CVDPREVENT Online Methodology Annex v2

November 2024

Introduction

This document provides the methodology for the CVDPREVENT Audit. It explains how the data is collected and analysed.

How is the data collected?

CVDPREVENT data is collected by NHS England by their General Practice Extraction Service (GPES). GP practices are invited to participate in CVDPREVENT through their GP clinical system via the Calculating Quality Reporting Service (CQRS). By accepting the invitation, practices opt into GP system supplier-developed queries to automatically extract their data, based on the CVDPREVENT business rules set.

Data is not extracted about patients who do not consent for the use of GP patient identifiable data as part of national audits or collections, referred to as a Type 1 Opt-out. The selection of clinical data from those patients who do consent is collected using SNOMED codes, as defined in the CVDPREVENT business rules set.

System suppliers go through a testing and certification process with GPES prior to the final data extraction to highlight any issues with the data and ensure the data that was extracted was acceptable and aligned to the specification.

The extracted data are cleaned and supplied to NHS England who then provide the pseudo-anonymised audit data securely to CVD intelligence team at OHID (CVD- I team) in flat file format.

Data confidentiality and security

NHS England have established the CVDPREVENT audit, which will support the implementation of the NHS Long Term Plan under which cardiovascular disease is identified as a clinical priority. The direction has been raised in accordance with sections 254(1) and 254(6) of the 2012 Health and Social Care Act to establish and operate the collection and analysis of information described in the Cardiovascular Disease Prevention Audit Direction Specification. Further details on the direction can be found in the Cardiovascular Disease Prevention Audit Directions 2020.

The CVDPREVENT audit is a collaboration between NHS England, CVD-I team, NHS Benchmarking Network and the Healthcare Quality Improvement Partnership. These are audit partners specified within the Cardiovascular Disease Prevention Audit Direction.

Data collected by NHS England is analysed by OHID's CVD-I. Patient-identifiable data, that may identify an individual, such as a name, date of birth or postcode is removed and pseudo-anonymised by NHS England before it transfers to OHID for analysis. This means that personal data are replaced with an arbitrary unique identifier ensuring no patient can be identified whilst allowing records to be linked across different extracts.

Data processing

Cleaning and Validation

Data are automatically extracted from GP systems in England via system supplier developed queries and GPES. Data submissions are initially checked and certified by GPES using test submissions before data was received into NHS England.

Once data is supplied to CVD- I team, the audit team undertake several checks and cleaning processes. Duplicate patient records caused by people changing addresses, moving GP practice, mislabelled records or other errors were deleted. For these patients the latest record from the Journal table (see section 'Production of analysis database') was obtained and used to update the relevant patient information included in the patient record table whilst older records were deleted.

No further data alteration or removal of data was undertaken after patient deduplication. The selection of valid patient characteristics, date ranges and plausible clinical values were applied at the time of indicator construction.

Production of analysis database

Data is provided to CVD- I team in two compressed files. The files are loaded into a secure SQL database with managed access available only to the CVD- I team.

The database includes a patient demographic table and a clinical table. The anonymised patient table includes demographic information for age, sex, ethnicity, lower super output area (LSOA11) and the patient practice. This table also includes the cohort indicator. For further details please see the CVDPREVENT quick guide.

The second file contains the clinical records and readings for each patient. Each patient can have multiple clinical records and will include metabolic and physical readings, tests, referrals and drug and therapy prescriptions for all patients.

Patients are collected in three cohorts:

Cohort 1 – patients with a high-risk condition for CVD

Cohort 2 – patients with CVD

Cohort 3 – patients not in cohorts 1 or cohort 2 but who have a high reading that may be indicative of CVD or a high-risk condition.

Data collected for patients in cohorts 1 and 2 are the same. There is less information collected for patients in cohort 3. For more details about the cohorts please see the CVDPREVENT quick guide or the CVDPREVENT business rules.

Indicator production

The CVDPREVENT business rules set populates the audit with a subset of all the recorded clinical coding for selected patients.

If you require further information on the primary care SNOMED codes used for the calculation of any of the indicators please contact ncvin-ohid@phe.gov.uk.

Audit analysis is undertaken by producing indicators, which present a count of a recorded condition, a clinical treatment process or outcome. The audit cannot produce complex measures such as change in patient treatment patterns over the audit period. Indicators are defined as proportions of patients meeting the criteria against the eligible patient group. Wherever possible indicators are presented with indicator sub-analyses by sex, age group, ethnicity or deprivation. Prevalence Indicators are also additionally calculated with age-adjustment to enable comparisons between areas or deprivation groups where the age structure of patients is controlled. These are referred to as age-standardised in the data explorer.

Clinical treatment indicators that require readings or measures, for example blood pressure values, are restricted within agreed plausible value ranges, as agreed with the audit clinical lead. Plausible values are available on request from ncvin-ohid@dhsc.gov.uk for each indicator.

