

Aniek F. Markus¹, Julieta Politi¹, Anum Zahra¹, Maarten van Kessel¹, Adam Black¹, Ger Inberg¹, Berta Raventós¹, Maxim Moinat¹, Daniel R. Morales², María-Clara Restrepo-Méndez², Katia Verhamme^{1,3}, Talita Duarte-Salles^{1,4}

¹Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands

²Data Analytics and Methods Task Force, European Medicine Agency, Amsterdam, The Netherlands

³Department of Infection Control & Epidemiology, OLV Hospital, Aalst, Belgium

⁴Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain

BACKGROUND

Routinely-collected data are commonly used in pregnancy research by applying rule-based algorithms to reliably extract pregnancy episodes. Accurate identification of pregnancy episodes is essential for advancing maternal and child health outcomes and informing regulatory decisions. However, validation is often missing or limited to certain databases, limiting our understanding on the quality of the extracted pregnancy episodes. For pregnancy research at scale, we need a standardised quality assessment framework that can be applied across heterogeneous databases.

OBJECTIVES

1. To develop a standardised quality assessment framework for evaluating the accuracy of derived pregnancy episodes.
2. To validate the framework across datasets within the DARWIN EU® network to assess readiness for multi-dataset studies and regulatory use.

METHODS

We designed a framework including four quality dimensions of pregnancy episodes (see Table 1 for metrics):

1. **Episode frequency:** plausibility and consistency of counted pregnancies within and across databases,
2. **Episode duration:** accuracy of gestational length and alignment with known clinical timelines per pregnancy outcome,
3. **Episode construction:** completeness and correctness of the algorithmic output,
4. **Episode outcomes:** capture of key maternal and fetal outcomes linked to each pregnancy episode.

Table 1. Metrics to assess internal consistency and plausibility of pregnancy episodes.

Metric	Purpose	Expected Output
Episode Frequency		
Number of pregnancy episodes and pregnant individuals	Evaluate pregnancy frequency overall and per individual	Total episodes, total individuals, distribution of pregnancies per person (min, max, median, IQR), total number of first pregnancy episodes, total number of subsequent pregnancy episodes (pregnancies with ≥1 preceding episode for the same individual)
Maternal age distribution	Assess pregnancy distribution by age	Distribution of maternal age at pregnancy start (min, max, median, IQR)
Pregnancy incidence rate	Evaluate number of pregnancies relative to database population	Pregnancies per 1,000 individuals of female sex (at birth) aged 12–55
Episode Duration		
Gestational duration distribution	Assess plausibility of pregnancy length	Min, max, median, IQR; count and % of episodes <1 day, >308 days
Distribution of gestational age at anchor event (e.g., delivery, first antenatal visit, ultrasound after first trimester, gestational diabetes mellitus)	Assess whether anchor events occur at plausible gestational ages	Histogram/density plots of gestational age at event; % preterm, term, post-term
Gestational week bins	Summarise gestational age into clinically relevant categories	% in bins: <12 12–27+6 (extremely preterm) 28–31+6 (very preterm) 32–36+6 (moderate to late preterm) 37–40 -41+6 (term) >42+ (post-term)
Seasonal and temporal patterns	Assess expected timing patterns, detect artifacts	Monthly/annual trends of pregnancy start and end
Episode Construction		
Overlapping pregnancy episodes	Identify episodes that occur simultaneously within individuals	Count and % of overlapping episodes; individual-level review
Missing or inconsistent dates	Identify technical errors in episode construction	Count and % of episodes with missing or reversed dates
Control variables not used in the algorithm (e.g., breech presentation, preterm labour, polyhydramnios)	Validate episode timing using external events not used in construction	Number of occurrences, number or occurrences overlapping with pregnancy episodes, percent overlap
Episode Outcomes		
Mode of delivery	Evaluate the proportion of delivery modes	% vaginal and c-section deliveries
Pregnancy outcome events	Evaluate the proportion of delivery outcomes	Proportion of pregnancy outcome events and distribution of duration by pregnancy outcome type

RESULTS

The pregnancy episode quality framework was applied across 15 DARWIN EU® databases. Preliminary findings demonstrate substantial heterogeneity in the ability to detect pregnancy episodes across the network. Standardised output across databases allowed to identify and prioritise the most high-impact issues based on feasibility and impact, supporting algorithm refinements as shown in Figure 1.

Episode duration and episode construction were recognized as key areas for refinement, as implausible gestational lengths (> 44 weeks) were observed in up to 20% of episodes across all databases and pregnancy overlap occurred in more than 5% of individuals with multiple pregnancies in nine databases.

