoking, blood pressure, and body mass index. This likely reflects better health monitoring among veterans and thus, differences between veterans and non-veterans should be interpreted with caution. For example, alcohol use findings appear to be influenced by selection and recording biases, whereby patients are more likely to be assessed, and their alcohol use recorded when there is clinical concern. As a result, the data may overrepresent those with more severe drinking problems rather than the wider population.

The findings suggested that veterans had higher prevalence of physical conditions such as lower back pain, osteoarthritis, chronic obstructive pulmonary disease and myocardial infarction. Mental health conditions were also more common, especially depression and post-traumatic stress disorder (PTSD). Furthermore, living with multiple long-term conditions was more common among veterans.

For England, linking CPRD with hospital data improved completeness of ethnicity data and increased detection of several conditions, showing veterans have higher rates of hospitalisation, especially for musculoskeletal and cardiovascular conditions.

This is the first large-scale approach to identifying veterans in UK primary care data.

It overcomes many limitations of previous veteran research by enabling population-level analysis and long-term follow-up, without relying solely on self-reports or small regional samples. However, challenges remain, including inconsistent coding of veteran status, missing information in some records, and limited dataset linkage outside of England.

What does this mean?

This innovative method enables a new era of UK veterans' health research using primary care data. CPRD can now be used to monitor health trends, identify emerging needs, and evaluate interventions for veterans with a level of detail and scale previously unavailable. The findings can inform NHS planning, guide government policy through the Office for Veterans' Affairs, and support targeted services by charities and specialist healthcare providers. By creating a transferable and validated approach, this project ensures that future studies can build on its foundations, driving improvements in health outcomes for the UK's veteran community.

Recommendations

CPRD provides an important opportunity to inform practice and policy by identifying patterns of physical and mental health needs among veterans and highlighting where they face higher risks or barriers to care. At the practice level, encouraging GP practices to adopt veteran-friendly approaches, including consistent recording of veteran status and appointing veteran leads (staff member responsible for championing veteran care), can help ensure veterans are recognised within primary care. This can build trust in services by showing their backgrounds are understood, and improve referrals to appropriate support, such as specialist NHS veteran services or charities.

Research opportunities include developing detailed health profiles across the UK, analysing under researched groups, such as women and ethnic minorities, and linking CPRD with wider datasets to fill evidence gaps, strengthen prevention, and evaluate policy initiatives.

Report

1. Introduction

Despite being the first port of call for most health concerns, UK general practice data remains an underused source of insight into the needs of military veterans[1]. General practitioners (GPs) manage around 90% of all physical and mental health conditions (General Medical Council, 2019), yet there is limited evidence on how veterans engage with primary care after leaving service. Accessing these healthcare records offers an unrivalled opportunity to evaluate the health status and needs of veterans after leaving military service.

The US, in comparison, has been at a relative advantage in terms of its access to administrative health data. This is in part due to the Department of Veterans' Affairs (VA), which provides VA healthcare. VA healthcare data has been used extensively to understand the needs of veterans presenting to primary care services within the VA health system (Hynes et al., 2021). This data has been linked to non-VA health datasets such as health insurance records, birth records, and drug prescription monitoring data (Hynes et al., 2018). US National health surveys are also more likely to include a veteran indicator than in the UK, so they offer ample research opportunities. Data from the Centers for Disease Control and Prevention (CDC)'s[2] Behavioural Risk Factor Surveillance System (BRFSS) showed that veterans had higher rates of obesity, skin cancer, coronary heart disease, chronic obstructive pulmonary disease (COPD), arthritis and kidney disease (Betancourt et al., 2023).

In the UK, we mostly rely on evidence from cohort studies. Findings from the latest phase (2022-2023) of a UK health and wellbeing cohort of serving personnel (N=1,157) and veterans (N=2,944) who served in the Armed Forces during the Iraq and Afghanistan conflicts indicated that 9% met criteria for probable Post-Traumatic Stress Disorder (PTSD), while 28% reported symptoms of common mental disorders (Sharp et al., 2024). There has been some research in the UK which has taken advantage of health records. For example, Bergman and colleagues have used National Health Service (NHS) Scotland hospital (secondary care) data and disease-specific registers to identify health trends among Scottish veterans, reporting an elevated risk of cardiovascular disease compared to civilians, as well as a higher likelihood of type 2 diabetes among veterans diagnosed with PTSD (Bergman et al., 2022). While hospital data provides insight into more acute and severe health issues, it captures only a fraction of the healthcare needs veterans experience. GP data, or primary care data, is essential for understanding the broader spectrum of health concerns, including chronic conditions, mental health issues, and early-stage illnesses that may never lead to hospital admission but significantly affect quality of life and long-term wellbeing.

Additionally, primary care records provide a more comprehensive and detailed view of a patient's health. A single consultation may include multiple diagnostic codes, capturing co-existing physical and mental health conditions. Given that

[1]In the UK, a veteran is officially defined as anyone who has served for at least one day in His Majesty's Armed Forces, either as a regular or reservist (Office for Veterans' Affairs, 2020).

[2]The CDC is the U.S public health agency responsible for disease prevention, control and health promotion. It functions as the US's first line of defence against global health threats and maintains a network of country and regional offices in over 60 countries (Centers for Disease Control and Prevention, 2025).

physical and mental health conditions rarely occur in isolation, it is only through examining both in tandem that we can develop tailored strategies to address the complex and often interrelated needs of veterans (Cassell et al., 2018; Guthrie et al., 2012; Yoon et al., 2014). Primary care records, therefore, offer a cost- and time-effective alternative to traditional research methods (e.g. cross-sectional or cohort studies), with advantages including freedom from recall bias and near real-time availability, both vital in the fast-changing field of physical and mental health needs and multimorbidity[3].

However, there are notable challenges in using these data for veteran research. These include obtaining consent to access primary care records, establishing infrastructure for secure data management, extracting relevant information, ensuring data are suitable for analysis, and crucially, reliably identifying patients with a history of military service. The UK only introduced a military service question in the Census in 2021, meaning robust estimates of the veteran population are not available before that date and in healthcare records, although an Armed Forces service flag exists and GP practices are encouraged to record this information, it is not consistently captured (Finnegan & Randles, 2023b).

To address these challenges, we explored the utility and feasibility of the Clinical Practice Research Datalink (CPRD) for veteran health research. CPRD provides longitudinal anonymised patient record data of 65 million patients who have accessed primary care services across the UK (England, Scotland, Wales and Northern Ireland) – including 19 million currently registered patients – and is representative of age, gender, and ethnicity distribution in the UK population (Clinical Practice Research Datalink, 2025a).

1.1 Project aim

The study examined the feasibility and utility of using CPRD for veteran health research, and specifically, to identify and characterise the physical and mental health needs of veterans and compare these to non-veterans.

1.2 Research questions (RQ)

RQ1. How feasible and useful is CPRD for identifying veterans within the UK primary care system and selecting a suitable non-veteran comparison group?

RQ2. What are the key data quality and data management challenges identified through a validation exercise aimed at identifying and characterising the physical and mental health needs of veterans in primary care?

RQ3. To what extent do the socio-demographic, mental health, and physical health profiles of veterans accessing primary care differ from those of the general primary care population?

RQ4. What added benefit does linking CPRD with other datasets provide in terms of improving detection of physical and mental health needs and enhancing the completeness of key socio-demographic characteristics?

RQ5. What are the key data quality and data management challenges associated with identifying and characterising the physical and mental health needs of veterans in primary care?

RQ6. What clinical practice, policy and research opportunities are enabled by using the CPRD and linked datasets in the context of veterans' health, and what are the strengths and limitations of these data sources?

[3] *The presence of two or more long-term health conditions in an individual.*

2. Methods

2.1 Project Advisory Board

A Project Advisory Board (PAB) was established to provide strategic guidance, oversight, and support throughout the study. The PAB brought together a diverse group of experts, including frontline clinicians with direct experience supporting veterans in NHS and specialist mental health services, senior representatives from leading veteran charities, the Royal British Legion and Combat Stress, who offered insight into the lived experiences and needs of the veteran community, and academics with backgrounds in statistics, occupational health, and data-driven research. The PAB was actively engaged throughout the project, advising on study design, interpretation of findings, and dissemination strategies, and played a critical role in ensuring the research remained meaningful and applicable to real-world policy and practice.

2.2 Data sources

Clinical Research Practice Datalink (CPRD)

The NHS is one of the largest publicly funded healthcare system in the world. In 2024, around 64 million people in England were registered with a GP (general practitioner) (NHS Digital, 2024). When people visit their GP or receive hospital care, information about their symptoms, test results, diagnoses, and prescriptions is recorded in electronic health records (EHRs). CPRD is a trusted research service that collects anonymised EHRs from 2,400 GP practices across the UK (Clinical Practice Research Datalink, 2025a). It holds historical records on 65 million patients, including around 19 million patients (28% of UK population) who are currently registered with a GP (Clinical Practice Research Datalink, 2025a). For GPs in England, patient data is linked to other national datasets, such as hospital records and ONS death registrations, allowing researchers to study patient care and outcomes across the NHS.

Previous studies have shown that CPRD data is reliable and accurate for researching a wide range of health conditions (Dregan et al., 2011, 2012, 2014, 2019). For this project, we used data from the September 2024 release[4] of CPRD Aurum and CPRD Gold.

CPRD Aurum

CPRD Aurum contains anonymised health records primarily from GP practices in England and a few in Northern Ireland using EMIS Web software. Observations are coded using SNOMED CT[5], Read, and local EMIS codes. It includes data from over 51 million patients, with around 17 million (25% of the UK population) currently registered across 1,600 practices (Clinical Practice Research Datalink, 2025c). The database captures key information such as symptoms, diagnoses, prescriptions, tests, and referrals. All six research questions were addressed using this dataset.

CPRD Gold

CPRD Gold contains anonymised health records from GP practices using Vision software. Observations are coded using Read v2 codes. CPRD Gold covers all four UK nations. It includes historical data from over 22 million patients, with about 2.3 million patients (4% of the UK population) currently registered at 291 practices from Scotland (72%),

[4]Data is collected by practices daily and monthly builds of the dataset are generated and made available for researchers (Herrett et al., 2015).

[5]SNOMED CT and Read Codes are standardised clinical coding systems used in UK primary care to record information such as diagnoses, symptoms, and treatments in patients EHRs. Read Codes were developed for the NHS while SNOMED CT (Systematised Nomenclature of Medicine – Clinical Terms) is the newer, internationally recognised terminology.

Wales (19%), and Northern Ireland (9%) (Clinical Practice Research Datalink, 2025d). The database holds information on diagnoses, treatments, test results, and referrals. Since this dataset is smaller and data from the devolved nations cannot be linked to other datasets[6], only research questions 1, 5, and 6 were able to be assessed using both CPRD Aurum and Gold.

Linkages

CPRD linkage data is mainly available for England. Since the majority of CPRD Aurum data is collected from England, only this dataset was used for the analyses that included linkage.

Index of Multiple Deprivation

This dataset uses the postcode of English GP practices or patient residence postcodes to provide a neighborhood level measure of socioeconomic status. It links the postcodes to an Index of Multiple Deprivation (IMD) score. These scores reflect factors like income, education, health, housing, crime, and environment (Department for Communities and Local Government, 2015).

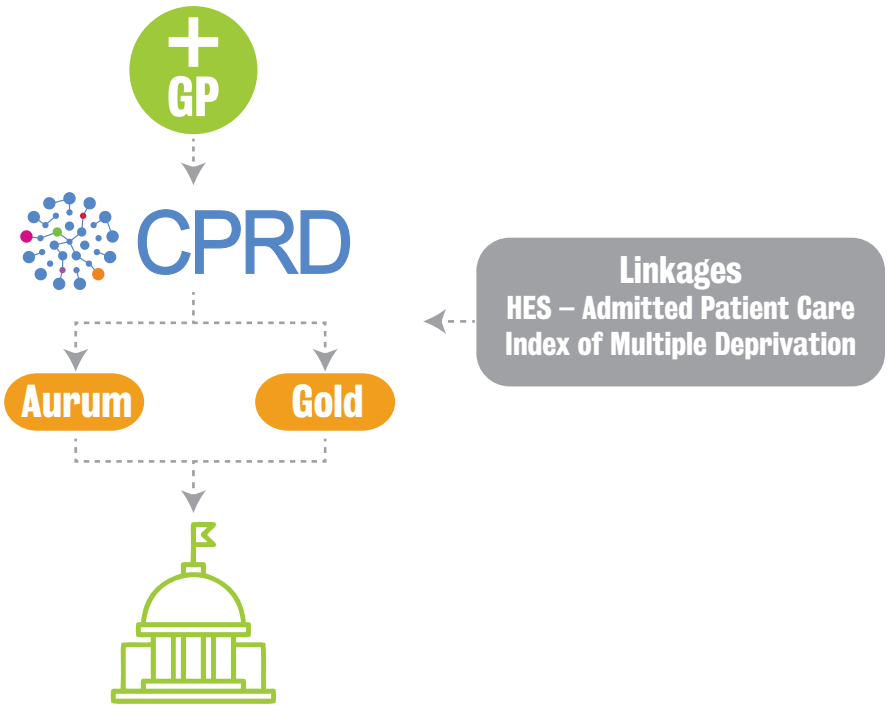
Hospital Episode Statistics Admitted Patient Care (HES APC)

This dataset records hospital admission information for patients treated in NHS hospitals across England. It includes details like admission and discharge dates, reasons for admission, specialist care provided, and any procedures performed using ICD-10 codes. For this study, the data covered admissions from April 1997 to March 2023.

Ethical approval

This project was covered within CPRD's legal and ethical framework (protocol ID is 22_001829).

Figure 1. CPRD database and linkages



[6]Linkage of patient level data is carried out by NHS England and thus is only available for patients at practices in England which have consented to participate in the linkage scheme (Clinical Practice Research Datalink, 2025b)

CPRD Aurum vs CPRD Gold: Data differences and analytical implications

CPRD Aurum and CPRD Gold differ in their structure, coding systems, and data coverage, which shaped what analyses could be conducted in each. Aurum is the larger dataset, mainly covering practices across England and supporting linkage to external datasets such as hospital records and socioeconomic indices. This made it suitable for analyses requiring linked data and adjustment for key variables. In contrast, Gold primarily covers practices in Scotland, Wales, and Northern Ireland, with data linkage opportunities available for a subset of practices in England. While Gold provides valuable opportunities for cross-national comparisons, its limited data linkage opportunities and differences in coding and data structure meant that only research questions 1, 5, and 6 were able to be assessed using both CPRD Aurum and Gold. Research questions 2 to 4 were assessed using CPRD Aurum only.

2.3 Sample

Veterans

The main factor we looked at in this study was whether someone was a veteran or not. Veteran status (veteran vs. non-veteran) was identified using SNOMED or Read medical codes, which are standard systems used to record patient information in medical records. **Importantly, veteran identification in primary care relies on patient self-report**.

Since 2011, the UK Department of Health required that a Read Code be applied to medical records to indicate a "history relating to military service" (Royal College of General Practitioners, 2011). However, as described by Finnegan and colleagues, in practice, multiple military codes exist (Finnegan et al., 2018). Although considerable work has gone into standardising terminology, for example by deprecating several codes in the terminology browser and issuing guidance, these steps do not correct legacy records.

We used a list of medical codes that had been used in previous research identifying veterans in secondary mental healthcare records and reviewed common terms found in the scientific literature (Leightley et al., 2023; Mark et al., 2020). We also searched CPRD's medical dictionaries for any terms that might relate to military or veteran status. The resulting list of codes was reviewed by members of the PAB (see section 2.1) to make sure they accurately reflected veteran status.

In total, 121 codes were used to identify veterans in the CPRD Aurum dataset, and 35 codes were used for the CPRD Gold dataset (see section 6). **For veterans, the date we used as reference point – referred here as the "index date" – was defined as the date on which a veteran-related term first appeared in the patient's record.** Where possible, we also categorised the codes by branch of military service: Royal Navy (including Royal Marines), Army, or Royal Air Force (RAF). Additionally, we identified veterans who served during the National Service era. This was compulsory military service introduced in 1949 requiring all able-bodied men aged 18 to 21 years to serve in the UK Armed Forces for at least 18 months (extended to two years in 1950), followed by four years on the reserve list. The policy ended in 1960, National Service era veterans were classified as such if they were aged 25 or over in 1963 when the last National Service person was discharged (Hickman, 2004).

For each person, follow-up began at the 'start date', defined as the latest of their GP registration date, index date (i.e. when a veteran-related term first appeared in a patient's record), or study start date (01 January 1987). Follow-up ended at the 'end date', defined as the earliest of their registration end date (e.g., moved to a different GP practice), last CPRD collection date for their practice, their date of death, or the study end date (31 July 2024). Because it is unlikely that younger individuals are veterans, only those aged 18 and older were included in the analysis.

Non-veterans

A comparison non-veteran group was identified in CPRD Aurum as this was the largest dataset and allowed for linkage to other datasets. For each veteran, two non-veteran individuals were selected for comparison (see section 6, 'Manual for researchers' for more information on choosing a comparison group). These non-veterans were registered at the same general practice as the veteran and were matched on age, gender, and index date. This approach helps ensure that any differences observed between the groups are less likely to be due to age, gender, the specific practice they attend, or their registration date. **For the non-veterans the "index date" was the date they first registered with their GP.**

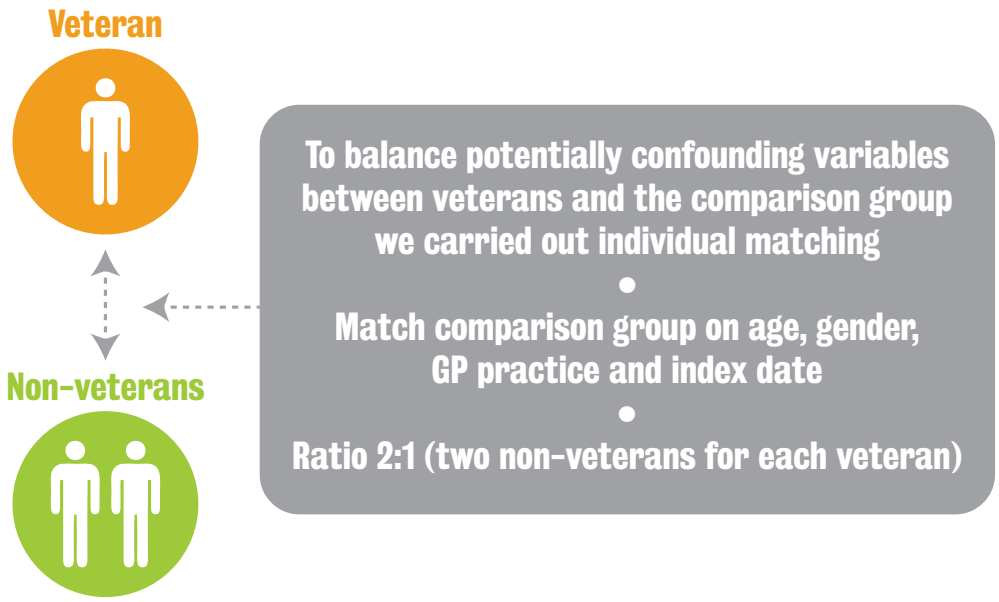
Key term: matching

The veteran population differs from the general population on key characteristics such as age and gender, whereby veterans are predominantly men (85%) and older (30% are aged 80 or over) than the general population, reflecting the National Service era (Office for National Statistics, 2023a). Age and gender are strongly associated with the risk of developing certain health conditions and patterns of healthcare utilisation. Veterans also differ from the general population in their contact with NHS primary care. During service, UK Armed Forces personnel receive healthcare through the Defence Medical Services rather than the NHS and therefore may have fewer NHS primary care records prior to discharge. Furthermore, GP practices vary in recording practices, patient demographics and healthcare access which could explain differences in the incidence and prevalence of health outcomes.

Our goal was to understand whether veterans have different health outcomes compared to non-veterans. To make this comparison as fair as possible, we used a method called matching. This means that for each veteran in our data, we found one or two people who had their same age and gender, attended the same GP practice and had the same index date. By doing this, we reduce the chances that any observed differences in their health profile are due to these factors rather than military service itself.

