Elsevier Editorial System(tm) for Vaccine
Manuscript Draft

Manuscript Number:

Title: ADVANCE system testing: can coverage of pertussis vaccination be estimated in European countries using electronic healthcare databases: an example

Article Type: SI: ADVANCE

Keywords: Acellular pertussis vaccination; whole cell pertussis vaccination; vaccination coverage estimation; proof-of-concept study; database characteristics; benchmarking

Corresponding Author: Dr. Hanne-Dorthe Emborg,

Corresponding Author's Institution: Statens Serum Institut

First Author: Hanne-Dorthe Emborg

Order of Authors: Hanne-Dorthe Emborg; Johnny Kahlert; Toon Braeye; Jorgen Bauwens; Kaatje Bollaerts; Giorgia Danieli; Talita Duarte-Salles; Steffen Glismann; Consuelo Huerta; Simon de Lusignan; Elisa Martín-Merino; Chris McGee; Ana Correa; Lara Tramontan; Daniel Weibel; Miriam Sturkenboom

Abstract: Introduction
The Accelerated Development of VAccine beNefit-risk Collaboration in Europe (ADVANCE) is a public-private collaboration aiming to develop and test a system for rapid benefit-risk (B/R) monitoring of vaccines, using existing healthcare databases in Europe. The objective of this paper was to assess the feasibility of using electronic healthcare databases to estimate dose-specific acellular pertussis (aP) and whole cell pertussis (wP) vaccine coverage.
Methods
Seven electronic healthcare databases in four European countries (Denmark (n=2), UK (n=2), Spain (n=2) and Italy (n=1)) participated in this study. Children were included from birth and followed up to age six years. Vaccination exposure was obtained from the databases and classified by type (aP or wP), and dose. Coverage was estimated using period prevalence. For the 2006 birth cohort, two estimation methods for pertussis vaccine coverage, period prevalence and cumulative incidence were compared for each database.
Results
The majority of the 2,575,576 children included had been vaccinated at the country-specific recommended ages. Overall, the estimated coverage was 88-97% in Denmark (birth cohorts from 2003 to 2014), 96-100% in the UK (2003-2014), 95-98% in Spain (2004-2014) and 94% in Italy (2006-2007). The estimated coverage per birth cohort in Denmark and the UK differed by 1-6% compared with national estimates, with our estimates mostly higher. The estimated coverage in Spain differed by 0-2% with no consistent over- or underestimation. In Italy, the estimates were 3% lower compared with the national estimates. Except for Italy, for which the two coverage estimation methods generated the same results, the estimated cumulative incidence coverages were consistently 1% to 10% lower than period prevalence estimates.

Conclusion
This study showed that it was possible to provide reliable estimates of pertussis immunisation coverage from the electronic healthcare databases included, and that the estimates were comparable with the national estimates.

Dr Gregory A Poland
Editor-in-Chief, Vaccine

26 November 2018

Dear Dr Poland

We are pleased to submit our paper 'ADVANCE system testing: can coverage of pertussis vaccination be estimated in European countries using electronic health data: an example' to your Journal, Vaccine for the ADVANCE supplement. This paper describes the feasibility to estimate aP and wP pertussis vaccine coverage based on electronic health care databases. It is the fourth of the series of 10 papers that will be included in the supplement.

On behalf of all co-authors

Senior scientist, Hanne-Dorthe Emborg

Biostatistician Johnny Kahlert

*Author Agreement

We, the undersigned, Hanne-Dorthe Emborg and Johnny Kahlert declare that all authors have seen and approved the final version of the manuscript being submitted. We warrant that the article is our original work that has not been previously published and is not under consideration for publication elsewhere

Senior Scientist Hanne-Dorthe Emborg

Biostatistician Johnny Kahlert

***Suggested Reviewers**

Name	Institute	email
Øystein Rolandsen Riise	Norwegian Institute of Public Health	oysteinrolandsen.riise@fhi.no
Linda Y. Fu	Goldberg Center for Community, Pediatric Health, Children's National Health Center	lfu@childrensnational.org
Heidi Theeten	University of Antwerp	heidi.theeten@uantwerpen.be
Elizabeth T. Luman	Centers for Disease Control and Prevention	ECL7@cdc.gov
Pierre Van Damme	University of Antwerp	pierre.vandamme@uantwerpen.be
Phung Lang	University of Zurich	phung.lang@uzh.ch
Philip J. Smith	Centers for Disease Control and Prevention, National Center for Immunizations and Respiratory Disease, Immunization Services Division	philipsmith@alumni.albany.edu
Lone Simonsen	Department of Science and Environment, Roskilde University, Roskilde DK-4000, Denmark	lonesimo@ruc.dk
Daniel Levy-Bruhl	Santé publique France, French National Public Health Agency, Saint-Maurice	Daniel.Levy-Bruhl@santepubliquefrance.fr

***Declaration of Interest Statement**

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Hanne-Dorthe Emborg, Toon Braeye, Jorgen Bauwens, Kaatje Bollaerts, Giorgia Danieli, Talita Duarte-Salles, Consuelo Huerta Elisa Martin, Chris McGee, Ana Correa and Lara Tramontan declared no conflicts of interest. Johnny Kahlert declared that although he does not personally receive fees, honoraria, or grants he is employed at Department of Clinical Epidemiology, Aarhus University Hospital that receives research grants from various pharmaceutical companies administered by Aarhus University. Steffen Glismann is employed by the GSK group of companies and holds company shares. Simon de Lusignan declared he has received funding through his University to conduct enhanced surveillance of influenza vaccine (GSK), and is a member of Seqirus and Sanofi Pasteur advisory boards for which he received personal payment within the limits defined by his university. Daniel Weibel declared that he has received consultancy fees from GSK unrelated to the submitted work. Miriam Sturkenboom declared that she has received grants from Novartis, CDC and Bill & Melinda Gates Foundation, for work unrelated to the submitted work.

Highlights

- Feasibility of estimating pertussis vaccination covering using 7 European healthcare databases
- The majority of children were vaccinated at the recommended age of vaccination
- Two estimation methods provided comparable coverage estimates
- Benchmarking using national coverage estimates showed comparable results
- The approached used provided reliable estimates for pertussis vaccination coverage

Abstract

Introduction

The Accelerated Development of VAccine beNefit-risk Collaboration in Europe (ADVANCE) is a public-private collaboration aiming to develop and test a system for rapid benefit-risk (B/R) monitoring of vaccines, using existing healthcare databases in Europe. The objective of this paper was to assess the feasibility of using electronic healthcare databases to estimate dose-specific acellular pertussis (aP) and whole cell pertussis (wP) vaccine coverage.

Methods

Seven electronic healthcare databases in four European countries (Denmark (n=2), UK (n=2), Spain (n=2) and Italy (n=1)) participated in this study. Children were included from birth and followed up to age six years. Vaccination exposure was obtained from the databases and classified by type (aP or wP), and dose. Coverage was estimated using period prevalence. For the 2006 birth cohort, two estimation methods for pertussis vaccine coverage, period prevalence and cumulative incidence were compared for each database.

Results

The majority of the 2,575,576 children included had been vaccinated at the country-specific recommended ages. Overall, the estimated coverage was 88-97% in Denmark (birth cohorts from 2003 to 2014), 96-100% in the UK (2003-2014), 95-98% in Spain (2004-2014) and 94% in Italy (2006-2007). The estimated coverage per birth cohort in Denmark and the UK differed by 1-6% compared with national estimates, with our estimates mostly higher. The estimated coverage in Spain differed by 0-2% with no consistent over- or underestimation. In Italy, the estimates were 3% lower compared with the national estimates. Except for Italy, for which the two coverage estimation methods generated the same results, the estimated cumulative incidence coverages were consistently 1% to 10% lower than period prevalence estimates.

Conclusion

This study showed that it was possible to provide reliable estimates of pertussis immunisation coverage from the electronic healthcare databases included, and that the estimates were comparable with the national estimates.