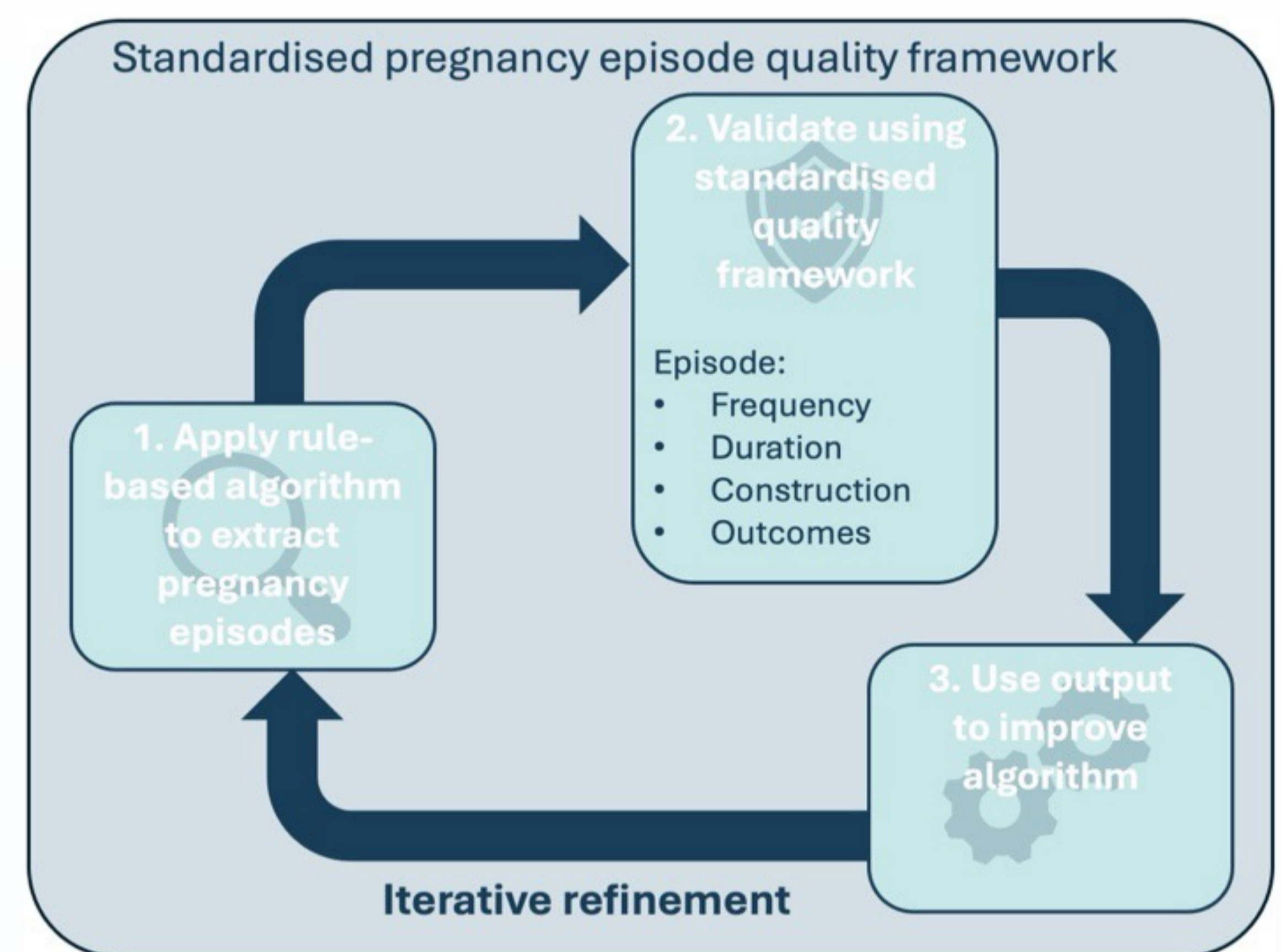


Figure 1. Proposed framework to evaluate and improve the accuracy of derived pregnancy episodes.

CONCLUSIONS

The proposed framework enables a standardised, transparent assessment of the performance of pregnancy episode algorithms across heterogeneous real-world data mapped to the OMOP Common Data Model. It supports targeted fixes to improve derived pregnancy episodes across databases by providing insight into differences between health care systems and data granularity. Future work will define minimum quality requirements, assess performance over time, and benchmark the results against national statistics.

FUNDING

This study (EUPAS100000812) was funded by EMA and performed via DARWIN EU®. The study funder was involved in revising the study protocol and the objectives. Data partners' role is only to execute code at their data source. They do not have an investigator role. This communication represents the views of the DARWIN EU® Coordination Centre only and cannot be interpreted as reflecting those of the EMA or the European Medicines Regulatory Network.

DISCLOSURE

This presentation represents the views of the DARWIN EU® Coordination Centre only and cannot be interpreted as reflecting those of the European Medicines Agency or the European Medicines Regulatory Network.

CONTACT

Aniek Markus (a.markus@darwin-eu.org)

Julieta Politi (j.politi@darwin-eu.org)