The choice of matching variables depends on data availability and on the research question of each study. In our case, matching was conducted on age, gender, GP practice, and index date. However, it is important to note that other factors, such as ethnicity, are also associated with the prevalence of certain conditions. For example, diabetes shows higher rates in some ethnic groups (Goff, 2019). Thus, future comparative analyses should match or adjust for such variables, where data availability allows.

Figure 2. Matching



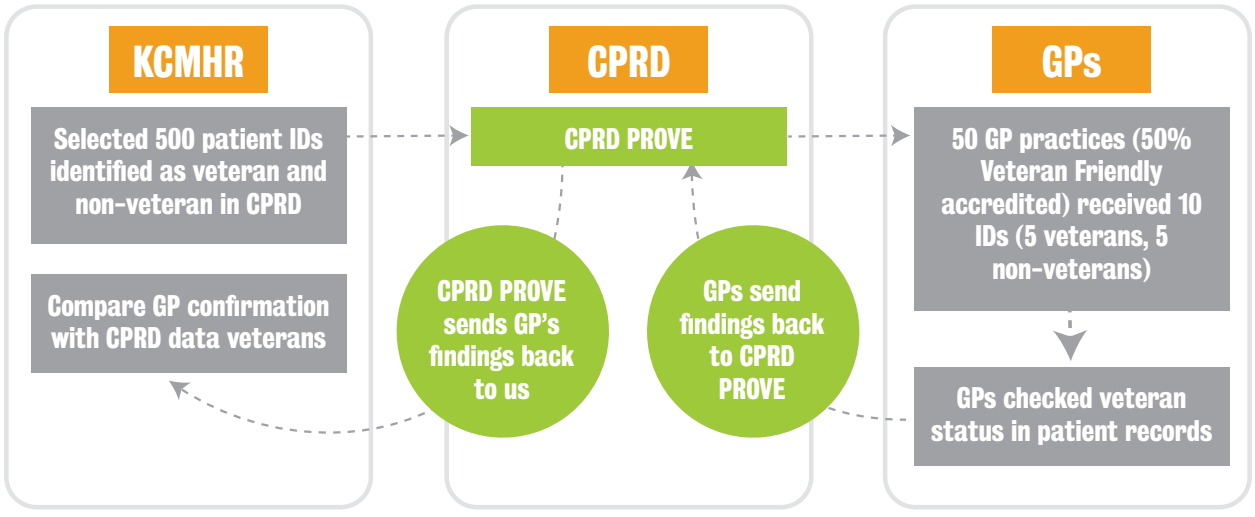
Veteran status validation study

We carried out a validation study using CPRD's Providing Online Verification of EHR (PROVE) service to identify the accuracy of veteran status in CPRD (Clinical Practice Research Datalink, 2024). This service allows researchers to contact general practices to confirm details in patients' records.

We aimed to receive confirmation of veteran status for 100 patients (identified as veterans and non-veterans). However, since we expected that not all practices would respond, we "oversampled" by sending 500 patient IDs to the CPRD PROVE service to increase the chances of reaching our target sample. The PROVE service contacted 50 general practices across England on our behalf and sent each practice 10 patient IDs (5 veterans and 5 non-veterans), based on the information we extracted from CPRD. Half of these practices were accredited as veteran friendly. GPs were asked to confirm whether the patients listed were veterans or not, based on the information in their practice records.

All practices were also asked a practice level question which enquired the estimated number of veterans registered which could be classified into 8 bandings: Up to 10; 10 to <25; 25 to <50; 50 to <75; 75 to <100; 100 to <500; 500 to <1,000; >1000.

Figure 3. Validation study



2.4 Primary care and linked data

Demographic characteristics

The study looked at key demographic information, including age (18 and over), gender (men or women), ethnicity (Asian, Black, Mixed, Other, or White), UK region (based on the location of the person's GP), and socioeconomic status (based on linked IMD data).

Follow-up characteristics

Index date

As mentioned above, UK Armed Forces personnel receive health services through the Defence Medical Services rather than the NHS and therefore may have fewer NHS primary care records prior to discharge. Thus, the different index date definitions ensure the observation period aligns with veterans' status and is comparable between veterans and non-veterans. As indicated in section 2.3, index date for veterans was defined as the date on which a veteran-related term first appeared in the patient's record, whereas for non-veterans, the index date was defined as the date when the patient registered with the GP.

Yearly consultations

Consultations to primary care were summarised as the average number of consultations per patient per year. Consultation rates were created to compare consultations between veterans and non-veterans.

Mortality

Date of death was extracted from CPRD to explore differences in mortality between veterans and non-veterans.

Physical and mental health diagnoses

All diagnoses and health risk factors included in this project were selected based on previous evidence (Bergman et al., 2022; Betancourt et al., 2023; Finnegan & Salem, 2024; Goodwin et al., 2013; Katon et al., 2018; Rona et al., 2013; Sharp et al., 2019) and were informed by discussions with the PAB (see section 2.1). These were derived from medical codes recorded in the EHR by GPs, using SNOMED codes developed previously, as reported in previous research (Dregan et al., 2011, 2012).

The study's main outcomes were a set of common physical and mental disorders including Alzheimer's disease, breast cancer, coronary heart disease (CHD), chronic obstructive pulmonary disease (COPD), hearing loss, lower back pain, myocardial infarction, osteoarthritis, prostate cancer, anxiety, depression, and PTSD.

For a sub-set of outcomes namely COPD, depression, lower back pain, myocardial infarction, osteoarthritis and PTSD, secondary care data was used (HES APC linkage) to identify cases recorded in hospital but not necessarily in primary care.

Health risk factors

Using primary care data, health risk factors were identified. These included alcohol consumption severity (low risk, hazardous, harmful, and dependence), smoking status (never, ex-smoker, and smoker), body mass index (underweight, normal weight, overweight, moderate obesity, and severe obesity), and blood pressure (normal blood pressure, pre-hypertension[7], and hypertension).

2.5 Statistical analysis

Descriptive statistics were used to compare the socio-demographic, administrative, mental and physical health characteristics of veterans accessing primary healthcare with a non-veteran comparison group.

We assessed the baseline distribution of age, gender, ethnicity, social deprivation (IMD), health risk factors, and medical history between veterans and non-veterans. We also evaluated completeness of these data by veteran status.

Subsequently, we ran adjusted regression

[7]Pre-hypertensive, also known as high normal blood pressure and borderline hypertensive, is a medical classification indicating higher than normal blood pressure but not yet meeting the threshold for hypertension.

models[8] to examine the association between veteran status (predictor) and each health condition (outcome). Given that health conditions considered in this project are relatively common and patients were followed-up for different times, Poisson regressions with robust standard errors was used to estimate prevalence ratios[9]. These models also account for the fact that patients tend to be clustered within practices. Some of the socio-demographic variables (ethnicity, smoking, BMI) had various degree of missingness. To account for potential biases in estimates due to missing data, we used multiple imputations by chained equations (MICE)[10]. The imputation model included relevant military, demographic and outcome variables (White et al., 2011). A total of 50 imputed dataset were generated.

[8]A regression model is a statistical approach to assess the impact of a risk factor on a health outcome (for example, how having served in the Armed Forces might affect physical health), while also considering other factors, such as age and sex, that could play a role.

[9]Poisson regression with robust standard errors is a type of statistical data modelling approach that can be used to compare the relative risk or prevalence for common health conditions between two or more groups.

[10]Multiple imputation is a method used to handle missing data by creating several complete versions of the dataset, where missing values are filled using statistical models.

3. Findings

.3.1 RQ1: Feasibility of identifying veterans (CPRD Aurum and Gold) and matching non-veterans in CPRD (CPRD Aurum)

How feasible and useful is CPRD for identifying veterans within the UK primary care system and selecting a suitable non-veteran comparison group?

Veteran profile in CPRD Aurum

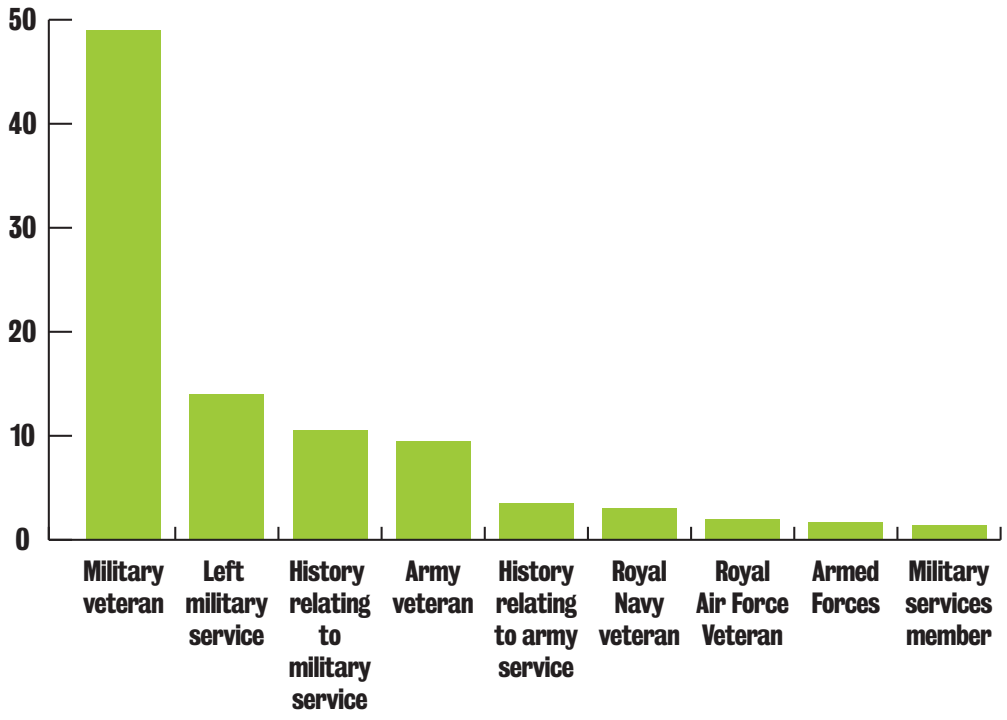
A total of 121 codes were included for extraction from CPRD Aurum and were assigned a category based on branch of service where available (Royal Navy [including Royal Marines], Army, and Royal Air Force) and based on the certainty of veteran status: "definite" if it clearly stated veteran status, and "probable", if it related to a military role or active duty in the military (see section 6). Veterans could have more than one term in their records.

A total of 122,484 veterans were identified and 95% of these were identified with definite terms.

Most common terms

Figure 4 displays the most common terms used to refer to veterans in the CPRD Aurum dataset. The term "Military veteran" is by far the most frequently used, appearing in nearly 50% of cases. This is followed by "Left military service" and "History relating to military service," both used in just over 10% of cases. Other terms such as "Army veteran," "History relating to Army service," and references to specific branches like "Royal Navy veteran" and "Royal Air Force veteran" are used far less frequently, each appearing in fewer than 5% of cases.

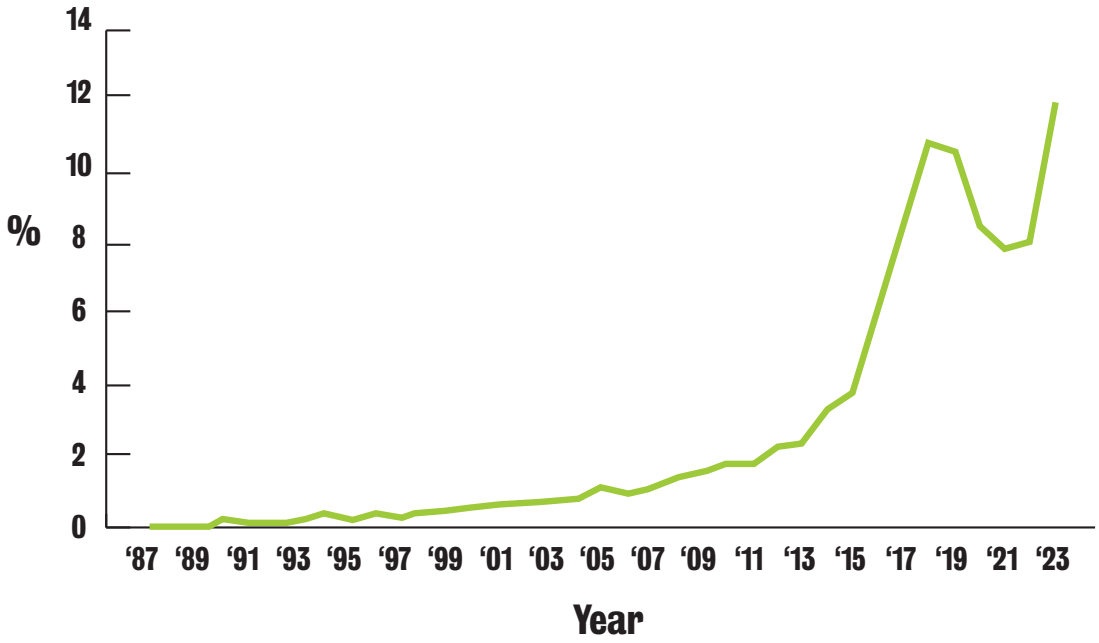
Figure 4. Most common veteran terms in CPRD Aurum



Year of identification

Figure 5 shows the percentage of cases by year of first identification from 1987 to 2023. Identification rates remained consistently low from 1987 to around 2013, with less than 2% of cases identified annually. From 2014 onwards, there was a marked increase, peaking in 2018 at just over 10%. This pattern aligns with the UK Department of Health's 2011 directive to apply a code for veterans which may have influenced coding practices in subsequent years, and the introduction of the Veteran Friendly GP Practice Accreditation in 2018 which might have provided a further boost (Finnegan et al., 2022; Royal College of General Practitioners, 2011). After this peak, rates declined slightly but remained significantly higher than earlier years, stabilising at around 7 to 13% between 2020 and 2023. The data suggests that system-level changes might have increased awareness of veteran identification trends at a primary care level.

Figure 5. Year of veteran identification in CPRD Aurum



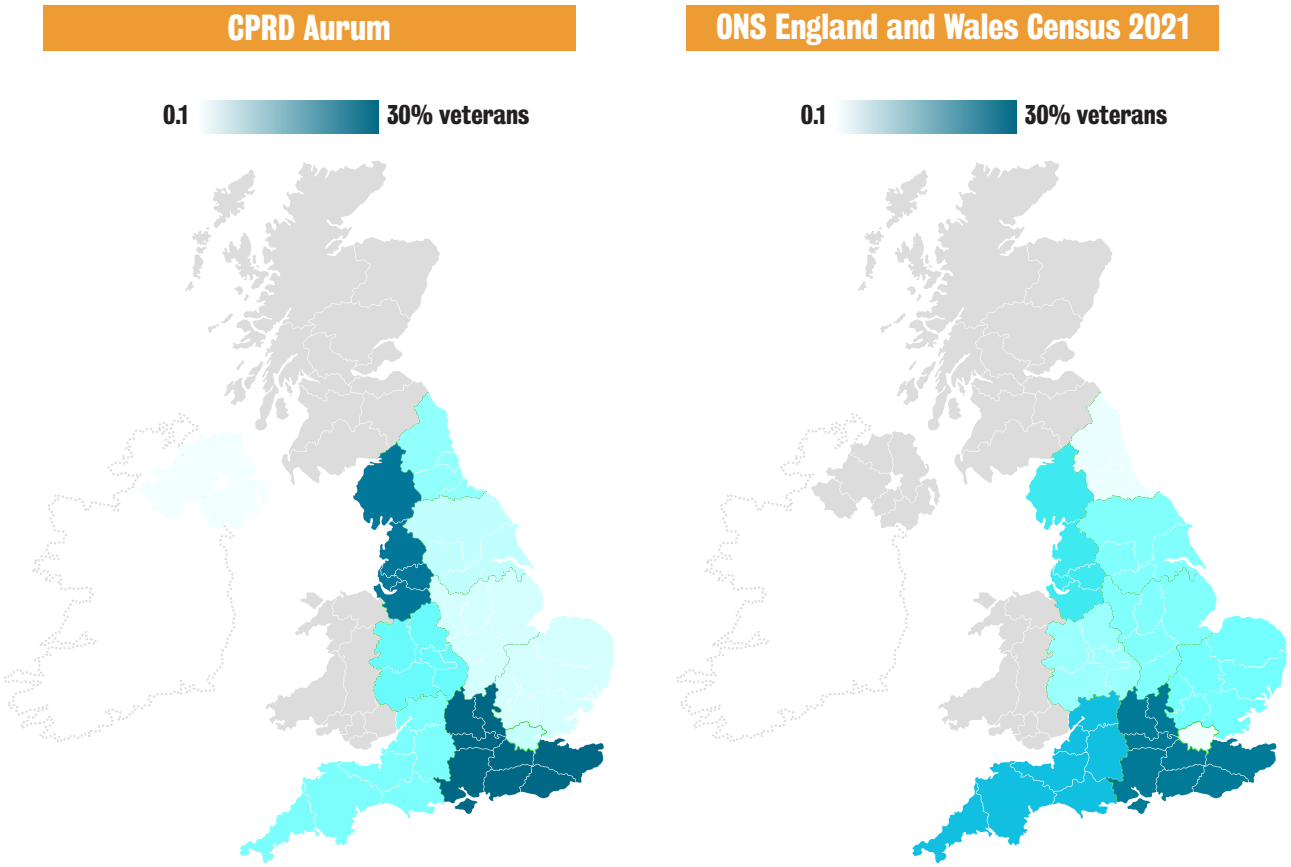
Distribution of CPRD Aurum patients identified as veterans across the UK

Figure 6 shows the regional distribution of veterans identified in CPRD Aurum dataset across England and Northern Ireland, as well as the regional distribution of veterans according to the 2021 Census for England and Wales (Office for National Statistics, 2023a). The blue gradient represents the percentage of the sample identified in each region, ranging from 0.1% (light blue) to 20% in the Census and 30% in CPRD (dark blue).

For the CPRD data, in England the South East (29.03%) and the North West (27.42%) contained the largest percentage of veterans, whereas the East Midlands (2.59%) and East of England (2.31%) showed the lowest percentage. Very few veterans were identified by CPRD data in Northern Ireland (0.10%). Scotland and Wales are shaded in grey, indicating that data from these nations are not included in the CPRD Aurum coverage. The map highlights substantial geographic variation in veteran representation across England.

Compared to the 2021 Census data (Office for National Statistics, 2023a), veterans in CPRD Aurum were slightly overrepresented in the North West (27.42% vs 13.16%) and South East (29.03% vs 18.25%) while they were underrepresented in other regions, particularly in the East of England (2.31% vs 11.16%), East Midlands (2.59% vs 9.81%), Yorkshire and The Humber (4.20% vs 10.65%) and the South West (10.26% vs 15.24%). The North East (8.41% vs 6.29%) and London (3.77% vs 5.80%) show a closer alignment.