ADVANCE system testing: can coverage of pertussis vaccination be estimated in European countries using electronic healthcare databases: an example

Hanne-Dorthe Emborg[a,*,#], Johnny Kahlert[b,#], Toon Braeye[c], Jorgen Bauwens [d,e,f], Kaatje Bollaerts[g], Giorgia Danieli[h,i], Talita Duarte-Salles[j], Steffen Glismann[k], Consuelo Huerta[l], Simon de Lusignan[m,n], Elisa Martín-Merino[l], Chris McGee[m,n], Ana Correa[m,], Lara Tramontan[h,i], Daniel Weibel[o,p] Miriam Sturkenboom[g,p,q]

[a]Statens Serum Institut, Artillerivej 5, DK-2300 Copenhagen, Denmark (hde@ssi.dk)

[b]Aarhus University Hospital, Olof Palmes Alle 43-45, DK-8200 Aarhus, Denmark (jok@clin.au.dk)

[c]Sciensano Rue Juliette Wytsmanstraat 14, 1050, Brussels Belgium (toon.braeye@sciensano.be)

[d]University Children's Hospital Basel, PO Box, CH 4033 Basel, Switzerland (j.bauwens@brightoncollaboration.org)

[e]University of Basel, Switzerland (j.bauwens@brightoncollaboration.org)

[f]Brighton Collaboration Foundation, Switzerland (j.bauwens@brightoncollaboration.org)

[g]P-95 Epidemiology and Pharmacovigilance, Leuven, Belgium (kaatje.bollaerts@p-95.com; m.c.j.sturkenboom@umcutrecht.nl)

[h]Consorzio Arsenàl.IT, Veneto Region, Italy (gdanieli@consorzioarsenal.it; ltramontan@consorzioarsenal.it)

[i]PEDIANET, Padova, Italy (gdanieli@consorzioarsenal.it; ltramontan@consorzioarsenal.it)

[j]Intitut Universitari d'Investigació en Atenció Primària Jordi Gol (IDIAP Jordi Gol), Barcelona, Spain (tduarte@idiapjgol.org)

[k]GSK, Av. Fleming 20, 1300, Wavre, Belgium (steffen.x.glismann@gsk.com)

[l]BIFAP database, Spanish Agency of Medicines and Medical Devices, Madrid Spain (chuerta@aemps.es; emartinm@aemps.es)

[m]University of Surrey, Guildford, Surrey GU2 7XH, UK (s.lusignan@surrey.ac.uk; c.mcgee@surrey.ac.uk; accorrea1@googlemail.com)

[n]Royal College of General Practitioners Research and Surveillance Centre, 30 Euston Square, London NW1 2FB, UK (s.lusignan@surrey.ac.uk)

[o]Erasmus University Medical Center, PO box 2014, 3000 CA, Rotterdam, the Netherlands (d.weibel@vaccinegrid.org; m.c.j.sturkenboom@umcutrecht.nl)

[p]VACCINE.GRID Foundation, Spitalstrasse 33, Basel, Switzerland (d.weibel@vaccinegrid.org; m.c.j.sturkenboom@umcutrecht.nl)

[q]Julius Global Health, Julius Center, University Medical Center Utrecht, Heidelberglaan 100, The Netherlands

* Corresponding author. Hanne-Dorthe Emborg Tel +45 32683937; fax +45 32683868

# These authors contributed equally to the work and share first authorship

ADVANCE system testing: can coverage of pertussis vaccination be estimated in European countries using electronic healthcare databases: an example

Hanne-Dorthe Emborg[a,*,#], Johnny Kahlert[b,#], Toon Braeye[c], Jorgen Bauwens [d,e,f], Kaatje Bollaerts[g], Giorgia Danieli[h,i], Talita Duarte-Salles[j], Steffen Glismann[k], Consuelo Huerta[l], Simon de Lusignan[m,n], Elisa Martin[l], Chris McGee[m,n], Ana Correa[m,], Lara Tramontan[h,i], Daniel Weibel[o,p] Miriam Sturkenboom[g,p,q]

[a]Statens Serum Institut, Artillerivej 5, DK-2300 Copenhagen, Denmark (hde@ssi.dk)

[b]Aarhus University Hospital, Olof Palmes Alle 43-45, DK-8200 Aarhus, Denmark (jok@clin.au.dk)

[c]Sciensano Rue Juliette Wytsmanstraat 14, 1050, Brussels Belgium (toon.braeye@sciensano.be)

[d]University Children's Hospital Basel, PO Box, CH 4033 Basel, Switzerland (j.bauwens@brightoncollaboration.org)

[e]University of Basel, Switzerland (j.bauwens@brightoncollaboration.org)

[f]Brighton Collaboration Foundation, Switzerland (j.bauwens@brightoncollaboration.org)

[g]P-95 Epidemiology and Pharmacovigilance, Leuven, Belgium (kaatje.bollaerts@p-95.com; m.c.j.sturkenboom@umcutrecht.nl)

[h]Consorzio Arsenàl.IT, Veneto Region, Italy (gdanieli@consorzioarsenal.it; ltramontan@consorzioarsenal.it)

[i]PEDIANET, Padova, Italy (gdanieli@consorzioarsenal.it; ltramontan@consorzioarsenal.it)

[j]Intitut Universitari d'Investigació en Atenció Primària Jordi Gol (IDIAP Jordi Gol), Barcelona, Spain (tduarte@idiapjgol.org)

[k]GSK, Av. Fleming 20, 1300, Wavre, Belgium (steffen.x.glismann@gsk.com)

[l]BIFAP database, Spanish Agency of Medicines and Medical Devices, Madrid Spain (chuerta@aemps.es; emartinm@aemps.es)

[m]University of Surrey, Guildford, Surrey GU2 7XH, UK (s.lusignan@surrey.ac.uk; c.mcgee@surrey.ac.uk; accorrea1@googlemail.com)

[n]Royal College of General Practitioners Research and Surveillance Centre, 30 Euston Square, London NW1 2FB, UK (s.lusignan@surrey.ac.uk)

[o]Erasmus University Medical Center, PO box 2014, 3000 CA, Rotterdam, the Netherlands (d.weibel@vaccinegrid.org; m.c.j.sturkenboom@umcutrecht.nl)

[p]VACCINE.GRID Foundation, Spitalstrasse 33, Basel, Switzerland (d.weibel@vaccinegrid.org; m.c.j.sturkenboom@umcutrecht.nl)

[q]Julius Global Health, Julius Center, University Medical Center Utrecht, Heidelberglaan 100, The Netherlands

* Corresponding author. Hanne-Dorthe Emborg Tel +45 32683937; fax +45 32683868

# These authors contributed equally to the work and share first authorship

Abbreviations used

ADVANCE: Accelerated Development of VAccine beNefit-risk Collaboration in Europe;

aP: acellular pertussis;

AUH: Aarhus University Hospital Denmark;

BIFAB: Database for Pharmacoepidemiological Research in Primary Care Spain;

B/R: benefit-risk;

CumInc: cumulative incidence;

DTP: diphtheria tetanus pertussis;

DT: diphtheria tetanus;

DTaP: diphtheria tetanus acellular pertussis;

ECDC: European Centre for Disease Prevention and Control;

PEDIANET: Family Paediatrician Database Veneto vaccine registry Italy;

POC: Proof of Concept;

PP_{FU}: Period Prevalence;

RCGP RSC: Royal College of General Practitioners Research and Surveillance Centre UK;

SIDIAP: Information System for Primary Care Research Spain;

SSI: Statens Serum Institut Denmark;

THIN: The Health Improvement Network UK;

wP: whole cell pertussis.

Abstract

Introduction

The Accelerated Development of VAccine beNefit-risk Collaboration in Europe (ADVANCE) is a public-private collaboration aiming to develop and test a system for rapid benefit-risk (B/R) monitoring of vaccines, using existing healthcare databases in Europe. The objective of this paper was to assess the feasibility of using electronic healthcare databases to estimate dose-specific acellular pertussis (aP) and whole cell pertussis (wP) vaccine coverage.

