



Study Report

P3-C2-004

DARWIN EU® - Monitoring prescription of medicines for public health emergencies at risk of shortages

Authors: M. Pineda-Moncusí, Y. Guo

25/09/2025

Version 3.0

CONTENTS

1. TITLE	5
2. DESCRIPTION OF STUDY TEAM.....	5
3. DATA SOURCES	6
4. ABSTRACT	7
5. LIST OF ABBREVIATIONS	9
6. AMENDMENTS AND UPDATES	10
7. MILESTONES	10
8. RATIONALE AND BACKGROUND.....	11
9. RESEARCH QUESTION AND OBJECTIVES.....	11
10. RESEARCH METHODS.....	13
10.1. Study type and study design.....	13
10.2. Study setting and data sources	13
10.2.1. Data sources selected.....	13
10.2.2. Rationale for the data sources selected	13
10.2.3. Data sources description:	14
10.3. Study period	17
10.4. Follow-up.....	17
10.5. Study population with in and exclusion criteria	18
10.6. Variables	18
10.6.1. Exposure	18
10.6.2. Outcomes	19
10.6.3. Other covariates, including confounders, effect modifiers, and other	19
10.7. Study size	19
10.8. Patient privacy protection	20
10.9. Statistical methods	20
10.9.1. Federated network analyses	20
10.9.2. Statistical model specification and assumptions of the analytical approach considered	20
10.9.3. Methods to derive parameters of interest.....	20
10.9.4. Main Statistical Methods.....	21
Population-level drug utilisation study.....	21
Time series modelling.....	22
10.9.5. Evidence synthesis.....	22
11. DATA MANAGEMENT	23
12. QUALITY CONTROL	23
13. RESULTS	24
13.1. Participants.....	24
13.2. Main results.....	24
13.2.1. Model diagnostics.....	24
13.2.2. ARIMAX model forecasting 12 months of antibiotic use	27
CDW Bordeaux.....	27
H12O	28
IMASIS.....	29

InGef RDB.....	31
IPCI	32
IQVIA DA Germany	33
IQVIA LPD Belgium.....	34
14. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS	36
15. DISCUSSION	36
15.1. Key results	36
Amoxicillin	36
Amoxicillin/clavulanate	36
Azithromycin.....	36
Benzylpenicillin.....	36
Cefotaxime.....	37
Ceftriaxone	37
Cefuroxime	37
Clarithromycin	37
Meropenem.....	37
Phenoxymethylpenicillin	38
Piperacillin/tazobactam.....	38
15.2. Limitations of the research methods	38
15.2.1. Model diagnostics and limitations:	38
15.2.2. General limitations:	39
15.2.3. Database-specific limitations:	39
15.3. Interpretation	39
15.4. Generalisability	40
16. CONCLUSION.....	40
17. REFERENCES	42
18. ANNEXES.....	43

Study title	DARWIN EU® - Monitoring prescription of medicines for public health emergencies at risk of shortages																																						
Study report version identifier	3.0																																						
Date of last version of report	25/09/2025																																						
EU PAS number	EUPAS1000000473																																						
Active substance	<table><thead><tr><th>Antibiotic</th><th>Class</th><th>ATC code</th></tr></thead><tbody><tr><td>Amoxicillin</td><td>Penicillins with extended spectrum</td><td>J01CA04</td></tr><tr><td>Amoxicillin/clavulanate</td><td>Combinations of penicillins</td><td>J01CR02</td></tr><tr><td>Azithromycin</td><td>Macrolides</td><td>J01FA10</td></tr><tr><td>Benzylpenicillin (Penicillin G)</td><td>Beta-lactamase sensitive penicillins</td><td>J01CE01</td></tr><tr><td>Cefotaxime</td><td>Third-generation cephalosporins</td><td>J01DD01</td></tr><tr><td>Ceftriaxone</td><td>Third-generation cephalosporins</td><td>J01DD04</td></tr><tr><td>Cefuroxime</td><td>Second-generation cephalosporins</td><td>J01DC02</td></tr><tr><td>Clarithromycin</td><td>Macrolides</td><td>J01FA09</td></tr><tr><td>Meropenem</td><td>Carbapenems</td><td>J01DH02</td></tr><tr><td>Phenoxymethylpenicillin (Penicillin V)</td><td>Beta-lactamase sensitive penicillins</td><td>J01CE02</td></tr><tr><td>Piperacillin/tazobactam</td><td>Combinations of penicillins</td><td>J01CR05</td></tr></tbody></table> ATC = Anatomical Therapeutic Chemical			Antibiotic	Class	ATC code	Amoxicillin	Penicillins with extended spectrum	J01CA04	Amoxicillin/clavulanate	Combinations of penicillins	J01CR02	Azithromycin	Macrolides	J01FA10	Benzylpenicillin (Penicillin G)	Beta-lactamase sensitive penicillins	J01CE01	Cefotaxime	Third-generation cephalosporins	J01DD01	Ceftriaxone	Third-generation cephalosporins	J01DD04	Cefuroxime	Second-generation cephalosporins	J01DC02	Clarithromycin	Macrolides	J01FA09	Meropenem	Carbapenems	J01DH02	Phenoxymethylpenicillin (Penicillin V)	Beta-lactamase sensitive penicillins	J01CE02	Piperacillin/tazobactam	Combinations of penicillins	J01CR05
Antibiotic	Class	ATC code																																					
Amoxicillin	Penicillins with extended spectrum	J01CA04																																					
Amoxicillin/clavulanate	Combinations of penicillins	J01CR02																																					
Azithromycin	Macrolides	J01FA10																																					
Benzylpenicillin (Penicillin G)	Beta-lactamase sensitive penicillins	J01CE01																																					
Cefotaxime	Third-generation cephalosporins	J01DD01																																					
Ceftriaxone	Third-generation cephalosporins	J01DD04																																					
Cefuroxime	Second-generation cephalosporins	J01DC02																																					
Clarithromycin	Macrolides	J01FA09																																					
Meropenem	Carbapenems	J01DH02																																					
Phenoxymethylpenicillin (Penicillin V)	Beta-lactamase sensitive penicillins	J01CE02																																					
Piperacillin/tazobactam	Combinations of penicillins	J01CR05																																					
Medicinal product	NA																																						
Research question and objectives	This study aimed to characterise the monthly use (prescription/dispensation) of 11 antibiotics to understand trends, cycles and seasonality in the use of those medicines; and to forecast the prescription rates of such medicines under assumed scenarios, which could help anticipate and prevent potential shortages, or manage them. 1) To estimate monthly number of prescriptions of the 11 selected medicines during a 10-year period from the most recent data available (e.g., 2015-2024) overall and stratified by age and sex, in each of the databases. 2) To conduct time series modelling by fitting an ARIMA model to data generated in objective 1 for short-term (12-month) forecasting.																																						
Countries of study	Belgium, Finland, France, Germany, Netherlands and Spain																																						
Authors	Marta Pineda Moncusí Yuchen Guo																																						

1. TITLE

DARWIN EU® - Monitoring prescription of medicines for public health emergencies at risk of shortages

2. DESCRIPTION OF STUDY TEAM

Study team roles	Names	Organisations
Principal Investigator	Marta Pineda Moncusí	University of Oxford
Data Scientist	Yuchen Guo	University of Oxford
	Edward Burn	University of Oxford
Epidemiologist	Daniel Prieto Alhambra	University of Oxford/Erasmus MC
Study Manager	Natasha Yefimenko	Erasmus MC
Data partner name*	Data Partner member names	Organisations
Local Study Coordinator/Data Analyst	Romain Griffier Guillaume Verdy	CDWBordeaux
	Juan Luis Cruz Bermudez Noelia Garcia Barrio Javier de la Cruz Bertolo	H12O
	Angela Leis Miguel Angel Mayer Pujadas Juan Manuel Ramirez Anguita	IMASIS
	Josephine Jacob Raeleesha Norris Annika Vivirito Alexander Harms	InGef RDB
	Katia Verhamme	IPCI
	Dina Vojinovic Isabella Kaczmarczyk James Brash	IQVIA DA Germany/IQVIA LPD Belgium

*Data partners' role was only to execute code at their data source, review and approve their results. These people do not have an investigator role. Data analysts/programmers do not have an investigator role and thus declaration of interests (DOI) for these people was not needed.

3. DATA SOURCES

Table 1. Description of the selected data sources.

Country	Name of Database	Health Care setting	Type of Data	Number of active subjects	Last month of observed data used in ARIMA(X)
France	CDW Bordeaux	Hospital	EHR	2 million	01/11/2024
Spain	H120	Hospital	EHR	0.295 million	01/03/2025
Spain	IMASIS	Hospital	EHR	0.6 million	01/06/2024
Germany	InGef RDB	Primary and Hospital Care	EHR	10 million	01/09/2024
Netherlands	IPCI	Primary Care	EHR	2.9 million	01/06/2024
Germany	IQVIA DA Germany	Primary & Outpatient Care	EHR	41.9 million	01/09/2024
Belgium	IQVIA LPD Belgium	Primary Care	EHR	8.5 million	01/09/2024

ARIMA(X) = Autoregressive Integrated Moving Average [ARIMA] with an exogenous variable [ARIMAX], CDW = Clinical Data Warehouse, DA = Disease Analyzer, EHR = Electronic Health record, H120 = Hospital Universitario 12 de Octubre, IMASIS= Institut Municipal Assistència Sanitària Information System, IPCI = Integrated Primary Care Information, InGef RDB = Institut für angewandte Gesundheitsforschung, LPD = Longitudinal Patient Data.

4. ABSTRACT

Title

DARWIN EU® - Monitoring prescription of medicines for public health emergencies at risk of shortages

Rationale and background

The EMA is responsible for monitoring medicine shortages that might lead to a crisis situation, as well as reporting shortages of critical medicines during public health emergencies. Such shortages would make it difficult or impossible to meet the treatment needs of individual patients or populations.

The Agency also has the mandate to coordinate responses of EU / EEA countries to shortages of critical medical devices and in-vitro diagnostics in crisis situations.

Scientific and commercial data on monthly prescriptions of medicines that may be critical in PHE can help in understanding trends and seasonal variations. In conjunction with time series and forecasting models, as well as data on medicines supply, such data will contribute to the on-going efforts of the Agency to better monitor and coordinate its response to shortages of critical medicines.

This study aimed to generate the number of monthly prescriptions of selected medicines over the last 10 years of available data and to conduct time series modelling by fitting Autoregressive Integrated Moving Average (ARIMA) models to such data to then forecast short-term (12-months) prescription numbers.

Research question and objectives

- 1) To estimate monthly number of prescriptions or dispensations of the 11 selected medicines during the last 10 years of available data, stratified by age and sex, in each of the databases.
- 2) To conduct a time series modelling by fitting an ARIMA model to data generated in objective 1 for short-term (12-month) forecasting

Research Methods

Study design

- Population-level cohort (Objective 1, Population-level drug utilisation study)
- Time series modelling (Objective 2, Time series modelling based on Objective 1 findings)

Population

All individuals present in the database in the last 10 years of available data were included in the analysis. For this population, monthly of use of antibiotics was estimated.

Data sources

- 1) CDW Bordeaux (France, Secondary Care Database)
- 2) H12O (Spain, Secondary Care Database)
- 3) IMASIS (Spain, Secondary Care Database)
- 4) InGef RDB (Germany, Primary and Secondary Care Database)
- 5) IQVIA DA Germany (Primary and Outpatient care Database)
- 6) IQVIA LPD Belgium (Primary Care Database)
- 7) IPCI (Netherlands, Primary Care Database)

Variables

Drugs of interest: list of 11 antibiotics (amoxicillin, amoxicillin/clavulanate, azithromycin, benzylpenicillin, cefotaxime, ceftriaxone, cefuroxime, clarithromycin, meropenem, phenoxymethylpenicillin and piperacillin/tazobactam) that may be critical in PHE.

Calendar month for stratification.

Data analyses

Population-level drug utilisation study on antibiotics: estimation of the number of monthly prescriptions/dispensations for each antibiotic of interest in each database of the study.