The case finder indicators include data from cohort 3 which includes patients who have no coded diagnosis of any of the six high-risk conditions or existing CVD but have an entry in their record that suggests that they are at risk of developing or may have an undiagnosed high-risk condition.

Patients who had missing data in any of the variables for age, sex or LSOA11 were excluded from the calculation of the final indicators. A 'missing' ethnic category was created for ethnic display of the indicators in order to show the impact of the large amount of missing ethnicity data.

Defining the indicators

Some indicators were defined prior to receipt of the data using the CVDPREVENT business rules set with a number based on the original work from the PRIMIS feasibility study[1]. Priority was given to indicators relating directly to the national ABC (atrial fibrillation (AF), blood pressure and high cholesterol) priorities.

All indicators were described and defined by the combined CVDPREVENT team (CVD- I team and NHSBN) with clinical overview from the CVDPREVENT clinical lead, input from the CVDPREVENT Audit Steering Group and methodological support from OHID's Indicator Methodology Review Group.

Data Quality

The extract taken in March 2020 did not include data from all the practices signed up to CVDPREVENT because of an issue with one system supplier. The practice coverage (for analytical purposes) is therefore calculated as the number of practices for which data has been received as a proportion of practices in England.

The population coverage was calculated from the GP list size of contributing practices against the total GP patient registered population in England. List sizes are taken from the time period closest to the CVDPREVENT audit extraction date.

For more information on the specific dates of extraction, list size or Indices of Multiple Deprivation used for each publication of CVDPREVENT data, please contact ncvin-ohid@dhsc.gov.uk.

1 The University of Nottingham: PRIMIS team (2018). CVDPREVENT: A National Primary Care Audit. Feasibility Report. https://www.bhf.org.uk/for-professionals/healthcare-professionals/innovation-in-care/cvdprevent

Reporting

Geographies

Patient demographic data in the audit includes a GP practice code, and the administrative geographic area where the patient's residence is located called a lower super output area (LSOA11) derived from a matching post code by NHSE.

CVDPREVENT reports indicators by organisational area by mapping the patient practice with the most recent publicly available lookups linking GP practices to both PCN and ICBs.

For more information on how mapping is performed please contact ncvin-ohid@dhsc.gov.uk.

Age groups

The reporting age groups were selected to align with other key CVD related age groupings including the age specific hypertension treatment targets and the NHS Health Check attendees. They also provide some granularity while avoiding small number disclosure issues.

Ethnicity

NHS England supplied a lookup for individual SNOMED ethnicity codes to the wider census groups. These groups were used to aggregate and report the standard ethnic categories reported in the audit data. These were White, Black, Asian, Mixed, Other, Not stated and Missing. The ethnic group category 'Not stated' represents a group of people who do not wish to state their ethnic group whereas 'Missing' describes where there is no record of ethnic group.

Deprivation

Each patient has an LSOA11 derived from their residence. LSOA11s were matched to the Index of Multiple Deprivation 2019 (IMD 2019), which is published at LSOA11 geography for all of England.

IMD 2019 is a summary index score for England published by the Ministry of Housing, Communities & Local Government, which relates a small geographical area called a lower super output area (LSOA11). The IMD is based on 7 deprivation domain indicators that provide relative measures of deprivation for small areas across England. Each LSOA11 is mapped geographically to the statutory geographical boundary of an ICB and higher geographies and data aggregated at these levels into 5 equal groups, or quintiles, based on their rank within England. Quintile number 1 represents the most deprived quintile and quintile 5 represents the least deprived quintile. In some ICB areas this means that highest and lowest quintiles have limited data, due to unequal relative deprivation within the area.

PCNs are not geographically defined, so no IMD matching is available and deprivation analysis is not possible using this method.

Note that deprivation reporting is based on geographic matching of patients, whereas all persons, sex, age and ethnicity reporting are based on practice matching.

Deprivation analysis excludes "non-England" residents. Any presentation of data by deprivation groups only includes patients who are resident within England, so the small number of practice patients who reside in Wales or Scotland were excluded for indicators displayed by deprivation.

Time Series

Changes to NHS England code sets may cause minor differences between data extracts. The code sets are clusters of SNOMED codes used within the business rules and are constantly reviewed by NHS England's clinicians. There may be inclusions or deletions to already existing SNOMED codes and clusters. There is also likely to be variation in the way coding occurs within practices and codes may therefore be erroneous, biased or even missing. It will be necessary to be mindful of the potential limitation of inconsistent coding practices and policy changes when interpreting results.

Data Validation

Analytical outputs from CVDPREVENT are checked against other data sources before publication.

