C-042

Relative contribution of primary care, hospital, and mortality data for myocardial infarction and stroke events case ascertainment in the UK

Katrina Wilcox Hagberg[1], Rebecca Persson[1], David Neasham[2], Michael Grayer[2], Swati Sakhuja[3], Francesco Giorgianni[2], Nafeesa Dhalwani[3], George Kafatos[2]

[1]**Boston Collaborative Drug Surveillance Program (BCDSP)**, Lexington, MA, USA; [2]**Amgen Limited**, Uxbridge, UK [3]**Amgen Inc**, Thousand Oaks, CA, USA

BACKGROUND

- Missing or misclassification of outcomes may bias estimates of cardiovascular risk and treatment benefit in real world data (RWD) studies, particularly in studies of patients with prior cardiovascular events
- In England, myocardial infarction (MI) and stroke events are expected to be treated in-hospital and captured in Hospital Episode Statistics (HES) data
- The extent to which MI/stroke events may be missing in HES is unclear, potentially leading to under-ascertainment of cases in studies relying on hospital data alone
- Office of National Statistics (ONS) death registry data records medical information captured on death certificates
- Clinical Practice Research Datalink (CPRD) Aurum captures events recorded in primary care settings
  - Incident (first ever) MI and stroke events in CPRD are of high validity[1-5]
  - To accurately identify new events among patients with a prior history in CPRD, careful distinction must be made between coding of follow-up care for a previous event and a subsequent new ("recurrent") event

OBJECTIVES

In a secondary prevention population with elevated LDL-C/non-HDL-C:

1. Identify new MI and stroke events using a) hospital (HES), b) primary care (CPRD Aurum), and c) mortality (ONS) data sources
2. Develop and apply an algorithm to identify incident and recurrent MI and stroke events in CPRD Aurum
3. Describe the relative contribution of HES, CPRD Aurum, and ONS data sources to the identification of MI and stroke events

METHODS

Study Population

- Among CPRD patients (aged ≥40 years) with HES/ONS linkage eligibility (2013–2017), we identified a cohort with atherosclerotic cardiovascular disease (ASCVD), defined by prior MI, stroke, or peripheral arterial disease, high LDL-C, and prior use of statins or ezetimibe
- Index on an LDL-C ≥70 mg/dL or non–HDL-C ≥100 mg/dL; last date of follow-up 31 March 2021

Outcome Definitions

Step 1. MI and stroke event selection in each data source, defined as:

HES: After index, selected the first MI or stroke diagnosis (ICD-10) recorded as the reason for hospitalization

ONS: Selected an MI or stroke diagnosis (ICD-10) recorded as the cause of death (primary or any secondary)

CPRD Aurum:
- Classified patients based on presence or absence of a pre-index MI/stroke diagnosis in CPRD Aurum (MedCode) or HES (i.e., Prior Event, No Prior Event)
- Algorithms were developed to 1) identify incident (first ever) diagnosis and 2) differentiate new MI/stroke events from diagnoses recorded as part of follow-up visits related to an earlier pre-index event (**Figures 1** and **2**).

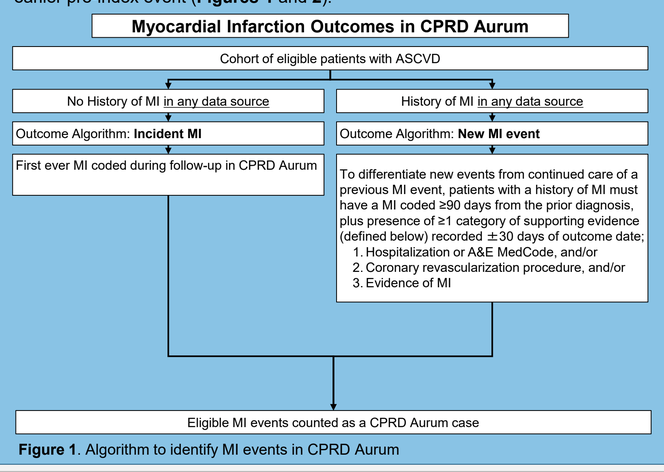


Figure 1. Algorithm to identify MI events in CPRD Aurum

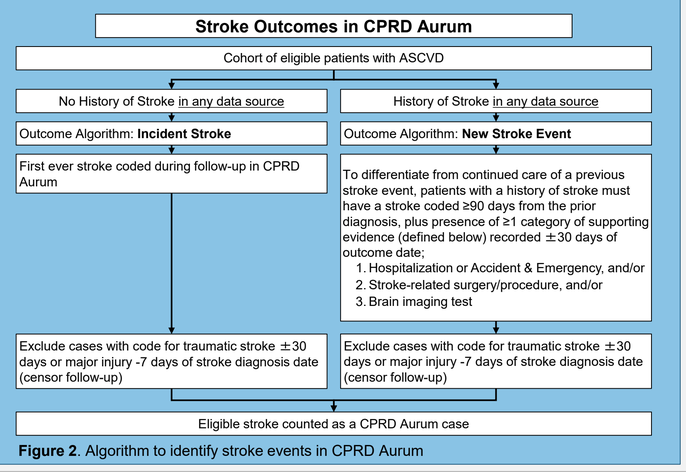


Figure 2. Algorithm to identify stroke events in CPRD Aurum

Step 2. Adjudication of MI and stroke events across all data sources
- Event adjudication was conducted across HES, ONS, and CPRD Aurum to avoid double counting
  - First eligible date regardless of data source was selected as the qualifying event
  - Traumatic strokes were excluded based on presence of trauma or injury-related censoring event in any data source
- The resulting events and dates relative to index were incorporated into a survival analysis (not discussed in this poster)