Time series modelling: forecast of the 12-month monthly prescription/dispensation of antibiotic use after the end of available data using seasonal Auto-Regressive Integrated Moving Average (ARIMA) models, and ARIMA models incorporating the effect of the COVID-19 pandemic as an exogenous variable (ARIMAX).

Results

ARIMAX models projected an increase in the use of amoxicillin/clavulanate in InGef and IQVIA DA Germany but a decrease in CDW Bordeaux, an increase in azithromycin use in CDW Bordeaux, a decrease in cefotaxime use in IMASIS, an increase in ceftriaxone use in InGef, IPCI and IQVIA DA Germany but a decrease in CDW Bordeaux, a decrease in clarithromycin use in CDW Bordeaux and InGef, a decrease in phenoxymethylpenicillin use in CDW Bordeaux and in IQVIA DA Germany, and a decrease in piperacillin/tazobactam use in CDW Bordeaux in the following 12 month from the last observations. The remaining antibiotics that were forecasted were projected to maintain similar prescription/dispensation values as the ones observed before, which could include fluctuations that followed either a non-seasonal trend or a seasonality pattern.

Therefore, the French hospital CDW Bordeaux is projected to reduce their use of antibiotics in 6 out of the 8 antibiotics forecasted. Furthermore, databases from the same country did not necessarily exhibit the same trends of use. For instance, the two Spanish hospitals projected different trends of use for cefotaxime, and the two German databases projected different trends of use for clarithromycin and phenoxymethylpenicillin.

Based on the model diagnostics, ARIMAX presented a marginal better fit than ARIMA models. ARIMAX models demonstrate adequate levels of goodness of fit, which validates the reliability of their respective coefficients. A number of observations were out of the 95% predictive intervals of the train and tune values but the majority of the ARIMAX models presented a good fitting and captured the seasonality pattern of the antibiotics typically used for bacterial infections frequent in winter, such as amoxicillin or azithromycin. Azithromycin use presented a seasonality pattern in most of the primary care databases. Similarly, amoxicillin use presented a seasonality pattern in primary care databases whilst hospitals exhibited a linear non-seasonal pattern. Conversely, amoxicillin/clavulanate use only presented a seasonality pattern in one of the primary care databases. No apparent patterns were observed when comparing hospital and primary care seasonality trends for the remaining antibiotics. These generally exhibited a linear non-seasonal pattern, and few even exhibited a stationary trend. Additionally, azithromycin in CDW Bordeaux, ceftriaxone and meropenem in IMASIS, benzylpenicillin and meropenem in InGef, cefuroxime in IPCI and benzylpenicillin in IQVIA DA Germany models failed to capture seasonality and autocorrelation.

The forecasted monthly prescriptions/dispensations can be a useful tool to anticipate a minimum stock of the different antibiotics of interest necessary to cover the study population needs during the next 6 to 12 months.

5. LIST OF ABBREVIATIONS

Acronyms/terms	Description
ARIMA	Autoregressive Integrated Moving Average predictive models
ARIMAX	ARIMA models with an eXogenous variable which added an extra variable for the effect of the COVID-19 pandemic
ATC	Anatomical Therapeutic Chemical
CDM	Common Data Model
CDW	Clinical Data Warehouse
CI	Confidence Interval
DA	Disease Analyzer
DARWIN EU®	Data Analysis and Real-World Interrogation Network
DUS	Drug Utilization Study
EEA	European Economic Area
EHR	Electronic Health Records
EMA	European Medicines Agency
EU	European Union
GP	General Practitioner
H12O	Hospital Universitario 12 de Octubre
IMASIS	Institut Municipal Assistència Sanitària Information System
InGef RDB	Institut für angewandte Gesundheitsforschung
IPCI	Integrated Primary Care Information
LPD	Longitudinal Patient Data
OMOP	Observational Medical Outcomes Partnership
PHE	Public Health Emergencies

6. AMENDMENTS AND UPDATES

DOCUMENT HISTORY

Version	Date	Description
V1.0	30/04/2025	First Draft
V2.0	20/05/2025	Inclusion of changes to address EMA comments
V3.0	30/09/2025	Catalogue version, including formatting quality check

7. MILESTONES

Study deliverable	Timelines (planned)	Timelines (actual)
Draft Study Protocol	15/01/2025	15/01/2025
Final Study Protocol	January 2025	17/02/2025
Creation of Analytical code	February 2025	03/03/2025 Diagnostics 11/03/2025 Main Study code
Execution of Analytical Code on the data	March 2025	15 March-18 April 2025
Draft Study Report	April 2025	30/04/2025
Final Study Report	-	20/05/2025

8. RATIONALE AND BACKGROUND

The extended mandate of EMA, reinforcing the role of the Agency in crisis preparedness and management of medicinal products and medical devices, became applicable on 1st March 2022 ([Regulation on EMA's extended mandate becomes applicable | European Medicines Agency \(europa.eu\)](#)).

EMA is now responsible for monitoring medicine shortages that might lead to a crisis situation, as well as reporting shortages of critical medicines during public health emergencies (PHE). Such shortages would make it difficult or impossible to meet the treatment needs of individual patients or populations.(1) The Agency also has the mandate to coordinate responses of EU / EEA countries to shortages of critical medical devices and in-vitro diagnostics in crisis situations.

Scientific and commercial data on monthly prescriptions of medicines that may be critical in PHE can help in understanding trends and seasonal variations. In conjunction with time series and forecasting models, as well as data on medicines supply, such data will contribute to the on-going efforts of the Agency to better monitor and coordinate its response to shortages of critical medicines.

This study aimed to generate the number of monthly prescriptions of selected medicines over the last 10 years of available data and to conduct time series modelling by fitting Autoregressive Integrated Moving Average (ARIMA) models to such data to then forecast short-term (12-months) prescription numbers.

9. RESEARCH QUESTION AND OBJECTIVES

Table 2. Primary and secondary research questions and objective.

A. Primary research question and objective

Objective:	To estimate monthly number of prescriptions or dispensations of the 11 selected medicines during the last 10 years of available data, stratified by age and sex, in each of the databases.
Hypothesis:	Not applicable
Population (mention key inclusion-exclusion criteria):	The study cohort comprised all individuals present in the database in a 10-year period from the most recent data available.
Exposure:	Eleven antibiotics identified as potentially critical in public health emergencies (PHE)
Comparator:	None
Outcome:	None
Time (when follow up begins and ends):	Follow-up started on a pre-specified calendar time point, e.g., 1st of January, for each month between the 10-year study period for the calculation of the monthly number of prescriptions/dispensations. End of follow-up was defined as the earliest of loss to follow-up, end of data availability, death, or end of the 10-year study period.

Setting:	Inpatient and outpatient setting using data from the following data sources: 1) CDW Bordeaux (France, Secondary Care Database) 2) H12O (Spain, Secondary Care Database) 3) IMASIS (Spain, Secondary Care Database) 4) InGef RDB (Germany, Primary and Secondary Care Database) 5) IQVIA DA Germany (Primary and Outpatient care Database) 6) IQVIA LPD Belgium (Primary Care Database) 7) IPCI (Netherlands, Primary Care Database)
Main measure of effect:	Monthly number of antibiotic prescriptions/dispensations

B. Secondary research question and objective

Objective:	To conduct a time series modelling by fitting an ARIMA model to data generated in objective 1 for short-term (12-month) forecasting
Hypothesis:	Not applicable
Population (mention key inclusion-exclusion criteria):	The study cohort comprised the monthly number of prescriptions/dispensations of the 11 selected medicines during the last 10 years of available data, from the primary research question.
Exposure:	Eleven antibiotics identified as potentially critical in PHE
Comparator:	None
Outcome:	None
Time (when follow up begins and ends):	We used the overall (i.e., not stratified nor sensitivity analysis) monthly number of prescriptions/dispensations from the last 10 years of available data calculated in objective 1 to forecast the following 12 months.
Setting:	Inpatient and outpatient setting using data from the following data sources: 1) CDW Bordeaux (France, Secondary Care Database) 2) H12O (Spain, Secondary Care Database) 3) IMASIS (Spain, Secondary Care Database) 4) InGef RDB (Germany, Primary and Secondary Care Database) 5) IQVIA DA Germany (Primary and Outpatient care Database) 6) IQVIA LPD Belgium (Primary Care Database) 7) IPCI (Netherlands, Primary Care Database)
Main measure of effect:	Forecast of the 12-month number of prescriptions/dispensations of antibiotic use after the end of study period

10. RESEARCH METHODS

10.1. Study type and study design

Retrospective cohort studies were conducted using routinely collected health data from 7 databases. The study comprised two consecutive parts:

1. A population-based cohort study was conducted to address objective 1, assessing the monthly number of prescriptions/dispensations of the 11 selected antibiotics.
2. A time series modelling study based on the Population-level drug utilisation study on antibiotics was conducted to address objective 2, fitting the time series data generated in objective 1 into an ARIMA model to then forecast prescription rates for the subsequent 12 months.

10.2. Study setting and data sources

10.2.1. Data sources selected

Inpatient and outpatient setting using data from the following data sources:

- 1) Clinical Data Warehouse of Bordeaux University Hospital [CDW Bordeaux] (France, Secondary Care Database)
- 2) Hospital Universitario 12 de Octubre [H12O] (Spain, Secondary Care Database)
- 3) Institut Municipal Assistència Sanitària Information System [IMASIS, Parc Salut Mar Barcelona - Hospital del Mar] (Spain, Secondary Care Database)
- 4) InGef Research Database [InGef RDB, Institut für angewandte Gesundheitsforschung] (Germany, Primary and Secondary Care Database)
- 5) IQVIA Disease Analyzer Germany [IQVIA DA Germany] (Primary and Outpatient care Database)
- 6) IQVIA Longitudinal Patient Data Belgium [IQVIA LPD Belgium] (Primary Care Database)
- 7) Integrated Primary Care Information [IPCI] (Netherlands, Primary Care Database)

Information on data source(s) used is described in [Table 1](#).

10.2.2. Rationale for the data sources selected

In this routinely collected study, we selected 7 databases onboarded for DARWIN EU® in 2025, representing 5 European countries. The selection of databases for this study was based on availability of data until at least June 2024. All selected databases fulfilled the criteria required for a population-level drug utilisation study, as they systematically and adequately captured all drug prescriptions. Additionally, the proposed data sources captured the relevant data necessary to answer the research questions of this study, they had recent data covering the 10-year study period, and they covered different healthcare settings where the drugs of interest might have been used (i.e., primary and secondary care).

Two databases included records from primary care (IPCI and IQVIA LPD Belgium), three from hospital (IMASIS, CDW Bordeaux and H12O), one included records from primary and outpatient care (IQVIA DA Germany), and one included records from primary and hospital care (InGef RDB). Since certain medications observed in primary care might not have been observed at secondary care, and vice versa, given the nature of their own indication, we included data sources from both health care settings, and the administration path (primary and/or secondary care) had been considered when interpreting the results. We selected these three outpatient databases (IPCI, IQVIA LPD Belgium, and IQVIA DA Germany, covering primary care and outpatient specialist care) as many antibiotics, particularly beta-lactam antibiotics/combinations and macrolides, were expected to be prescribed in that ambulatory health care setting. In addition, IMASIS,

CDW Bordeaux, H120, and InGef RDB included in- and outpatient records from the hospital, where antibiotics used for the treatment of more severe conditions, including the third-generation cephalosporins and carbapenems, were expected to be administered during hospitalisation. These data sources also captured data on antibiotics that were expected to be initiated and prescribed for outpatient use following hospital discharge. With that, the diversity of healthcare systems enabled us to capture all the antibiotics of interest with a relevant number of prescriptions. While two of the databases included in the study were from the same country, potentially overlapping information of patients (IQVIA DA Germany and InGef RDB), we initially expected to provide a different coverage, with IQVIA DA Germany capturing a larger proportion of the German population and providing a broader and more representative coverage of prescriptions in the German primary care, whereas we considered InGef RDB particularly relevant for medications administered in hospital settings, which we did not expect to be captured by IQVIA DA Germany.

