

Research Letter

# Pregnancy-Related Clinical Codes in Unlikely Populations in Primary Care

Helen Curtis, DPhil; Peter Inglesby, MPhil; Ben Goldacre, MRCPsych; Rose Higgins, MPhil; Amir Mehrkar, MRCP; Sebastian Bacon, BA; Brian MacKenna, MPharm

Bennett Institute for Applied Data Science, Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Site, Oxford, England, United Kingdom

**Corresponding Author:**

Helen Curtis, DPhil  
Bennett Institute for Applied Data Science, Nuffield Department of Primary Care Health Sciences  
University of Oxford, Radcliffe Observatory Site  
Woodstock Road  
Oxford, England OX2 6GG  
United Kingdom  
Phone: 44 1865617855  
Email: [helen.curtis@phc.ox.ac.uk](mailto:helen.curtis@phc.ox.ac.uk)

## Abstract

We discovered that the official list of clinical codes for pregnancy during the COVID-19 pandemic identified some unlikely pregnancies (for example, in older men), principally due to a code describing a specific fetal position (“knee presentation”), which notably lacks “fetal” in the code description. This is an informative example of commonly overlooked problems in creating and using clinical data.

*JMIR Med Inform* 2026;14:e89620; doi: [10.2196/89620](https://doi.org/10.2196/89620)

**Keywords:** electronic health records; pregnancy records in primary care; pregnancy; clinical codelists; SNOMED CT; Systematized Nomenclature of Medicine – Clinical Terms

## Introduction

During the COVID-19 pandemic, we monitored for the uptake of COVID-19 vaccines in OpenSAFELY, a trusted research environment covering 95% of the general practitioner (GP)-registered population in England. This included identifying patients with a code related to pregnancy or delivery leading up to the vaccination or vaccine eligibility period. We noticed that some codes occurred in people of male sex or those beyond reproductive age. We investigated which codes these were and their differential use in the 2 main primary care electronic health record (EHR) software systems, EMIS and TPP.

## Methods

### Overview

Our population included all patients registered with either an EMIS or TPP practice in OpenSAFELY on March 31, 2020, of male or female sex (at birth) and aged 16-119 years. For each EHR, we counted patients in each age/sex group (age

groups: 16-39, 40-69, and 70+ years) with any SNOMED CT (Systematized Nomenclature of Medicine – Clinical Terms) codes indicative of pregnancies and births as per the official national UK specification for COVID vaccine eligibility/analysis [1]. This included pregnancy, delivery, fetal presentations and abnormalities (full list available at OpenCodelists [2]). We counted those recorded between January 2020 and the latest date available at the time of the study (March 2021) and calculated the rate per 1000 for each age/sex group. We extracted the latest relevant code recorded per patient, and compared the recording rate per code between EMIS vs TPP practices per age/sex group. Extracting the latest code per patient pragmatically meets the privacy-preserving OpenSAFELY requirement of producing one-row-per-patient, where looping through 2000+ codes would present disproportionate processing burden. Age bands were chosen as applicable to the broader analysis exploring differences in all vaccine-eligibility criteria. The full code is available on GitHub [3].

## Ethical Considerations

This study was approved by the Health Research Authority (Research Ethics Committee reference 20/LO/0651) and by the London School of Hygiene and Tropical Medicine Ethics Board (reference 21863). For further details on information governance and ethical approval, please see [Multimedia Appendix 1](#).

pregnancy/delivery codes in EMIS and 0.3 per 1000 in TPP ([Table 1](#)); in females aged 70+ years, there were 0.1 per 1000 in EMIS and 0.5 per 1000 in TPP. The most common pregnancy/delivery clinical code in TPP among males and older females was 249098007, “knee presentation” (0.2 per 1000 patients aged 40-69 years, 0.1 per 1000 patients over 70 years, and 0.1 per 1000 males aged 16-39 years; [Table 1](#)). It was not found in older females or males in EMIS.