Comparison with QOF indicators

CVDPREVENT indicators were compared to QOF indicators where they were similar. In order to do this, practices that contributed both to QOF and CVDPREVENT were matched and where business rules were similar for certain indicators, patient counts and treatments were compared. Both the counts of numerators and proportions for similar clinical indicators in hypertension, atrial fibrillation and chronic kidney disease, were matched. For hypertension and atrial fibrillation CVDPREVENT prevalence, all-age list sizes were used as denominators to more closely resemble QOF. Some of the initial CVDPREVENT GP recorded prevalence and clinical treatment indicators are similar to the prevalence and treatment measures in the QOF reporting mechanism, however, the two processes are not identical, and this results in differing results for each system. CVDPREVENT and QOF indicators differ for several reasons:

- CVDPREVENT has more than 96% coverage (at June 2024) of GP practices, but this means not all GP practices within an area (e.g. ICB) will report to CVDPREVENT and some practices not included in the audit may report to QOF.

- There are no date limiters on either QOF or the audit for finding SNOMED codes. QOF uses the latest date of diagnosis to extract patients, whilst the audit uses the earliest date of the diagnosis.

- CVDPREVENT indicators are collected on different dates to the QOF indicators. Even when the audit data collection period is the same as the QOF period, differences in date of collection mean that the GP list sizes between QOF and CVDPREVENT may differ markedly for some practices. For this reason, direct comparisons between patient numbers with similar indicators between QOF and CVDPREVENT is not advisable.

- CVDPREVENT does not collect information from patients within practices who opt out of audit data collection, but these patients are included in QOF indicator counts.

- CVDPREVENT reports on ages 18 and over for its prevalence indicators. This is not the case for the equivalent QOF indicators which include all ages.

- For the treatment indicators QOF has the option to remove people either not clinically eligible for treatment or who decline treatment from the numerator and denominator. CVDPREVENT does not collect all codes which GPs use to indicate that the patient may not benefit from treatment or have opted out of treatment through personal choice. This means that some patients who are not clinically eligible for treatment may appear in the numerator and the denominator of the treatment indicators.

Other validation

Indicators that did not relate to QOF indicators were externally validated using a variety of published references and data sources. These sources included the Indicators No Longer in QOF (INLIQ), the National CKD Audit and research publications. Where it was not possible to find existing data to externally validate newly developed indicators, advice and data checks were sought from partners.

Disclosure control

We applied NHS England disclosure rules that were defined for Hospital Episode Statistics (HES) in 2019. These rules round sub-national counts and denominators to the nearest 5 and suppresses any which are between 1 and 7. Displayed proportions are rounded to one decimal place.

Additional analysis for Annual reports

Adjusted effect of age on deprivation

Initial analysis of prevalence indicators by deprivation quintile was likely to be influenced by the known difference in age structure of areas with different levels of deprivation. Areas with higher deprivation tend to have younger populations and the high-risk conditions for CVD are often more prevalent with increasing age. To account for this and allow a comparison of prevalence of deprivation data both between and within areas, directly age-standardised prevalence estimates were created for each deprivation group.

NHS England publish GP list sizes by patient residence using their LSOA11 area, but not by patient residence and age. For round one synthetic denominator populations were created by deprivation group and age to use in the calculations of the standardised figures by using a technique called iterative proportional fitting. This technique used GP list sizes by patient residence and GP list sizes by age combined with the age and deprivation population distribution of mid-year estimates published by the Office for National Statistics. For round two synthetic denominators were created by applying the proportional age-splits of deprivation data within the Office for National Statistics mid-year estimates directly to the GP list size data.

Comorbidities reporting (First Annual Audit Report)

Comorbidity analysis was based on all people with GP recorded hypertension in the audit. All people in the audit with a GP record of hypertension were identified and the presence of a GP record of any of the following conditions was then flagged. CHD; stroke; PAD; aortic aneurism; heart failure; dyslipidaemia; CKD; non- diabetic hyperglycaemia (NDH); diabetes mellitus (DM); AF and obesity.

Each person was classified by the number of comorbidities that are recorded from zero up to 10 comorbidities. Counts of people with hypertension and the differing numbers of comorbidities were generated. Proportions of people with different numbers of comorbidities were derived using all people with GP recorded hypertension as the denominator.

The hypertension recorded audit population was split into different subgroups of people based on age, sex and ethnicity. Proportions for each subgroup were derived using the same methods as above, to describe variation in the number of comorbidities by these factors.

Impact on CVD prevention of the Covid-19 pandemic reporting (Second Annual Audit report)

Investigation of the blood pressure and cholesterol readings by month over the two year period required the analysis of both CVDPREVENT extracts (i.e. to the end of March 2020 and the end of March 2021). To ensure that the data is comparable a sample of the practices present in both extracts was taken. The sample included approximately half the practices in England. For ease of interpretation the data has been displayed on the same graph.