The proposed data sources captured the relevant data necessary to answer the research questions of this study, they had recent data covering the 10-year study period, and they covered different healthcare settings where the drugs of interest might have been used (i.e., primary and secondary care). Moreover, 4 of the 7 databases (CDW Bordeaux, IPCI, IMASIS and IQVIA DA Germany) were included for a previous DARWIN EU study which focussed on the use of antibiotics of the WHO Watch list (EUPAS103381).

10.2.3. Data sources description:

[Clinical Data Warehouse of Bordeaux University Hospital \[CDW Bordeaux\] \(France, Secondary Care Database\)](#)

The clinical data warehouse of the Bordeaux University Hospital comprised electronic health records on more than 2 million patients with data collection starting in 2005. The hospital complex was made up of three main sites and comprised a total of 3,041 beds (2021 figures). The database included information about the person (demographics), visits (inpatient and outpatient), conditions and procedures (billing codes), drugs (outpatient prescriptions and inpatient orders and administrations), measurements (laboratory tests and vital signs), and dates of death (in or out-hospital death). Drug information was primarily derived from the EHR administration record, followed by EHR prescription. Administration data consisted of information filled out principally by nurses when a medication was administered to the patient. Administration and prescription records may have contained some discrepancies, such as administrations without corresponding prescriptions or prescriptions that were not necessarily administered. Conversely, data from critical care units (ICUs) was managed using a separate software, and at present, only prescription data was integrated into the OMOP-CDM. The data from the hospital production information system were loaded daily into a CDW in i2b2 format. A specific ETL from i2b2 to OMOP had been set up to standardize the data in OMOP-CDM format. Currently, this ETL from i2b2 to OMOP was launched manually when needed (this ETL could be scheduled regularly if necessary). The data was integrated into the OMOP Common Data Model version 5.3.1 and was stored in Oracle version 19c.

[Hospital Universitario 12 de Octubre \[H120\] \(Spain, Secondary Care Database\)](#)

The main source of data for the OMOP database was the Electronic Health Record (EHR) of the Hospital Universitario 12 de Octubre, which had a reference population of about 450,000 patients. The data source contained information from the different healthcare domains (laboratory, prescriptions, treatments, administrative, diagnoses, etc.). In addition, information was also obtained from other data sources such as the pathological anatomy system, which provided information about sample analysis, and the cost system, containing information on the cost associated with a contact with the hospital. In the case of drug data, it was derived from multiple sources within the EHR, including structured data from inpatient and emergency prescriptions and natural language processing of EHR notes. Work for the inclusion of further data is ongoing, among others, radiological information, or PROMs.

[Institut Municipal Assistència Sanitària Information System \[IMASIS, Parc Salut Mar Barcelona - Hospital del Mar\] \(Spain, Secondary Care Database\)](#)

The Institut Municipal Assistència Sanitària Information System (IMASIS) contained the Electronic Health Record (EHR) system of Parc de Salut Mar Barcelona (PSMar), which is a complete healthcare services organisation. This information system included and shared the clinical information of two general hospitals (Hospital del Mar and Hospital de l'Esperança), one mental health centre (Centre Dr. Emili Mira) and one social-healthcare centre (Centre Fòrum), including emergency room settings, which offered specific and varied services in the Barcelona city area (Spain). IMASIS included clinical information on more than 1 million patients with at least one diagnosis and who had used the services of this healthcare system since 1990 and from different settings such as admissions, outpatients, emergency room, and major ambulatory surgery. Diagnoses were coded using The International Classification of Diseases ICD-9-CM and ICD-10-CM. The average follow-up period per patient in years was 6.37 (SD±6.82). Most drug data mapped to the OMOP-CDM in IMASIS was obtained from EHR administration records, which served as a primary source of medication-related information. These records were typically documented by nurses at the time of medication administration to ensure accurate tracking of drug usage in clinical settings. The recorded data included details such as the drug name, dosage, route of administration, and the time of administration. This process helped maintain consistency and reliability in capturing real-world medication exposure, supporting both clinical decision-making and research initiatives. IMASIS-2 was the anonymized relational database of IMASIS, which was used for mapping to OMOP, including additional sources of information such as the Tumours Registry.(2)

[Institut für angewandte Gesundheitsforschung Research Database \[InGef RDB\] \(Germany, Primary and Secondary Care Database\)](#)

The InGef RDB database comprised anonymized longitudinal claims data of about 10 million individuals across more than 70 statutory health insurance providers (SHIs) throughout Germany. Data were longitudinally linked over a period of currently ten years. Patients could be traced across healthcare sectors. All patient-level and provider-level data in the InGef RDB research database were anonymised to comply with German data protection regulations and German federal law. German SHI claims data available in the InGef RDB database included information on demographics (year of birth, gender, death date if applicable, region of residence on administrative district level); hospitalizations; outpatient services (diagnoses, treatments; specialities of physicians); dispensing of drugs; dispensing of remedies and aids; and sick leave and sickness allowance times. In addition, costs or cost estimates from the SHI perspective were available for all important cost elements. All diagnoses in Germany were coded using the International Classification of Diseases, version 10 in the German Modification (ICD-10-GM). The persistence (membership over time) was rather high in the InGef RDB database: During a time period of 5 years (2009 to 2013), 70.6% of insurance members survived and remained insured with the same SHI without any gap in their observational time. Persons leaving one of the participating SHIs and entering another participating SHI could be linked during yearly database consistency updates and were thus not lost over time. The InGef RDB database was dynamic in nature, i.e. claims data were updated in an ongoing process, and new SHIs might join or leave the database. By law, only the last 10 years of data were allowed to be used. At every new release, this window shifted, dropping older data and adding new data. Additionally, drug data in InGef RDB captured dispensation of medicines, with the records of the specific date of dispensation driven from the claims data-based dispensing record. All drugs that were reimbursed by the German statutory health insurance were captured within the data. However, a minority of prescription records from outpatient procedure codes (OPS-code), which represented procedures where a drug was administered, and which had been mapped to the observation table instead of the drug table, did not have a specific date associated and therefore were not included in this study. Additionally, over-the-counter medication could not be found in the database.

There was some overlap between the InGef RDB and IQVIA DA Germany databases. However, the IQVIA DA Germany database provided broader coverage of prescriptions in primary care, as it captured a larger proportion of the German population, whereas the InGef RDB database was particularly relevant for medications administered in hospital settings, which were not captured by IQVIA DA Germany.

[***IQVIA Disease Analyzer Germany \[IQVIA DA Germany\] \(Germany, Primary and Secondary Care Database\)***](#)

DA Germany was collected from extracts of patient management software used by GPs and specialists practicing in ambulatory care settings. Data coverage included more than 34M distinct person records out of a total population of 80M (42.5%) in the country and was collected from 2,734 providers. Patients visiting more than one provider were not cross-identified for data protection reasons and therefore were recorded as separate in the system. Dates of service included from 1992 through the present. Observation time was defined by the first and last consultation dates. Germany has no mandatory GP system and patient have free choice of specialist. As a result, data were collected from visits to General (28.8%), Orthopaedic Surgery (13.4%), Otolaryngology (11.8%), Dermatology (11.2%), Obstetrics/Gynaecology (7.7%), various Neurology and Psychiatry (6.2%), Paediatric (7.0%), Urology (4.6%), Cardiology (3.7%), Gastroenterology (3.5%), Pulmonary (1.5%), and Rheumatology (0.7%) practices. Drugs were recorded as prescriptions of marketed products, and the information was derived mainly from the EHR prescription issue record, followed by the EHR administration record.

There was some overlap between IQVIA DA Germany and InGef RDB databases. However, the IQVIA DA Germany database provided broader coverage of prescriptions in primary care, as it captured a larger proportion of the German population, whereas the InGef RDB database was particularly relevant for medications administered in hospital settings, which were not captured by IQVIA DA Germany.

[***IQVIA Longitudinal Patient Data Belgium \[IQVIA LPD Belgium\] \(Belgium, Primary Care Database\)***](#)

IQVIA LPD Belgium is a computerised network of GPs who contributed to a centralised database of anonymised data of patients with ambulatory visits. Currently, around 300 GPs from 234 practices were contributing to the database, covering 1.1M patients from a total of 11.5M Belgians (10.0%). The database covered time from 2005 through the present. Observation time was defined by the first and last consultation dates. Drug information was derived from GP prescriptions, which were systematically recorded in the EHR prescription issue record. Drugs obtained over the counter by the patient outside the prescription system were not reported. No explicit registration or approval was necessary for drug utilisation studies.

[***Integrated Primary Care Information \[IPCI\] \(Netherlands, Primary Care Database\)***](#)

The Integrated Primary Care Information (IPCI) database is a longitudinal observational database containing routinely collected data from computer-based patient records of a selected group of GPs throughout the Netherlands (N=723). IPCI was started in 1992 by the department of Medical Informatics of the Erasmus University Medical Center in Rotterdam with the objective to enable better post-marketing surveillance of drugs. The current database included patient records from 2006 on, when the size of the database started to increase significantly. In 2016, IPCI was certified as a Regional Data Center. Since 2019, the data was also standardised to the OMOP-CDM, enabling collaborative research in a large network of databases within the Observational Health Data Sciences and Informatics (OHDSI) community. The primary goal of IPCI was to enable medical research. In addition, reports were generated to inform GPs and their organizations about the provided care. Contributing GPs were encouraged to use this information for their internal quality evaluation. Since IPCI is a registration of GP systems, the majority of drug records were provided directly from GP prescriptions, with the occasional addition of a few pharmacy dispensation records. Drugs that were captured as prescription records contained the information of the product, quantity, dosing directions, and strength. Additionally, prescription records might have contained information on the

indication of use, although this information was not always complete. Drugs not prescribed in the GP setting might have been underreported.

10.3. Study period

The study comprised the last 10 years of available data counting backwards from the respective latest date of data lock of the respective databases (**Table 3**).

Table 3. Operational Definition of Time 0 (index date) and other primary time anchors.

Study population names	Time Anchor Description (e.g., time 0)	Number of entries	Type of entry	Washout window	Care Setting ¹	Code Type	Diagnosis position	Incident with respect to...	Measurement characteristics/validation	Source of algorithm
All patients from the database eligible for the study	Patient present in the database during the study period (the 10-year period from the latest available data)	Multiple	Incident	n/a	IP and OP	n/a	n/a	Specific antibiotics	n/a	n/a

¹ IP = inpatient, OP = outpatient, OT = other, n/a = not applicable

10.4. Follow-up

The number of monthly prescriptions/dispensations was obtained directly from the numerators of the incidence rates estimates. To estimate the incidence rates, we required the appropriate population and their contributed observation time to first be identified. Study individuals in the denominator population began contributing person time on the respective date of the latest of the following: 1) study start date 2) date at which the observation period started. Individuals stopped contributing person time at the earliest date of the following: 1) end of available data in each of the data sources or 2) date at which the observation period of the specific person ended.

An example of entry and exit into the denominator population is shown in **Figure 1**. In this example, person IDs 1 and 3 were included as denominators at the study start date as both were being observed in the database from a prior date. Person ID 2 entered the study at the date of starting their period of observation, which was later than the study start date. Person IDs 1 and 2 were followed until the study end date (end of available data in each of the data sources) whilst person ID 3 left when exiting the database (the end of the observation period). Lastly, person ID 4 had two observation periods in the database. The first period contributed time from the study start until the end of the observation period, and the second period started contributing time again on the date their second observation period start and exited at the study end date.