## Results

The total population was 47 million adults aged 16+ years. In males, there were 0.1 per 1000 patients with recent

**Table 1.** Rates of patients having recent pregnancy/delivery codes, per 1000 registered patients, and their latest corresponding code, in EMIS and TPP practices in England, by sex and age band, 2020-21. Total registered patients as of March 31, 2021, is also shown for context. Codes are arranged into groups covering pregnancy, delivery/postpartum, and miscarriage. The rates per individual code do not represent prevalence as only the latest single code per person is counted (eg, gestational diabetes mellitus will often be followed by another code, such as a birth record). Codes with maximum rates <0.05 across all groups are excluded.

Codes and GP <sup>a</sup> system supplier	Patient sex and age band (years)					
	Male			Female		
	16-39	40-69	70+	16-39 <sup>b</sup>	40-69 <sup>b</sup>	70+
Total registered patients (millions)						
EMIS	5.27	6.23	1.99	5.19	6.00	2.41
TPP	3.73	4.58	1.58	3.66	4.44	1.90
Recent pregnancy/delivery (rate per 1000)						
EMIS	0.1	0.1	0.1	100.4	8.2	0.1
TPP	0.2	0.3	0.4	99.4	6.3	0.5
SNOMED <sup>c</sup> code (description)						
Pregnancy						
77386006 (pregnant)						
EMIS	<0.05	<0.05	— <sup>d</sup>	25.9	1.6	<0.05
TPP	<0.05	<0.05	<0.05	26.4	1.2	<0.05
11687002 (gestational diabetes mellitus)						
EMIS	—	—	—	<1.9	0.3	—
TPP	—	<0.05	—	2.1	0.3	—
Delivery and postpartum						
309469004 (spontaneous vertex delivery)						
EMIS	—	—	—	4.6	0.2	—
TPP	—	—	—	7.8	0.3	—
169826009 (single live birth)						
EMIS	<0.05	—	—	5.1	0.3	—
TPP	—	—	—	7.6	0.4	—
48782003 (delivery normal)						
EMIS	—	—	—	6.0	0.3	—
TPP	<0.05	—	<0.05	4.6	0.2	<0.05
177184002 (normal delivery procedure)						
EMIS	—	—	—	5.0	0.3	<0.05
TPP	—	—	—	<2.0	<0.2	—
274130007 (emergency cesarean section)						
EMIS	—	—	—	3.1	0.3	—
TPP	—	—	—	2.1	<0.2	—
177141003 (elective cesarean section)						

Codes and GP <sup>a</sup> system supplier	Patient sex and age band (years)					
	Male			Female		
	16-39	40-69	70+	16-39 <sup>b</sup>	40-69 <sup>b</sup>	70+
EMIS	—	—	—	2.6	0.3	—
TPP	—	—	—	2.0	0.2	—
200144004 (deliveries by cesarean)						
EMIS	—	—	—	1.9	<0.2	—
TPP	—	—	—	<2.0	<0.2	—
11466000 (cesarean section)						
EMIS	—	—	—	<1.9	<0.2	—
TPP	—	—	—	2.0	0.2	—
249098007 (knee presentation)						
EMIS	—	—	—	<1.9	<0.2	—
TPP	0.1	0.2	0.1	<2.0	0.2	0.1
133906008 (postpartum care)						
EMIS	—	—	—	8.7	0.6	—
TPP	—	—	—	<2.0	<0.2	—
Miscarriage						
156073000 (complete miscarriage)						
EMIS	—	—	—	<1.9	<0.2	—
TPP	—	—	—	2.0	0.3	—
17369002 (miscarriage)						
EMIS	<0.05	—	—	4.4	0.9	—
TPP	—	—	—	2.6	0.4	—

<sup>a</sup>GP: general practitioner.

<sup>b</sup>Extracted counts were limited to the top 10 codes per age/sex group per GP system supplier and codes with  $\geq 10$  occurrences, so rates sum to less than the overall rates. This also means that some counts are shown as less than (“<”) the rate of the 10th most common code.

<sup>c</sup>SNOMED: Systematized Nomenclature of Medicine.