Methods

Seven electronic healthcare databases in four European countries (Denmark (n=2), UK (n=2), Spain (n=2) and Italy (n=1)) participated in this study. Children were included from birth and followed up to age six years. Vaccination exposure was obtained from the databases and classified by type (aP or wP), and dose. Coverage was estimated using period prevalence. For the 2006 birth cohort, two estimation methods for pertussis vaccine coverage, period prevalence and cumulative incidence were compared for each database.

Results

The majority of the 2,575,576 children included had been vaccinated at the country-specific recommended ages. Overall, the estimated coverage was 88-97% in Denmark (birth cohorts from 2003 to 2014), 96-100% in the UK (2003-2014), 95-98% in Spain (2004-2014) and 94% in Italy (2006-2007). The estimated coverage per birth cohort in Denmark and the UK differed by 1-6% compared with national estimates, with our estimates mostly higher. The estimated coverage in Spain differed by 0-2% with no consistent over- or underestimation. In Italy, the estimates were 3% lower compared with the national estimates. Except for Italy, for which the two coverage estimation methods generated the same results, the estimated cumulative incidence coverages were consistently 1% to 10% lower than period prevalence estimates.

Conclusion

This study showed that it was possible to provide reliable estimates of pertussis immunisation coverage from the electronic healthcare databases included, and that the estimates were comparable with the national estimates.

Keywords: Acellular pertussis vaccination; whole cell pertussis vaccination; vaccination coverage estimation; proof-of-concept study; database characteristics; benchmarking

1. Introduction

Whole cell pertussis (wP) vaccines have been available since the 1940s and were effective in reducing the number of pertussis cases and mortality [1, 2]. However, due to common minor adverse reactions and less common severe systemic reactions to wP, acellular pertussis (aP) vaccines were developed and used from the mid-1990s [1]. Many countries replaced wP with aP, and Poland is the only country in Europe where wP vaccine is still included in the childhood vaccination programme [3].

World-wide, countries provide annual reports on national pertussis vaccine coverage estimates to WHO/UNICEF's Vaccine Preventable Diseases Monitoring System [4]. Electronic registration of vaccination is becoming more widespread in Europe, allowing countries to share vaccine coverage data for further analysis. In a survey by the European Centre for Disease Prevention and Control (ECDC) in 2016, 16 out of 27 EU/EEA countries reported that they had a national or sub-national vaccination information system, 5 countries reported they were piloting a system and 6 countries said they had plans to set up a system in the future [5]. Collations of data from these registries and other sources have shown that pertussis vaccine coverage in Europe is generally above 90%, although coverage has dropped in some countries in some years [6, 7].

The Accelerated Development of VAccine beNefit-risk Collaboration in Europe (ADVANCE) is a public-private collaboration aiming to develop and test a system for rapid benefit-risk (B/R) monitoring of vaccines, using existing healthcare databases in Europe. A series of proof of concept (POC) studies were designed to assess the processes and system proposed for generating data on vaccination coverage, benefits and risks required to perform B/R monitoring. As a preparatory step to these studies, a systematic approach was used to characterise and assess the eligibility of these healthcare databases for their use in coverage and B/R studies [8].

The objective of this paper was to determine the feasibility of using electronic healthcare databases to estimate dose-specific vaccination coverage by age and its variation across birth cohorts, using aP and wP vaccination coverage as an example.

2. Material and methods

2.1. Databases

Seven of the 19 electronic European healthcare databases that were suitable and whose owners agreed to participate in the ADVANCE POC studies in 2016 were included [8]. There was one regional and one national hospital discharge database linked to vaccination registries from Denmark (Aarhus University Hospital: AUH and Statens Serum Institut: SSI) , one multiregional and one regional primary healthcare record database from Spain (Database for Pharmacoepidemiological Research in Primary Care: BIFAP and the Information System for Primary Care Research: SIDIAP), two national primary care medical record databases from the UK (The Health Improvement Network: THIN and Royal College of General Practitioners Research and Surveillance Centre: RCGP RSC) and one regional family paediatrician database linked to the Veneto vaccine registry from Italy (PEDIANET). Details about the extraction, management, transformation, sharing, and analyses of the data using the ADVANCE system workflows and methodology (common protocol, common data model and common analytics) can be found in another paper in this supplement [9]. In response to a survey, the database owners provided information on database characteristics such as representativeness, origin of data, population size, data-availability period, switch from wP to aP vaccines, historical pertussis vaccine schedules, availability of ATC-codes and doses, how information about dose was recorded and rounding rules for birth dates for privacy (**Table 1**).

2.2. Study population

The eligible population comprised all children from birth to five years old, registered in any of the participating databases and identifiable through a unique anonymised patient-ID and

with at least one day of follow-up during the overall study period: 1st January 1990 to 31st December 2015. Children were eligible for inclusion in the study population if they entered the database at the age of one month or younger during the study period (defined as start of follow-up). We defined end of follow-up, as the date of whichever of the following events occurred first: receipt of their pre-school-entry pertussis booster, 6th birthday, end of study (varied across databases - see country specific study periods in **Table 1**), transferring out from the catchment of the database, or death. Children with incomplete dates for birth, start or end of follow-up were excluded. Rounding of birth dates was allowed.

2.3. Exposure

The exposure of interest was vaccination with any pertussis-containing vaccine in the study population during follow-up. When the vaccine type (aP or wP) could not be determined reliably, it was coded uP (unknown). If the dose number was not recorded in the database, it was derived based on the chronological sequence of administered doses and the age of the child. A child was assumed to be vaccinated on the day the vaccine dose was recorded. All records with missing patient-ID, dates or vaccine type were excluded from analysis.

2.4. Statistical methods

Vaccination coverage (per dose) was estimated as the percentage of the children in the study population who had received the specific vaccine dose by a certain age. Pertussis vaccine coverage was estimated by dose and by age (in weeks) in each birth cohort, using period prevalence, taking into consideration any children lost-to-follow-up (PP_{FU}). The PP_{FU}-estimate for children at a certain age (in weeks) was the number of children vaccinated with the first dose (D1), second dose (D2), and third dose (D3), respectively, divided by the total number of children in follow-up at that age (in weeks). For example, in the 2012 birth cohort, at five weeks of age, the number of children vaccinated with the first dose (D1), second dose (D2) and third dose (D3), respectively, was divided by the total number of children in the

2012 birth cohort and being follow-up at five weeks of age. At six weeks of age the total number of children still in the population and vaccinated with D1, D2 and D3 respectively, was divided by the total number of children still being followed-up in the 2012 birth cohort aged six weeks old etc. Thus the numerator and the denominator could decrease over time as children left the database. To compare between different methods to estimate pertussis coverage, for the 2006 birth cohort in each database the PP_{FU} coverage estimates were compared with the cumulative incidence (CumInc) of pertussis vaccination. The cumulative incidence was estimated for each birth cohort as the number of all vaccinated children at a certain age in weeks divided by the number of eligible children, i.e. those at the start of the follow-up period.

The age at vaccination per birth cohort, dose and vaccine type were estimated and presented in Cleveland dot plots (8) as 10%, 50% (median) and 90% quantiles [10].

To assess the validity of the coverage estimates obtained in this study, the estimates were compared with the national coverage estimates that have been published by the public health institutes in each of the four countries that the databases originated from.

We did not report wP for the THIN-database before 2000, because birthdates were rounded to 1st of July at extraction, leading to inaccurate age of vaccination.

3. Results

3.1. Study population

The seven databases that participated showed large variation in their overall population size (0.0097- 13.6 millions) and availability of data during the overall study period: this varied from 2 to 26 years (**Tables 1 and 2**). The total study population across all databases comprised 2.575 million children aged <6 years (**Table 2**).