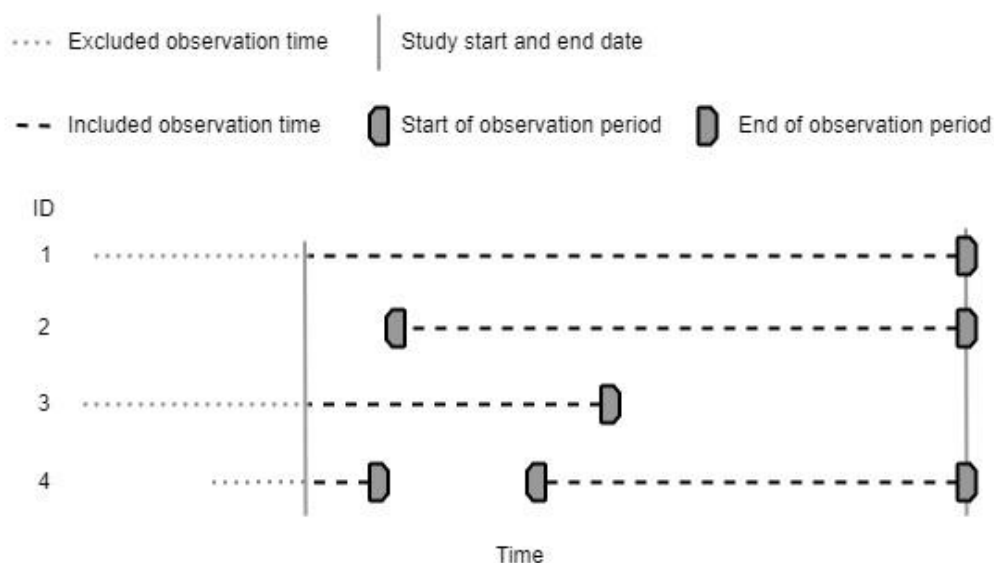


Figure 1. Included observation time for the denominator population.

10.5. Study population with in and exclusion criteria

The study cohort comprised all individuals present in the database during the study period (the last 10 years of available data counting backwards from the respective latest date of data lock of the respective databases).

The operational definitions of the in and exclusion criteria are presented in [Table 4](#).

Table 4. Operational definitions of inclusion criteria.

Criterion	Details	Order of application	Assessment window	Care Settings ¹	Code Type	Diagnosis position	Applied to study populations:	Measurement characteristics/validation	Source for algorithm
Observation period in the database during the 10-year period from the latest available data	All individuals present in the database during the study period	After	N/A	PC, SC and Combination of both	N/A	N/A	All individuals within the selected databases	N/A	N/A

¹ PC = Primary Care, SC = Secondary Care, OT = other, n/a = not applicable

10.6. Variables

10.6.1. Exposure

For this study, the exposure of interest was use (during study period) of antibiotics identified as potentially critical in PHE. This list of antibiotics (with respective ATC code) is described in [Table 5](#). Details of exposure are described in [Table 6](#).

Table 5. Exposure of interest.

Antibiotic	Class	ATC code
Amoxicillin	Penicillins with extended spectrum	J01CA04
Amoxicillin/clavulanate (Amoxicillin + Clavulanic acid)	Combinations of penicillins	J01CR02
Azithromycin	Macrolides	J01FA10
Benzylpenicillin (Penicillin G)	Beta-lactamase sensitive penicillins	J01CE01
Cefotaxime	Third-generation cephalosporins	J01DD01
Ceftriaxone	Third-generation cephalosporins	J01DD04
Cefuroxime	Second-generation cephalosporins	J01DC02
Clarithromycin	Macrolides	J01FA09
Meropenem	Carbapenems	J01DH02
Phenoxymethylpenicillin (Penicillin V)	Beta-lactamase sensitive penicillins	J01CE02
Piperacillin/tazobactam (Piperacillin + Tazobactam)	Combinations of penicillins	J01CR05

ATC = Anatomical Therapeutic Chemical

Table 6. Operational definitions of exposure.

Exposure group name(s)	Details	Wash out window	Assessment Window	Care Setting ¹	Code Type	Diagnosis position	Applied to study populations:	Incident with respect to...	Measurement characteristics/validation	Source of algorithm
Antibiotics identified as potentially critical in PHE – users cohort	Preliminary code lists provided in Appendix	n/a	Monthly	PC, SC and Combination of both	RxNorm	n/a	All individuals present in the database during the study period	n/a	n/a	n/a

¹ PC = Primary Care, PHE = Public Health Emergencies, SC = Secondary Care, OT = other, n/a = not applicable

10.6.2. Outcomes

None.

10.6.3. Other covariates, including confounders, effect modifiers, and other

Covariates for stratification in population-level drug utilisation study:

- Age: age was stratified in three groups (<18, 18-64, 65+).
- Time period: Monthly
- Sex

10.7. Study size

No sample size was calculated for this study. The expected number of prescriptions/dispensations across data sources was expected to be roughly between <1000 and 25 million.

10.8. Patient privacy protection

To prevent confidentiality issues, cell counts lower than 5 were reported as “<5”.

10.9. Statistical methods

This section described the details of the analysis approach and rationale for the choice of analysis, with reference to the *D1.3.8.1 Draft Catalogue of Data Analysis* which described the type of analysis in function of the study type.

The analysis included calculation of population-based incidence rates as described in section 10.9.4.1 – *Population-level drug utilisation study on antibiotics*; and the fitting of the data generated in objective 1 into ARIMA models as described in section 10.9.4.2 – *Time series modelling*.

10.9.1. Federated network analyses

Analyses were conducted separately for each database. Before study initiation, test runs of the analytics were performed on a subset of the data sources or on a simulated set of patients, and quality control checks were performed. Once all the tests were passed, the final package was released in the version-controlled Study Repository for execution against all the participating data sources.

The data partners locally executed the analytics against the OMOP-CDM in R Studio and reviewed and approved the by default aggregated results before returning them to the Coordination Centre. Sometimes, multiple execution iterations were performed, and additional fine-tuning of the code base was needed. A service desk was available during the study execution for support.

The study results of all data sources were checked, after which they were made available to the team in the Digital Research Environment, and the Dissemination Phase could start. All results were locked and timestamped for reproducibility and transparency.

10.9.2. Statistical model specification and assumptions of the analytical approach considered

Users cohorts

Users (for incidence rates estimation) were selected based on their use (prescription or dispensation) of the respective drug of interest after the start of the study and/or in a predefined time window.

Missing Data

Individuals who were lost to follow-up or had an end of observation period before the end of the study period had missing follow-up between the time of loss to follow-up/end of observation period and the end of the study period (see the corresponding definitions in Section 10.4), and data on the prescription/dispensation of the medicines of interest were missing. For these individuals, the time at risk of experiencing the event of interest (prescription of one of the medicines listed in Section 10.6.1) was censored at the time of loss-to-follow-up/end of observation period, and the estimated prescription rates implicitly assumed censoring occurred at random.

R-packages

We used the R package “IncidencePrevalence” for the population-level estimation of drug utilisation. Additionally, we used the R package “forecast” for the ARIMA modelling.

10.9.3. Methods to derive parameters of interest

Calendar time

Calendar time was based on monthly periods of the index prescription/dispensation in the population level analysis (e.g., 01/01/2015 – 31/01/2015, 01/02/2015 – 30/02/2015, etc.).

Age

Age at index date was calculated using January 1st of the year of birth as proxy for the actual birthday. The following age groups were used for stratification: <18, 18-64, 65+.

Sex

Results were presented stratified by sex.

10.9.4. Main Statistical Methods

Population-level drug utilisation study

Monthly number of prescriptions/dispensations were obtained directly from the numerators of the incidence rates estimates (i.e., these estimates were produced by the IncidencePrevalence R package automatically after the extraction of the relevant data for the numerator and the denominator of the incidence rate). Incidence calculations were conducted separately for each antibiotic of interest using the following methodology described below, which has been included for the sake of completeness. However, please note in this study only the numerator (monthly number of prescriptions/dispensations) was reported and analysed in the second objective. The estimated incidence rates were not reported.

Incidence calculations

Monthly incidence rates of the antibiotics of interest were calculated as the number of **users** per 100,000 person-years of the population at risk of getting exposed during the period for each calendar year. Those study individuals who entered the denominator population then contributed time at risk up to their first use (prescription or dispensation) during the study period. If they did not have a drug exposure, they contributed time at risk as described above in section 10.4 (Follow-up). Incidence rates were given together with 95% Poisson exact confidence intervals.

An illustration of the calculation of incidence of antibiotic use is shown below in **Figure 2**.

Patient IDs 1, 2, and 3 contributed time at risk between the study start and end date except when they became incident users of antibiotics. For Patient IDs 2 and 5, use of antibiotics before the study start was not counted as incident use. Patient IDs 4 and 5 were not seen to use antibiotics between the study start and end date and so contributed time at risk but no incident outcomes.

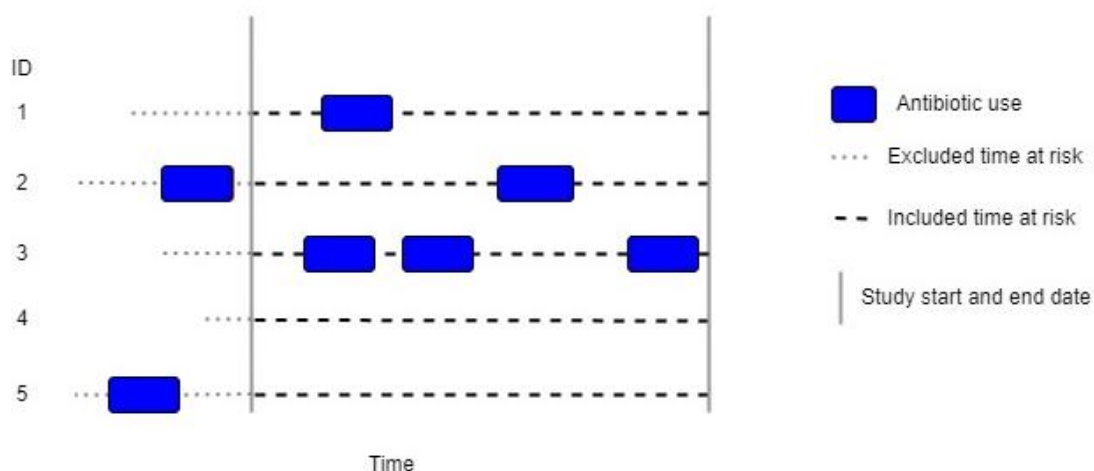


Figure 2. Incidence example for antibiotic use.

Time series modelling

We conducted time series modelling where AutoRegressive Integrated Moving Average (ARIMA) models were used to model trends and seasonality over time for the overall number of prescriptions/dispensations (i.e., the numerator of the IR estimates from the main analysis, without age and sex stratification). Modelling was performed only when no more than one month's value was censored, with censored values replaced by 5 to avoid skewing the results.

These models were characterized by varying combinations of autoregressive (AR) and moving average (MA) parameters, possibly on the integrated time series.

The two types of ARIMA models used were:

1. Seasonal AutoRegressive Integrated Moving Average (ARIMA):

These models were characterized by varying combinations of seasonal autoregressive (AR) and moving average (MA) parameters, possibly on the integrated time series.

2. Seasonal AutoRegressive Integrated Moving Average with an eXogenous variable (ARIMAX):

These models were characterized by varying combinations of seasonal autoregressive (AR) and moving average (MA) parameters, possibly on the integrated time series. Additionally, an extra variable was incorporated to account for the effect of the COVID-19 pandemic from January 2020 to December 2021, transforming the models into ARIMAX models to better capture this external influence.

The model selection and evaluation process were outlined in the following sub-sections:

1. Train: First, ARIMA(X) models were fit using the `auto.arima()` function from the forecast R package, which considered stationarity, seasonality, and trend testing. A range of different combinations of AR and MA parameters were assessed, and the best parsimonious model was chosen based on the lowest Akaike's Information Criterion (AIC), together with diagnostic tools (e.g., plots of residuals).
2. Validation: The model's predictive accuracy was assessed using out-of-sample validation techniques. Evaluation metrics included the Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). For the validation, we removed the last 12 months of available data, fit ARIMA(X) models to the remaining data, predicted the next 12 months, and compared the predictions with the actual observations. We noted that while the ARIMA(X) models fitted may well have been different from the ones obtained using all available data, the obtained metrics still informed the validity of the general approach used to estimate the ARIMA(X) models.
3. Forecasting: The selected ARIMA(X) model(s) were employed to forecast the numerator of the IR of prescriptions or dispensations for the next 12 months after the end of data availability.