Step 3. Evaluate relative contribution of cases by data source

Analyses

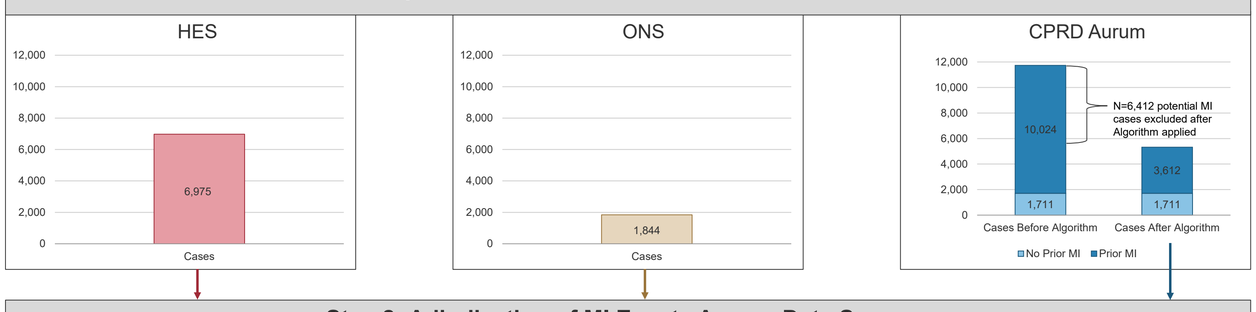
- In CPRD Aurum, describe N (%) of cases identified before and after application of algorithm among patients with a history of MI/stroke
- Evaluate the relative contribution of eligible events in HES, Aurum, and ONS

RESULTS

- 97,626 ASCVD patients met criteria for inclusion: 65.9% male; mean age 68.2 ±9.9 years; 87.2% White; median follow-up 5.5 (IQR 3.8, 7.2) years

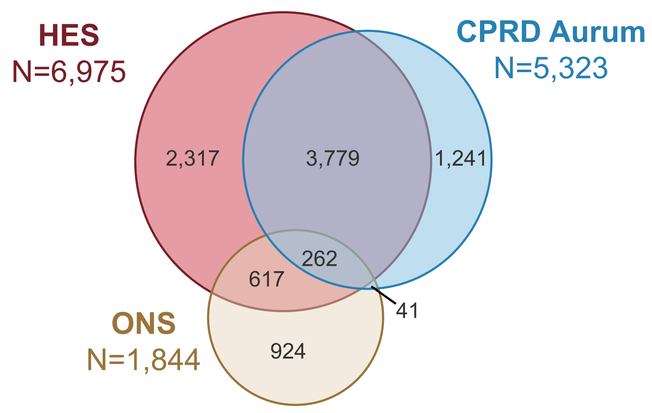
Myocardial Infarction

Step 1: MI Event Selection in Each Data Source



Step 2: Adjudication of MI Events Across Data Sources

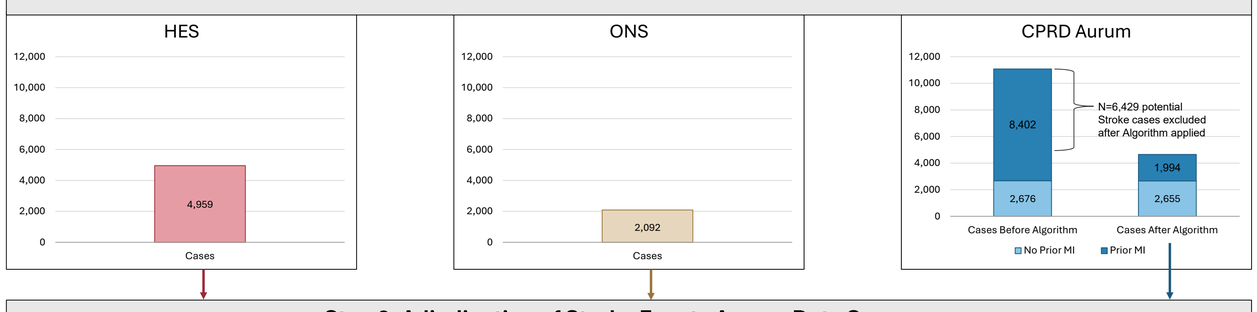
Step 3: Relative Contribution of Final MI Cases by Data Source (N=9,181)