3.2. Pertussis vaccination data

In the BIFAP, AUH and SSI databases the vaccine dose was not recorded reliably for all records and had to be derived. In Spain, aP vaccines were first introduced in 1997 and the switch from wP to aP vaccines occurred gradually until 2004. From 2005 onwards only aP vaccines were available, and hence we assume that wP vaccines were not used from this point. In the other three countries the switch occurred within one calendar year (**Table 1**). Only the UK databases could provide a large proportion of data on wP vaccination coverage. Across all birth cohorts within a database and across databases within a country, the majority of children were vaccinated at the recommended age of vaccination (**Figure 1**). However, the 90% quantile for age at vaccination was higher than the recommended age of vaccination in some cases, indicating that a certain percentage of individuals in the birth cohorts were vaccinated late.

3.3. Pertussis vaccination coverage for dose 3

The coverage estimates (PP_{FU}) for dose 3 (D3) are summarised by database, birth cohort, type of vaccine (aP, wP) and age in Figure 2. The coverage started to increase in all databases at the age when the D3 was recommended in the country. For example, the aP D3 for children is recommended at 12 months old in Denmark and the D3 coverage estimates in AUH and SSI were close to zero until just before the children were 12 months old, then increased rapidly after they were 12 months old to above 80% at 15 months of age in all birth cohorts. In general, the observed age of the D3-vaccination was similar across birth cohorts from the same database. Following the steep increase in coverage estimates, little change was observed up to the end of the follow-up period. The differences in coverage reached at the end of follow-up at 72 months old were between 0% and 8% between birth cohorts within the same database, except for the RCGP RSC and THIN databases. In these databases the differences were 21% and 27%, respectively, with the largest difference observed at the end of follow-up in the older birth cohorts (**Figure 2**). Spikes at the end of follow-up were due to small sample

sizes, especially in the younger birth cohorts. The switch from wP to aP occurred in 2004 in the UK, therefore, the 2004 birth cohort from RCGP RSC is the only one with a substantially lower coverage for both aP and wP in 2004, compared with the other cohorts that included the switch from wP to aP. This was not observed in the THIN database since the vaccine type could not be determined during the switch in 2004.

3.4. Comparison of vaccination coverage estimation methods

We estimated coverage for the 2006 birth cohort in each database as an example, to compare the results with the estimation methods (PP_{FU} and CumInc). The CumInc estimates were consistently 1% to 10% lower at the end of follow-up in all databases, except for PEDIANET, where the two methods generated the same results (**Figure 3**).

3.5. Comparison with national coverage rates

Overall, the estimated coverage was 88-97% (birth cohorts from 2003 to 2014) in Denmark, 96-100% (2003-2014) in the UK, 95-98% (2004-2014) in Spain and 94% (2006-2007) in Italy (**Table 3**). The estimated coverage in the Danish SSI and AUH databases differed by 1-4% and 2-6%, respectively, compared with the national estimates. In the UK THIN and RCGP RSC databases the estimated coverage differed by 2-5% and 1-6%, respectively, compared with the national estimates. In both Denmark and UK, our estimates were almost always higher. The estimated coverage in the Spanish BIFAP and SIDIAP databases differed both by 0-2% compared to the national estimates with no clear direction of the deviation. The coverage estimates from the Italian PEDIANET database were 3% lower than the national estimates.

4. Discussion

This study showed that it was possible to provide reliable estimates on pertussis vaccination coverage using data in the seven participating healthcare databases in four countries. The results showed that the ages when the pertussis doses were administered were comparable

243 across up to 26 birth cohorts within the same database. We observed a steep increase in the
244 period coverage prevalence estimates for dose 3 at almost the same age in all birth cohorts
245 within each database. In addition, the observed age at vaccination was consistent with the
246 recommended age for vaccination, as defined in the national guidelines, and the overall
247 coverages obtained at the end of follow-up (72 months) were similar to the national coverage
248 estimates, which demonstrates the feasibility of obtaining accurate estimates for vaccination
249 coverage using data from healthcare databases.
250 In addition to the information on the median age at vaccination, we also observed variability
251 in age of vaccination, which confirms a delay in vaccination for a part of the population that
252 has been reported in several countries [11-14]. The more recent birth cohorts had shorter
253 follow-up time than the earlier birth cohorts, because of the retrospective nature of the study,
254 which led to unequal truncation of follow-up time. For example: at the end of the study period
255 the oldest child in 2014 birth cohort would be 730 days old (December 2015), and
256 consequently, the median age for the third dose in this birth cohort would be expected to be
257 lower than in the earlier birth cohorts because data for any children with late vaccination
258 would be truncated in the later birth cohorts.
259 A review of the databases that were considered for the inclusion in this proof of concept study
260 revealed differences in how data about the vaccine and the vaccine dose administered were
261 recorded, and the level of information provided, in terms of the codes used and what free text
262 was used [8]. Despite these differences, it was generally possible to identify what pertussis
263 vaccine was administered and when, derive the relevant doses and to obtain coverage
264 estimates (overall, by birth cohort and age) close to the national estimates provided by public
265 healthcare authorities. In the PRISM programme in the USA, DTP, DT and DTaP vaccination
266 coverage was estimated to be 76% for the third dose, using data from three claims databases,

compared with the estimated 91% vaccination coverage obtained through the National Immunization Survey, which is a bigger difference than we observed [15].

The high level of concordance between our coverage estimates and the national estimates was not expected because the inclusion criteria for this study (see Methods section) differed from those used for the national coverage estimation. We followed the birth cohorts up to 72 months (6^{th} birthday) whereas most countries usually select an age closer to the recommended age of vaccination to be able to estimate vaccination coverage in relation to the national recommendations for vaccination (see references for national estimates in Table 3 for further details). In addition, ecological data (i.e. total number of administered vaccines divided by the total number of children registered in the census) are used for national estimations, which could potentially result in larger deviations from the patient-level estimations used in our study.

The databases included in our study varied in size, geographical coverage and healthcare setting. The national registries have a more stable population with fewer individuals leaving or entering the database at different time points, compared with, for example GP databases, where the turnover of patients is expected to be higher, resulting in incomplete follow-up for a proportion of the population [8]. This could result in biased prevalence rates, if considered as complete follow-up, and thus compromise the coverage estimates [16]. In our study, we estimated vaccination coverage as a period prevalence, taking into consideration loss-to-follow-up. This approach can result in decreasing coverage estimates when vaccinated children leave the database faster than unvaccinated children and increasing coverage estimates when unvaccinated children leave the database faster than vaccinated children. This could occur if vaccination schedules differ across regions in the same country, for example in a multi-regional database setting. However, this is less likely to be important across regions with homogeneous vaccination schedules.

Our results showed that vaccination coverage at the end of follow-up, estimated using PP_{FU} or CumInc were comparable (0-10% difference) for the 2006 birth cohorts (Figure 3). This suggests that incomplete follow-up did not have a large impact on the coverage estimation in this study, possibly because incomplete follow-up was limited in this comparison. The inclusion of children who entered the database only within one month of birth limited the left-censoring, which may reduce bias in coverage estimates since the youngest children are also those who receive pertussis vaccine [17]. This inclusion criterion could increase bias in coverage estimation since children who have their first GP visit before one month of age might be those who are more likely to be vaccinated and, also, perhaps more likely to be vaccinated in compliance with the national guidelines [16]. Some of the participating databases round birth dates to the 1^{st} or the 15^{th} of the birth month which means that some children will only have two weeks to be enrolled in the study population and others will have six weeks, which could potentially exclude a large number of children. This could also result in misclassification of age. If vaccination adherence and registration in the database after one month of age were dependent variables, the two coverage estimation methods would provide biased, overestimated or underestimated coverage estimates, depending on the direction of the dependency.

It was possible to estimate aP-containing vaccine coverage rates in all seven participating databases using the two estimation methods. However, it was only possible to estimate wP-containing vaccine coverage before and during the switch from wP to aP in two databases due to low numbers of registered wP vaccinations in the remaining five databases where the switch occurred prior to the study period.

In conclusion, we identified heterogeneity in the characteristics of the databases, which lead to challenges in defining inclusion criteria and taking incomplete follow-up into account, and thus for estimating pertussis vaccination coverage. We handled these elements in a

317 homogenous manner across countries and were, therefore able to provide reliable pertussis
318 coverage estimates.
319

Role of funding source

The research leading to these results received support from the Innovative Medicines Initiative Joint Undertaking under the ADVANCE grant agreement n° 115557 through financial contributions from the European Union's Seventh Framework Programme (FP7/2007-2013) and in kind contribution from participating EFPIA companies.