10.9.5. Evidence synthesis

Results from analyses described in Section 10.9 are presented separately for each database and no meta-analysis of results were conducted.

11. DATA MANAGEMENT

All databases had been mapped to the OMOP common data model. This enabled the use of standardised analytics and tools across the network since the structure of the data, and the terminology system were harmonised. The OMOP CDM was developed and maintained by the Observational Health Data Sciences and Informatics (OHDSI) initiative and was described in detail on the wiki page of the CDM: <https://ohdsi.github.io/CommonDataModel> and in The Book of OHDSI: <http://book.ohdsi.org>. The analytic code for this study was written in R. Each data partner executed the study code against their database containing patient-level data and then returned the results set, which only contained aggregated data. The results from each of the contributing data sites were then combined in tables and figures for the study report.

12. QUALITY CONTROL

General database quality control

A number of open-source quality control mechanisms for the OMOP CDM had been developed (see Chapter 15 of The Book of OHDSI <http://book.ohdsi.org/DataQuality.html>). In particular, it was expected that data partners had run the OHDSI Data Quality Dashboard tool (<https://github.com/OHDSI/DataQualityDashboard>). This tool provided numerous checks relating to the conformance, completeness, and plausibility of the mapped data. Conformance focused on checks that described the compliance of the representation of data against internal or external formatting, relational, or computational definitions, completeness in the sense of data quality was solely focused on quantifying missingness, or the absence of data, while plausibility sought to determine the believability or truthfulness of data values. Each of these categories had one or more subcategories and were evaluated in two contexts: validation and verification. Validation related to how well data aligned with external benchmarks with expectations derived from known true standards, while verification related to how well data conformed to local knowledge, metadata descriptions, and system assumptions.

Study specific quality control

When defining drug cohorts, non-systemic products were excluded from the list of included codes summarised on the ingredient level. A pharmacist reviewed the codes of the antibiotics of interest.

When defining cohorts for indications, a systematic search of possible codes for inclusion was identified using the CodelistGenerator R package (<https://github.com/darwin-eu/CodelistGenerator>). This software allowed the user to define a search strategy and using this, it then queried the vocabulary tables of the OMOP common data model to find potentially relevant codes. In addition, the CohortDiagnostics R package (<https://github.com/OHDSI/CohortDiagnostics>) was run if needed to assess the use of different codes across the databases contributing to the study and identify any codes potentially omitted in error.

The study code was based on two R packages currently being developed to (1) estimate Incidence and Prevalence and (2) characterise drug utilisation using the OMOP common data model. These packages included numerous automated unit tests to ensure the validity of the codes, alongside software peer review and user testing. The R package was made publicly available via GitHub.

13. RESULTS

All results for each individual drug and database can be found on a shiny app at <https://data-dev.darwin-eu.org/P3-C2-004-RR/>, including counts of antibiotic records and subjects [Cohort count tab], number of monthly observed prescriptions/dispensation vs forecasted estimates by ARIMA(X) models [Predictions tab], and model parameters and diagnostics [Diagnostics tab].

13.1. Participants

Table 7 summarises the number of subjects with the outcomes of interest and which antibiotic per database was used for forecast modelling.

13.2. Main results

We generated times series data (i.e., the monthly prescriptions/dispensations of the antibiotic of interest in each of the databases) up to the date displayed in the column "Last month of observed data used in ARIMA(X)" from **Table 1**.

We fitted the number of the monthly prescriptions/dispensations of the antibiotic of interest in each of the databases of the study selected for forecasting (as detailed in **Table 7**), into seasonal ARIMA and ARIMAX (ARIMA models with an external variable for the effect of the COVID-19 pandemic between 1st January 2020 and 1st December 2021) models.

The initial model parameters were estimated automatically by the auto.arima function, which used a maximum sum of $p + q + P + Q = 5$, where p is the order of the autoregressive process (AR), q is the order of the moving average process (AR) and P and Q are their seasonal counterparts, respectively. Based on these diagnostics, the models were fine-tuned until an acceptable fit was achieved.

13.2.1. Model diagnostics

Full model diagnostics results available in the shiny app tabs "Diagnostics" (including MAE, MAPE, model parameters and AIC), "Scatterplots", "QQ-plots" and "ACF PACF Plots".

The median MAE and MAPE evaluation metrics of all the models were 18.23 and 18.18, respectively. When MAE and MAPE metrics between ARIMA and ARIMAX were compared per drug and database, there was no many differences between them (see an example with amoxicillin and clarithromycin metrics in **Table 8**) with the exception of specific cases that could favoured the use of the extra variable accounting for the COVID-19 pandemic effect, such as amoxicillin models in INGEF (e.g., MAPE ARIMA vs ARIMAX: 21.6 vs 16.3, respectively), or could favoured not having such term, such as clarithromycin models in INGEF (e.g., MAPE ARIMA vs ARIMAX: 21.0 vs 45.1, respectively).

QQ-plots of the residuals demonstrate that most models for different drugs have normally distributed residuals while some models exhibit outliers resulting in deviations from normality. The ACF and PACF plots of the residuals indicated that at least 95% of the values across the plotted lags fell within the 95% confidence intervals, showing no significant autocorrelation and supporting a reliable model fit. Specifically for amoxicillin in IPCI, the ACF plot showed only one significant lag at zero (the series correlated with itself), implying no residual moving average components. The PACF plot showed no significant lags, suggesting no remaining autoregressive structure.

The models reported below focus on ARIMAX. Equivalent results from ARIMA models are available in the shiny app.

Table 7. Number of records and drug users per antibiotic cohort during the overall study period. Usage of the antibiotics of interest in each database was forecasted when study cohort had no more than one month's value was censored to ensure the results were not skewed by replacing the censored values with a value of 5.

CDM name	Variable name	Antibiotic cohort										
		amoxicillin	amoxicillin/clavulanate	azithromycin	benzylpenicillin	cefotaxime	ceftriaxone	cefuroxime	clarithromycin	meropenem	phenoxymethylpenicillin	piperacillin/tazobactam
CDW Bordeaux	N records	706,484	477,851	19,285	302*	4877*	401,635	5876*	10,781	32,342	40,277	291,067
	N subjects	114,551	84,922	4,254	120*	1465*	69,364	2254*	1,558	4,343	3,316	35,728
H120	N records	408,191	53,179	28,210	4402*	4,218	173,794	8,478	688*	32,287	521*	8855*
	N subjects	146,220	41,445	20,656	4177*	3,362	65,563	7,735	604*	21,500	198*	6758*
IMASIS	N records	560,218	531,727	49,620	8,107	62,584	138,731	11,413	4357*	82,954	-	146,236
	N subjects	107,198	100,238	20,774	2,688	10,694	29,706	5,659	2166*	7,455	-	15,745
INGEF	N records	7,353,476	2,205,576	2,408,263	14,137	540*	16,965	3,236,472	952,746	1,797	1,471,900	983*
	N subjects	3,964,617	1,572,648	1,550,287	6,775	418*	12,844	2,053,183	719,102	998	1,090,837	679*
IPCI	N records	1,576,173	571,325	368,632	8,918	391*	7,875	3,540	160,976	549*	5912*	195*
	N subjects	812,758	369,515	214,460	1,981	345*	6,193	2,233	115,561	213*	4568*	59*
IQVIA DA Germany	N records	3,554,892	842,966	1,441,432	13,395	402*	10,007	1,791,889	825,069	732*	1,022,745	241*
	N subjects	2,171,222	626,389	910,593	7,071	292*	7,376	1,180,654	567,051	330*	717,744	208*
IQVIA LPD Belgium	N records	632,998	279,058	191,961	1243*	-	1032*	55,256	66,247	-	1770*	-
	N subjects	306,037	157,265	114,739	389*	-	869*	37,241	48,377	-	1593*	-

- : counts were <5 and censored and therefore antibiotic use was not forecasted. **In bold**: antibiotics that were forecasted. *In italics*: antibiotics that were not forecasted in the respective database. * Time series data for the antibiotic of interest had more than one month with a censored value in the database of interest and therefore, it was not forecasted in the database of interest. Abbreviations: CDW = Clinical Data Warehouse, DA = Disease Analyzer, H120 = Hospital Universitario 12 de Octubre, IMASIS= Institut Municipal Assistència Sanitària Information System, IPCI = Integrated Primary Care Information, InGef RDB = Institut für angewandte Gesundheitsforschung, LPD = Longitudinal Patient Data, N = number.

Table 8. Mean Absolute Error and Mean Absolute Percentage Error evaluation metrics, model parameters, and Akaike's Information Criterion between ARIMA and ARIMAX models forecasting amoxicillin and clarithromycin use.

CDM name	Antibiotic cohort	Model	MAE value	MAPE value	p	d	q	P	D	Q	AIC
CDW Bordeaux	amoxicillin	ARIMA	370.0	12.6	0	1	4	1	0	0	1540.5
CDW Bordeaux	amoxicillin	ARIMAX	362.1	12.3	0	1	4	1	0	0	1542.2
H12O	amoxicillin	ARIMA	315.0	35.4	0	1	0	1	0	0	1702.7
H12O	amoxicillin	ARIMAX	315.4	35.5	0	1	0	1	0	0	1703.4
IMASIS	amoxicillin	ARIMA	219.6	9.9	1	1	1	2	0	0	1600.9
IMASIS	amoxicillin	ARIMAX	200.8	9.1	1	1	1	3	0	0	1603.2
INGEF	amoxicillin	ARIMA	13874.3	21.6	0	1	2	0	1	1	2063.3
INGEF	amoxicillin	ARIMAX	11489.6	16.3	1	0	2	0	1	1	2085.2
IPCI	amoxicillin	ARIMA	1753.4	13.6	2	0	0	0	1	1	1747.5
IPCI	amoxicillin	ARIMAX	1771.8	13.8	0	0	4	0	1	1	1744.7
IQVIA DA Germany	amoxicillin	ARIMA	3830.2	15.6	1	0	1	0	1	1	1970.4
IQVIA DA Germany	amoxicillin	ARIMAX	3390.1	13.6	1	0	2	0	1	1	1971.9
IQVIA LPD Belgium	amoxicillin	ARIMA	844.1	38.8	1	0	2	0	1	1	1693.9
IQVIA LPD Belgium	amoxicillin	ARIMAX	892.3	39.1	1	0	2	0	1	1	1685.0
CDW Bordeaux	clarithromycin	ARIMA	11.1	46.6	3	0	0	NA	NA	NA	1032.9
CDW Bordeaux	clarithromycin	ARIMAX	9.8	43.6	3	0	0	1	0	0	1034.5
INGEF	clarithromycin	ARIMA	1210.9	21.6	2	1	1	2	1	0	1731.5
INGEF	clarithromycin	ARIMAX	2721.3	45.1	1	0	1	0	1	1	1751.1
IPCI	clarithromycin	ARIMA	55.1	8.2	1	1	1	1	0	1	1320.4
IPCI	clarithromycin	ARIMAX	56.4	8.4	1	1	1	1	0	1	1320.6
IQVIA DA Germany	clarithromycin	ARIMA	796.5	26.0	1	0	0	0	1	1	1617.0
IQVIA DA Germany	clarithromycin	ARIMAX	860.0	28.3	1	0	0	0	1	1	1618.6
IQVIA LPD Belgium	clarithromycin	ARIMA	101.6	55.8	2	1	1	1	1	1	1277.7
IQVIA LPD Belgium	clarithromycin	ARIMAX	119.8	46.9	2	0	0	3	1	0	1290.5

AIC = Akaike's Information Criterion, ARIMA = AutoRegressive Integrated Moving Average models, ARIMAX = ARIMA models with an exogenous variable which added an extra variable for the effect of the COVID-19 pandemic,

ARIMA(X) terms: p = number of autoregressive terms, d = number of differencing operations needed to make the time series stationary, q = number of moving average terms, P = number of seasonal autoregressive terms, D = number of seasonal differencing required to remove seasonal trends, Q = number of seasonal moving average terms based on seasonal lags;

CDW = Clinical Data Warehouse, DA = Disease Analyzer, H12O = Hospital Universitario 12 de Octubre, IMASIS= Institut Municipal Assistència Sanitària Information System, IPCI = Integrated Primary Care Information, InGef RDB = Institut für angewandte Gesundheitsforschung, LPD = Longitudinal Patient Data, MAE = Mean Absolute Error, MAPE = Mean Absolute Percentage Error.

