

International Journal of Population Data Science

Journal Website: www.ijpds.org



Exploring different models for the standardisation of data for the Born and Bred in (BaBi) Meta-cohort

Sabah Shaib¹, Alex Newsham¹, Hollie Henderson¹, Dan Mason¹, Kath Roberts², Kayley Ciesla¹, and Sally Bridges¹

¹Bradford Institute for Health Research, Bradford, United Kingdom

²University of York, York, United Kingdom

Objectives

To assess, evaluate and determine the most appropriate data model for implementation by the BaBi Network, which will collaborate to create a meta-cohort with a standardised dataset. A common data model serves as a blueprint for the standardisation of data across multiple sites and organisations.

Conclusion

Opting for a bespoke common data model would benefit all BaBi sites including those who lack resources, skills and capabilities. However, this doesn't rule out migrating the bespoke BaBi common data model into OMOP at a future date, if this is deemed beneficial for future data collaboration and sharing.

Methods

BaBi is an evolving multi-site cohort study which involves recruiting pregnant women and their babies with the focus on linking together routine data across organisations. Online resources were consulted including attendance at a webinar on an established common data model (the Observational Medical Outcomes Partnership (OMOP) model) to explore the practical applications of a common data model in research contexts. Discussions were conducted with BaBi sites to assess their skills, capacity, and infrastructure, providing a deeper understanding of how to implement a data model that would be feasible for all parties while also meeting the brief.

Results

Despite significant advantages offered by the OMOP model, including data standardisation, this model was found to be unsuitable for standardising datasets for the BaBi meta-cohort. It was determined that OMOP may introduce unnecessary complexity especially where BaBi sites are in their early stages. BaBi sites have varying technical skills and capabilities, and we want to avoid imposing excessive technical demands at this stage. Since sites join our cohort at different times, not all will initially have a dedicated data specialist, though they are expected to hire one as they expand. Therefore, we recommend the implementation of a bespoke common data model given that it can be tailored to the project's specific needs while accounting for variations in skills and capacity across the network.