Conflicts of interest

Hanne-Dorthe Emborg, Toon Braeye, Jorgen Bauwens, Kaatje Bollaerts, Giorgia Danieli, Talita Duarte-Salles, Consuelo Huerta Elisa Martin, Chris McGee, Ana Correa and Lara Tramontan declared no conflicts of interest. Johnny Kahlert declared that although he does not personally receive fees, honoraria, or grants he is employed at Department of Clinical Epidemiology, Aarhus University Hospital that receives research grants from various pharmaceutical companies administered by Aarhus University. Steffen Glismann is employed by the GSK group of companies and holds company shares. Simon de Lusignan declared he has received funding through his University to conduct enhanced surveillance of influenza vaccine (GSK), and is a member of Seqirus and Sanofi Pasteur advisory boards for which he received personal payment within the limits defined by his university. Daniel Weibel declared that he has received consultancy fees from GSK unrelated to the submitted work. Miriam Sturkenboom declared that she has received grants from Novartis, CDC and Bill & Melinda Gates Foundation, for work unrelated to the submitted work.

Acknowledgements

The authors would like to acknowledge Klára Berencsi who contributed to the project at AUH (Denmark); Rachel Byford, Mariya Hriskova, Filipa Ferreira, Ivelina Yonova, Sameera Pathirannehelage and Harshana Liyanage who contributed to the project at Surrey University / RCGP RSC (UK); Ana Llorente who contributed to the project at BIFAP (Spain); Yesika Díaz who contributed to the project at SIDIAP (Spain). They also acknowledge that, during preparation of this paper, Palle Valentiner-Branth (SSI, Denmark) and Signe Sørup (AUH, Denmark) provided general insights on the use of vaccines and the relevant literature. Lastly, they acknowledge Lina Titievsky (Pfizer, USA) who contributed as work package co-lead and through her assistance in writing the study report and Margaret Haugh (MediCom Consult, Villeurbanne, France) who provided editorial services for the study report and for this paper.

Disclaimer

The results described in this publication are from the proof of concept studies conducted as part of the IMI ADVANCE project with the aim of testing the methodological aspects of the design, conduct and reporting of studies for vaccine benefit-risk monitoring activities. The results presented relate solely to the methodological testing and are not intended to inform regulatory or clinical decisions on the benefits and risks of the exposures under investigation. This warning should accompany any use of the results from these studies and they should be used accordingly. The views expressed in this article are the personal views of the authors and should not be understood or quoted as being made on behalf of or reflecting the position of the agencies or organisations with which the authors are affiliated.

References

[1] Edwards KM, Decker MD. Chapter 44 - Pertussis Vaccines. In: Plotkin SA, Orenstein WA, Offit PA, Edwards KM, editors. Plotkin's Vaccines (Seventh Edition): Elsevier; 2018. p. 711-61.e16.

[2] Chow MYK, Khandaker G, McIntyre P. Global childhood deaths from pertussis: a historical review. Clin Infect Dis. 2016;63:S134-S41.

[3] European Centre for Disease Prevention and Control. Vaccine Scheduler [Last accessed 26 October 2018]. Available from: https://vaccine-schedule.ecdc.europa.eu/Scheduler/ByDisease?SelectedDiseaseId=3&SelectedCountryIdByDisease=-1.

[4] World Health Organisation. WHO vaccine-preventable diseases: monitoring system 2017 [Last accessed 18 June 2017]. Available from: http://apps.who.int/immunization_monitoring/globalsummary.

[5] Derrough T, Olsson K, Gianfredi V, Simondon F, Heijbel H, Danielsson N, et al. Immunisation information systems - useful tools for monitoring vaccination programmes in EU/EEA countries, 2016. Euro Surveill. 2017;22.

[6] World Healh Organisation Europe. European Health Information Gateway; % of infants vaccinated against pertussis 2018 [Last accessed 26 October 2018]. Available from: https://gateway.euro.who.int/en/indicators/hfa_608-7180-of-infants-vaccinated-against-pertussis/.

[7] Heininger U, Andre P, Chlibek R, Kristufkova Z, Kutsar K, Mangarov A, et al. Comparative epidemiologic characteristics of pertussis in 10 Central and Eastern European Countries, 2000-2013. PloS One. 2016;11:e0155949.

[8] Sturkenboom M, Weibel D, van der Aa L, Braeye T, Gheorge M, Becker B, et al. ADVANCE database characterization and fit for purpose assessment for multi-country studies on the coverage, benefits and risks of vaccinations. Vaccine. 2018;Paper 3 in Supplement.

[9] Sturkenboom M, van der Aa L, Bollaerts K, Emborg HD, Ferreira G, Gino R, et al. The ADVANCE distributed network system for evidence generation on vaccines coverage, benefits and risks based on electronic health care data. Vaccine. 2018;Paper 2 in supplement.

[10] Cleveland WS, McGill R. Graphical perception: theory, experimentation and application to the development of graphical methods. J Am Stat Assoc. 1984;79:531-54.

[11] Lernout T, Theeten H, Hens N, Braeckman T, Roelants M, Hoppenbrouwers K, et al. Timeliness of infant vaccination and factors related with delay in Flanders, Belgium. Vaccine. 2014;32:284-9.

[12] Luman ET, Barker LE, Shaw KM, McCauley MM, Buehler JW, Pickering LK. Timeliness of childhood vaccinations in the United States: days undervaccinated and number of vaccines delayed. JAMA. 2005;293:1204-11.

[13] Pedersen KB, Holck ME, Jensen AKG, Suppli CH, Benn CS, Krause TG, et al. How are children who are delayed in the Childhood Vaccination Programme vaccinated: A nationwide register-based cohort study of Danish children aged 15-24 months and semi-structured interviews with vaccination providers. Scand J Public Health. 2018:1403494818786146.

[14] Riise ØR, Laake I, Bergsaker MAR, Nøkleby H, Haugen IL, Storsæter J. Monitoring of timely and delayed vaccinations: a nation-wide registry-based study of Norwegian children aged < 2 years. BMC Pediatr. 2015;15.

[15] Baker MA, Nguyen M, Cole DV, Lee GM, Lieu TA. Post-licensure rapid immunization safety monitoring program (PRISM) data characterization. Vaccine. 2013;31:K98-K112.

412 [16] Lanes S, Quinlan SC, Mast TC, Greenland S, Holick CN. Assessing bias in
413 administrative database studies of RotaTeq vaccine completion due to exclusion of subjects
414 with incomplete follow-up. Emerg Themes Epidemiol. 2015;12:5.
415 [17] Cain KC, Harlow SD, Little RJ, Nan B, Yosef M, Taffe JR, et al. Bias due to left
416 truncation and left censoring in longitudinal studies of developmental and disease processes.
417 Am J Epidemiol. 2011;173:1078-84.
418
419

Figure captions

Figure 1: Median age in weeks, at first (D1), second (D2) and third (D3) dose of whole cell (wP) or acellular (aP) pertussis vaccine and the 10% and 90% quantiles for age at vaccination for the entire birth cohort. A: aP vaccination from AUH and SSI in Denmark; B and C: wP and aP vaccination from RCGP RSC and THIN in the UK; D: aP vaccination from BIFAP and SIDIAP in Spain; E: aP vaccination from PEDIANET in Italy. The recommended ages of vaccination are indicated by the vertical lines and shaded areas. We show every second or third birth cohort only due to limited space. The SSI database provided data until September 2014, so that none of the children born in 2014 had reached the age of 12 month when the aP dose 3 is recommended in Denmark; thus this dose is missing for the 2014 birth cohort.