13.2.2. ARIMAX model forecasting 12 months of antibiotic use

The plots shown in this section (**Figure 3** to **Figure 7**) display the time series data (i.e., the monthly prescriptions/dispensations of the antibiotic of interest in each of the databases) in orange, the train and tune estimates using grey dots, and the values forecasted by the model for the next 12 months using black dots. The vertical black dashed lines indicate the start of the tune and forecast periods, respectively.

We focused on reporting the values from January 2019 onwards to improve the visualisation of the forecasted estimates, but all estimates for the observed and train periods are shown in the shiny app.

CDW Bordeaux

Values and predictive intervals from the train and tune periods generally captured the observed number of monthly prescriptions/dispensations, with the exception of certain observations, such as the upsurge in clarithromycin use noted in December 2023 (observed vs trained value: 163 vs 69 [95%CI: 20-118], respectively); or the peak in phenoxymethylpenicillin observed in October 2024 (observed vs trained value: 352 vs 189 [95%CI: 19-360], respectively).

However, although the observed values were encompassed within the model's predictive interval, the tune estimates for azithromycin did not follow the observed fluctuations; instead, they exhibited a steady increase that was projected to continue into the forecast period. Looking at the model parameters for azithromycin, we observed a non-seasonal ARIMAX(0,1,2) model. The absence of autoregressive terms ($p = 0$) indicated that past values of the series were not directly used for prediction, while the first-order differencing ($d = 1$) captured a steady linear trend over time, driven by the last estimated changes and the external covariate to account for the effect of the COVID-19 pandemic.

Amoxicillin alone was expected to exhibit similar values in 2025 as in 2024, with increased use during the winter months (peaking in December 2024 at 3754 prescriptions/dispensations [95%CI: 3008-4501]), followed by a decline over the spring and summer (with the lowest value forecasted for September 2025 at 3386 [95%CI: 2436-4336]).

Meropenem was projected to display a stable seasonal pattern throughout 2025, with forecasted values ranging between 271 [95%CI: 93-450] and 326 [95%CI: 153-498] monthly prescriptions/dispensations.

Phenoxymethylpenicillin tune and forecasted estimates decreased from 198 [95%CI: 79-317] (observed was 165) in June 2024 to 158 [95%CI: -100 to 417] monthly prescriptions/dispensations in November 2025. However, the last three observed prescriptions/dispensations in September–November 2024 showed an increased use of phenoxymethylpenicillin (ranged between 266 and 352).

In contrast, amoxicillin/clavulanate was expected to continue its moderate downward trend, decreasing from 1086 observed prescriptions/dispensations in November 2024 to a forecasted 909 [95% CI: 171–1647] in September 2025. Similarly, ceftriaxone, piperacillin/tazobactam and clarithromycin were projected to show decreased use after December 2024 - January 2025. (**Figure 3**)

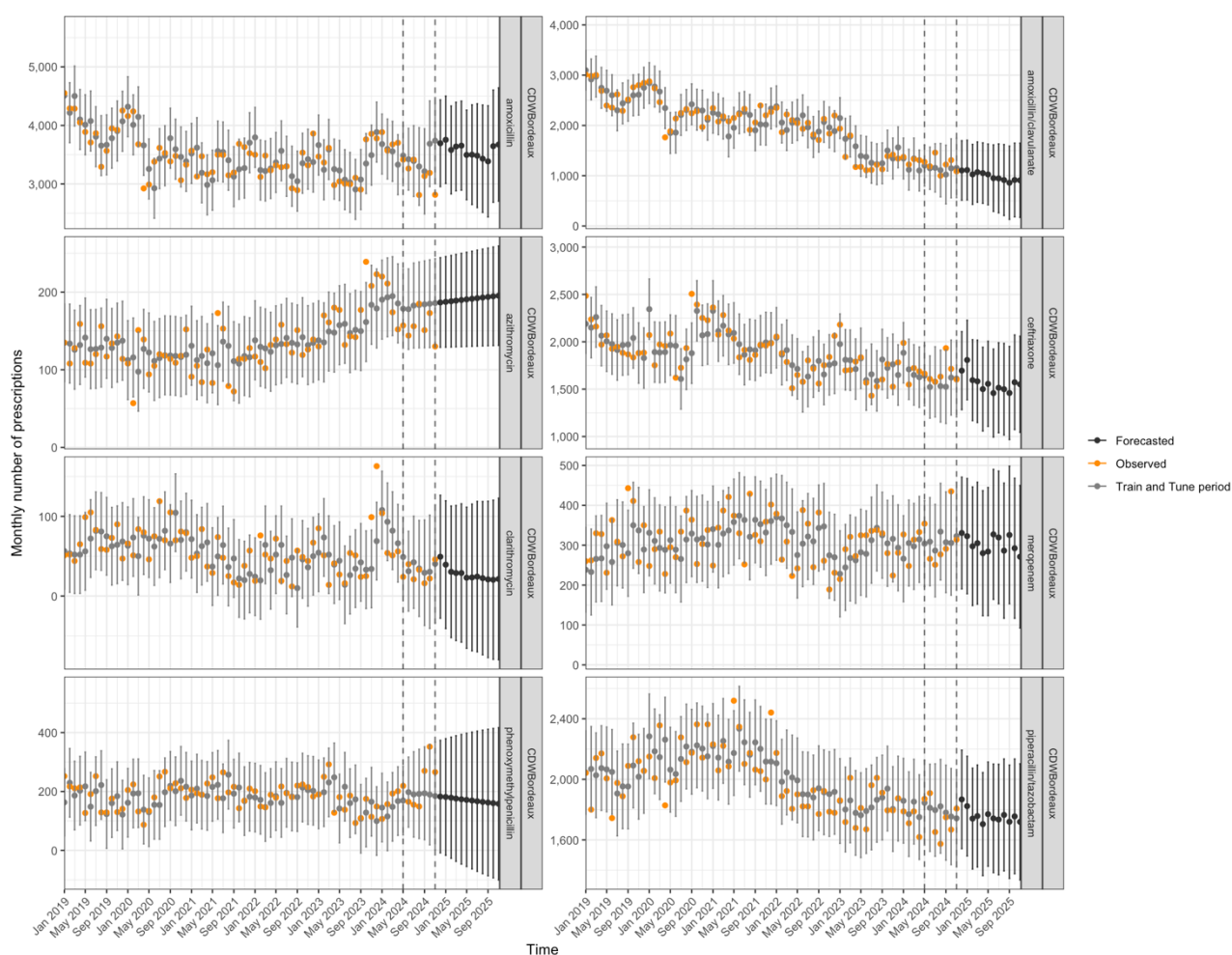


Figure 3. CDW Bordeaux models: forecasted monthly prescriptions/dispensations of the antibiotic of interest with ARIMAX (in black) vs observed values (in orange). The train and tune values from the ARIMAX are represented in grey. The vertical dotted lines divide time into three sections from left to right: the trained, tuned and forecasted periods. The ARIMAX values are presented with 95% confidence intervals.

H120

Values and predictive intervals from the train and tune periods generally captured the observed number of monthly prescriptions/dispensations, except for certain observations such as outliers around March 2020 or the surge in use of cefotaxime and meropenem noted in December 2022 (observed vs trained values: 69 vs 44 [95%CI: 27-60] for cefotaxime, and 335 vs 272 [95%CI: 226-317] for meropenem, respectively).

Observed prescriptions/dispensations of amoxicillin and ceftriaxone presented a substantial drop in use between February and March 2024 (from 4876 to 924, and 1867 to 1153, respectively). Forecasted values for amoxicillin prescriptions/dispensations were projected to fall below 0 from March 2025, but the high 95%CI ranged between 2042 and 4616. Ceftriaxone monthly use was projected to fluctuate between 886 and 924 from March 2025 onwards. These drops in amoxicillin and ceftriaxone use were likely caused by an incomplete recording from unstructured sources (see more details in [15.2.3 Database-specific limitations](#)).

The observed monthly prescriptions/dispensations of amoxicillin/clavulanate and azithromycin increased during the tune period. Although these trends were captured within the predictive intervals, the model projected a stable use that persisted throughout the forecast period (around 522 prescriptions/

dispensations from July 2024 onwards for amoxicillin/clavulanate, and around 280 for azithromycin from October 2024).

Observed prescriptions/dispensations of cefotaxime between October 2024 and January 2025 were lower than the tune estimates. The model forecasted a decline from March to July 2025 (49 [95%CI: 31-68] to 37 [95%CI: 19-56]), followed by an increase through to December 2025 (46 [95%CI: 25-66]). Lastly, use of cefuroxime and meropenem were expected to remain around 53 and 315 monthly prescriptions/dispensations, respectively, from January 2025 onwards. (**Figure 4**)

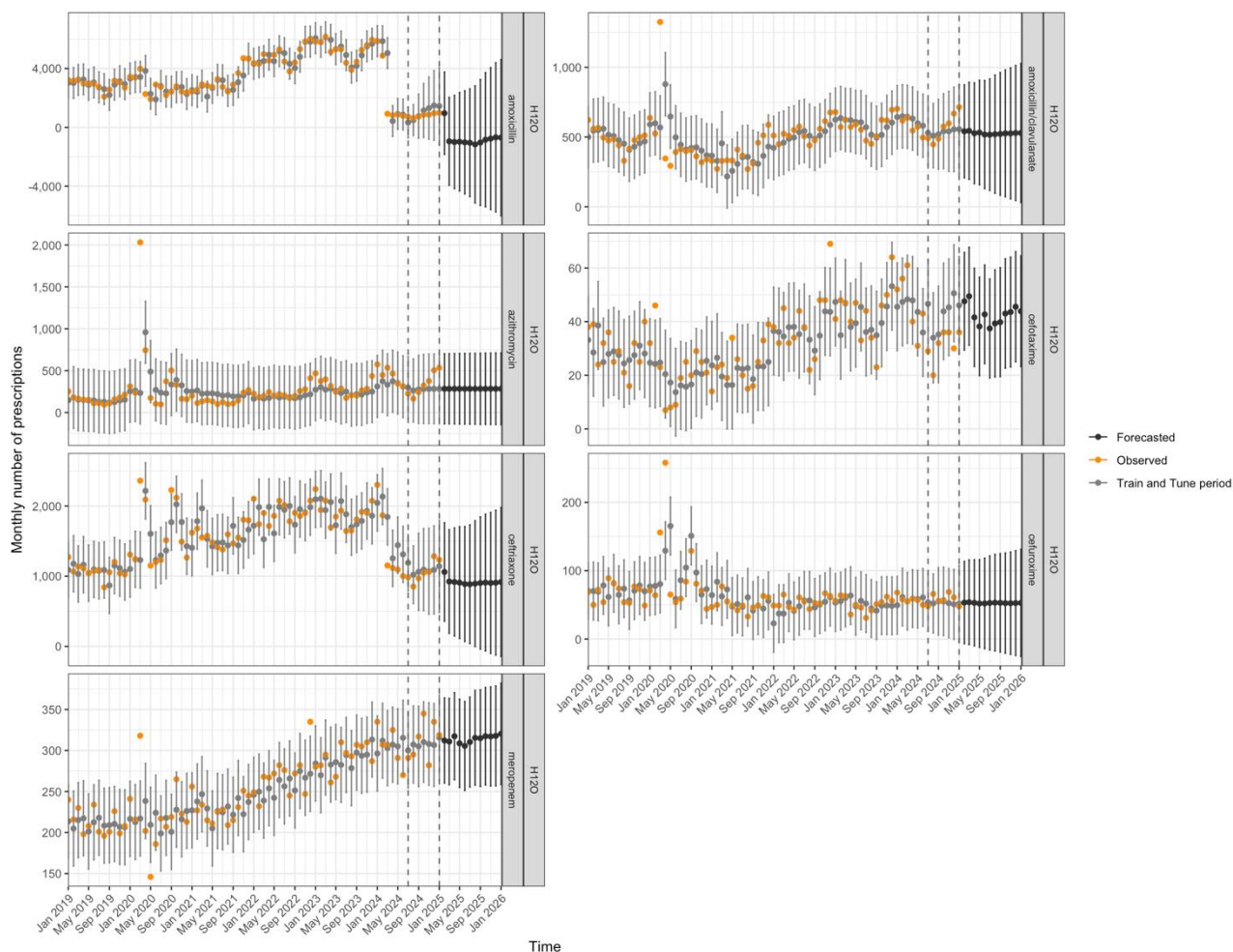


Figure 4. H120 models: forecasted monthly prescriptions/dispensations of the antibiotic of interest with ARIMAX (in black) vs observed values (in orange). The train and tune values from the ARIMAX are represented in grey. The vertical dotted lines divide time into three sections from left to right: the trained, tuned and forecasted periods. The ARIMAX values are presented with 95% confidence intervals.