<sup>d</sup>Dashes represent codes not found.

For context, there were 100.4 and 99.4 patients with recent pregnancy/delivery codes per 1000 females aged 16-39 years in EMIS and TPP practices, respectively, and 8.2 vs 6.3 per 1000 females aged 40-69 years in EMIS vs TPP.

## Discussion

### Principal Findings

Between 0.1 and 0.4 per 1000 patients in English primary care who were assigned male sex at birth had a recent code indicative of pregnancy/delivery in their GP record, highest in practices using TPP software and increasing by age group. Additionally, up to 0.5 per 1000 registered female patients aged 70 years and over had a recent code indicative of pregnancy or birth.

The most common pregnancy/delivery code in older adults and males in TPP practices was 249098007, “knee presentation.” Despite its description not being explicitly related to delivery, in the SNOMED CT hierarchy, it is a child code of “malpresentation of fetus,” that is, the knee of the fetus emerging first in labor (a rare occurrence). This may be due to its incorrect inclusion in some local prespecified clinical templates created by users for knee-related patient

consultations in some areas using TPP software. National primary care SNOMED CT counts indicate that use of this code is increasing, from 4290 times in 2020/21 to 9350 in 2024/25 [4]. It is possible that other codes are added for similar reasons (eg, gestational diabetes being part of a diabetes template). Other pregnancy-related codes may be erroneously recorded (through free text input) when a relevant pregnancy is mentioned in a consultation, for example, that of the patient’s partner. Interoperable messaging standards that automatically update GP records with coded information [5] have also suffered from pregnancy coding errors [6]. New records being created after gender transition [7] could account for some apparent male pregnancies but would not explain recent pregnancies in the 70+ years group, nor the prevalence of knee presentation codes.

These findings have implications for researchers developing codelists for research; sensitivity analysis of the cohort selected is important to confirm consistency with the use of clinical codes by users of clinical systems. An event such as birth is likely to generate multiple codes per patient, many of which will be widely used. When using very wide codelists to select patients or conditions, users may wish to check the proportion with a single (potentially obscure) code recorded and lacking any other positive indicators.

## Conclusion

Clinical codes may not always reflect the expected condition. Secondary users of clinical data should check carefully that

the clinical codes included in their algorithms accurately identify the intended concept.

## Acknowledgments

We are very grateful for all the support received from the EMIS and TPP Technical Operations teams throughout this work, and for generous assistance from the information governance and database teams at National Health Service (NHS) England and the NHS England Transformation Directorate. We acknowledge all other members of the OpenSAFELY Collaborative who have contributed to the development of the platform.

## Funding

Authors received funding from a National Institute for Health and Care Research (NIHR) Programme Development Grant (NIHR206954) during the submission of this work. The OpenSAFELY platform is principally funded by grants from National Health Service (NHS) England [2023-2025], the Wellcome Trust (222097/Z/20/Z [2020-2024] and 311535/Z/24/Z [2025-2031]), and the Medical Research Council (MRC) (MR/V015737/1 [2020-2021]). Additional contributions to OpenSAFELY have been funded by grants from the MRC via the National Core Study program Longitudinal Health and Wellbeing strand (MC\_PC\_20030, MC\_PC\_20059 [2020-2022]) and the Data and Connectivity strand (MC\_PC\_20058 [2021-2022]), the NIHR and the MRC via the CONVALESCENCE programme (COV-LT-0009, MC\_PC\_20051 [2021-2024]), and NHS England via the Primary Care Medicines Analytics Unit [2021-2024]. The views expressed are those of the authors and not necessarily those of the NIHR, NHS England, UK Health Security Agency (UKHSA), the Department of Health and Social Care, or other funders. Funders had no role in the study design, collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

## Data Availability

Please see [Multimedia Appendix 1](#) for further information on data sharing, software, and reproducibility.