Figure 2. Pertussis coverage for dose 3 estimated as period prevalence (PP_{FU}) by database, birth cohort, type of vaccine (wP or aP) and age. A: AUH, Denmark, aP; B: SSI, Denmark, aP; C: RCGP RSC, UK, wP; D: RCGP RSC, UK E: THIN, UK, wP; F: THIN, UK, aP;, aP; G: BIFAP, Spain, aP; H: SIDIAP, Spain, aP; I:PEDIANET, Italy, aP.

Figure 3. Comparison of the monthly dose 3 vaccination coverage for 2006 birth cohorts using period prevalence (PP_{FU}) and cumulative incidence (CumInc) methods from birth up to 6th birthday. A: AUH and B: SSI from Denmark; C: RCGP RSC and D: THIN from the UK; E: BIFAP and F: SIDIAP from Spain; G: PEDIANET from Italy.

442 **Table 1:** Characteristics of the seven participating databases

Country	**Denmark**		**Spain**		**United Kingdom**		**Italy**
Database	**AUH**	**SSI**	**BIFAP**	**SIDIAP**	**THIN**	**RCGP RSC**	**PEDIANET**
Representativeness	Regional Whole	National Whole	Multi-regional Subset	Regional Subset	National Subset	National Subset	Regional Subset
Origin of data	Record linkage between different registries		GP and primary care paediatricians	GP and paediatricians	GP	GP	Family paediatricians
Birth cohorts available for coverage estimation	2002-2015	1997-2014*	2002-2014**	2005-2015	1990-2015	1990-2015	2006-2007
Switch from wP to aP	1997	1997	1997-2004	1997-2004	2004	2004	1995
Percentage wP of all vaccinations recorded	0	2.1	1.7	0	64.3	47.7	0
ATC code available (%)	0	100	100***	0	0	0	0****
Dose recorded (% missing)	Yes (35.4)	Yes (2.7)	Yes (23)	Yes (0)	Yes (<0.01)	-	Yes (0)
Dose derived (% missing)	Yes (0)	Yes (2.8)	Yes (2)	-	-	-	None
Rounding of birthdates	None	None	None	Rounded to 1st of month	Rounded to 1st of month	Rounded to 1st of month	Rounded to 15th of month

443

** The study period was 1997 to September 2014*

*** The 2002 and 2003 birth cohorts were excluded from the coverage analyses since the vaccine type was unknown for the majority of the vaccines administered*

****ATC code was derived based on the described antigen combinations, marketed vaccines at every calendar year of vaccination and age at vaccination according to the rules in the national scheme.*

***** Could be derived from coded association information*

444 *AUH: Aarhus University Hospital Denmark; SSI, Statens Serum Institut Denmark; BIFAP: Database for Pharmacoepidemiological Research in Primary Care Spain;*
445 *SIDIAP, Information System for Primary Care Research Spain; THIN, The Health Improvement Network UK; RCGP RSC: Royal College of General Practitioners Research*
446 *and Surveillance Centre UK; PEDIANET: Family Paediatrician Database Veneto vaccine registry Italy.*

447

448

Table 2: Attrition table of seven databases (AUH and SSI from Denmark, RCGP RSC and THIN from UK, BIFAP and SIDIAP from Spain and PEDIANET from Italy)

	Denmark		UK		Spain		Italy	Total
	AUH	SSI	THIN	RCGP RSC	BIFAP	SIDIAP	PEDIANET	
Number of persons (all ages)	1,725,165	7,512,032	13,646,770	3,017,610	7,541,864	7,096,695	9,708	40,549,844
Number of persons born during the birth years of interest (1990-2015)	499,318	1,822,953	1,616,311	860,411	1,467,618	1,774,085	9,708	8,050,404
Number of persons without follow-up (<1 day of follow-up)	0	31,434	59,933	0	23	92,932	0	184,322
Number of persons with start of follow-up after the age of 1 month	310,765	571,787	1,206,165	814,171	1,174,425	1,213,193	0	5,290,506
Number of persons eligible for analysis	188,553	1,219,732	350,213	46,240	293,170	467,960	9,708	2,575,576

AUH: Aarhus University Hospital, Denmark; SSI, Statens Serum Institut, Denmark; RCGP RSC: Royal College of General Practitioners Research and Surveillance Centre, UK; THIN, The Health Improvement Network, UK; BIFAP: Database for Pharmacoepidemiological Research in Primary Care Spain; SIDIAP, Information System for Primary Care Research Spain; PEDIANET: Family Paediatrician Database Veneto vaccine registry Italy.

Table 3. Comparison of estimated database -specific dose 3 pertussis coverage (%) based on period prevalence (PP_{FU}) at the end of follow-up (after 72 months) with the national pertussis coverage estimates reported by a national authority

Birth cohort	Denmark			UK			Spain			Italy	
	National*	AUH	SSI	National*	RCGP RSC	THIN	National*	BIFAP	SIDIAP	National*	PEDIANET
2003	88	92	90			98	98			97	
2004	87	92	89			98	97	96		97	
2005	86	90	88		97	98	96	96	98	96	
2006	87	92	89	93	98	98	98	96	97	97	94
2007	88	93	91	93	97	98	96	97	96	97	94
2008	89	94	91	93	99	98	97	97	96	97	
2009	89	94	92	95	98	98	96	96	96	96	
2010	90	94	92	94	98	98	97	97	96	96	
2011	89	95	92	95	96	99	97	98	96	96	
2012	91	94	92	95	98	99	96	95	96	96	
2013	91	97	90	94	100	98	96	97	95	96	
2014	91	89		95	98	97	97	97	95	95	

Denmark:

https://www.ssi.dk/Smitteberedskab/Sygdomsovervaagning/VaccinationSurveillance.aspx?xaxis=Cohort&vaccination=3&sex=3&landsdel=100&show=&datatype=Vaccination&extendedfilters=False#HeaderTe

xt

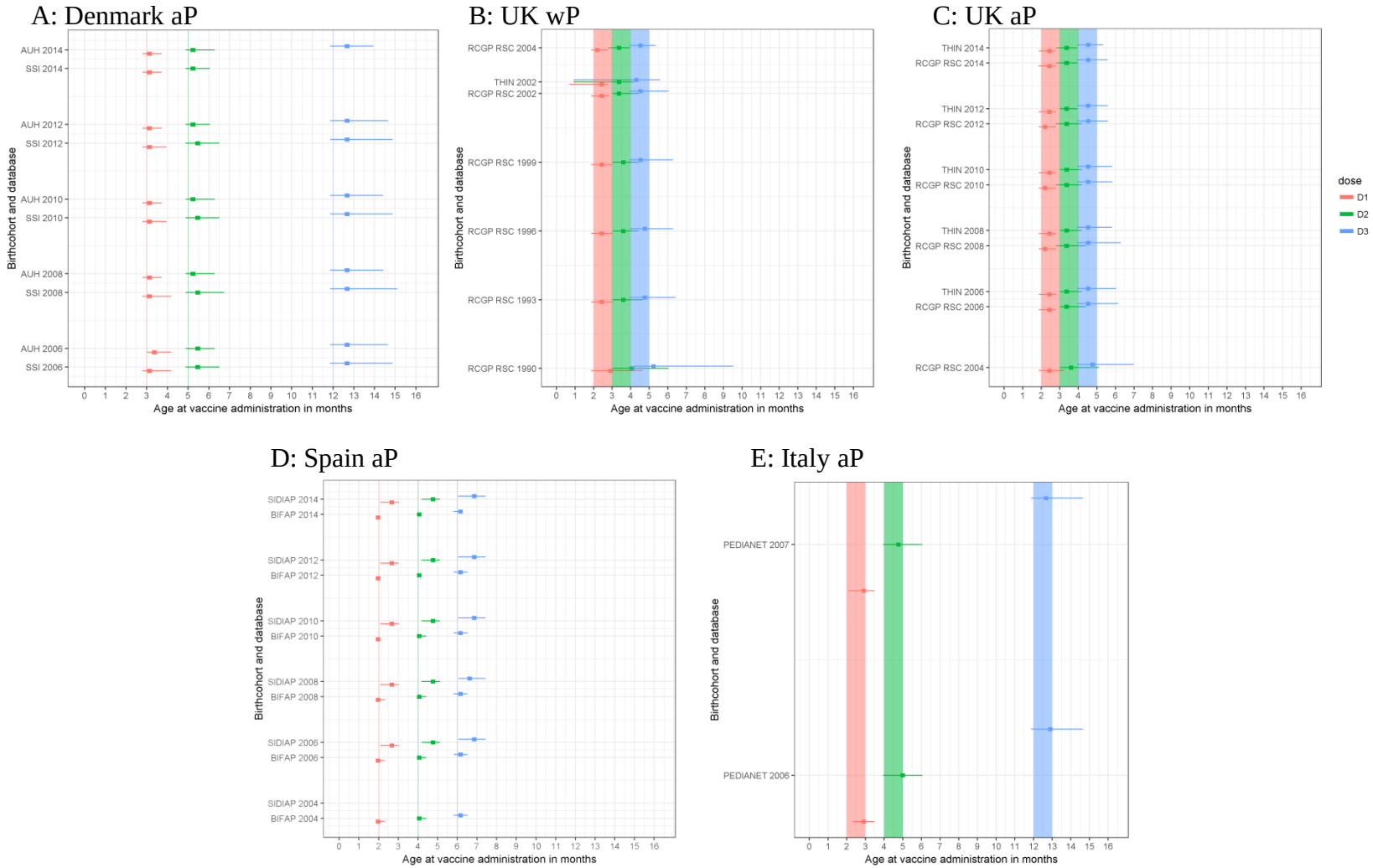
UK: https://www.gov.uk/government/publications/cover-of-vaccination-evaluated-rapidly-cover-programme-annual-data