Figure 6. Heatmap for percentage of veterans (CPRD Aurum and 2021 Census) in the UK



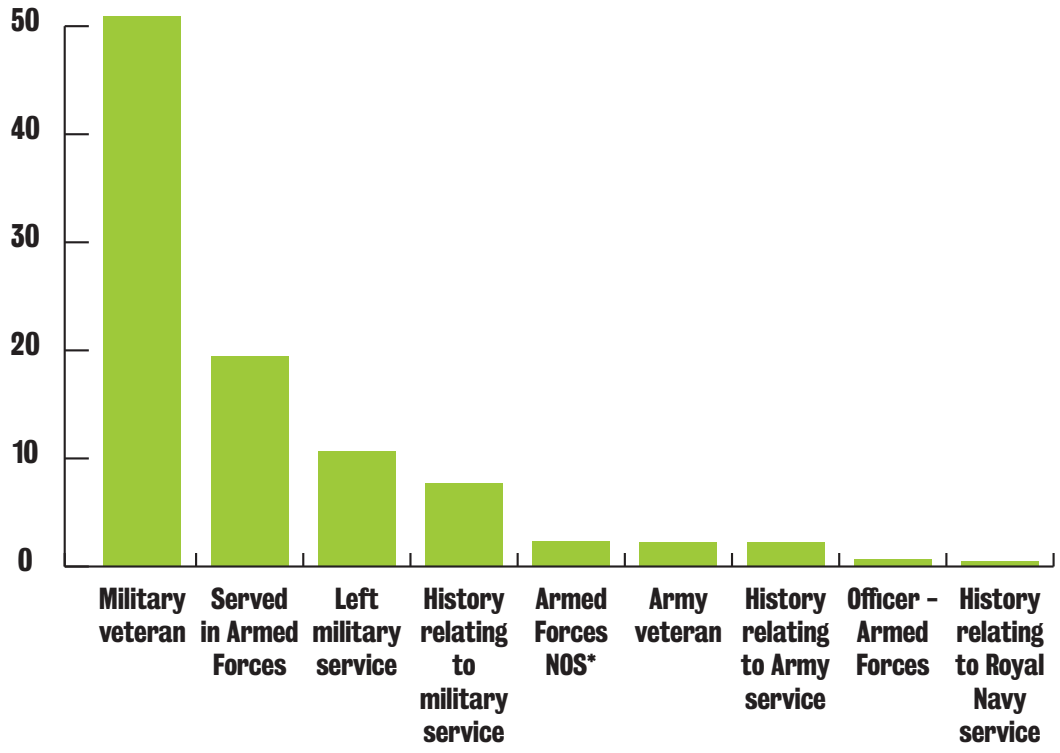
Veteran profile in CPRD Gold

A total of 35 codes were included for extraction from CPRD Gold. Like the CPRD Aurum terms, these were also assigned a category based on branch of service where available and categorised as "definite" or "probable" terms (see section 6). Veterans could have more than one term in their records. A total of 15,973 veterans were identified and 95% of these were identified with definite terms.

Most common terms

Figure 7 displays the most commonly used terms to describe veterans in the CPRD Gold dataset. Similar to the CPRD Aurum dataset, "Military veteran" is the dominant term, appearing in over 50% of cases. This is followed by "Served in armed forces" (around 20%) and "Left military service" (just over 10%). Other terms such as "History relating to military service," "Armed forces NOS" (Not Otherwise Specified), and "Army veteran" are used far less frequently, each appearing in fewer than 10% of records.

Figure 7. Most common veteran terms in CPRD Gold

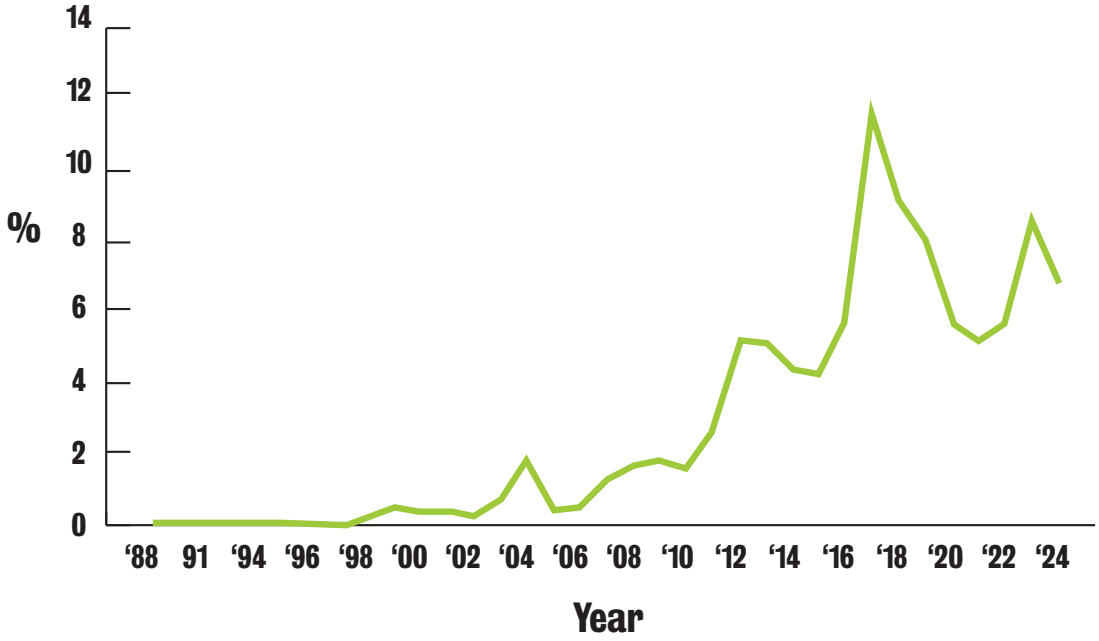


NOS- Not Otherwise Specified

Year of identification

Figure 8 shows the percentage of the total number of veteran patients who were first recorded in each year between 1988 and 2024. From 1988 until the early 2000s, annual identification percentages were close to zero, with only small fluctuations, including a modest rise in 2004. Rates remained low until around 2010, after which there was a steady increase, reaching about 5% in 2012. This early uplift aligns with the UK Department of Health's 2011 directive to apply a code to identify veterans, which likely began to improve recording practices, as noted in the veteran profile for CPRD Aurum above where a similar post-2011 inflection is noted (Royal College of General Practitioners, 2011). Following a slight dip in 2014, identifications rose sharply to a peak of around 12% in 2017, the highest annual proportion in the series. Although the proportion declined thereafter, it remained well above pre-2010 levels, fluctuating between approximately 5% to 8% in most years. Notably, 66% of all veterans in the 2024 dataset were first identified after 2016, indicating that the majority of identifications have occurred in recent years.

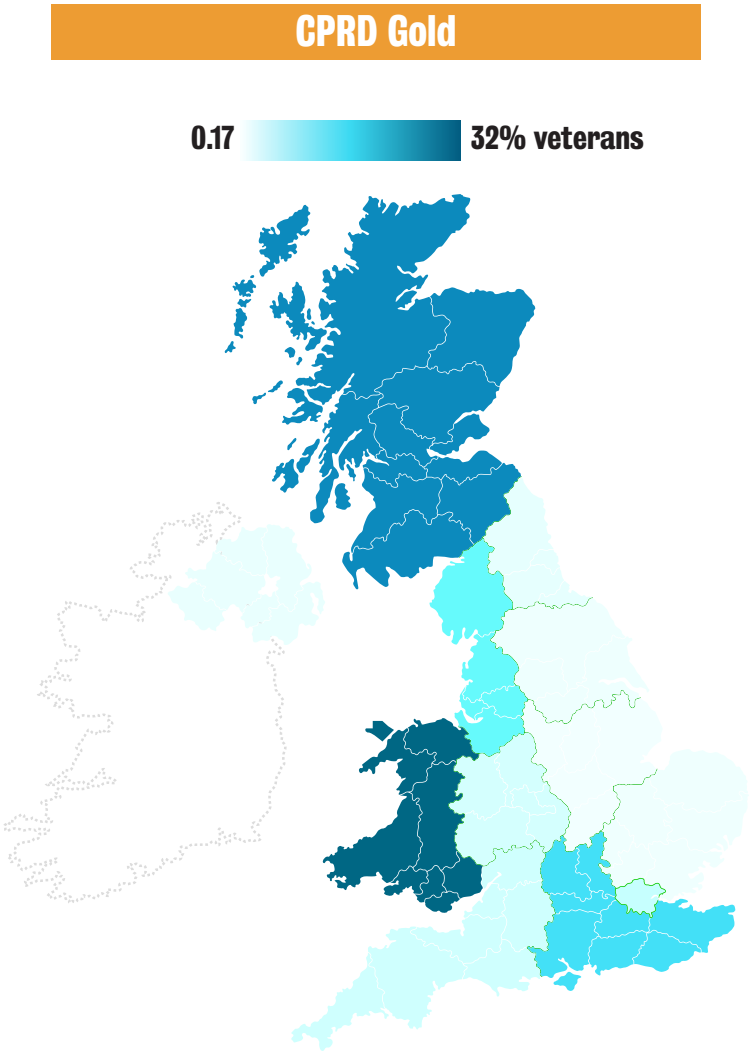
Figure 8. Year of veteran identification in CPRD Gold



Distribution of CPRD Gold patients identified as veterans across the UK

Figure 9 illustrates the distribution of UK veterans captured in the CPRD Gold dataset. The map covers England, Wales, Northern Ireland and Scotland, with a blue gradient representing veteran coverage ranging from 0.17% (light blue) to 32% (dark blue). The distribution of veterans across England for this dataset mirrors that of CPRD Aurum, with most veterans concentrated in the South East and North West. However, over half of the veteran sample in CPRD Gold is distributed between Wales and Scotland, with 31.75% in Wales and 27.32% in Scotland.

Figure 9. Heatmap for percentage of veterans (CPRD Gold) in each UK region



Veteran profiles for England, Wales, and Scotland

Figure 10 compares the demographic characteristics of veterans captured in CPRD Aurum (England) and CPRD Gold (Wales and Scotland) with official reference data from the 2021 Census for England and Wales (Office for National Statistics, 2023a) and the 2022 Scotland's Census (Scotland's Census, 2024). It is organised into three panels by country (England, Wales, Scotland), and three rows by characteristic: age, gender, and service branch.

Age

Across all three nations, the CPRD datasets showed a smaller proportion of veterans aged 80+, relative to the 2021 and 2022 Censuses. In England, CPRD Aurum, 47.03% of veterans were in the 16 - 49 age category whereas in the Census, 20.95% were in this age category. A similar pattern is seen in Wales and Scotland (CPRD Gold). This suggests that CPRD records might skew toward younger veterans and may not fully capture the oldest cohort. However, since veteran identification within primary care records relies on self-identification or disclosure, some older veterans, having left service decades ago, may choose not to disclose their veteran status.

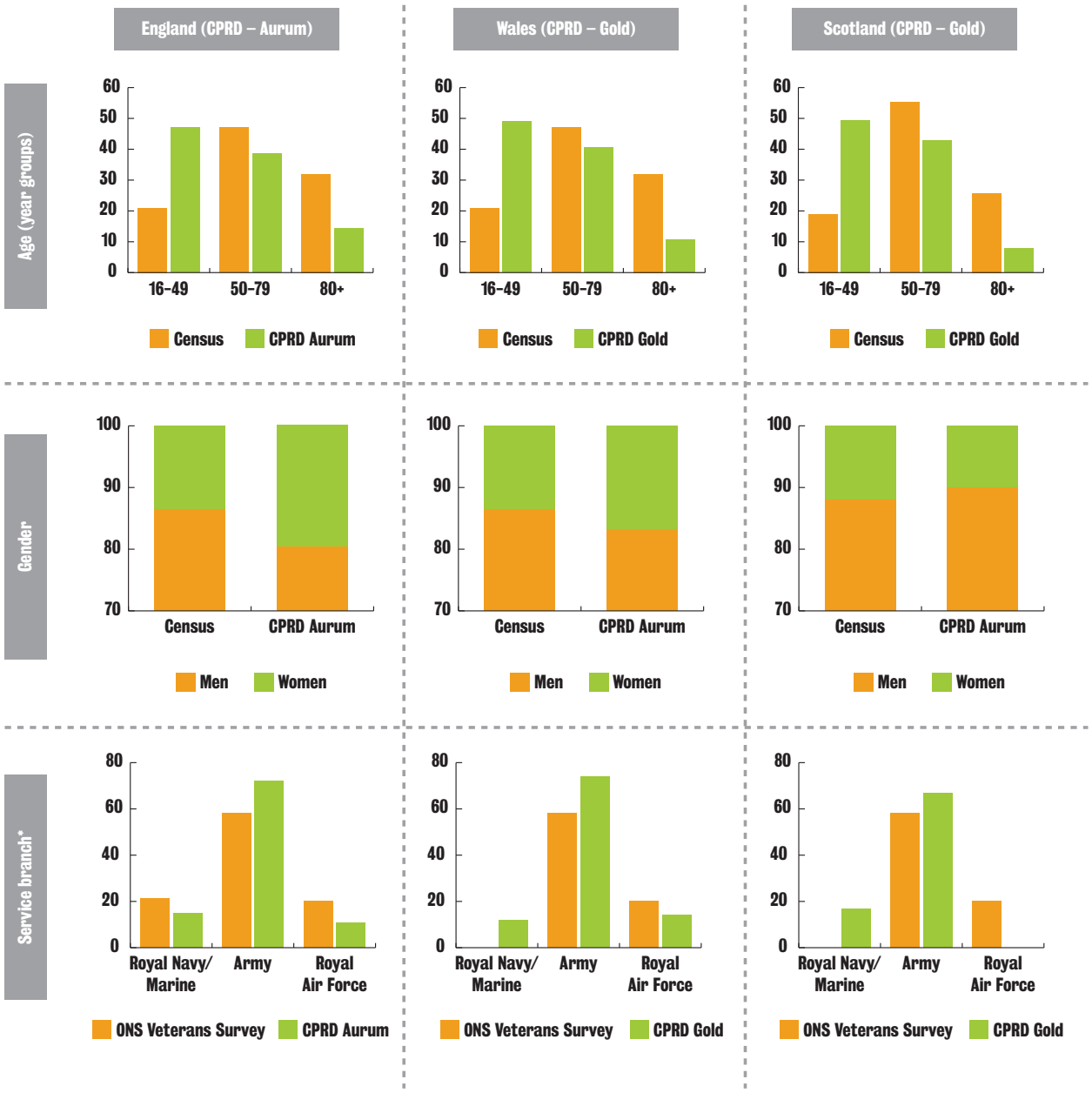
Gender

In England and Wales, CPRD data showed a higher proportion of women veterans than recorded in the 2021 Census. For example, in England, women make up 19.66% of the CPRD Aurum population compared to 13.57% in the Census. In Scotland however, the gender split is similar between the 2022 Census and CPRD Gold (12.00% vs 10.22%).

Service branch

Across England, Wales, and Scotland veterans from the Army are overrepresented in the CPRD datasets, while the Royal Navy/Marines and Royal Air Force and are underrepresented relative to the Veterans' Survey of 2022 (Office for National Statistics, 2023b). For example, in England, around 72.23% of CPRD Aurum veterans served in the Army compared to around 58.30% in the Veterans' Survey. The same pattern is seen in Wales and Scotland, though to a slightly lesser extent.

Figure 10. Veteran profiles for England, Wales, and Scotland



Census: ONS England and Wales Census 2021 and the 2022 Scotland's Census

**Only 32% of CPRD Aurum and 7% of CPRD Gold veteran sample had service branch infromation*

Matching non-veterans for CPRD Aurum

To minimise confounding[11], one or two non-veterans were matched to each veteran on age, gender, GP practice, and index date (see section 2.3). Due to sample size and linkage availability (key objective for research question 4) the subsequent analyses were carried out for CPRD Aurum only.

Matching variables

The study sample included 122,484 veterans and 244,573 matched non-veterans, totalling 367,057 individuals, all identified in CPRD Aurum. Matching was achieved at 2:1 ratio for 122,089 veterans, and 1:1 on 395 veterans due to a lack of eligible matches in the non-veteran pool, particularly at some variable extremes (e.g., youngest and oldest age groups). The veteran and non-veteran groups were well matched on age, gender, index date, and GP practice, with nearly identical distribution across the age at index, gender, year of index and region variables (see Table 1).

Most of the sample were identified as veterans or registered with their GP between 2011 and 2020 (57.17%). The mean age at index for both groups was 53 years, and the age distribution was well balanced, with approximately 27% of the sample being under 35 and around 30% being over the age of 65 in both groups.

In the matched sample, women represented 19.67% of both veterans and non-veterans. Since the veteran population is predominantly comprised of men (Office for National Statistics, 2023a), the matching process reduced the proportion of women in the non-veteran group and thus, the overall proportion of women in the sample is lower than what would typically be seen in the general primary care population where women tend to be slightly overrepresented (Hobbs et al., 2016). Similarly, because GP practice was used for matching, the regional distribution of the non-veterans mirrored that of veterans. Most of the sample was in England, with the largest proportion in the South East (29%) and North West (27%). Only 0.10% of the sample was in Northern Ireland.

[11]Confounding occurs when something (e.g., age) other than the factor we are interested in (e.g. veteran status) influences the results, making it look like there is a difference when it may just be due to that other factor. For example, veterans are on average older than non-veterans, and older age is itself liked to poorer health. Thus, if age is not accounted for through techniques such as matching or statistical adjusting, any difference we see in health outcomes might be due to age rather than veteran status

Table 1. CPRD Aurum sample (veteran and non-veteran) matching variables descriptives (N=367,057)

	NON-VETERANS N=244,573		VETERANS N=122,484	
	N	% or Mean (SD)	N	% or Mean (SD)
Year of index				
1987-1990	1,546	0.63	776	0.63
1991-2000	9,671	3.95	4,843	3.95
2001-2010	26,975	11.03	13,499	11.02
2011-2020	139,872	57.19	70,019	57.17
2021-2024	66,509	27.19	33,347	27.23
Age at index (years)		52.55 (20.99)		52.54 (20.99)
18-25	26,352	10.77	13,200	10.78
26-35	39,470	16.14	19,772	16.14
36-45	35,359	14.46	17,713	14.46
46-55	36,918	15.09	18,492	15.10
56-65	33,964	13.89	17,008	13.89
66-75	26,950	11.02	13,494	11.02
76-85	29,122	11.91	14,579	11.90
86-95	15,354	6.28	7,684	6.27
96+	1,084	0.44	542	0.44
Gender				
Men	196,472	80.33	98,404	80.34
Women	48,101	19.67	24,080	19.66
Region				
North East	20,584	8.42	10,303	8.41
North West	66,975	27.38	33,589	27.42
Yorkshire and the Humber	10,273	4.20	5,139	4.20
East Midlands	6,335	2.59	3,171	2.59
West Midlands	29,153	11.92	14,594	11.92
East of England	5,643	2.31	2,825	2.31
London	9,237	3.78	4,622	3.77
South East	71,018	29.04	35,552	29.03
South West	25,113	10.27	12,568	10.26
Northern Ireland	242	0.10	121	0.10

SD: Standard Deviation

Table 2. CPRD Aurum sample (veteran and non-veteran) descriptives (N=367,057)

	NON-VETERANS N=244,573		VETERANS N=122,484	
	N	% or Median [IQR]	N	% or Median [IQR]
Follow-up time (years)*		3.00 [1.16-5.73]		2.91 [1.24-5.57]
Consultations		6.50 [4.00-11.00]		10.12 [6.00-16.71]
Mortality				
Deceased	17,175	7.02	12,816	10.46
Year of death				
1990-2000	768	4.47	61	0.48
2001-2010	3,063	17.83	939	7.33
2011-2020	8,781	51.13	5,687	44.37
2021-2024	4,563	26.57	6,129	47.82
Ethnicity				
White	162,864	81.48	99,381	87.60
Black	4,093	2.05	1,798	1.58
Asian	11,224	5.62	2,986	2.63
Mixed	2,802	1.40	924	0.81
Other	18,909	9.46	8,357	7.37
Missing	44,681	-	9,038	-
Service branch				
Royal Navy (incl Marines)	-	-	5,723	14.76
Army	-	-	28,015	72.23
Royal Air Force	-	-	4,170	10.75
Multiple mentioned	-	-	876	2.26
Missing	-	-	83,700	-
National Service Era				
Post-National Service+	-	-	103,297	84.34
National Service	-	-	19,187	15.66

**Start date calculated as maximum value between registration date, index date, and protocol start date (01 of January 1987) and end date calculated as minimum value between registration end date, last CPRD collection date for the practice, date of death, or protocol end date (31 of July 2024).*
+Defined as Post-National Service if veterans were aged 25 or over in 1963 when the last National Service person was discharged.