- 6,975 MI events identified using HES alone
- 1,844 MI events identified in ONS alone
- In CPRD Aurum alone:
  - Before the algorithm was applied, 11,735 MI events identified
  - After application of the algorithm, 5,323 MI events eligible for inclusion
- After adjudication using all data sources, MI event identification increased to 9,181, representing an increase of 2.3% of the cohort classified as cases compared to using HES alone
- Of 9,181 MI events, 76.0% were in HES, 58.0% in Aurum, and 20.1% in ONS
  - 12.0% recorded in CPRD Aurum, but not HES
  - 10.1% in ONS only (not HES or CPRD Aurum)

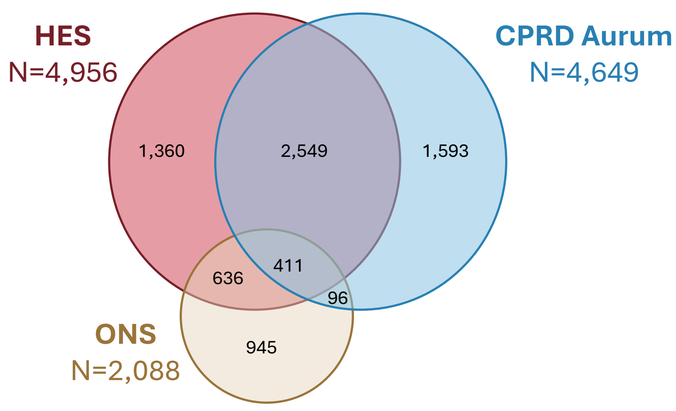
Stroke

Step 1: Stroke Event Selection in Each Data Source



Step 2: Adjudication of Stroke Events Across Data Sources

Step 3: Relative Contribution of Final Stroke Cases by Data Source (N=7,590)



- 4,959 stroke events identified using HES alone, and before excluding traumatic stroke events
- 2,092 stroke events identified in ONS alone, and before excluding traumatic stroke events
- In CPRD Aurum alone:
  - Before the algorithm was applied, 11,078 stroke events identified
  - After application of the algorithm, 4,649 stroke events eligible for inclusion
- After adjudication using all data sources, stroke event identification increased to 7,590, representing an increase of 2.7% of the cohort classified as cases compared to using HES alone
- Of 7,590 stroke events, 65.3% were in HES, 61.2% in CPRD Aurum, and 27.5% in ONS
  - 22.3% recorded in CPRD Aurum, but not HES
  - 12.5% in ONS only (not HES or CPRD Aurum)

CONCLUSIONS

- Case ascertainment using CPRD Aurum with linkage to HES and ONS data is high, enabling more complete capture of cardiovascular events across care settings
- For patients with a prior MI or stroke event, use of the algorithms to identify subsequent MI or stroke events in CPRD Aurum may reduce case misclassification

LIMITATIONS

- Study population was a secondary prevention population with elevated LDL-C/non-HDL-C; thus, generalizability may be limited
- Event linkage across data sources relied on pre-specified time windows, which may have resulted in some event misclassification
- The algorithms for recurrent MI and stroke identification were not formally validated against a clinical reference standard; therefore, some residual event misclassification may remain

REFERENCES

[1]Herrett E. doi: 10.1136/bmj.f2350. [2]Arana A, doi: 10.1002/pds.5150. [3]Morgan A, doi: 10.3399/BJGPO.2020.0117. [4]Ruigómez A, doi: 10.1007/s40264-021-01044-4. [5]Varas-Lorenzo C, doi: 10.1161/01.cir.101.22.2572.

See full abstract at BCDSP.org

DISCLOSURES

This study is based in part on data from the Clinical Practice Research Datalink obtained under license from the UK Medicines and Healthcare products Regulatory Agency. The data is provided by patients and collected by the NHS as part of their care and support. The interpretation and conclusions contained in this study are those of the authors alone. Hospital Episode Statistics (HES) and Office of National Statistics death registration data Copyright © (2026), re-used with the permission of The Health & Social Care Information Centre. All rights reserved. This study was approved by Research Data Governance (RDG). This study was approved by CPRD Research Data Governance (protocol 23_003305).

This study was funded by Amgen Ltd. Authors retain full and scientific control over the content of this presentation. KWH and RP received contractual research grants to their institution from Amgen and UCB during the conduct of the study. DN, FG, and GK are employees of Amgen Ltd and own shares of Amgen Inc.; SS and ND are employees of Amgen Inc and own shares of Amgen Inc. MG is a contract worker on assignment at Amgen Ltd. KWH is an unpaid volunteer of Clinical Practice Research Datalink Data Quality Advisory Group. The authors declare no other conflicts of interest in this work.

Presented at the 42nd International Conference on Pharmacoepidemiology and Therapeutic Risk Management (ICPE), 29 August – 2 September 2026, Milan, Italy
Correspondence: Katrina Wilcox Hagberg, khagberg@bcdsp.org