Spain: http://www.msssi.gob.es/profesionales/saludPublica/prevPromocion/vacunaciones/docs/TosFerina.pdf

Italy: http://www.epicentro.iss.it/temi/vaccinazioni/dati_Ita.asp#pertosse

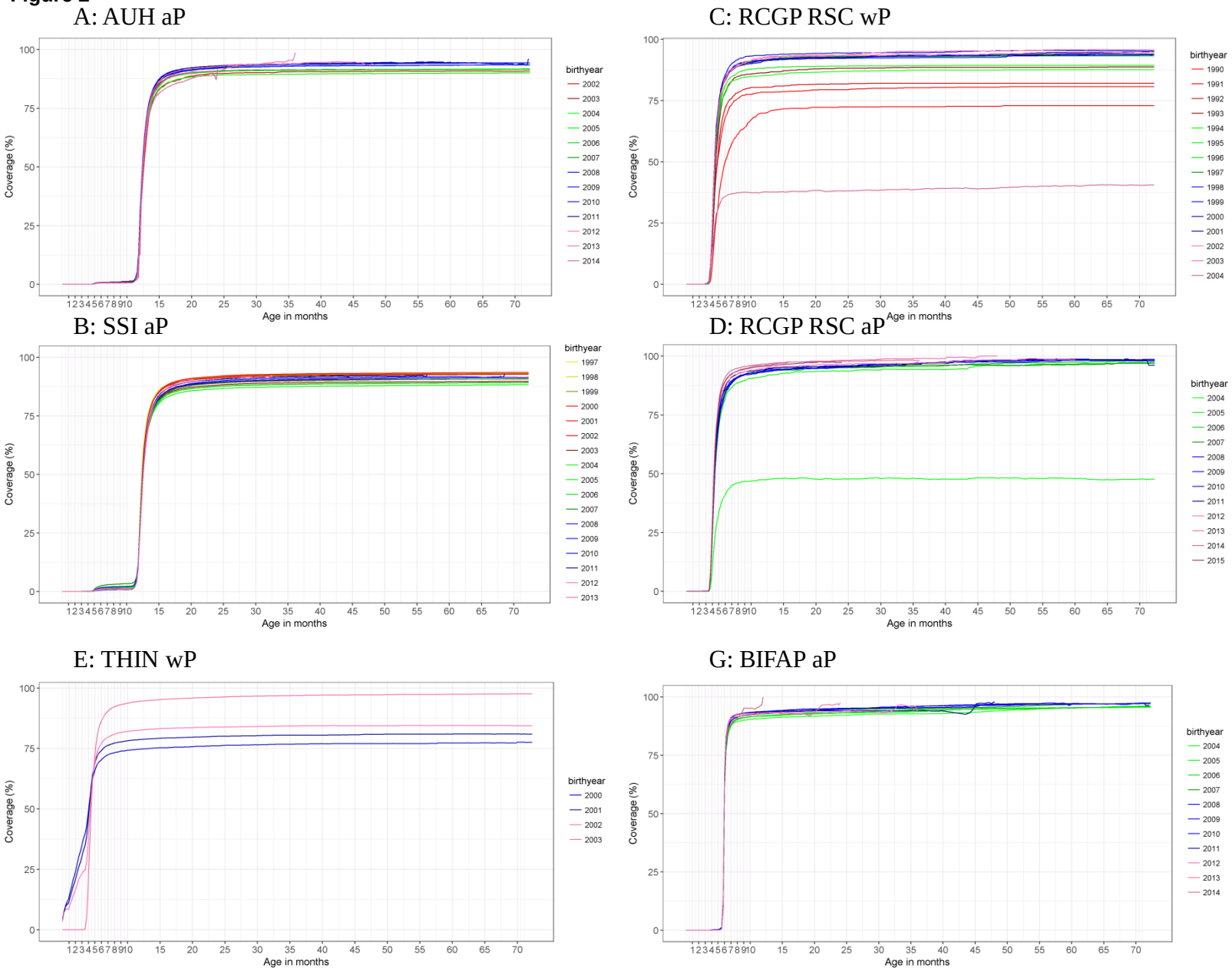
AUH: Aarhus University Hospital Denmark; SSI, Statens Serum Institut Denmark; RCGP RSC: Royal College of General Practitioners Research and Surveillance Centre UK; THIN, The Health Improvement Network UK; BIFAP: Database for Pharmacoepidemiological Research in Primary Care Spain; SIDIAP, Information System for Primary Care Research Spain; PEDIANET: Family Paediatrician Database Veneto vaccine registry Italy.

Figure 1

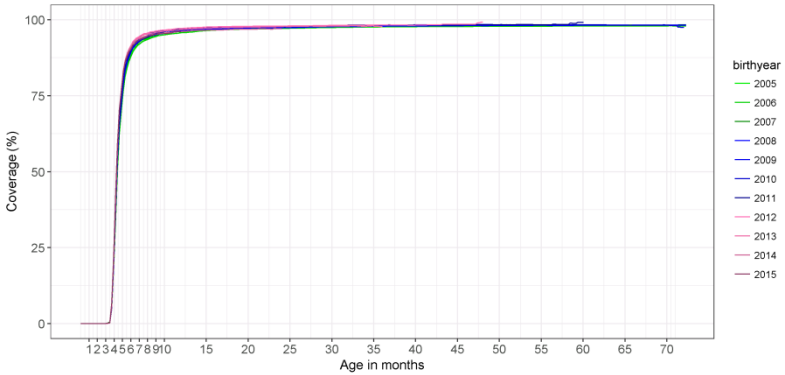


AUH: Aarhus University Hospital Denmark; SSI, Statens Serum Institut Denmark; THIN, The Health Improvement Network UK; RCGP RSC: Royal College of General Practitioners Research and Surveillance Centre UK; BIFAP: Database for Pharmacoepidemiological Research in Primary Care Spain; SIDIAP, Information System for Primary Care Research Spain; PEDIANET: Family Paediatrician Database Veneto vaccine registry Italy.