IQR: Interquartile Range (25th percentile and 75th percentile)
Proportions calculated for patients with complete data on variables

Demographic and administrative characteristics of veteran and non-veteran sample

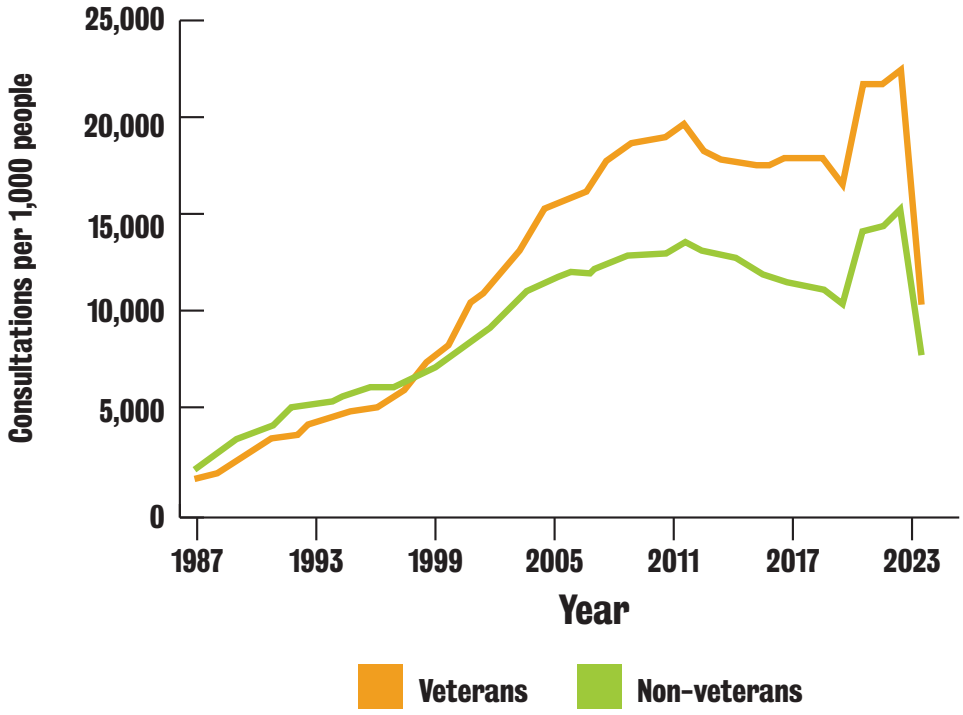
On average, veterans and non-veterans were followed up for approximately three years (see table 2). Here, 'follow-up' refers to the length of time individuals contributed to data to CPRD based on their start and end dates (see section 2.3). However, there was considerable variability in follow-up duration for both groups indicating some individuals had brief follow-up periods (≤one year), while others were followed for longer (≥six years).

Veterans had a higher number of GP consultations over the follow-up period compared to non-veterans (10.12 vs 6.50), reflecting greater healthcare needs and/or differences in help-seeking behaviours. Figure 11 shows consultation rates (per 1,000 people) over time (1987–2023) for veterans (orange line) and non-veterans (green line). Veterans' rates were lower than non-veterans up to the late 1990s, crossing over around 1999. Thereafter, veterans' rates rose steeply. From 2012 onwards, the two series follow a similar trajectory with rates rising steadily until 2011, stabilising thereafter, and peaking again during the COVID-19 pandemic.

Veterans showed a higher overall mortality than non-veterans (10.46% vs 7.02%). While most deaths in both groups occurred between 2011 and 2020, a larger proportion of veteran deaths were recorded more recently (2021 to 2024). This finding is at odds with previous evidence showing a lower mortality risk among veterans relative to the general population (Oster et al., 2017). However, it is in line with US evidence suggesting a recent erosion of this healthy solider effect which might result from cohort differences in deployment frequency, survival from injuries, and access to benefits after service (Bollinger et al., 2015; Cassidy et al., 2023; Oster et al., 2017).

Similar to the 2021 Census where 96.40% of veterans identify as white vs 83.10% of the non-veteran population, the ethnicity distribution was less diverse among veterans (Office for National Statistics, 2023a). Compared to non-veterans, veterans were more likely to be White (87.60% vs 81.48%) and less likely to be from other ethnic groups, especially Black (1.58% vs 2.05%) and Asian (2.63% vs 5.62%). The proportion of missing ethnicity data was much lower among

Figure 11. Annual consultation rates per 1,000 people among veterans and non-veterans



veterans (7.38%) compared to non-veterans (18.27%). This difference could bias prevalence comparisons for conditions influenced by ethnicity; thus, comparisons should be adjusted for ethnicity. Due to the large sample size, absolute counts in the non-White groups are sufficient to support these adjusted and stratified analyses.

Among veterans, only 32% had service branch information. The veteran sample was predominantly Army veterans (72.23%), followed by Royal Navy (14.76%) and Royal Air Force (10.75%) veterans. Around 2% of the veteran sample had multiple service branch codes in their EHR.

Summary

Analysis of CPRD Aurum and Gold demonstrates that veterans can be identified successfully in UK primary care records using a set of military service-related terms, most of which are "definite" indicators of veteran status (95% in both datasets). In CPRD Aurum, 122,484 veterans were identified, with recording rates rising sharply after 2014, possibly reflecting improved recognition and coding practices. In CPRD Gold we identified 15,973 veterans, particularly from Wales and Scotland. Compared to Census data, CPRD veterans were younger and included a higher proportion of women.

CPRD Aurum supported the creation of a large non-veteran comparison group (n=244,573), matched on age, gender, GP practice, and index date. Veterans had more GP consultations and were slightly less ethnically diverse. These findings highlight CPRD's practical value for veteran health research and for selecting suitable comparison groups in primary care datasets.

3.2 RQ2: Validation findings: data quality and management challenges (CPRD Aurum)

What are the key data quality and data management challenges identified through a validation exercise aimed at identifying and characterising the physical and mental health needs of veterans in primary care?

PROVE study results

From the 500 patient IDs (including both veterans and non-veterans) identified in CPRD and sent to the PROVE service, we received responses for 95 participants. However, GPs could only access valid patient records for 61 patients (see Figure 12 for study flowchart).

Figure 12. Flowchart of CPRD data sent to PROVE and GP data received

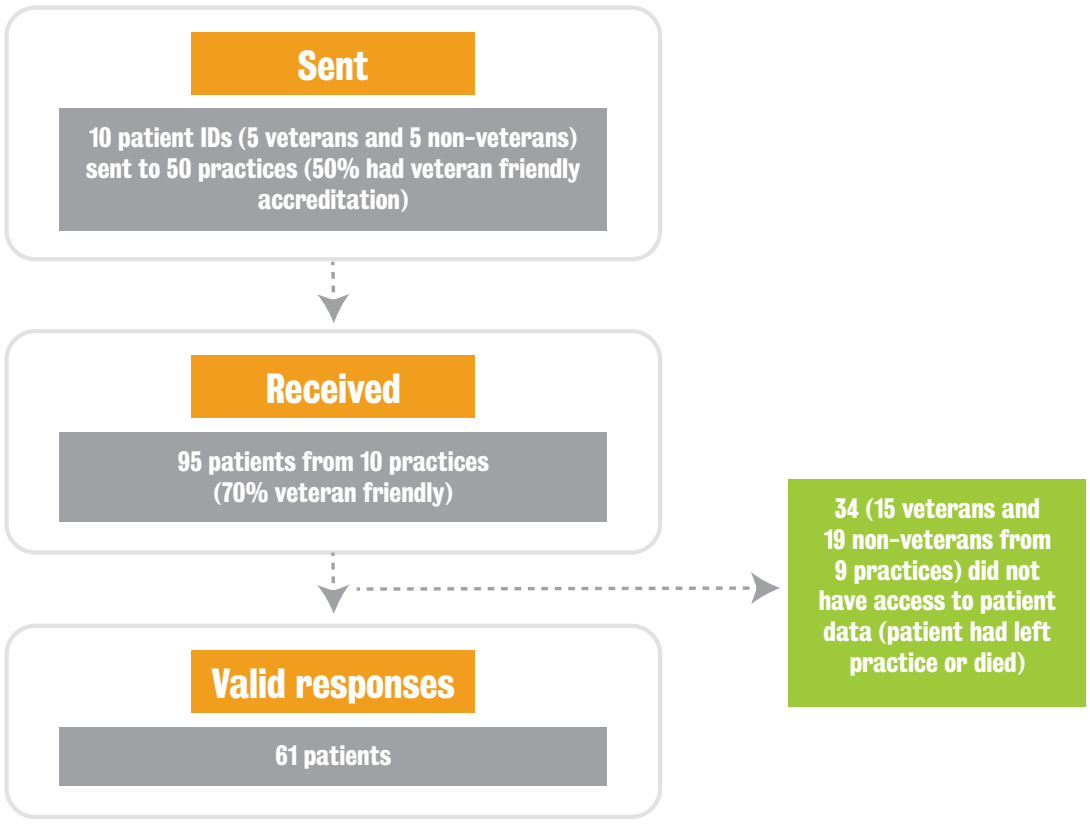


Table 3 shows how well veteran status in CPRD matched what GPs reported. While most veterans were correctly identified in both sources, 31% (n=11) of patients classified as veterans in CPRD were reported by GPs as non-veterans. However, all 11 false positives contained definite terms related to military service in CPRD (e.g., 'left military service' and 'military veteran'), suggesting these are likely to be genuine veterans. All non-veterans were correctly classified by both sources.

Within this small sample, agreement between CPRD and GP records appeared substantial, with a Cohen's Kappa score of 0.65. However, because the sample was limited, these results should be interpreted with caution.

We saw no strong evidence that veteran-friendly practices recorded veteran status more accurately, although they were more likely to respond to our request.

Additionally, a practice level question was included which enquired on the total estimated number of patients who were known to be veterans in their GP practice. Half of the practices reported that the estimated number of veterans registered in their practice ranged between 100 to 500 (see Table 4 for results). Only one non-accredited practice reported having over 1,000 veterans registered.

Table 3. Agreement between CPRD and GP records on veteran status

	GP Positive N (%)	GP Negative N (%)	Row total N
CPRD Positive	24 (69%)	11 (31%)	35
CPRD Negative	0 (0%)	26 (100%)	26
	24	37	61

Table 4. Estimated number of veterans registered in GP practice

	Not accredited	Veteran friendly	Total
Bandings	N	N	N
25 - <50	1	1	2
75 - <100	0	1	1
100 - <500	1	4	5
500 - <1,000	0	1	1
>1,000	1	0	1
Total	4	7	10

Only bandings with response are reported

Data quality and data management challenges

The validation study helped identify several challenges with the quality and management of data related to veteran status in CPRD.

Gaps and uncertainty in veteran status recording

In about one-third of cases where CPRD identified someone as a veteran, the GP said the patient was not a veteran. This could be because GPs did not have time to look through the full records, or because the information was not easy to find. However, all of these "false positives" had definite military-related terms in CPRD, which suggests they were likely veterans.

Low GP response rate and difficulty accessing records

Despite using the CPRD PROVE service and offering incentives to participating practices, the study achieved a low response rate, with only 10 of 50 practices responding. Of the 95 patient IDs returned, 34 were unusable due to restricted access to records for patients who had died or moved practices. This likely reflects the heavy workload in general practice and indicates that future validation studies may encounter similar challenges.

No standardised code for veteran status in CPRD

There is no specific, standard code for identifying veterans in primary care (Finnegan & Randles, 2023b). Instead, we relied on a list of possible military-related medical codes to determine veteran status for patients who have disclosed being veterans (see section 2.3). While this approach required careful code selection and review, it proved highly effective since 95% of our veteran sample was identified using definite terms. This minimises the likelihood of incorrectly classifying non-veterans as veterans. However, some veterans may not have such terms recorded in their primary care data and would be misclassified as non-veterans.

Summary

The validation exercise showed that there was substantial agreement between CPRD and the data recorded in the underlying GP records. However, one-third of veterans identified in CPRD were not confirmed by GPs, possibly due to difficulty locating the relevant information. Completeness was limited by practical barriers such as low GP response rates and restricted access to patient records for patients who had died or moved practices.

3.3 RQ3: Comparative health and demographic profiles of veterans and non-veterans (CPRD Aurum)

To what extent do the socio-demographic, mental health, and physical health profiles of veterans accessing primary care differ from those of the general primary care population?

Health risk factors

Figure 13 compares health risk factors between veterans and non-veterans. It is organised into four panels, each indicating the prevalence of a different risk factor: alcohol consumption severity, smoking status, blood pressure, and BMI.

Alcohol consumption severity

Alcohol consumption was assessed by examining scores in the Alcohol Use Disorders Identification Test (AUDIT), the Fast Alcohol Screening Test (FAST), and the Severity of Alcohol Dependence questionnaire (SADQ). Alcohol consumption data were available for 47.33% of veterans and 15.99% of controls.

Veterans were more likely to be low-risk drinkers (13.98% vs 5.51% non-veterans) and slightly more likely to report hazardous use (2.17% veterans vs 0.99% non-veterans), however, harmful use (12.45% veterans vs 17.69% non-veterans) and alcohol dependence (70.87% veterans vs 75.81% non-veterans) were more prevalent among non-veterans than veterans.

The finding that the majority of veterans and non-veterans fall into the "alcohol dependence" category compared to the other three categories might reflect selection and recording biases. Alcohol assessments are generally recorded in response to clinical concerns, meaning that individuals perceived to have problematic alcohol use are more likely to be assessed and therefore captured in the dataset, potentially introducing selection bias. Similarly, help-seeking behaviour plays a crucial role. Those experiencing more severe alcohol problems are likely to present to primary care and to be screened. Thus, the sample with alcohol severity data might represent a clinically identified subset of patients rather than representing the broader population of veterans and non-veterans.

Smoking status

For smoking status, compared to non-veterans, veterans were less likely to have never smoked (veterans 45.66% vs 51.74% non-veterans) and more likely to be ex-smokers (33.87% veterans vs 27.28% non-veterans), with current smoking rates similar between groups (~21%).

Blood pressure

Blood pressure measurements showed veterans were slightly more likely than non-veterans to fall in the pre-hypertensive range (Diastolic 80–89 or Systolic 120–139: 48.22% veterans vs 45.82% non-veterans) but to have similar prevalence of hypertension (Diastolic ≥ 90 or Systolic ≥ 140: 33.71% veterans vs 34.24% non-veterans).

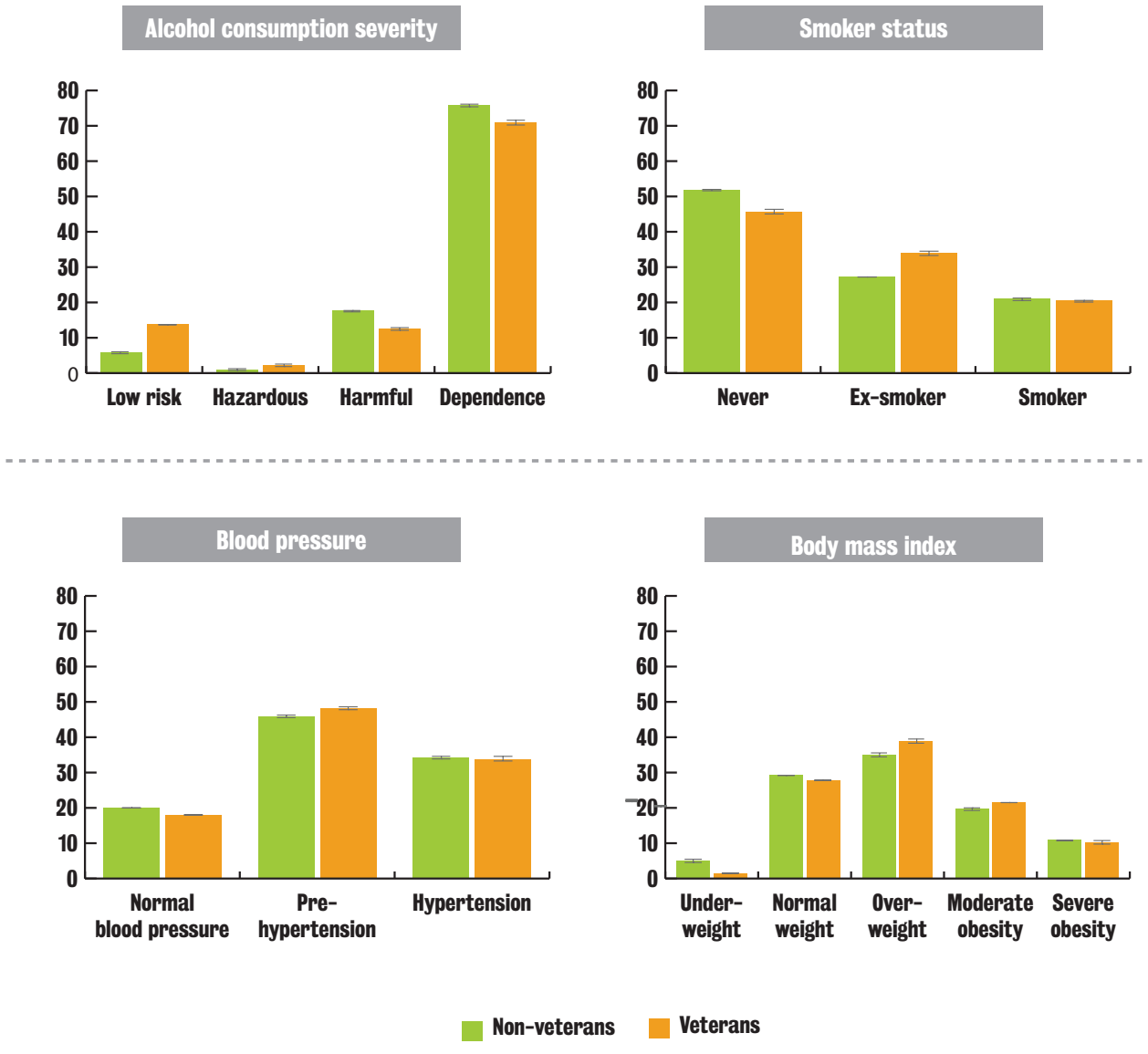
Body mass index

For BMI, veterans were less likely than non-veterans to be underweight (BMI ≤18.49: 1.48% veterans vs 5.15% non-veterans) and more likely or be overweight (BMI 25 to 29.99: 38.86% veterans vs 35.01% non-veterans) or have obesity (BMI 30 to 34.99: 21.50% veterans vs 19.74% non-veterans). The prevalence of severe obesity was similar for veterans and non-veterans (BMI ≥35: 10.32% veterans vs 10.87% non-veterans).

Missing data

Across all four domains of health risk factors, veterans had more complete health data recorded than non-veterans, suggesting more complete health monitoring in this group. For example, missing data on alcohol consumption was 52.67% among veterans compared to 84.01% among non-veterans. Similarly, 32.45% of veterans were missing BMI data versus 65.04% of non-veterans; 20.89% of veterans were missing smoking status compared to 57.50% of non-veterans; and 26.66% of veterans were missing blood pressure measurements compared to 54.69% of non-veterans. Thus, comparisons may reflect differences in recording practices rather than true differences in health. For example, if non-veterans are missing information on health risk factors, the non-veterans who do have this information are more likely to be those with health concerns, which could exaggerate differences between the two groups. This means that any observed patterns should be interpreted with caution, as they may be influenced by the way information is captured in primary care records rather than underlying health status alone.

Figure 13. Veteran and non-veteran health risk factors



Error bars represent 95% confidence intervals.

Prevalence of mental and physical health conditions in veterans and non-veterans

Figure 14 shows the prevalence of several physical and mental health conditions amongst veterans and matched non-veterans. **Due to the matching, these estimates are adjusted for age, gender, GP practice and index date. No other variables are included in the estimations presented in this section**. All observed differences were statistically significant.

Physical health conditions

The most prevalent conditions in both veterans and non-veterans were lower back pain (11.72% veterans vs 6.53% non-veterans) and osteoarthritis (9.94% veterans vs 4.67% non-veterans), followed by hearing loss (6.78% veterans vs 3.08% non-veterans), chronic obstructive pulmonary disease (COPD) (5.92% veterans vs 2.87% non-veterans), and coronary heart disease (CHD) (5.46% veterans vs 2.79% non-veterans). In general, these conditions were more common among veterans than non-veterans. Although other health conditions were comparatively less prevalent in both groups, their rates were almost doubled among veterans compared to non-veterans, including prostate cancer (2.39% veteran vs 1.19% non-veterans), myocardial infarction (heart attack) (1.85% veterans vs 1.02% non-veterans), Alzheimer's disease (1.82% veterans vs 0.65% non-veterans), and breast cancer (0.43% veterans vs 0.29% non-veterans).