## Conflicts of Interest

BG has received research funding from the Bennett Foundation, the Laura and John Arnold Foundation, the National Health Service (NHS) National Institute for Health Research (NIHR), the NIHR School of Primary Care Research, NHS England, the NIHR Oxford Biomedical Research Centre, the Mohn-Westlake Foundation, NIHR Applied Research Collaboration Oxford and Thames Valley, the Wellcome Trust, the Good Thinking Foundation, Health Data Research UK, the Health Foundation, the World Health Organization, UKRI MRC, Asthma UK, the British Lung Foundation, and the Longitudinal Health and Wellbeing strand of the National Core Studies program; he has previously been a non-executive director at NHS Digital. He also receives personal income from speaking and writing for lay audiences on the misuse of science. BM is also employed by NHS England, working on medicines policy and is the clinical lead for primary care medicines data. The other authors declare no conflicts of interest.

## Multimedia Appendix 1

Supplementary information on information governance and ethical approval as well as data availability.

[\[DOCX File \(Microsoft Word File\), 18 KB-Multimedia Appendix 1\]](#)

## References

1. Curtis HJ, Inglesby P, Morton CE, et al. Trends and clinical characteristics of COVID-19 vaccine recipients: a federated analysis of 57.9 million patients' primary care records in situ using OpenSAFELY. *Br J Gen Pract*. Jan 2022;72(714):e51-e62. [doi: [10.3399/BJGP.2021.0376](https://doi.org/10.3399/BJGP.2021.0376)] [Medline: [34750106](https://pubmed.ncbi.nlm.nih.gov/34750106/)]
2. PRIMIS. Pregnancy or delivery codes recorded in the 8.5 months before audit run date. OpenCodelists. 2021. URL: <https://www.opencodelists.org/codelist/primis-covid19-vacc-uptake/pregdel/v1.2.1/> [Accessed 2026-06-22]
3. Curtis H. OpenSAFELY PRIMIS codelist prevalence. GitHub. 2021. URL: <https://github.com/opensafely/primis-codelist-prevalence> [Accessed 2025-08-11]
4. Wiedemann M, Tamborska A, Higgins R, et al. General practice (SNOMED CT). OpenCodeCounts. 2025. URL: <https://bennetxford.github.io/opencodecounts/articles/app.html> [Accessed 2025-08-11]
5. GP Connect: Update Record. NHS England Digital. URL: <https://digital.nhs.uk/services/gp-connect/gp-connect-in-your-organisation/gp-connect-update-record> [Accessed 2025-12-09]
6. Pregnant patients received 'inappropriate' medication due to pharmacy coding errors. Patient Safety Learning - the hub. URL: <https://www.pslhub.org/blogs/entry/8697-pregnant-patients-received-%E2%80%9Binappropriate%E2%80%99-medication-due-to-pharmacy-coding-errors/> [Accessed 2025-12-09]

7. Process for registering a patient gender re-assignment. Primary Care Support England (PCSE). 2023. URL: <https://pcse.england.nhs.uk/sites/default/files/2023-08/process-for-registering-a-patient-gender-re-assignmentv10.pdf> [Accessed 2025-12-09]

---

## Abbreviations

**EHR:** electronic health record

**GP:** general practitioner

**SNOMED CT:** Systematized Nomenclature of Medicine – Clinical Terms

---

*Edited by Arriel Benis; peer-reviewed by Aaron Jones, Cheryl McQuire; submitted 15.Dec.2025; final revised version received 14.Apr.2026; accepted 11 Jun.2026; published 30.Jun.2026*

*Please cite as:*

*Curtis H, Inglesby P, Goldacre B, Higgins R, Mehrkar A, Bacon S, MacKenna B*

*Pregnancy-Related Clinical Codes in Unlikely Populations in Primary Care*

*JMIR Med Inform 2026;14:e89620*

*URL: <https://medinform.jmir.org/2026/1/e89620>*

*doi: [10.2196/89620](https://doi.org/10.2196/89620)*

© Helen Curtis, Peter Inglesby, Ben Goldacre, Rose Higgins, Amir Mehrkar, Sebastian Bacon, Brian MacKenna. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 30.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.