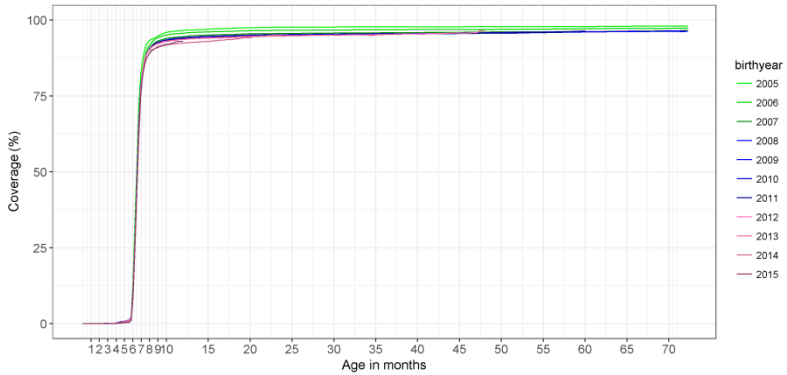
Figure 2


A: AUH aP
C: RCGP RSC wP
B: SSI aP
D: RCGP RSC aP
E: THIN wP
G: BIFAP aP
Coverage (%)
Age in months
birthyear

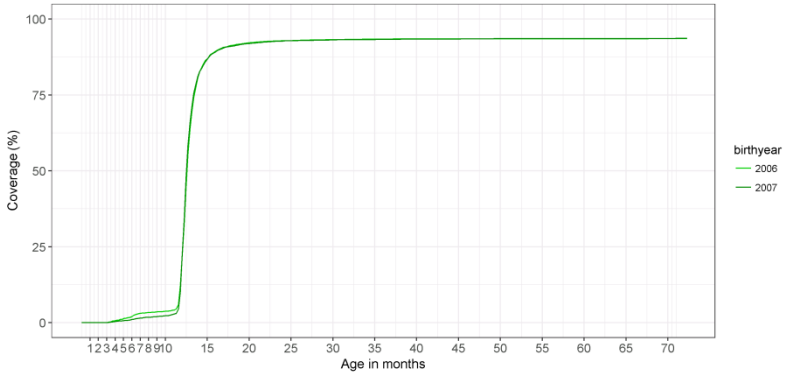
F: THIN aP



H: SIDIAP aP

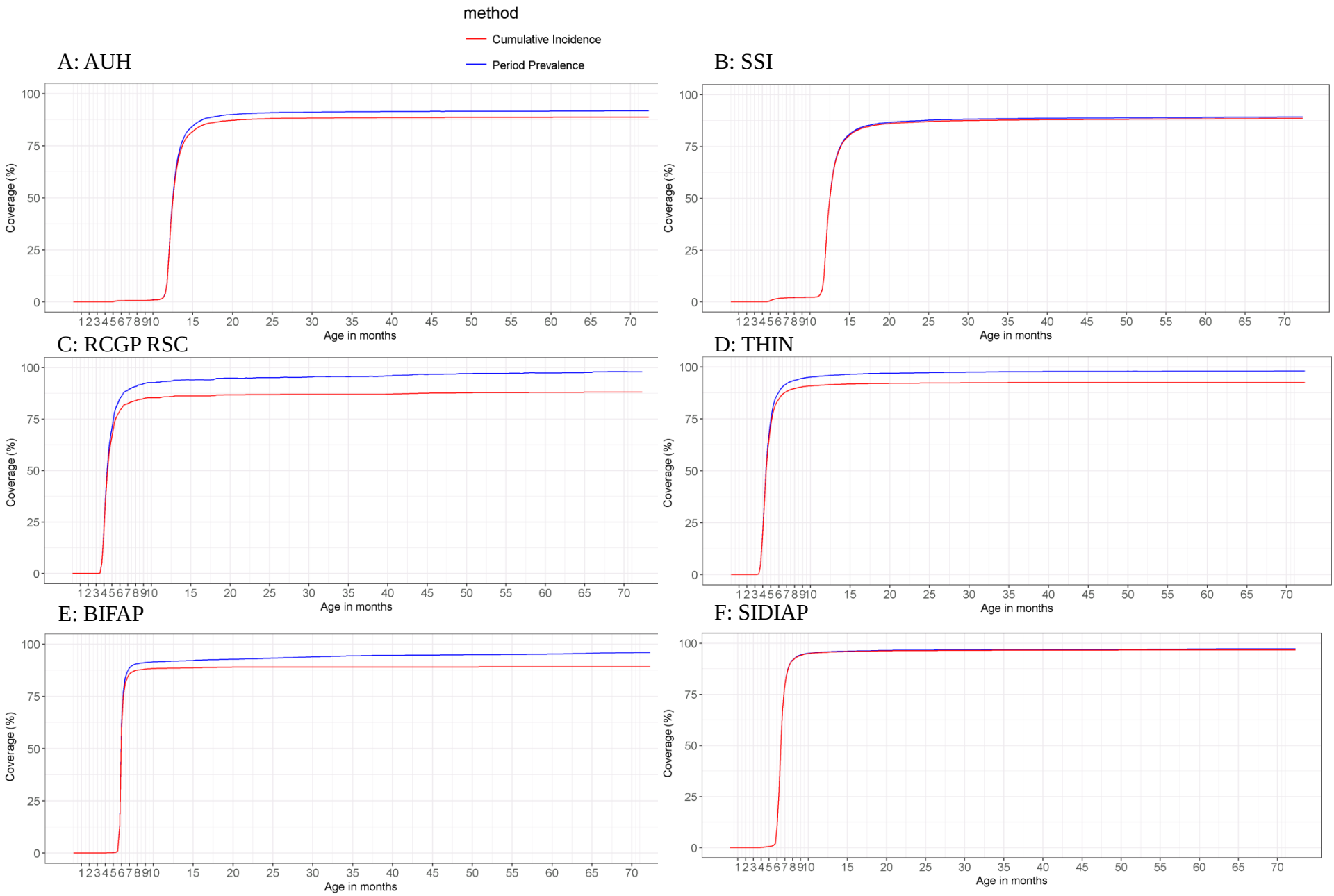


I: PEDIANET aP

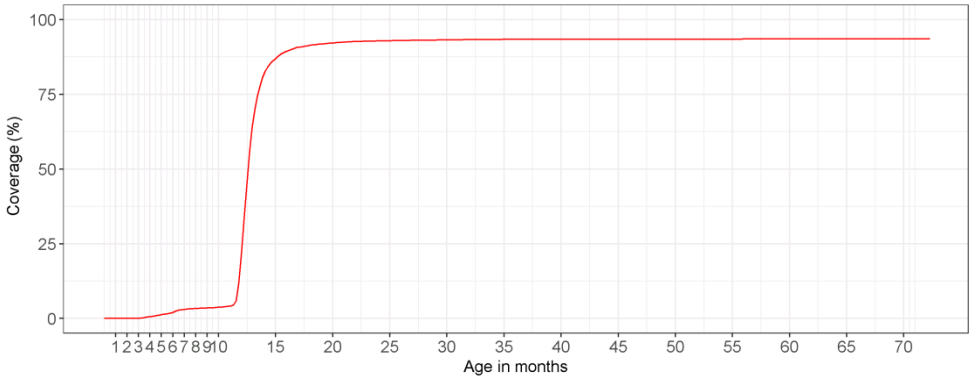


AUH: Aarhus University Hospital Denmark; SSI, Statens Serum Institut Denmark; THIN, The Health Improvement Network UK; RCGP RSC: Royal College of General Practitioners Research and Surveillance Centre UK; BIFAP: Database for Pharmacoepidemiological Research in Primary Care Spain; SIDIAP, Information System for Primary Care Research Spain; PEDIANET: Family Paediatrician Database Veneto vaccine registry Italy

Figure 3


method
Cumulative Incidence
Period Prevalence
A: AUH
B: SSI
C: RCGP RSC
D: THIN
E: BIFAP
F: SIDIAP
Coverage (%)
Age in months

G: PEDIANET



AUH: Aarhus University Hospital Denmark; SSI, Statens Serum Institut Denmark;; THIN, The Health Improvement Network UK; RCGP RSC: Royal College of General Practitioners Research and Surveillance Centre UK BIFAP: Database for Pharmacoepidemiological Research in Primary Care Spain; SIDIAP, Information System for Primary Care Research Spain; PEDIANET: Family Paediatrician Database Veneto vaccine registry Italy