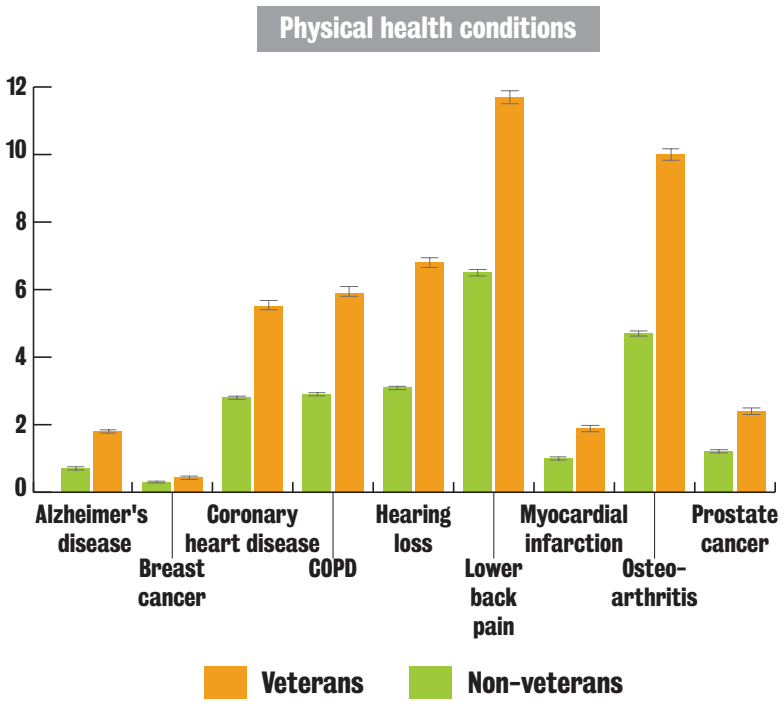
Mental health conditions

Mental health conditions were more common among veterans, including depression (10.78% veterans vs 5.71% non-veterans), anxiety (8.98% veterans vs 5.86% non-veterans), and PTSD (3.06% veterans vs 0.18% non-veterans).

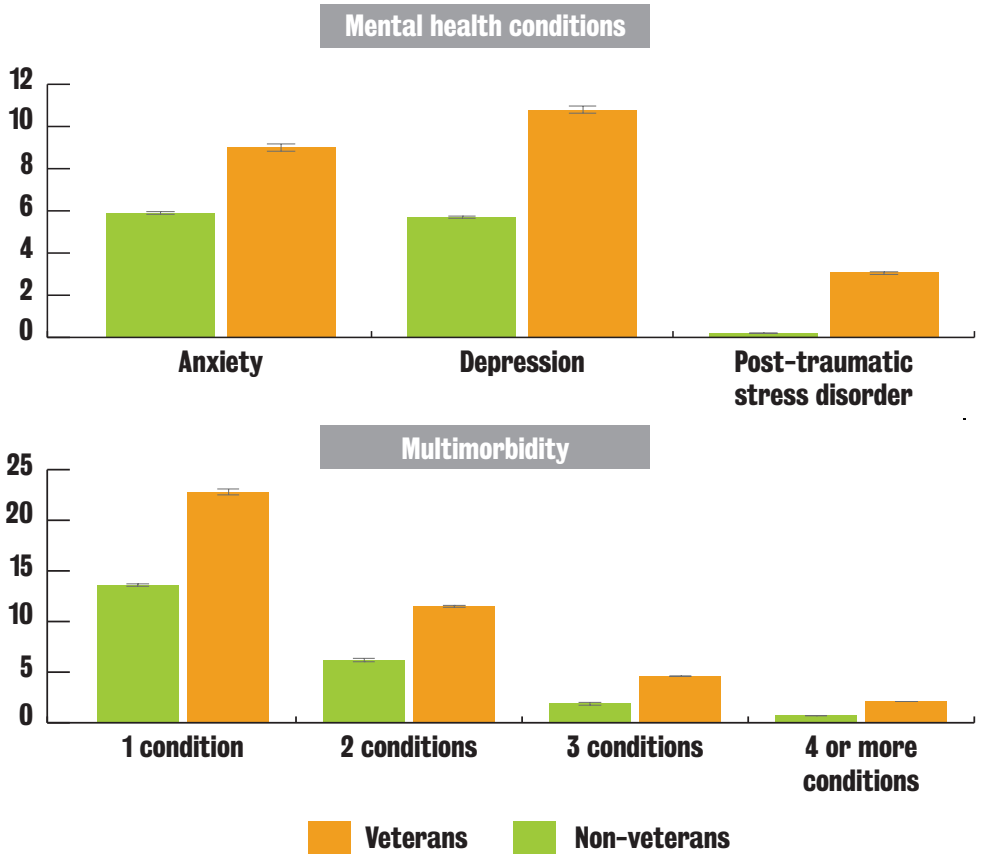
Multimorbidity

Multimorbidity was substantially more common in the veteran group, with 11.53% of veterans having two conditions versus 6.20% in non-veterans, and with 2.01% of veterans having four or more physical or mental health conditions versus 0.70% in non-veterans.

Figure 14. Physical and mental health of veterans and non-veterans, including multimorbidity



Error bars represent 95% confidence intervals.



Impact of veteran status on health outcomes: adjusted regression analysis

Poisson regressions with robust standard errors were used to estimate prevalence ratios (PR) across the imputed datasets. See the Appendix for demographic characteristics and prevalence ratios for the sample with complete data.

As presented in Table 5, veteran status was consistently associated with an increased prevalence of all physical and mental health conditions examined. Adjustment for ethnicity and further adjustment for BMI and smoking attenuated most associations. In the fully adjusted models, veterans had higher prevalence of COPD (PR=1.83, 95% CI 1.77-1.89), lower back pain (PR=1.74, 95% CI 1.70-1.78), myocardial infarction (PR=1.66, 95% CI 1.57-1.76), and osteoarthritis (PR=1.99, 95% CI 1.94-2.04). Similarly, elevated prevalence of depression persisted after adjustment (PR=1.83, 95% CI 1.79-1.88). The association with PTSD remained higher than for other conditions, with veterans showing substantially higher prevalence of PTSD even after adjustment (PR=16.43, 95% CI 14.89-18.13). However, this higher prevalence of PTSD may partly reflect trauma-related diagnostic overshadowing[12], which occurs when healthcare professionals misattribute symptoms to other aspects of a patient's clinical presentation (Wislocki & Zalta, 2025).

[12]Diagnostic overshadowing occurs when a clinician attributes a patient's symptoms to an existing diagnosis, characteristic, or background factor, leading to missed or incorrect identification of other underlying conditions

Summary

Compared to non-veterans, veterans were less likely to consume alcohol at harmful or dependent levels, less likely to have never smoked, and more likely to be ex-smokers. They also tended to have higher BMI and pre-hypertensive blood pressure levels, as well as more complete recording of health risk factors which may bias some of these observations. Additionally, veterans had consistently higher prevalence of physical health conditions, including lower back pain, osteoarthritis, hearing loss, COPD, and coronary heart disease. Mental health conditions were more common among veterans and multimorbidity was more prevalent compared to non-veterans.

Table 5. Poisson regression for the association between veteran status and physical and mental health conditions (N= 367,057)

	Model 1			Model 2		
	PR	95%	CI	PR	95%	CI
Chronic obstructive pulmonary disease	2.06	1.99	2.12	1.99	1.93	2.05
Depression	1.89	1.85	1.93	1.86	1.81	1.90
Lower back pain	1.79	1.76	1.84	1.79	1.75	1.83
Myocardial infarction	1.81	1.71	1.91	1.76	1.66	1.86
Osteoarthritis	2.13	2.08	2.18	2.08	2.03	2.14
Post-traumatic stress disorder	17.08	15.47	18.85	16.84	15.26	18.59

	Model 3		
	PR	95%	CI
Chronic obstructive pulmonary disease	1.83	1.77	1.89
Depression	1.83	1.79	1.88
Lower back pain	1.74	1.70	1.78
Myocardial infarction	1.66	1.57	1.76
Osteoarthritis	1.99	1.94	2.04
Post-traumatic stress disorder	16.43	14.89	18.13

PR: Prevalence Ratios; CI: Confidence Interval

Model 1: Unadjusted; Model 2: Model 1 + ethnicity; Model 3: Model 2 + BMI and smoking

3.4 RQ4: Benefits of data linkage for enhancing veteran health profiles (CPRD Aurum)

What added benefit does linking CPRD Aurum with other datasets provide in terms of improving detection of physical and mental health needs and enhancing the completeness of key socio-demographic characteristics?

Index of Multiple Deprivation

IMD data was available for 264,696 patients, representing 72% of the full sample of veterans and non-veterans. Following the exclusion of individuals with missing data in IMD, some matched pairs were disrupted, resulting in an unmatched sample of 264,696 individuals. The matched sample included 53,734 individuals who remained matched after listwise deletion based on IMD data availability. Table 6 presents the distribution of individuals across quintiles of the IMD for the veteran and non-veteran groups. Since the matching included GP practice, and GP practices have a fixed geographic location which itself may be situated in an area with higher or lower deprivation, this location can act as a proxy for the patient IMD. As a result, the proportions are relatively similar between groups across all quintiles, suggesting a comparable socioeconomic profile.

Hospital Episode Statistics Admitted Patient Care

From the total sample (N=367,057), HES APC data was available for 212,730 (58%) veterans and non-veterans. This data was only available for practices that have agreed to the linkage and for patients who have been to hospital.

Demographic factors (CPRD Aurum)

As mentioned earlier, key demographic factors (e.g., gender and age) are well recorded in the CPRD Aurum primary care data and as such, no missingness was identified (see Table 1, page 33). However, missing data in Aurum was identified in the ethnicity variable which is a key variable to consider when investigating health outcomes. Similar to the health risk factors (see section 3.3), veterans had less missing data in CPRD Aurum than non-veterans, possibly reflecting the higher consultation rates over the follow-up which might increase the likelihood of variables such as health risk factors or ethnicity being recorded.

Table 6. Veterans and non-veterans with IMD data

Index of Multiple Deprivation (quintiles)	Unmatched sample				Matched sample			
	Non-veteran		Veteran		Non-veteran		Veteran	
	N	%	N	%	N	%	N	%
1 (least depr.)	39,767	21.85	18,216	22.02	6,350	21.60	5,050	20.75
2	37,183	20.43	16,968	20.51	5,912	20.11	4,656	19.13
3	33,704	18.52	15,227	18.41	5,348	18.19	4,457	18.32
4	33,654	18.49	15,425	18.65	5,537	18.83	4,636	19.05
5 (most depr.)	37,672	20.70	16,880	20.41	6,253	21.27	5,535	22.75
Total	181,980		82,716		29,400		24,334	

Note: The unmatched sample (N=264,696) includes patients whose matching was disrupted due to missing data on the Index of Multiple Deprivation. The matched sample (N=53,734) includes only those individuals who remained matched (1:1 or 2:1) after excluding cases with missing IMD data.

Table 7. Comparison of ethnicity data availability in CPRD Aurum before and after backfilling with HES APC

Ethnicity	CPRD Aurum				CPRD Aurum + HES APC			
	Non-veteran		Veteran		Non-veteran		Veteran	
	N	%	N	%	N	%	N	%
White	162,864	66.59	99,381	81.14	183,371	74.98	104,166	85.04
Black	4,093	1.67	1,798	1.47	4,379	1.79	1,853	1.51
Asian	11,224	4.59	2,986	2.44	11,839	4.84	3,096	2.53
Mixed	2,802	1.15	924	0.75	3,089	1.26	957	0.78
Other	18,909	7.73	8,357	6.82	19,256	7.87	8,414	6.87
Missing	44,681	**18.27**	9,038	**7.38**	22,639	**9.26**	3,998	**3.26**

Demographic factors (CPRD Aurum + HES APC)

When CPRD Aurum ethnicity data is backfilled using HES APC data, we gained 27,082 observations (5,040 veterans and 22,042 non-veterans). This reduces the overall missing on ethnicity from 14.64% to 7.26% (Table 7).

Health risk factors

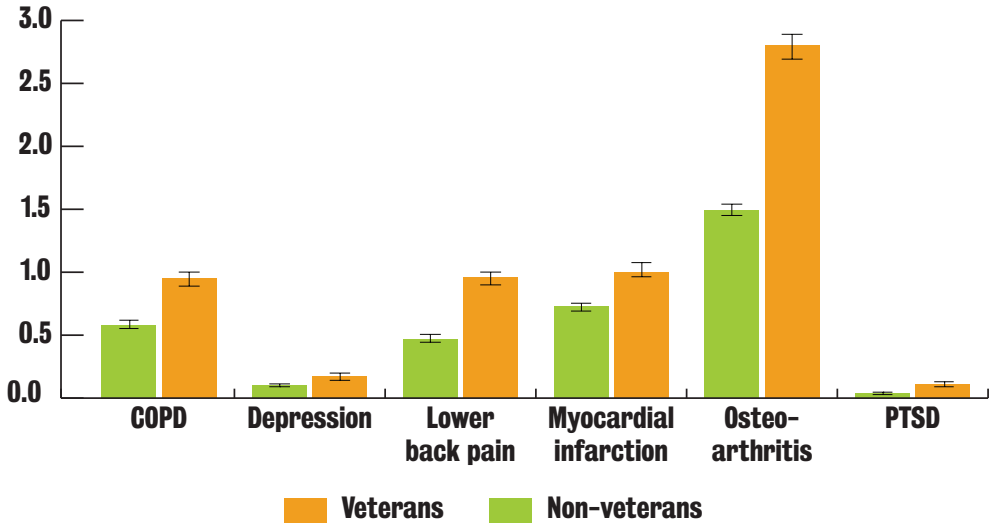
While HES APC data can be a useful supplementary source for enhancing the completeness of CPRD Aurum primary care data, particularly for health outcomes where hospitalisation records offer clear diagnostic indicators, it was found to be less suitable for backfilling information on certain risk factors. In this context, a key limitation is that HES APC data relies on ICD-10 codes, which are primarily designed to capture diagnoses that justify hospital admission or treatment. As such, they do not systematically record routine clinical measurements or behavioural risk factors.

Variables such as BMI, smoking status, blood pressure, and alcohol use are either poorly captured, only recorded in extreme or clinically significant cases, or absent altogether.

Health outcomes

Veterans consistently showed higher hospital episode rates than non-veterans, with the most pronounced differences observed for osteoarthritis, where the veteran rate is nearly double that of non-veterans (2.80% veterans vs 1.49% non-veterans) (Figure 15). Smaller differences are also seen for depression (0.17% veterans vs 0.10% non-veterans) and PTSD (0.11% veterans vs 0.04% non-veterans), with veterans still exhibiting slightly higher rates. Similar to the findings using primary care data (section 3.3), these results suggest that veterans experience greater hospitalisation for a range of physical and mental health conditions, particularly musculoskeletal and cardiovascular issues.

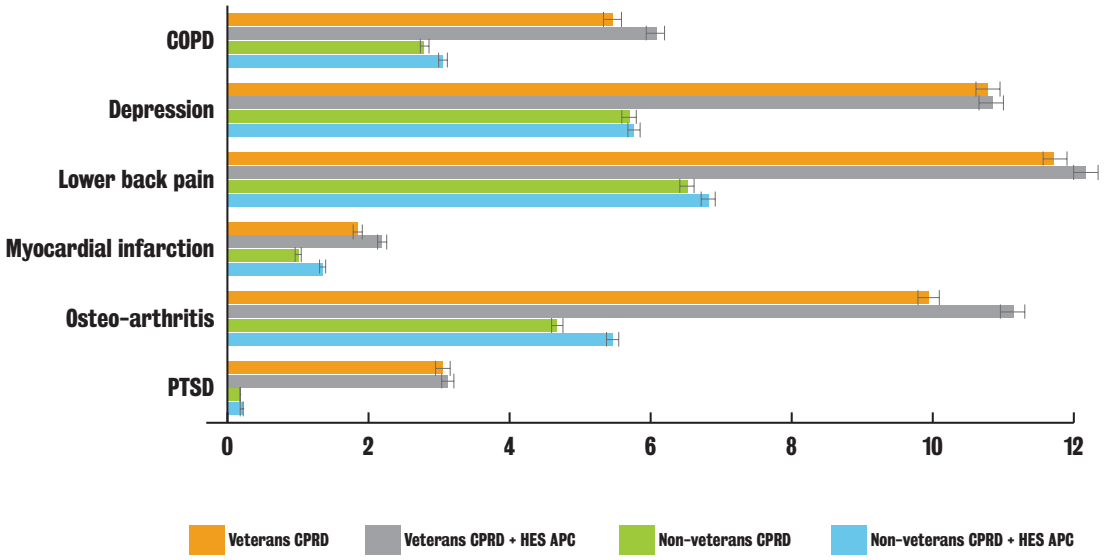
Figure 15. Diagnoses in Hospital Episode Statistics Admitted Patient Care dataset for veterans and non-veterans



Error bars represent 95% confidence intervals.

Across nearly all conditions, and particularly for physical health conditions, the inclusion of HES APC data increases the observed prevalence compared to primary care data alone (Figure 16). The largest differences are seen for lower back pain and osteoarthritis where veterans have substantially higher prevalence in both primary care and primary care combined with HES APC data. For example, the prevalence of osteoarthritis in veterans is 11.15% when including HES APC data, compared

Figure 16. Comparison of diagnoses between veterans and non-veterans in CPRD only and in CPRD linked to HES APC secondary care data



Error bars represent 95% confidence intervals.

to 9.94% in primary care alone. The prevalence of osteoarthritis in non-veterans is 5.47% when including HES APC data, compared to 4.67% in primary care data alone.

There was little change observed in the mental health variables, such as PTSD and depression, where the inclusion of HES data made negligible increases in rates of diagnosis. This could reflect the low rate of mental health data in the HES APC dataset (Degli Esposti et al., 2022; Royal College of General Practitioners, 2017). Overall, this figure underscores the added value of linked secondary care data in capturing health conditions, and it suggests that veterans face a higher prevalence of both physical and mental health diagnoses compared to their non-veteran counterparts. This profile may also be more severe given that HES APC data reflects individuals whose conditions require hospitalisation.

Impact of veteran status on health outcomes: adjusted regression analysis using linked datasets

The results of regressions (see Table 8) showed that when hospital data is included, the estimates of the association between veteran status and health conditions remained largely consistent to those found in primary care data only (see Table 5) in terms of direction and magnitude. For the association between veteran status and PTSD, the PR was lower in the combined dataset (CPRD + HES) than in the CPRD only dataset (13.99 vs 16.43).

Table 8. Poisson regression for the association between veteran status and physical and mental health conditions using CPRD Aurum and HES APC data (N= 367,057)

	Model 1			Model 2		
	PR	95%	CI	PR	95%	CI
Chronic obstructive pulmonary disease	1.99	1.93	2.06	1.92	1.86	1.98
Depression	1.88	1.84	1.93	1.85	1.81	1.89
Lower back pain	1.79	1.75	1.82	1.78	1.74	1.81
Myocardial infarction	1.62	1.54	1.71	1.58	1.50	1.66
Osteoarthritis	2.04	1.99	2.08	1.99	1.95	2.04
Post-traumatic stress disorder	14.50	13.25	15.88	14.30	13.06	15.65

	Model 3			Model 4		
	PR	95%	CI	PR	95%	CI
Chronic obstructive pulmonary disease	1.77	1.72	1.83	1.77	1.72	1.83
Depression	1.83	1.79	1.87	1.83	1.79	1.87
Lower back pain	1.73	1.69	1.77	1.73	1.69	1.77
Myocardial infarction	1.50	1.42	1.58	1.50	1.42	1.57
Osteoarthritis	1.90	1.86	1.95	1.90	1.86	1.95
Post-traumatic stress disorder	13.99	12.77	15.32	13.96	12.75	15.29

PR: Prevalence Ratios; CI: Confidence Interval

Model 1: Unadjusted association; Model 2: Model 1 + ethnicity; Model 3: Model 2 + BMI and smoking; Model 4: Model 3 + Index of Multiple Deprivation

Summary

Linking CPRD Aurum with IMD and HES APC data improved completeness of ethnicity data and enhanced detection of health conditions, particularly those requiring hospitalisation. Additional cases of musculoskeletal, cardiovascular, and few mental health conditions were identified using the linked data, with veterans consistently showing higher prevalence than non-veterans. Regression results were broadly consistent with the analysis using CPRD data only.

