Linkage of mother and child pairs in the information system for research in primary care (SIDIAP) in Catalonia

Dr Talita Duarte-Salles[1]; Leonardo Méndez-Boo[2]; Yesika Díaz[1]; Eduardo Hermosilla[1]; Maria Aragón[1]; Francesc Fina[2]; Daniel Prieto-Alhambra[3]; Manuel Medina-Peralta[2]; Bonaventura Bolíbar[1]; Maria Garcia-Gil[1]

1 Institut Universitari d'Investigació en Atenció Primària Jordi Gol (IDIAP Jordi Gol), Barcelona, Spain

2 Institut Català de la Salut, Barcelona, Spain

3 NDORMS, University of Oxford, Oxford, UK

Background

Studies on the benefits and risks of drugs during pregnancy are limited given that pregnant women are usually excluded from clinical trials for ethical reasons. The Information System for Research in Primary Care (SIDIAP) is a potential data source to investigate adverse drug reactions during pregnancy and early life. However, only a very small fraction of mother and child health records are directly linked in SIDIAP.

Objectives

To develop and validate an algorithm to link mother and child pairs in SIDIAP for children born from 2005 to 2016 in Catalonia.

Methods

SIDIAP contains anonymized information recorded in primary care centres of the Catalan Health Institute (CHI) for 5.8 million people since 2005, including clinical diagnoses, laboratory tests, prescribed/dispensed drugs, vaccinations, demographic/lifestyle information. 692.402 subjects born from 2005 to 2016 were included. The CHI has access to personal data and used the following information from women and children to create an algorithm of linkage: national insurance number and the co-insurance status (holder/beneficiary), family name, age (women aged 14-49 years), dates of birth and delivery. Descriptive comparisons of those linked and unlinked motherchild pairs were conducted. Children with direct linkage available in their records (n = 7.307) were used as the gold-standard of true matches in order to validate the created algorithm.

Results

The proposed algorithm established mother and child pairs for 67.8% (n = 241.859) of all boys and 68.1% (n = 228.687) of girls in SIDIAP. Of these, 382.082 (81.2%) pairs were linked using only national insurance number and co-insurance status. The mean of age (6.1 [SD = 3.3] vs 6.7 [SD = 3.5] years) and the MEDEA index of deprivation (22.0% vs 22.6% in the most deprived) was similar, while the proportion of foreigners (10.5% vs 20.8%) was lower in the linked compared with the unlinked children, respectively. The proposed algorithm established linkages for 5.358 children in the validation dataset, and 95.4% of those pairs were confirmed as true matches.

Conclusions

The linkage of mother and child pairs in this large electronic health records data base provides a powerful resource for investigating maternal and child health and a useful tool for studies on

pharmacovigilance and drug and vaccine safety among pregnant women. Additionally, the linked population is representative of the total children born from 2005 to 2016 in SIDIAP, although it underrepresents foreigner population.