IMASIS

Values and predictive intervals from the train and tune periods generally captured the observed number of monthly prescriptions/dispensations, with certain exceptions such as outliers around March 2020, and the discrepancy between 128 observed and 62 [95% CI: 19–106] forecasted benzylpenicillin prescriptions/dispensations in December 2022.

Regarding the forecasted values from July 2024 onwards, azithromycin and cefotaxime prescriptions/dispensations were projected to decrease (July 2024 vs June 2025: 230 [95%CI: -499 to 958] vs 104 [95%CI:

-700 to 908] for azithromycin, and 339 [95%CI: -9 to 668] vs 284 [95%CI: -221 to 789] for cefotaxime, respectively).

Amoxicillin, amoxicillin/clavulanate, benzylpenicillin and cefuroxime were projected to follow similar patterns to the previous observed year.

On the other hand, ceftriaxone and meropenem monthly prescriptions/dispensations were expected to remain stable (around 944 and 849, respectively). However, despite the observed values for ceftriaxone and meropenem were encompassed within the predictive intervals of the tune estimates, the tune values failed to reflect the fluctuations of the observed values. Ceftriaxone (ARIMAX(0,1,2)) and meropenem (ARIMAX(0,1,1)) exhibited non-seasonal models in IMASIS, which could not capture seasonality or autocorrelations.

Additionally, the fit for piperacillin/tazobactam during the tune period was poor, while observed prescriptions/dispensations increased (from 999 in December 2023 to values ranged 1148-1312 between January and May 2024), both the tune and forecasted values fluctuated between 900 and 1000 monthly prescriptions/dispensations. (Figure 5)

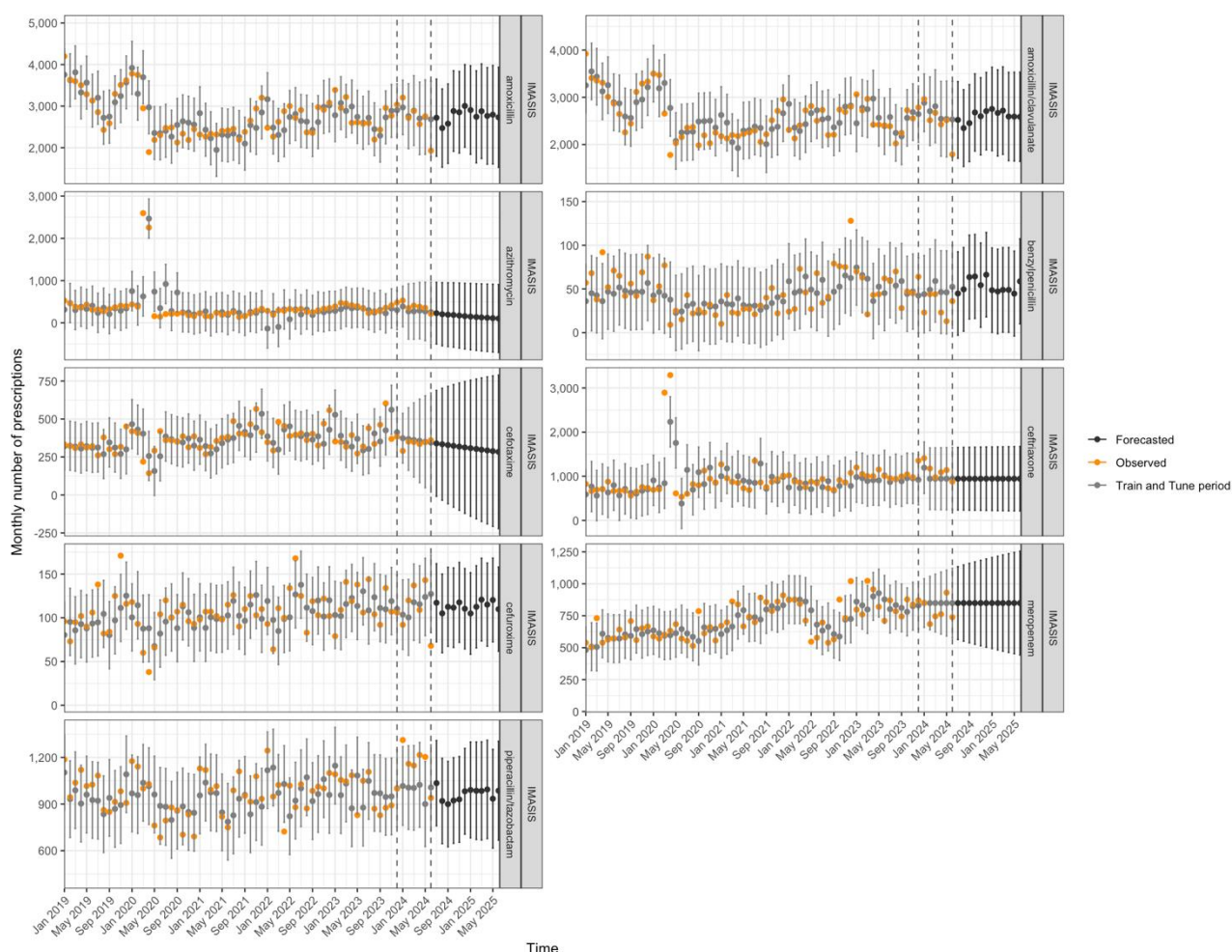


Figure 5. IMASIS models: forecasted monthly prescriptions/dispensations of the antibiotic of interest with ARIMAX (in black) vs observed values (in orange). The train and tune values from the ARIMAX are represented in grey. The vertical dotted lines divide time into three sections from left to right: the trained, tuned and forecasted periods. The ARIMAX values are presented with 95% confidence intervals.

InGef RDB

The model failed to adequately fit meropenem and benzylpenicillin, with forecasted values from March 2024 onwards remaining flat (around 16 and 133 monthly prescriptions/dispensations, respectively). Both exhibited non-seasonal models (ARIMAX(0,1,1)), which could not capture seasonality or autocorrelations. All other antibiotics presented a good fit during the train and tune periods, except for the last three months of the tuning period (July-September 2024) where some observed values were smaller than the predictive interval.

Regarding the forecasted values from October 2024 onwards, amoxicillin prescriptions/dispensations were expected to rise between October and December 2024 (from 83,283 [95%CI: 51,595-114,972] to 94,348 [95%CI: 61,016-127,679], respectively), followed by a decline through to August 2025 (63,878 [95%CI: 24,662-103,093]).

Observed prescriptions/dispensations of amoxicillin/clavulanate presented a bottom peak in August 2024 whilst tune values were stable (around 39,000 prescriptions/dispensations) and forecasted values increased (maximum forecasted value: 47,151 [95%CI: 34,146-60,157] prescriptions/dispensations in May 2025).

Forecasted values for azithromycin and cefuroxime were consistent with the trends observed in the previous year. Ceftriaxone prescriptions/dispensations were projected to increase, reaching up to 241 [95%CI: 174-308] in August 2025, whilst clarithromycin prescriptions/dispensations were projected to decrease from May 2025 onwards (e.g., prescriptions/dispensations observed in May 2024 vs forecasted in May 2025: 7025 vs 2966 [95%CI: -2567 to 8498], respectively). Phenoxymethylpenicillin was expected to fluctuate from October 2024 onwards, with values ranging between 8156 and 12,702 monthly prescriptions/dispensations. (**Figure 6**)

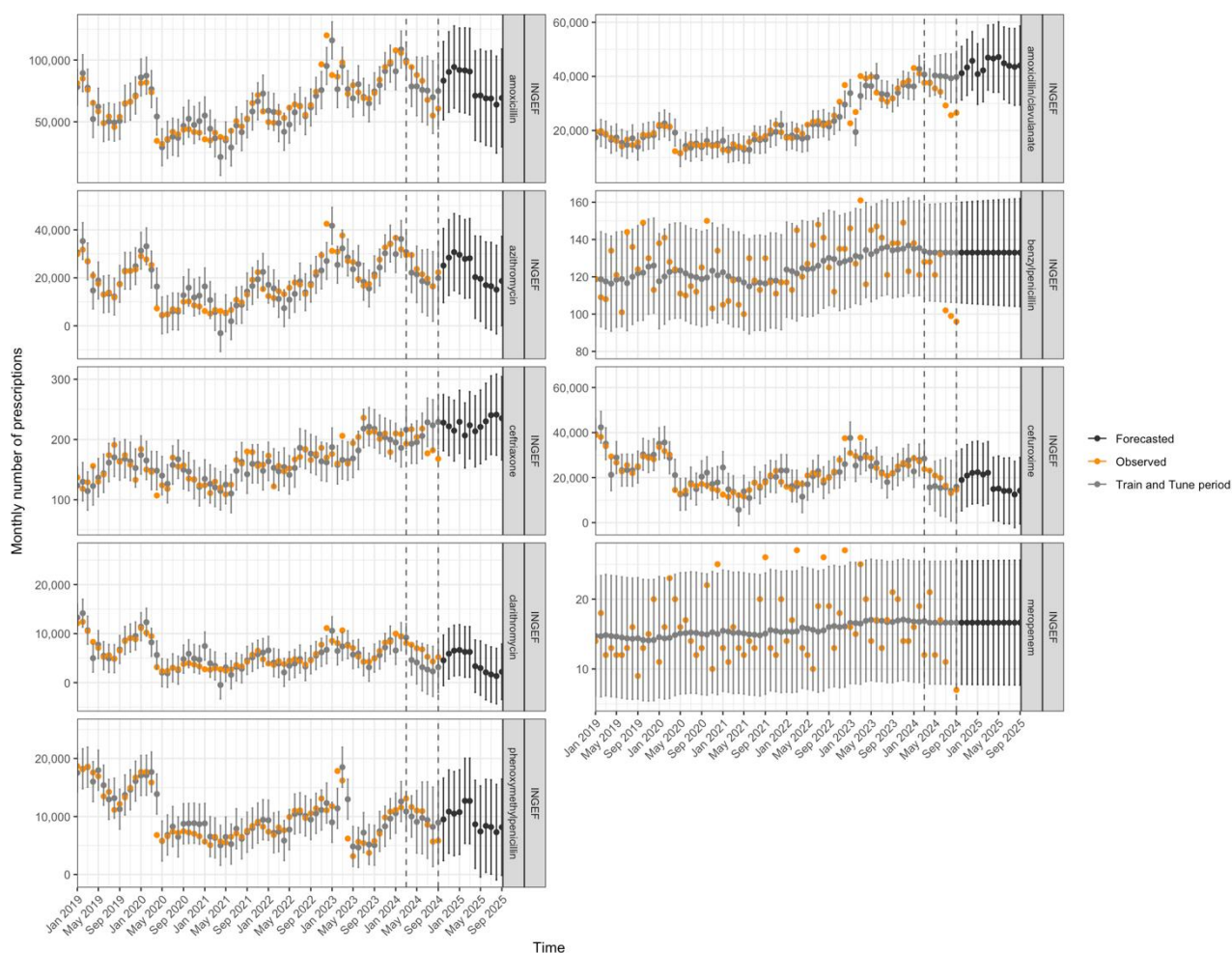


Figure 6. InGef RDB models: forecasted monthly prescriptions/dispensations of the antibiotic of interest with ARIMAX (in black) vs observed values (in orange). The train and tune values from the ARIMAX are represented in grey. The vertical dotted lines divide time into three sections from left to right: the trained, tuned and forecasted periods. The ARIMAX values are presented with 95% confidence intervals.