3.5 RQ5: Challenges in data quality and management for veteran identification and characterisation (CPRD Aurum and Gold)

What are the key data quality and data management challenges associated with identifying and characterising the physical and mental health needs of veterans in primary care?

Data quality

CPRD is a high-quality source of health data that is widely used in research (Clinical Practice Research Datalink, 2025a). It contains information about patients symptoms, tests results, diagnoses, referrals and prescriptions, capturing real-world care across the UK over time. To help ensure reliability, CPRD validates the quality of the data and only practices that reach an acceptable standard are included (Herrett et al., 2015; Wolf et al., 2019).

However, this information is collected for routine care and not for research purposes, which means that recording practices can vary between general practices or over time. National initiatives like the Quality and Outcomes Framework (QoF) encourage GPs to record key health information in a standardised way (NHS Digital, 2025b). However, coding practices in UK primary care have been shaped by QoF incentives, so items linked to QoF indicators tend to be more completely and consistently recorded than those outside of QoF. Furthermore, health conditions will not be recorded if patients do not consult their GP about them, as diagnosis in primary care are based on clinical assessment during consultations or following specialist referrals. Therefore, the absence of a diagnosis in the record may reflect patients not seeking care or delays in diagnosis, rather than the absence of the condition. Researchers must be careful when identifying conditions and populations, such as veterans, and be aware that some issues (e.g., missing data or misclassification) may in fact be present (Herrett et al., 2015; Wolf et al., 2019).

Missing information

Missing data in primary care records are a well-recognised challenge. Unless part of QoF, non-routine data such as lifestyle factors tend to be recorded only when clinically relevant rather than systematically across all consultations. However, data completeness has improved over time for some variables such as smoking status and BMI. Data completeness also depends on the patient's characteristics and healthcare engagement. For example, data on lifestyle factors is more complete among women, patients with chronic conditions, and individuals with more frequent consultations (Bhaskaran et al., 2013; Nicholson et al., 2019; Petersen et al., 2019). Even where recording is relatively high (e.g., blood pressure or smoking status), completeness may vary across subgroups and over time (Bhaskaran et al., 2013; Dhalwani et al., 2013).

In line with this, we found that veterans had a lower proportion of missing data compared to non-veterans. There are various reasons for this, for example, we found that veterans had higher consultation rates compared to non-veterans, and as such it is more likely that they will have additional information recorded (e.g., ethnicity, BMI, or smoking status). Alternatively, veterans may be more likely to attend for regular health screenings offered by the NHS compared to non-veterans, as they are used to undergoing regular, compulsory health screening during their time in the military. Although some of this information could be backfilled using linked datasets, for example supplementing CPRD ethnicity data

with what is captured in HES APC to improve completeness, some health risk factors could not be supplemented because they are poorly recorded or missing in secondary care datasets (see section 3.4).

For this report, we handled missing data using multiple imputation because complete-case analyses may introduce bias, particularly when missingness correlates with both exposure and outcome and because multiple imputation is recognised as the most appropriate method to handle missing data in primary care databases (Tolani et al., 2025). Evidence from studies using UK primary care data shows that multiple imputation outperforms other methods in reducing bias, particularly for key variables like BMI, smoking, and ethnicity, which are often incompletely recorded and correlated with consultation frequency and health status (Petersen et al., 2019; Tolani et al., 2025).

Inconsistent recording over time

Most veteran information in CPRD appears from 2010 onwards, likely because GPs only started being encouraged to record military service in recent years (Finnegan et al., 2022; Finnegan & Randles, 2023b). This means older records may not capture veteran status, making it harder to track long-term trends. Additionally, although CPRD Aurum data is available from 1987 onwards, the completeness of recorded information is considered more reliable from 1995 onwards (Wolf et al., 2019).

The accuracy of death data in CPRD has also improved over time. Exact agreement between CPRD and ONS mortality records on the date of death rose from 53% in 1998 to nearly 99% in 2013 (Gallagher et al., 2019).

Outcome code lists

CPRD does not provide standardised definitions for variables and therefore, researchers must create bespoke code lists or algorithms tailored to their study's aims to define populations, outcomes, exposures, and covariates. This method may result in differences in how key variables are operationalised across studies, which could contribute to variations in study exposures or outcomes (Matthewman et al., 2024). For this project we used medical codes from previous research on veterans in secondary mental healthcare, reviewed relevant terms from the literature records (Leightley et al., 2023; Mark et al., 2020), and searched CPRD's medical dictionaries for military or veteran-related terms. The final list was reviewed by the PAB (see Section 2.1) to ensure accuracy.

Data management

Problems linking different datasets

Although CPRD can be linked with HES APC and IMD data, not all GP practices contributing to CPRD have consented to the linkage. This means that analyses using linked datasets can only be carried out on a sub-sample of participants.

Some planned data linkages for this project could not be completed due to various factors. Identifiability concerns prevented the linkage of Rural-Urban Classification alongside IMD data. Data availability was also a limiting factor, with secondary mental health data no longer accessible via CPRD due to concerns over data quality. Additionally, although linkage to Cancer Registry was originally planned, this could not be achieved within the project timeline, since at the time of these analyses, it required 18 months from the point of application to delivery.

No central flag for veterans

The UK's Department of Health and Social Care indicates that the medical documentation of veterans should include a code indicating a "history relating to military service" (Royal College of General Practitioners, 2011). However, there are multiple military codes available (e.g., military veteran, left military service, history relating to military service, army veteran) (Finnegan et al., 2018). Although considerable work has gone into standardising terminology, for example by deprecating several codes in the terminology browser and issuing guidance, these steps do not correct legacy records.

Matching

To maximise statistical power and leverage the large CPRD patient pool, we initially aimed to match four non-veterans per veteran. However, the strict criteria, particularly matching on general practice and index date, limited eligible matches, especially in the oldest age groups (80 years and older). Consequently, most matches were achieved at a 2:1 ratio with a smaller subgroup matched 1:1 due to limited eligible non-veterans, particularly at the youngest and oldest ages.

Summary

Although CPRD is a robust, high-quality dataset that can be used to follow patients up over time, variability in GP recording practices, missing data on key variables (e.g., ethnicity, health risk factors), and inconsistent documentation of veteran status, especially before 2010, limit completeness and comparability over time. Veterans generally had fewer missing data than non-veterans, likely due to higher consultation rates and engagement with routine health checks, but certain risk factors remained poorly recorded even after linking with other datasets. Data linkage was further constrained by practice-level consent and data availability. Time and resource constraints also left some planned linkages incomplete due to lengthy approval processes. Lastly, demographic differences between veterans and non-veterans limited the intended matching ratio in some age groups.

3.6 RQ6: Clinical and research potential of CPRD for veterans' health (CPRD Aurum and Gold)

What clinical practice, policy and research opportunities are enabled by using the CPRD and linked datasets in the context of veterans' health, and what are the strengths and limitations of these data sources?

Clinical practice and policy opportunities

Insights from CPRD data can inform service planning and delivery, highlighting areas where veterans may have higher health needs or face barriers to care. For example, identifying differences in the prevalence of conditions such as PTSD, depression, or cardiovascular disease between veterans and the general population can support targeted interventions and guide resource allocation. CPRD data can be used to identify patterns of physical and mental health need and service utilisation among veterans, which could support more targeted delivery of veteran services such as Op COURAGE (NHS, 2024b), Op RESTORE (NHS, 2024a) and OVA's initiatives such as VALOUR (Ministry of Defence & Office for Veterans' Affairs, 2025). For instance, our finding that veterans have higher rates of depression and PTSD than non-veterans can support VALOUR's aim to strengthen coordination of mental health services, ensuring timely access to assessment, treatment, and specialist support tailored to veterans' needs, in line with the Armed Forces Covenant. It also offers the potential to monitor and optimise initiatives like the veteran friendly accreditation scheme (Armed Forces Covenant, 2025), whose impact could be evaluated beyond raising awareness, ensuring identification leads to tangible improvements in care and service access.

CPRD can provide information into health conditions that emerge or persist after military service, providing insights to refine pre-discharge assessments and post-service support, which would be relevant for the Ministry of Defence. The Office for Veterans' Affairs could use these findings to inform national policy decisions and guide resource allocation. Furthermore, considering the development of the next version of the Defence People Health and Wellbeing Strategy in 2026, there is interest within both the Ministry of Defence and the Office for Veterans' Affairs in this work. Additionally, military charities, such as the Royal British Legion, Help for Heroes, and Combat Stress, could use the evidence to inform their support offers, and advocate for specific healthcare interventions and associated funding.

Research opportunities

The identification of veterans in the CPRD dataset (i.e., Aurum and Gold) enables research opportunities to address critical evidence gaps in UK veteran health. This large cohort includes veteran data from across the UK allowing the creation of detailed health profiles for veterans from each nation, allowing for meaningful cross-country comparisons. Furthermore, for England, CPRD Aurum allows linkage to COVID-19, small area level data, NHS England data (e.g., HES APC, outpatient, accident and emergency, diagnostic), NHS England National Disease Registration Service (cancer diagnosis and treatment) and mortality data.

CPRD data supports robust analyses of physical and mental health outcomes, patterns of multimorbidity, and healthcare utilisation across diverse veteran subgroups, including women and ethnic minorities who are often underrepresented in research. This resource also facilitates exploration of service-related health trajectories, early detection of conditions, and evaluation of policy initiatives, ultimately improving healthcare equity for veterans.

Strengths and limitations

CPRD (Aurum and Gold) offers access to large-scale, real-world data from millions of patients across the UK, allowing for the study of long-term trends in health and healthcare use (Herrett et al., 2015; Wolf et al., 2019). Additionally, CPRD is widely regarded as a leading resource for health research and has supported thousands of peer-reviewed publications (Clinical Practice Research Datalink, 2025a). As the first point of contact for most health concerns, primary care data enables investigation into both physical and mental health outcomes. This is especially valuable in the context of veterans, who may present with a range of complex and co-occurring physical and mental health conditions.

Before this project, CPRD had not been used for veteran research with only few studies using primary care data (Finnegan & Randles, 2023a; Finnegan & Salem, 2024). Therefore, identifying veterans in CPRD offers the opportunity to address key evidence gaps by facilitating national-level research on the health of the UK veteran population. The ability to link CPRD data to other sources further enhances the richness of these analyses, helping to build a more complete picture of veterans' needs. This work marks substantial progress in understanding the health needs of veterans and sets a new benchmark for future research in this area.

However, these strengths must be considered alongside certain limitations. CPRD data are collected during routine care and not specifically for research, which means some details may be inconsistently recorded or missing. Moreover, while CPRD provides coverage across the UK, linkage to other datasets is mainly available for England; and CPRD Aurum's geographical coverage of England may not be fully representative of the national veteran population (see section 3.1), potentially limiting the generalisability of findings from studies using this dataset.

Additionally, recording practices related to health conditions, health risk factors and veteran status, can vary between general practices or over time. The use of code lists to determine health diagnoses and health risk factors can reduce consistency and affect comparability with other studies using code lists that might be more sensitive or specific (Matthewman et al., 2024).

Summary

The identification of veterans in the CPRD dataset represents an innovative step forward in veterans' health research, leveraging an unrivalled primary care dataset to provide robust comparative evidence between veterans and matched non-veterans. Thus, CPRD offers the opportunity to examine veterans' health at a national level, enabling detailed analysis of physical and mental health outcomes, healthcare use, and cross-country comparisons. These insights can inform clinical practice, policy, targeted intervention development and service evaluations.

4. Recommendations for future use of CPRD for veteran health research and implications for key stakeholders

Improving veteran identification in primary care

Our findings showed that veterans can be identified successfully in UK primary care records using a set of military service-related terms. Thus, GP practices should be encouraged to become veteran friendly and integrate a nationally consistent veteran term into veterans' EHRs to ensure they are identifiable in primary care and in CPRD.

Despite the existence of targeted NHS services for veterans such as Op COURAGE (NHS, 2024b), incomplete identification of veterans in GP records means some patients may miss out on appropriate referrals. Considering this, veteran leadership roles, such as a designated veteran lead GP or practice staff with military experience, could help build trust with veteran patients, increase likelihood of veterans' self-identifying as such, and improve the frequency with which patients are asked if they have ever served in HM Armed Forces. Practices should be encouraged to know how many veterans are in their patient population, to understand their demographic and health profiles, and assess GP awareness of existing veteran-specific NHS services so they can be appropriately signposted based on their needs.

Informing policy and prevention strategies

Our findings suggested differences in health risk factors and health needs of veterans compared to non-veterans. Further research on these differences should inform national policy updates, such as the Defence People Health and Wellbeing Strategy, the NHS Mental Health Strategy and the Men's Health Strategy, by ensuring they reflect the specific needs of the veteran population. These findings should also inform NHS service planning by identifying priority areas for resource allocation ensuring interventions are tailored to address the specific health burdens within veteran populations.

Additionally, evidence on health risk factors specific to veterans can point to opportunities for prevention and lifestyle interventions before health problems develop or worsen. Including prevention and intervention strategies during active service or immediately after leaving service could help reduce the risk of chronic disease after discharge, improve quality of life, and lower the future healthcare burden in both Defence and NHS settings.

Using CPRD and linked datasets for veteran health research

Our study demonstrated that while CPRD provides a rich source of primary care data, future veteran health research can be significantly enhanced through greater linkage with complementary datasets. Consideration should be given to repeating similar studies in the future, for example in 10 years' time, when a larger cohort of veterans will be identifiable within CPRD Aurum. Furthermore, since Aurum records are more complete from 1995 onwards, such repeated analyses would offer a stronger basis for examining trends over time.

GP practices contributing to CPRD should be encouraged to consent to data linkage to enhance the representativeness of the linked data. Furthermore, linking to mental health, hospital, and other administrative datasets would enable more complete understanding of social determinants of health, multimorbidity, and health behaviours. Extending primary care linkages to datasets covering education, employment, housing, and social relationships could also help to address evidence gaps around the wider social outcomes of military service.

To ensure national representativeness, future analyses should use both CPRD Aurum and CPRD Gold and assess demographic differences in veteran profiles across the UK. Research should prioritise understanding the impact of multimorbidity and health behaviours in veterans, especially in the context of prevention, and evaluating whether clinical interventions achieve comparable outcomes for veterans and non-veterans. There is also a need to address the under-researched health needs of women and ethnic minority veterans.

5. Conclusion

This project assessed the feasibility and utility of using CPRD for veteran health research. It sought to identify and characterise the physical and mental health needs of veterans accessing UK primary care, compare these to non-veteran group, and evaluate the value of linking CPRD with other datasets.

We have successfully developed and validated a method for identifying veterans in CPRD resulting in 138,457 (122,484 in Aurum and 15,973 in Gold) veterans identified in primary care records, established a matched non-veteran comparison group for CPRD Aurum, and conducted initial analyses of physical and mental health outcomes.

This project marks a key step forward in UK veteran health research. For the first time, a validated approach has been developed to identify veterans at scale within UK primary care records. By using CPRD, this work overcomes many of the limitations of previous veteran studies, which have often relied on self-reported information, have been restricted to certain geographical areas, or focused on specific service eras. The ability to identify veterans reliably in CPRD ensures large-scale population-based analyses than can follow health trajectories over time, explore variations across subgroups, and compare outcomes with non-veterans. The accompanying manual for researchers (see section 6) ensures that this method can be applied consistently in future studies, thereby enabling other researchers to make effective use of this rich data source.

6. Manual for researchers

This manual provides guidance on the procedures and data quality considerations involved in identifying veterans within the Clinical Practice Research Datalink (CPRD) (Clinical Practice Research Datalink, 2025a). It outlines a systematic method for veteran case identification and offers practical recommendations for ensuring data accuracy and completeness. The manual also provides guidance for selecting appropriate comparison groups. The aim of this manual is to support high-quality, replicable research on veteran health using routinely collected primary care data.

Background

UK Veterans

In the UK, veterans are defined as individuals who have served in the Armed Forces (either in a regular or reserve capacity) for a minimum of one day and have since returned to civilian life (UK Ministry of Defence, 2017). There are approximately 1.85 million UK Armed Forces veterans in England and Wales, which make up approximately 3.8% of the total population (Office for National Statistics, 2023a). This includes those who were conscripted under National Service, a system of mandatory military service that operated in the UK between 1949 and 1963. During this period, all able-bodied men aged 18 to 21 years were required to serve in the UK Armed Forces for up to two years, contributing to Britain's global military commitments in the post-war era. However, many former Armed Forces personnel in the UK do not define themselves as veterans (Burdett et al., 2013).

Primary care data

The National Health Service (NHS) is the publicly funded healthcare system of the UK, offering healthcare services across England, Wales, Scotland, and Northern Ireland. To access NHS care, patients need to register with a community-based practice, often called GP (for General Practitioner) practice or surgery. Within these primary care practices GPs are the first point of contact, providing healthcare consultations and referrals to wider healthcare services when necessary. All healthcare consultations and medical information related to the patient's care are recorded in primary care Electronic Health Records (EHR) by GPs.

CPRD provides researchers access to anonymised EHR collected from GP practices across the UK. Due to differences in structure and clinical coding in GP software systems, CPRD data is released in two separate databases. CPRD Gold relies on the Vision practice software, while CPRD Aurum draws from EMIS Web systems, with each maintaining separate but comparable structures. (Clinical Practice Research Datalink, 2025a; Herrett et al., 2015; Wolf et al., 2019).

Veteran identification in primary care data

CPRD does not provide standardised definitions for variables and therefore, researchers must create bespoke code lists or algorithms tailored to their study's aims to define populations, outcomes, exposures, and covariates. This method may result in differences in how key variables are operationalised across studies, which could contribute to variations in study exposures or outcomes.

The UK's Department of Health and Social Care indicates that the medical documentation of veterans should include a code indicating a "history relating to military service" (Royal

College of General Practitioners, 2011). However, there are multiple military codes available (e.g., military veteran, left military service, history relating to military service, army veteran), and while Health Education England advocates for the use of a single code, there is no national agreement on which code to apply (Finnegan et al., 2018).

Procedure

Understanding code lists

In CPRD, like any other primary care database, clinical information is coded using standardised medical coding systems, such as Read Codes and SNOMED-CT codes. These systems enable consistent recording of patient data across general practices. However, there is considerable variation in how different clinicians record the same observation, diagnosis or treatment. For example, a diagnosis of asthma might appear under multiple Read Codes such as 'H33 Asthma', 'H330z Extrinsic asthma', or 'H331 Intrinsic asthma'. Additionally, related symptoms or management terms like '66YJ Asthma annual review' or '663U Asthma management' might also appear in patients' records and might be relevant depending on the research question.

To account for this variation and ensure reliable data extraction, researchers use code lists which are curated sets of condition-specific codes that capture all relevant ways a concept might be recorded in the database. Code lists are essential tools for identifying study populations, clinical events, exposures, and covariates with consistency and reproducibility in research using routinely collected healthcare data (Davé & Petersen, 2009). The process of creating a list of key words has been compared to the process of matching Medical Subject Heading (MeSH) terms to identify relevant studies in a systematic review, where sensitivity and specificity need to be considered depending on the study's aims (Davé & Petersen, 2009). Additional guidance on creating code list has been provided by the NIHR (Matthewman et al., 2024).

In CPRD datasets, clinical events are stored using internal identifies, 'medcodeid' in CPRD Aurum and 'medcode' in CPRD Gold, which link to standard medical coding systems. In CPRD Aurum, each 'medcodeid' corresponds to a SNOMED-CT term representing a diagnosis, symptom, or procedure. On the other hand, in CPRD Gold, each 'medcode' maps to a Read Code. Researchers create code lists using SNOMED or Read Codes, and can use the CPRD code browsers to find the matching 'medcodeid' or 'medcode' values needed to extract relevant records form the clinical data files.

Veteran code lists

Code lists are commonly created by cross-referencing and updating published lists (Springate et al., 2014). Thus, the code list to identify veterans in CPRD was initially based on terms used in previous extraction of veteran codes form EHR (Leightley et al., 2023; Mark et al., 2020) and a review of the literature of commonly used terms for this population.

The medical dictionaries in the CPRD code browsers (version 3.0)[13] were searched for key words and synonyms relating to military and veteran status: *military*, *veteran*, *armed forces*, *combat*, *army*, *air force*, *raf*, *navy*, *marine*, *royal engineer* and *corps* where the asterisk represented a wildcard search operator. The terms *reserve*, *deploy, *service person*, *serving*, *soldier* returned no results or resulted in codes already defined using the terms listed first and were therefore excluded from the final search strategy.

Search terms were cross-checked against military and veteran-related terms in the NHS SNOMED-

[13]The CPRD Code Browsers for Aurum and Gold can be requested by contacting enquries@cprd.com. For more information on how to use the CPRD Code Browser, please refer to the Code Browser 3.0 User Guide.

CT Browser to assess completeness (NHS Digital, 2025a). All the selected codes were then deduplicated and manually inspected for terms that should be included or excluded from the code list. The resulting codes and associated terms were reviewed by a team of military health experts, which included researchers and clinicians, to identify relevant terms relating to veteran status. A total of 121 and 35 codes were included for extraction from Aurum (see Table 9) and Gold (see Table 10) respectively.

Data quality

CPRD provides a robust resource for epidemiological research, supported by systematic data quality assurance processes. The data are derived from routine primary care encounters which reflects longitudinal, real-world primary care activity across the UK. Data completeness and consistency are enhanced by national initiatives such as the Quality and Outcomes Framework (QOF), which incentivises the standardised recording of key clinical variables (NHS Digital, 2025b). Additionally, CPRD Gold and Aurum apply internal quality checks, with CPRD Aurum implementing over 900 automated checks on data integrity, structure, and referential consistency (Herrett et al., 2015; Wolf et al., 2019).

However, data entry is performed during routine care and is not specifically collected for research purposes. This can lead to variability in recording practices, possible misclassification, and missing data. Therefore, researchers must carefully construct case definitions and validate code lists to mitigate these issues and ensure reliable identification of populations, including veterans (Herrett et al., 2015; Wolf et al., 2019).

Identifying veterans in primary care data presents several challenges due to variability in clinical coding, incomplete documentation, and differences in data capture across practices and software systems. Although UK government guidance encourages the use of codes to record a patient's military history, implementation across general practices has been variable (Finnegan et al., 2018; Finnegan & Randles, 2023b). This may result in under-recording of veteran status in EHRs and to potential misclassification bias.

CPRD data are sourced from two different primary care software systems, EMIS Web and Vision, which use SNOMED-CT and Read Codes, respectively. While functionally similar, differences in terminology and available codes can affect the comprehensiveness and comparability of veteran identification across CPRD Aurum and CPRD Gold. Therefore, separate code lists must be developed for each database, using the relevant codes.

Furthermore, even within the same software system, clinicians may use different codes to document similar concepts. This leads to semantic heterogeneity and impacts case ascertainment. Therefore, the creation of comprehensive and validated code lists is critical to mitigate this variation. All code lists should undergo independent clinical review, and where feasible, be cross validated with published code lists.

Coding practices may vary locally, regionally or over time due to changes in clinical guidelines, software updates, or policy initiatives such as the veteran friendly GP Practice Accreditation Programme. Thus, researchers should consider time trends in code usage and conduct sensitivity analyses to assess the impact of these changes in veteran identification. Additionally, researchers should consider grouping terms as "definite" if it clearly stated veteran status, and "probable", if it related to a military role or active duty in the military. When comparing veteran to non-veteran groups, ensure geographical and temporal matching to minimise confounding due to systemic differences in coding practices.

Table 9. Codes (medcodeid and SNOMED-CT) used to identify veterans in CPRD Aurum

#	Term	medcodeid	SCTID
1	Active duty military	2995730017	702348006
2	Active serving member of the Army	2508511000000111	986471000000105
3	Active serving member of the Army - temp reg*	2508821000000115	986611000000102
4	Active serving member of the Army Reserve	2508701000000118	986551000000102
5	Active serving member of the Royal Air Force	2508551000000110	986491000000109
6	Active serving member of the Royal Air Force - tem reg*	1967001000006119	986631000000105
7	Active serving member of the Royal Air Force Reserve	2508741000000115	986571000000106
8	Active serving member of the Royal Marines	2508631000000119	986531000000109
9	Active serving member of the Royal Marines - tem reg*	1967021000006112	986671000000107
10	Active serving member of the Royal Marines Reserve	2510711000000115	987551000000100
11	Active serving member of the Royal Navy	2508591000000119	986511000000101
12	Active serving member of the Royal Navy - tem reg*	1967011000006116	986651000000103
13	Active serving member of the Royal Navy Reserve	2508781000000111	986591000000105
14	Acute posttraumatic stress disorder following military combat	1755921000006119	446175003
15	Adjutant General's Corps personnel	12118241000006112	1100041000000100
16	Administrative officer - RAF	8336001000006114	929201000000104
17	AIR FORCE	1790641000006116	1790641000006100
18	Aircraft ejection - military	947311000006114	947311000006105
19	Airman, aircrew - RAF	8336101000006113	929261000000100
20	Armed forces	395217018	720371000000100
21	Armed forces reservist	2390691000000115	933041000000109
22	Armed forces: non-comissioned	395216010	1100001000000110
23	Armed forces: non-commissioned	12118201000006110	1100001000000110
24	ARMY	1790921000006118	1790921000006100
25	Army - non-commissioned personnel	12118221000006117	1100021000000100

**temp reg = temporary registration*

#	Term	medcodeid	SCTID
26	Army officer	6004141000006118	310892005
27	Army personnel	3144411000006118	40045004
28	ARMY TRG	1790931000006115	1790931000006100
29	Army veteran	2509111000000114	986751000000107
30	Chronic post-traumatic stress disorder following military combat	1755931000006116	699241002
31	Commissioned officer - Royal Navy	8336641000006115	929761000000108
32	Commissioned officer medical - Royal Navy	8336751000006111	929861000000104
33	Delayed posttraumatic stress disorder following military combat	1755941000006114	446180007
34	Deserted from military service	412078014	276087003
35	Design draughtsman - Royal Engineers	8147411000006112	418041000000106
36	Dog trainer - Royal Army Veterinary Corps	8171171000006117	454661000000101
37	Duration of military service	4926971000006113	224359000
38	Electrical and mechanical draughtsman - Royal Engineers	8147531000006119	418221000000109
39	Engineer - Royal Navy	8336821000006114	929921000000102
40	Exposed to combat during military service	14581891000006110	1187605001
41	Finance and systems administrator - Adjutant General's Corps	6004191000006110	310897004
42	Finance and systems administrator - Adjutant General's Corps	12118261000006111	1100061000000110
43	General mechanic - RAF	8336511000006115	929641000000104
44	Heating, ventilation and air conditioning fitter - Royal Engineers	8148401000006110	420201000000107
45	History of serving in armed forces	14581821000006113	1187600006
46	History relating to Army service	661901000000115	440431002
47	History relating to military service	443668013	302121005
48	History relating to Royal Air Force service	662041000000111	352551000000102
49	History relating to Royal Navy service	661981000000113	352521000000107
50	Horse trainer - Royal Army Veterinary Corps	8144171000006115	412611000000103
51	Infantry soldier	12118601000006113	1100401000000100
52	Injury from military training/exercise	971251000006114	971251000006105
53	Joined military forces	354238011	236305007
54	Left military service	397791019	266964007
55	Left military service	14582051000006118	1187610002
56	Marine engineer - Royal Navy	8336841000006119	929941000000109
57	Marine engineer (non-commissioned) - Royal Navy	8337021000006118	930111000000100
58	Marine engineer officer	147331000006117	22835001
59	Medically discharged from military service	14582021000006110	1187609007
60	Merchant navy medical	943381000006111	943381000006107

#	Term	medcodeid	SCTID
61	Military administrative status	7999001000006110	14041000000105
62	Military clerk	12118231000006119	1100031000000100
63	Military services member	711241000006110	49588008
64	Military services member of ground forces	3144401000006116	40045004
65	Military veteran	1657891000000115	753651000000107
66	Movement controller - Royal Logistics Corps	8147841000006114	418561000000103
67	Musician - Royal Marines	8336611000006119	929731000000103
68	Nurse - RAF	8336281000006112	929421000000109
69	Officer - armed forces	443917013	302312000
70	Officer, armed forces NOS	395210016	302312000
71	Personnel administrator - RAF	8336371000006117	929501000000101
72	Pharmacy technician - RAF	8336311000006114	929451000000104
73	Pharmacy technician - Royal Army Medical Corps	8144031000006119	412461000000104
74	Plumber and pipe fitter - Royal Engineers	8178551000006114	466731000000101
75	Postal and courier operator - Royal Logistics Corps	8144451000006116	413071000000106
76	Princess Mary's Royal Air Force Nursing Service	8147361000006116	417991000000104
77	Regimental accountant - Adjutant General's Corps	12118351000006115	1100151000000110
78	Regimental administration officer finance - AGC*	12118361000006118	1100161000000110
79	Regt/corps/command	954031000006112	954031000006108
80	Royal Air Force administrative officer, education	8147261000006114	417831000000109
81	Royal Air Force administrative officer, secretarial	8137101000006116	401371000000106
82	Royal Air Force communications systems voice analyst	8147821000006119	418551000000101
83	Royal Air Force data analyst	8136671000006115	400901000000107
84	Royal Air Force electrical engineering technician, tele**	8145501000006115	414661000000105
85	Royal Air Force mechanical transport driver	8170681000006117	454101000000106
86	Royal Air Force musician	8137321000006111	401631000000100
87	Royal Air Force navigator, general duties	8144531000006110	413201000000106
88	Royal Air Force operations support manager	8170831000006119	454301000000108
89	Royal Air Force personnel administrator, automatic data processing	8180361000006111	470601000000108
90	Royal Air Force statistical analyst, automatic data processing	8163711000006115	443141000000109
91	Royal Air Force student dental technician	8144431000006111	413061000000104
92	Royal Air Force student laboratory technician	8137631000006116	403051000000104
93	Royal Air Force student pharmacy technician	8163131000006113	442861000000106
94	Royal Air Force student radiographer	8180871000006114	472251000000108
95	Royal Air Force telecommunications operator	8171371000006115	454921000000100

*AGC = Adjutant General's Corps;
**tele = telecommunications

#	Term	medcodeid	SCTID
96	Royal Air Force veteran	2509171000000116	986781000000101
97	Royal Army Dental Corps personnel	12118711000006115	1100511000000100
98	Royal Army Medical Corps personnel	6005031000006114	311127000
99	Royal Army Medical Corps personnel	12119061000006111	1100831000000110
100	Royal engineers diver	947431000006110	947431000006106
101	Royal Logistics Corps personnel	12119011000006113	1100781000000100
102	Royal Marines - non-commissioned personnel	8336561000006117	929691000000109
103	Royal Marines veteran	2509271000000111	986831000000106
104	Royal navy limitations	951831000006115	951831000006104
105	Royal Navy officer	8336631000006113	929751000000105
106	Royal Navy personnel	8336621000006110	929741000000107
107	Royal navy templates	947501000006118	947501000006102
108	Royal Navy veteran	2509231000000114	986811000000103
109	Seaman - Royal Navy	8336961000006111	930051000000102
110	Served in armed forces	1672001000006113	1672001000006110
111	Served in armed forces in combat	1672011000006111	419482003
112	Served in armed forces in combat	14581901000006114	1187605001
113	Served in military service	14581801000006115	1187600006
114	Serving in military service	14581861000006119	1187604002
115	Signposting to Veterans' Gateway	12618511000006114	1104161000000100
116	Soldier - home service force	12118451000006110	1100251000000100
117	Special Air Services soldier	12118421000006118	1100221000000110
118	Supply controller - Royal Logistics Corps	8147321000006110	417911000000108
119	Topographic technician - Royal Engineers	8137061000006119	401331000000109
120	Trainee - armed forces	248724019	159653008
121	Veterinary technician - Royal Army Veterinary Corps	8170871000006116	454341000000106

Table 10. Codes (medcode and Read Code) used to identify veterans in CPRD Gold

#	Term	medcode	Read Code
1	[X]Acute post-traumatic stress disorder follow military comb	101785	Eu43300
2	[X]Chron post-traumatic stress disorder follow military comb	101725	Eu43400
3	[X]Delayed post-traumat stress disorder follow military comb	113199	Eu43500
4	Active duty military	111470	095..00
5	Active serving member of Royal Navy - temporary registration	110715	912T.00
6	Active serving member of the Army	110378	950
7	Active serving member of the Army - temporary registration	112313	912R.00
8	Active serving member of the Army Reserve	111416	0Z70.00
9	Active serving member of the Royal Air Force	111618	951
10	Active serving member of the Royal Air Force Reserve	113944	0Z71.00
11	Active serving member of the Royal Marines	112539	953
12	Active serving member of the Royal Navy	112541	952
13	Active serving member of the Royal Navy Reserve	111276	0Z72.00
14	Armed forces NOS	8682	091..11
15	Armed forces NOS	40127	091Z.00
16	Armed forces reservist	110178	0Z7..00
17	Armed forces: non-comissioned	57998	091..00
18	Army veteran	110414	13Ji000
19	Deserted from military service	46454	13JR.17
20	History relating to Army service	95950	13q0.00
21	History relating to military service	101688	13JY.00
22	History relating to Royal Air Force service	96097	13q2.00
23	History relating to Royal Navy service	96254	13q1.00
24	Joined military forces	43054	091Z.11
25	Left military service	5617	13JR.00
26	Marine superintendant	73459	668
27	Member of armed forces	5671	912
28	Military veteran	101082	13Ji.00
29	Officer - armed forces	29626	06E..11
30	Officer, armed forces NOS	54323	06E..00
31	Royal Air Force veteran	110553	13Ji100
32	Royal Marines veteran	110957	13Ji300
33	Royal Navy veteran	110622	13Ji200
34	Served in armed forces	96075	13q3.00
35	Trainee - armed forces	43008	913


COLOURS

General population comparison group: matching non-veterans

The veteran population differs from the general population on key characteristics such as age and gender (Office for National Statistics, 2023a), both of which are strongly associated with the risk of developing certain health conditions and patterns of healthcare utilisation. Veterans also differ from the general population in their contact with NHS primary care. During service, UK Armed Forces personnel receive healthcare through the Defence Medical Services rather than the NHS and therefore may have fewer NHS primary care records prior to discharge. Furthermore, GP practices vary in recording practices, patient demographics and healthcare access which could explain health outcomes.

Thus, to reduce potential confounding and ensure comparability between veteran and non-veteran groups, we conducted exact matching (Iwagami & Shinozaki, 2022) on age, gender, general practice, and index date[14]. Given the large sample size available in CPRD, exact matching is feasible and allows for precise control of these key confounders. The choice of matching variables depends on data availability and on the research question of each study.

We initially attempted to carry out 4:1 matching, selecting four non-veterans for each veteran, to maximise statistical power and make full use of the large pool of eligible non-veterans available in CPRD. Increasing the number of non-veterans per case improves the precision of estimates and enhances the comparability between groups, particularly when the case group is relatively small (Wachoider et al., 1992). However, given the strict matching criteria, particularly the inclusion of general practice and index date, it was not always possible to identify four suitable non-veterans for each case without compromising match quality or excluding cases (veterans). This was especially challenging for the oldest age groups (80 years and older), where the veteran population is overrepresented, and where fewer non-veteran patients of similar age were available within the same general practice.

As a result, matching was achieved at a 2:1 ratio for the vast majority of the sample (122,089 participants), while a 1:1 ratio was used for a smaller subgroup (395 participants) due to a lack of eligible matches in the non-veteran pool, particularly at the extremes of certain variables such as age (e.g., the youngest and oldest age groups).

The choice of matching variables depends on data availability and on the research question of each study. In our case, matching was conducted on age, gender, GP practice, and index date. However, it is important to note that other factors, such as ethnicity, are also associated with the prevalence of certain conditions. For example, diabetes shows higher rates in some ethnic groups (Goff, 2019). Future comparative analyses may therefore benefit from matching or adjustment for such variables, where data availability allows.

Additionally, the choice of matching method depends on several factors, including the study objective, size and diversity of the non-veteran pool, as well as the number and distribution of covariates. Alternative matching approaches could also be considered in future research using CPRD data to match veterans and non-veterans. For example, propensity score matching (Iwagami & Shinozaki, 2022) allows for balancing multiple covariates simultaneously by estimating the probability of being a veteran based on observed characteristics. This method can increase the pool of eligible non-veterans by relaxing the requirement for exact matches, likely improving match rates and reducing the number of unmatched veterans, particularly at the extremes of age or in smaller GPs. Frequency matching (Gail, 2014) offers another option, whereby non-veterans are selected to ensure that the overall distribution of key variables (e.g., age, gender, general practice) mirrors that of the veteran group, without requiring individual-level matches.

[14]For veterans, the index date was defined as the date on which a veteran-related term first appeared in their primary care record, indicating their first identification as a veteran within the NHS. For non-veterans, the index date was defined as the date of first registration with the GP practice. Matching on index date helped ensure comparable periods of follow-up in NHS records between groups.

Appendix

We ran regression models to examine the association between veteran status and each health condition. Analyses included only individuals with complete outcome and covariate data (see Appendix Table 1). Since some matched patients had missing data in BMI, smoking status, and ethnicity, this resulted in the loss of matched pairs and disruption of the original matching. Unmatched controls (N=24,102) were excluded from the analysis. Poisson regressions with robust standard errors were used to estimate prevalence ratios. We first assessed the associations adjusting for the matching variables (age, gender, index date, and GP practice) and then adjusted for ethnicity and the health risk factors.

Compared to the findings from Table 5 (page 45) which used multiple imputation, the complete-case results from Appendix Table 2 underestimate associations between veteran status and health outcomes. This attenuation is consistent with evidence that missing data on lifestyle factors such as BMI and smoking are not random but more common among individuals who are healthier or who attend primary care less frequently (Dhalwani et al., 2013; Nicholson et al., 2019; Petersen et al., 2019). Thus, it is possible that the complete case analysis sample retains individuals with poorer health across both veteran and non-veteran, reducing observable differences and biasing the associations toward the null. This is supported by the observation that in the complete case sample, veterans and controls have more consultations than in the full matched sample (veterans: 11.65 median consultations in the complete case sample vs 10.12 in the full matched sample; controls: 10.32 median consultations in the complete case sample vs 6.50 in the full matched sample). In contrast, multiple imputation preserves the full matched sample and accounts for missingness patterns linked to demographic and clinical characteristics, leading to more plausible and consistent estimates.

Appendix Table 1. CPRD Aurum sample (veteran and non-veteran) variables descriptives for sample with complete data (N=118,674)

	NON-VETERANS N=44,198		VETERANS N=74,476	
	N	% / Median [IQR] / Mean (SD)	N	% / Median [IQR] / Mean (SD)
Follow-up time (years)*		3.40		3.69
		[3.07-7.88]		[1.72-6.39]
Consultations		10.32		11.65
		[6.80-15.74]		[7.20-18.25]
Mortality				
Deceased	4,209	9.52	7,218	9.69
Year of index				
1987-1990	142	0.32	515	0.69
1991-2000	838	1.90	2,651	3.56
2001-2010	5,876	13.29	8,499	11.41
2011-2020	29,836	67.51	45,051	60.49
2021-2024	7,506	16.98	17,760	23.85
Age at index (years)		55.63 (19.85)		54.80 (19.92)
18-25	3,402	7.70	6,148	8.26
26-35	5,268	11.92	9,356	12.56
36-45	5,792	13.10	10,362	13.91
46-55	7,379	16.70	12,373	16.61
56-65	7,291	16.50	12,092	16.24
66-75	6,094	13.79	9,929	13.33
76-85	6,366	14.40	9,887	13.28
86-95	2,498	5.65	4,134	5.55
96+	108	0.24	195	0.26
Gender				
Men	31,930	72.24	58,901	79.09
Women	12,268	27.76	15,575	20.91

SD: Standard Deviation; IQR: Interquartile Range.
**Start date calculated as maximum value between registration date, index date, and protocol start date (01 of January 1987) and end date calculated as minimum value between registration end date, last CPRD collection date for the practice, date of death, or protocol end date (31 of July 2024).*

	NON-VETERANS N=,44,198		VETERANS N=74,476	
	N	% / Median [IQR] / Mean (SD)	N	% / Median [IQR] / Mean (SD)
Ethnicity				
White	37,999	85.97	65,369	87.77
Black	603	1.36	1,097	1.47
Asian	1,742	3.94	1,993	2.68
Mixed	276	0.62	534	0.72
Other	3,578	8.10	5,483	7.36
Region				
North East	5,604	12.68	7,694	10.33
North West	13,199	29.86	20,225	27.16
Yorkshire and the Humber	1,964	4.44	3,047	4.09
East Midlands	1,210	2.74	2,066	2.77
West Midlands	5,334	12.07	8,787	11.80
East of England	830	1.88	1,635	2.20
London	1,258	2.85	2,456	3.30
South East	11,429	25.86	21,564	28.95
South West	3,338	7.55	6,948	9.33
Northern Ireland	32	0.07	54	0.07

SD: Standard Deviation
**Start date calculated as maximum value between registration date, index date, and protocol start date (01 of January 1987) and end date calculated as minimum value between registration end date, last CPRD collection date for the practice, date of death, or protocol end date (31 of July 2024).*

Appendix Table 2. Poisson regression for the association between veteran status and physical and mental health conditions (N=118,674)

	Model 1			Model 2		
	PR	95%	CI	PR	95%	CI
Chronic obstructive pulmonary disease	1.08	1.04	1.13	1.07	1.03	1.11
Depression	1.12	1.09	1.16	1.12	1.08	1.16
Lower back pain	0.99	0.96	1.02	0 .99	0.96	1.02
Myocardial infarction	0.93	0.86	1.00	0.93	0.86	1.00
Osteoarthritis	1.08	1.04	1.11	1.07	1.04	1.11
Post-traumatic stress disorder	8.61	7.29	10.17	8.55	7.24	10.09

	Model 3		
	PR	95%	CI
Chronic obstructive pulmonary disease	1.02	0.98	1.06
Depression	1.10	1.07	1.14
Lower back pain	0.98	0.95	1.00
Myocardial infarction	0.90	0.84	0.98
Osteoarthritis	1.04	1.01	1.08
Post-traumatic stress disorder	8.27	7.00	9.76

PR: Prevalence Ratios; CI: Confidence Interval

Model 1: gender, age, index date and GP practice; Model 2: Model 1 + ethnicity; Model 3: Model 2 + BMI and smoking

References

Armed Forces Covenant. (2024). Veteran-friendly accreditation for GP practices. https://www.armedforcescovenant.gov.uk/case-study/veteran-friendly-accreditation-for-gp-practices/

Bergman, B., Ross, D., & Mackay, D. (2022). Trends in Scottish Veterans' Health. The Scottish Veteran Health Research Group University of Glasgow. https://s31949.pcdn.co/wp-content/uploads/Trends-in-Scottish-Veterans-Health-Report-Digital-Copy-1.pdf

Betancourt, J. A., Dolezel, D. M., Shanmugam, R., Pacheco, G. J., Stigler Granados, P., & Fulton, L. V. (2023). The Health Status of the US Veterans: A Longitudinal Analysis of Surveillance Data Prior to and during the COVID-19 Pandemic. Healthcare (Switzerland), 11(14). https://doi.org/10.3390/healthcare11142049

Bhaskaran, K., Forbes, H. J., Douglas, I., Leon, D. A., & Smeeth, L. (2013). Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). BMJ Open, 3(9). https://doi.org/10.1136/bmjopen-2013-003389

Bollinger, M. J., Schmidt, S., Pugh, J. A., Parsons, H. M., Copeland, L. A., & Pugh, M. J. (2015). Erosion of the healthy soldier effect in veterans of US military service in Iraq and Afghanistan. Population Health Metrics, 13(1). https://doi.org/10.1186/s12963-015-0040-6

Burdett, H., Woodhead, C., Iversen, A. C., Wessely, S., Dandeker, C., & Fear, N. T. (2013). "Are You a Veteran?" Understanding of the Term "Veteran" among UK Ex-Service Personnel: A Research Note. Armed Forces and Society, 39(4), 751–759. https://doi.org/10.1177/0095327X12452033

Cassell, A., Edwards, D., Harshfield, A., Rhodes, K., Brimicombe, J., Payne, R., & Griffin, S. (2018). The epidemiology of multimorbidity in primary care: A retrospective cohort study. British Journal of General Practice, 68(669), e245–e251. https://doi.org/10.3399/bjgp18X695465

Cassidy, S. P., MacLean, A., & Denney, J. T. (2023). Military Service, Education, and Mortality Across Cohorts from World War II to the Post-Vietnam Era. Population Research and Policy Review, 42(4). https://doi.org/10.1007/s11113-023-09815-w

Centers for Disease Control and Prevention. (2025, April 16). About the Global Health Center. https://www.cdc.gov/global-health/about/index.html?utm_source=chatgpt.com

Clinical Practice Research Datalink. (2024, July 9). CPRD PROVE (PRoviding Online Verification of EHR). https://www.cprd.com/research-services/cprd-prove

Clinical Practice Research Datalink. (2025a). Clinical Practice Research Datalink (CPRD). Medicines & Healthcare Products Regulatory Agency. https://www.cprd.com/

Clinical Practice Research Datalink. (2025b). Linked data. https://www.cprd.com/data/linked-data

Clinical Practice Research Datalink. (2025c, September). CPRD Aurum September 2025 (Version 2025.09.001) [Data set]. Medicines and Healthcare Products Regulatory Agency. https://www.cprd.com/doi/cprd-aurum-september-2025

Clinical Practice Research Datalink. (2025d, September). CPRD GOLD September 2025 (Version 2025.09.001) [Data set]. Medicines and Healthcare Products Regulatory Agency. https://www.cprd.com/doi/cprd-gold-september-2025

Davé, S., & Petersen, I. (2009). Creating medical and drug code lists to identify cases in primary care databases. Pharmacoepidemiology and Drug Safety, 18(8), 704–707. https://doi.org/10.1002/pds.1770

Degli Esposti, M., Ziauddeen, H., Bowes, L., Reeves, A., Chekroud, A. M., Humphreys, D. K., & Ford, T. (2022). Trends in inpatient care for psychiatric disorders in NHS hospitals across England, 1998/99–2019/20: an observational time series analysis. Social Psychiatry and Psychiatric Epidemiology, 57(5), 993–1006. https://doi.org/10.1007/s00127-021-02215-5

Department for Communities and Local Government. (2015). The English Index of Multiple Deprivation (IMD) 2015-Guidance. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

Dhalwani, N. N., Tata, L. J., Coleman, T., Fleming, K. M., & Szatkowski, L. (2013). Completeness of Maternal Smoking Status Recording during Pregnancy in United Kingdom Primary Care Data. PLoS ONE, 8(9). https://doi.org/10.1371/journal.pone.0072218

Dregan, A., Charlton, J., Chowienczyk, P., & Gulliford, M. C. (2014). Chronic inflammatory disorders and risk of type 2 diabetes mellitus, coronary heart disease, and stroke : A population-based cohort study. Circulation, 130(10), 837–844. https://doi.org/10.1161/CIRCULATIONAHA.114.009990

Dregan, A., Matcham, F., Harber-Aschan, L., Rayner, L., Brailean, A., Davis, K., Hatch, S., Pariante, C., Armstrong, D., Stewart, R., & Hotopf, M. (2019). Common mental disorders within chronic inflammatory disorders: A primary care database prospective investigation. Annals of the Rheumatic Diseases, 78(5), 688–695. https://doi.org/10.1136/annrheumdis-2018-214676

Dregan, A., Moller, H., Murray-Thomas, T., & Gulliford, M. C. (2012). Validity of cancer diagnosis in a primary care database compared with linked cancer registrations in England. Population-based cohort study. Cancer Epidemiology, 36(5), 425–429. https://doi.org/10.1016/j.canep.2012.05.013

Dregan, A., Toschke, M. A., Wolfe, C. D., Rudd, A., Ashworth, M., & Gulliford, M. C. (2011). Utility of electronic patient records in primary care for stroke secondary prevention trials. BMC Public Health, 11. https://doi.org/10.1186/1471-2458-11-86

Finnegan, A., Jackson, R., & Simpson, R. (2018). Finding the Forgotten: Motivating Military Veterans to Register with a Primary Healthcare Practice. Military Medicine, 183(11–12), e509–e517. https://doi.org/10.1093/milmed/usy086

Finnegan, A., & Randles, R. (2023a). Prevalence of common mental health disorders in military veterans: using primary healthcare data. BMJ Military Health, 169(6), 523–528. https://doi.org/10.1136/bmjmilitary-2021-002045

Finnegan, A., & Randles, R. (2023b). Where are all the veterans? A mixed methods assessment of a systematic strategy to increase veteran registration in UK primary healthcare practices. BMJ Open, 13(6). https://doi.org/10.1136/bmjopen-2022-068904

Finnegan, A., & Salem, K. (2024). Comparative study of hypertension, diabetes, dementia and smoking in military veterans and non-veterans: A quantitative study using primary healthcare data. BMJ Military Health. https://doi.org/10.1136/military-2024-002817

Finnegan, A., Salem, K., Ainsworth-Moore, L., Randles, R., West, L., Simpson, R., & Grant, V. B. (2022). The Veteran Friendly Practice accreditation programme: a mixed-methods evaluation. BJGP Open, 6(3). https://doi.org/10.3399/BJGPO.2022.0012

Gail, M. H. (2014). Frequency matching. Wiley Statsref: Statistics Reference Online.

Gallagher, A. M., Dedman, D., Padmanabhan, S., Leufkens, H. G. M., & de Vries, F. (2019). 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. Pharmacoepidemiology and Drug Safety, 28(5), 563–569. https://doi.org/10.1002/pds.4747

General Medical Council. (2019). The state of medical education and practice in the UK. https://www.gmc-uk.org/-/media/documents/somep-2019---full-report_pdf-81131156.pdf

Goff, L. M. (2019). Ethnicity and Type 2 diabetes in the UK. In Diabetic Medicine (Vol. 36, Issue 8, pp. 927–938). Blackwell Publishing Ltd. https://doi.org/10.1111/dme.13895

Goodwin, L., Bourke, J. H., Forbes, H., Hotopf, M., Hull, L., Jones, N., Rona, R. J., Wessely, S., & Fear, N. T. (2013). Irritable bowel syndrome in the UK military after deployment to Iraq: What are the risk factors? Social Psychiatry and Psychiatric Epidemiology, 48(11), 1755–1765. https://doi.org/10.1007/s00127-013-0699-6

Guthrie, B., Barnett, K., Mercer, S. W., Norbury, M., Watt, G., & Wyke, S. (2012). Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet, 380, 37–43. https://doi.org/10.1016/S0140

Herrett, E., Gallagher, A. M., Bhaskaran, K., Forbes, H., Mathur, R., Staa, T. van, & Smeeth, L. (2015). Data Resource Profile: Clinical Practice Research Datalink (CPRD). International Journal of Epidemiology, 44(3), 827–836. https://doi.org/10.1093/ije/dyv098

Hickman, T. (2004). The call-up: a history of national service. Headline.

Hobbs, F. D. R., Bankhead, C., Mukhtar, T., Stevens, S., Perera-Salazar, R., Holt, T., & Salisbury, C. (2016). Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14. The Lancet, 387(10035), 2323–2330. https://doi.org/10.1016/S0140-6736(16)00620-6

Hynes, D. M., Edwards, S., Hickok, A., Niederhausen, M., Weaver, F. M., Tarlov, E., Gordon, H., Jacob, R. L., Bartle, B., Young, R., & Laliberte, A. (2021). Veterans' Use of Veterans Health Administration Primary Care in an Era of Expanding Choice. www.lww-medicalcare.com

Hynes, D. M., Maciejewski, M. L., & Atkins, D. (2018). HSR Commentary: Linking VA and Non-VA Data to Address Important US Veteran Health Services Research Issues. Health Services Research, 53, 5133–5139. https://doi.org/10.1111/1475-6773.13081

Iwagami, M., & Shinozaki, T. (2022). Introduction to Matching in Case-Control and Cohort Studies. Annals of Clinical Epidemiology, 4(2), 33–40.

Katon, J. G., Zephyrin, L., Meoli, A., Hulugalle, A., Bosch, J., Callegari, L., Galvan, I. V., Gray, K. E., Haeger, K. O., Hoffmire, C., Levis, S., Ma, E. W., McCabe, J. E., Nillni, Y. I., Pineles, S. L., Reddy, S. M., Savitz, D. A., Shaw, J. G., & Patton, E. W. (2018). Reproductive Health of Women Veterans: A Systematic Review of the Literature from 2008 to 2017. In Seminars in Reproductive Medicine (Vol. 36, Issue 6, pp. 315–322). Thieme Medical Publishers, Inc. https://doi.org/10.1055/s-0039-1678750

Leightley, D., Palmer, L., Williamson, C., Leal, R., Chandran, D., Murphy, D., Fear, N. T., & Stevelink, S. A. M. (2023). Identifying Military Service Status in Electronic Healthcare Records from Psychiatric Secondary Healthcare Services: A Validation Exercise Using the Military Service Identification Tool. Healthcare (Switzerland), 11(4). https://doi.org/10.3390/healthcare11040524

Mark, K. M., Leightley, D., Pernet, D., Murphy, D., Stevelink, S. A. M., & Fear, N. T. (2020). Identifying veterans using electronic health records in the United Kingdom: A feasibility study. Healthcare (Switzerland), 8(1). https://doi.org/10.3390/healthcare8010001

Matthewman, J., Andresen, K., Suffel, A., Lin, L.-Y., Schultze, A., Tazare, J., Bhaskaran, K., Williamson, E., Costello, R., Quint, J., & Strongman, H. (2024). Checklist and guidance on creating codelists for electronic health records research. NIHR Open Research, 4, 20. https://doi.org/10.3310/nihropenres.13550.1

Ministry of Defence & Office for Veterans' Affairs. (2025, May 23). VALOUR: information and next steps. https://www.gov.uk/guidance/valour-information-and-next-steps

NHS. (2024a, June 10). Veterans, service leavers, and non-mobilised reservists. https://www.nhs.uk/nhs-services/armed-forces-community/veterans-service-leavers-non-mobilised-reservists/

NHS. (2024b, July 18). Mental health support for veterans, service leavers and reservists. https://www.nhs.uk/nhs-services/armed-forces-community/mental-health/veterans-reservists/

NHS Digital. (2024, April 11). Patients registered at a GP practice, April 2024. https://digital.nhs.uk/data-and-information/publications/statistical/patients-registered-at-a-gp-practice/april-2024#

NHS Digital. (2025a). NHS Digital SNOMED CT Browser. https://termbrowser.nhs.uk/?

NHS Digital. (2025b, July 31). Quality and Outcomes Framework (QOF) online database. https://digital.nhs.uk/services/quality-and-outcomes-framework-qof-online-database

Nicholson, B. D., Aveyard, P., Bankhead, C. R., Hamilton, W., Hobbs, F. D. R., & Lay-Flurrie, S. (2019). Determinants and extent of weight recording in UK primary care: An analysis of 5 million adults' electronic health records from 2000 to 2017. In BMC Medicine (Vol. 17, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12916-019-1446-y

Office for National Statistics. (2023a). Characteristics of UK armed forces veterans, England and Wales: Census 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/armedforcescommunity/articles/characteristicsofukarmedforcesveteransenglandandwalescensus2021

Office for National Statistics. (2023b). Veterans Survey 2022, demographic overview and coverage analysis, UK. https://www.ons.gov.uk/peoplepopulationandcommunity/armedforcescommunity/articles/veteranssurvey2022demographicoverviewandcoverageanalysisuk/december2023#representativeness-of-the-veterans-survey-2022

Office for Veterans' Affairs. (2020). Veterans Factsheet 2020. www.gov.uk/government/statistics/annual-population-survey-uk-armed-forces-veterans-residing-in-great-britain-2017

Oster, C., Morello, A., Venning, A., & Lawn, S. (2017). The health and wellbeing needs of veterans: A rapid review. BMC Psychiatry, 17(1). https://doi.org/10.1186/s12888-017-1547-0

Palmer, L., Whyte, R., Leightley, D., Fear, N. T., & Gribble, R. (2025). Influences on alcohol and lifestyle behaviours among partners of UK (ex-)military personnel. https://s31949.pcdn.co/wp-content/uploads/LIFE-Q-Report-2025-v6.pdf?ref=kcmhr.org

Petersen, I., Welch, C. A., Nazareth, I., Walters, K., Marston, L., Morris, R. W., Carpenter, J. R., Morris, T. P., & Pham, T. M. (2019). Health indicator recording in UK primary care electronic health records: Key implications for handling missing data. Clinical Epidemiology, 11, 157–167. https://doi.org/10.2147/CLEP.S191437

Rona, R. J., Jones, M., Goodwin, L., Hull, L., & Wessely, S. (2013). Risk factors for headache in the UK military: Cross-sectional and longitudinal analyses. Headache, 53(5), 787–798. https://doi.org/10.1111/head.12101

Royal College of General Practitioners. (2017, September). Mental health in primary care. https://www.rcgp.org.uk/representing-you/policy-areas/mental-health-in-primary-care

Royal College of General Practitioners, T. R. B. L. C. Stress. (2011). Meeting the Healthcare Needs of Veterans A guide for general practitioners. https://www.nss.nhs.scot/publications/meeting-the-healthcare-needs-of-veterans/

Scotland's Census. (2024). Scotland's Census 2022 - UK Armed Forces veterans. https://www.scotlandscensus.gov.uk/2022-reports/scotland-s-census-2022-uk-armed-forces-veterans/

Sharp, M.-L., Busuttil, W., & Murphy, D. (2019). Examining physical health conditions and associations of pain, obesity, and function of UK Veterans diagnosed with PTSD and other mental health conditions. Journal of Military, Veteran and Family Health, 5(2), 75–87. https://doi.org/10.3138/jmvfh.2018-0051

Sharp, M.-L., Franchini, S., Jones, M., Leal, R., Wessely, S., Stevelink, S. A. M., & Fear, N. T. (2024). Office for Veterans' Affairs Final Report Health and Wellbeing Study of Serving and Ex-Serving UK Armed Forces Personnel: Phase 4.

Springate, D. A., Kontopantelis, E., Ashcroft, D. M., Olier, I., Parisi, R., Chamapiwa, E., & Reeves, D. (2014). ClinicalCodes: An online clinical codes repository to improve the validity and reproducibility of research using electronic medical records. PLoS ONE, 9(6). https://doi.org/10.1371/journal.pone.0099825

Tolani, E., Carpenter, M., Petersen, I., & Carpenter, J. (2025). Mind the Gaps: Literature Survey Reveals Shortcomings in Handling Missing Data in Clinical Practice Research Datalink (CPRD), a UK Primary Care Health Records Database. Clinical Epidemiology, Volume 17, 875–889. https://doi.org/10.2147/CLEP.S528401

UK Ministry of Defence. (2017). Veterans: Key Facts. www.armedforcescovenant.gov.uk

Wachoider, S., Silverman, D. T., Mclaughlin, J. K., & Mandel, J. S. (1992). Selection of Controls in Case-Control Studies III. Design Options. In American JounaJ of Epidemiology (Vol. 135, Issue 9).

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. Statistics in Medicine, 30(4), 377–399. https://doi.org/10.1002/sim.4067

Wislocki, K. E., & Zalta, A. K. (2025). Predictors of trauma-related diagnostic overshadowing bias. Behaviour Research and Therapy, 184. https://doi.org/10.1016/j.brat.2024.104651

Wolf, A., Dedman, D., Campbell, J., Booth, H., Lunn, D., Chapman, J., & Myles, P. (2019). Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. International Journal of Epidemiology, 48(6), 1740-1740G. https://doi.org/10.1093/ije/dyz034

Yoon, J., Zulman, D., Scott, J. Y., & Maciejewski, M. L. (2014). Costs associated with multimorbidity among VA patients. Medical Care, 52(3 SUPPL. 2). https://doi.org/10.1097/MLR.0000000000000061


SANCTE ET SAPIENTER