IPCI

Values and predictive intervals from the train and tune periods generally captured the observed number of monthly prescriptions with certain exceptions, such as outliers for ceftriaxone in September 2022 (observed vs trained value: 128 vs 102 [95%CI: 85-119], respectively) and in April 2024 (observed vs trained value: 72 vs 100 [95%CI: 82-119], respectively); or for azithromycin in February 2024 (observed vs trained value: 2818 vs 2028 [95%CI: 1472-2584]). However, the model performed poorly for cefuroxime (non-seasonal ARIMAX(0,1,1)), which exhibited a flat tune and forecasted value of 12 monthly prescriptions from January 2024 onwards.

Forecasted monthly prescriptions for amoxicillin, amoxicillin/clavulanate, azithromycin and clarithromycin followed similar patterns to those observed in the previous year. Forecasted monthly prescriptions for benzylpenicillin were projected around 34 [95%CI: 16-52] prescriptions from July 2024 onwards. Ceftriaxone use was expected to increase up to 119 [95%CI: 100-138] prescriptions in August 2024, and then moderately decrease, with a bottom peak at 109 [95%CI: 90-128] prescriptions in November 2024. (Figure 7)

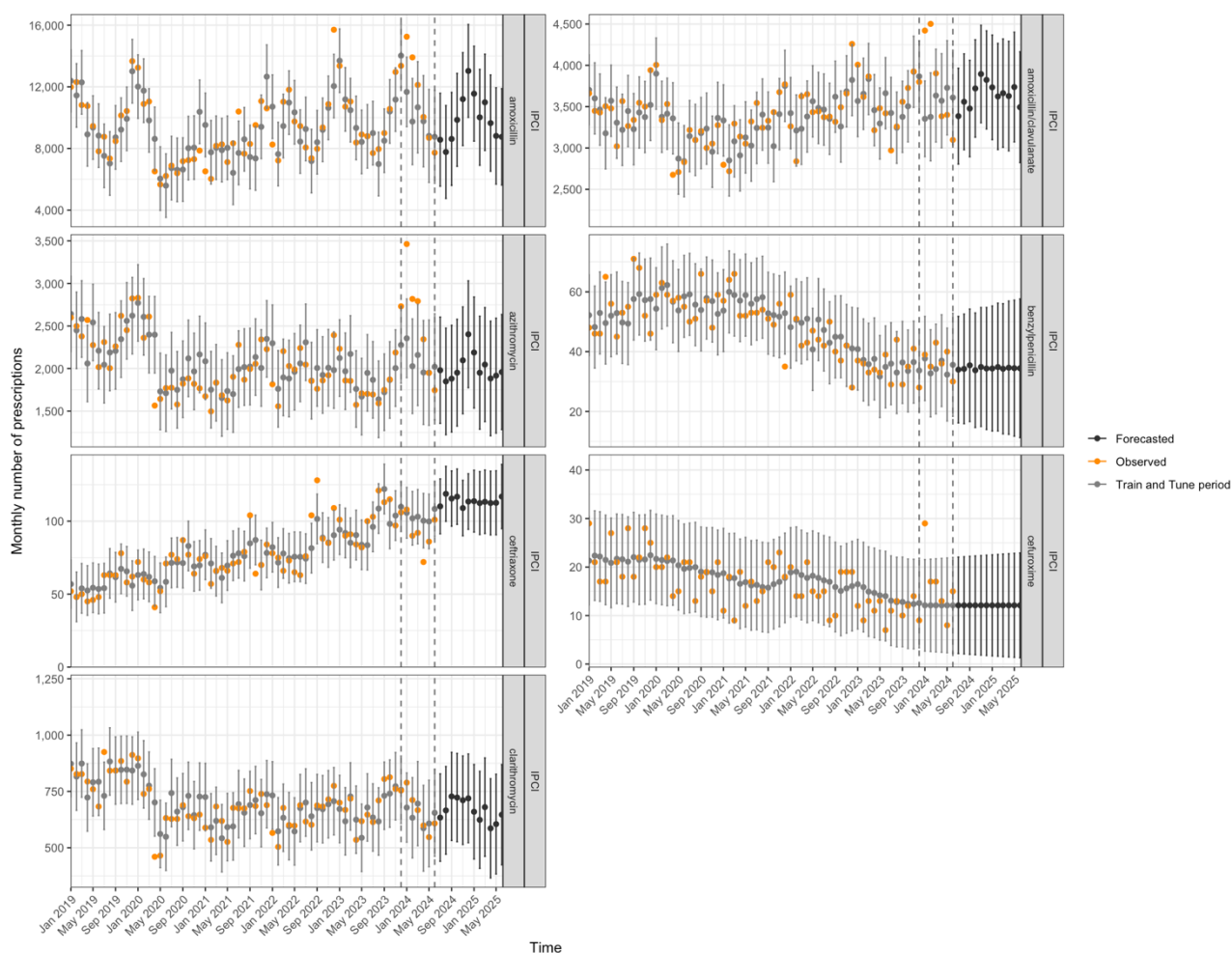


Figure 7. IPCI models: forecasted monthly prescriptions of the antibiotic of interest with ARIMAX (in black) vs observed values (in orange). The train and tune values from the ARIMAX are represented in grey. The vertical dotted lines divide time into three sections from left to right: the trained, tuned and forecasted periods. The ARIMAX values are presented with 95% confidence intervals.

IQVIA DA Germany

All observed monthly prescriptions/dispensations were captured within the predictive intervals from the train and tune periods. Forecasted monthly prescriptions/dispensations of amoxicillin and clarithromycin followed a similar pattern: an initial increase between October 2024 and March 2025, followed by a decline.

Amoxicillin/clavulanate was expected to present peaks of use in December 2024 and March 2025 (14,530 [95%CI: 10,173–18,888] and 14,544 [95%CI: 9,675–19,413] prescriptions/dispensations, respectively). Ceftriaxone prescriptions/dispensations were projected to increase from 87 [95%CI: 66–108] in October 2024 to 98 [95%CI: 73–123] in July 2025.

Cefuroxime use was expected to increase between December 2024 and March 2025 (monthly prescriptions/dispensations ranged between 11,449 and 12,952), followed by a decrease through to August 2025 (8106 [95%CI: -1501 to 17,713]).

Phenoxymethylpenicillin was projected to decrease between 3866 [95%CI: -268 to 8000] and 3305 [95%CI: -2553 to 9162] monthly prescriptions/dispensations from October 2024 to September 2025, respectively.

Conversely, the model performed poorly for benzylpenicillin (non-seasonal ARIMAX(0,1,1)), with tune and forecasted values from April 2024 onwards remaining flat at 53 monthly prescriptions/dispensations).

(Figure 8)

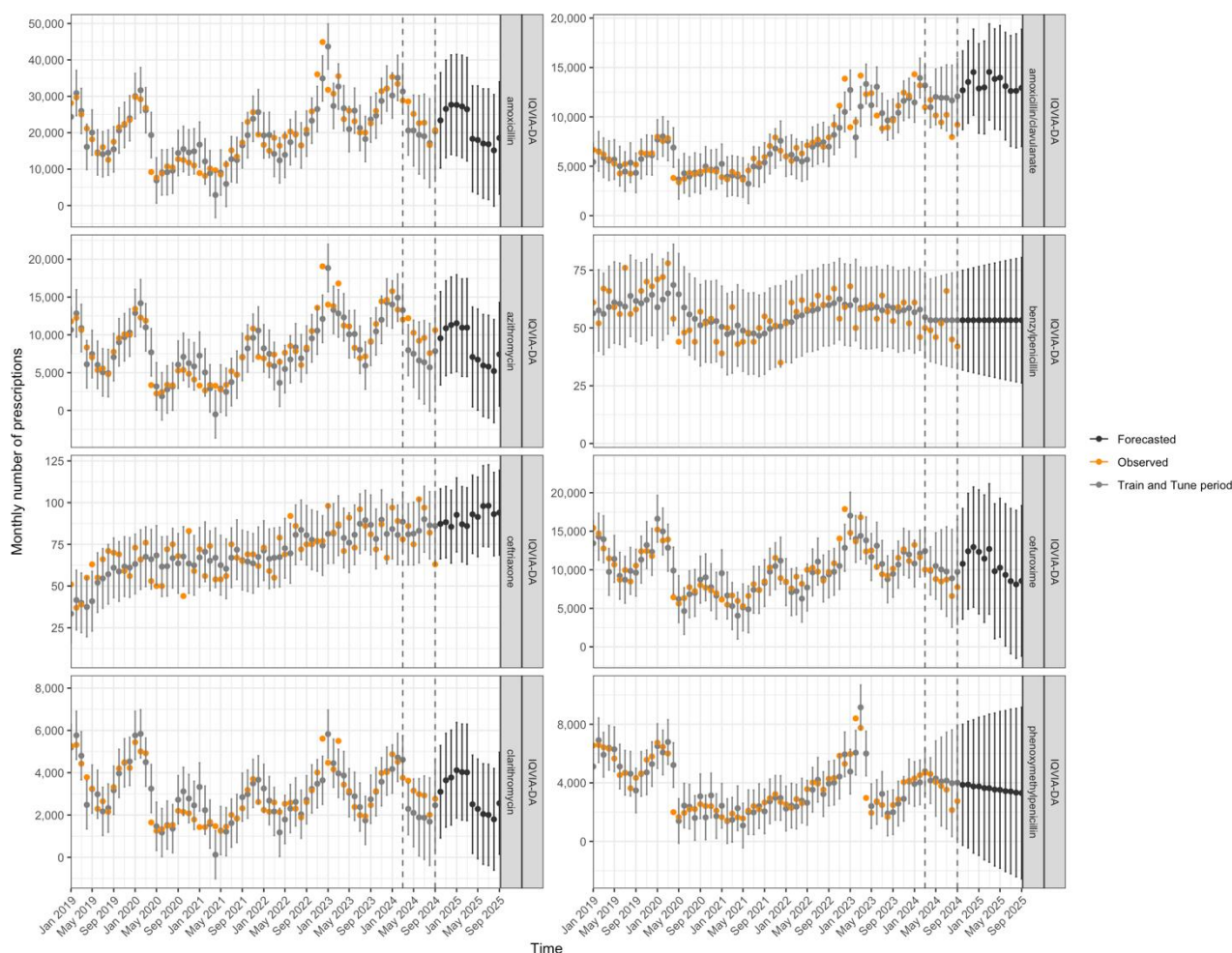


Figure 8. IQVIA DA Germany models: forecasted monthly prescriptions/dispensations of the antibiotic of interest with ARIMAX (in black) vs observed values (in orange). The train and tune values from the ARIMAX are represented in grey. The vertical dotted lines divide time into three sections from left to right: the trained, tuned and forecasted periods. The ARIMAX values are presented with 95% confidence intervals.

IQVIA LPD Belgium

Values and predictive intervals from the train and tune periods generally captured the observed number of monthly prescriptions with certain exceptions. Outliers were observed in September 2024 for amoxicillin (observed vs trained value: 1645 vs 4137 [95%CI: 2093-6180], respectively), amoxicillin/clavulanate (observed vs trained value: 788 vs 2040 [95%CI: 1226-2855], respectively) and azithromycin (observed vs trained value: 680 vs 1747 [95%CI: 1141-2352], respectively). Additionally, all antibiotics except cefuroxime presented an outlier in December 2022, with observed values falling below the predictive interval of the tune estimates (e.g., amoxicillin observed vs forecasted prescription: 9765 vs 6915 [95%CI: 5425-8406]).

All forecasted antibiotics projected similar patterns to the previous year, with an increase in use during the winter months and a decline in the summer months. (Figure 9)

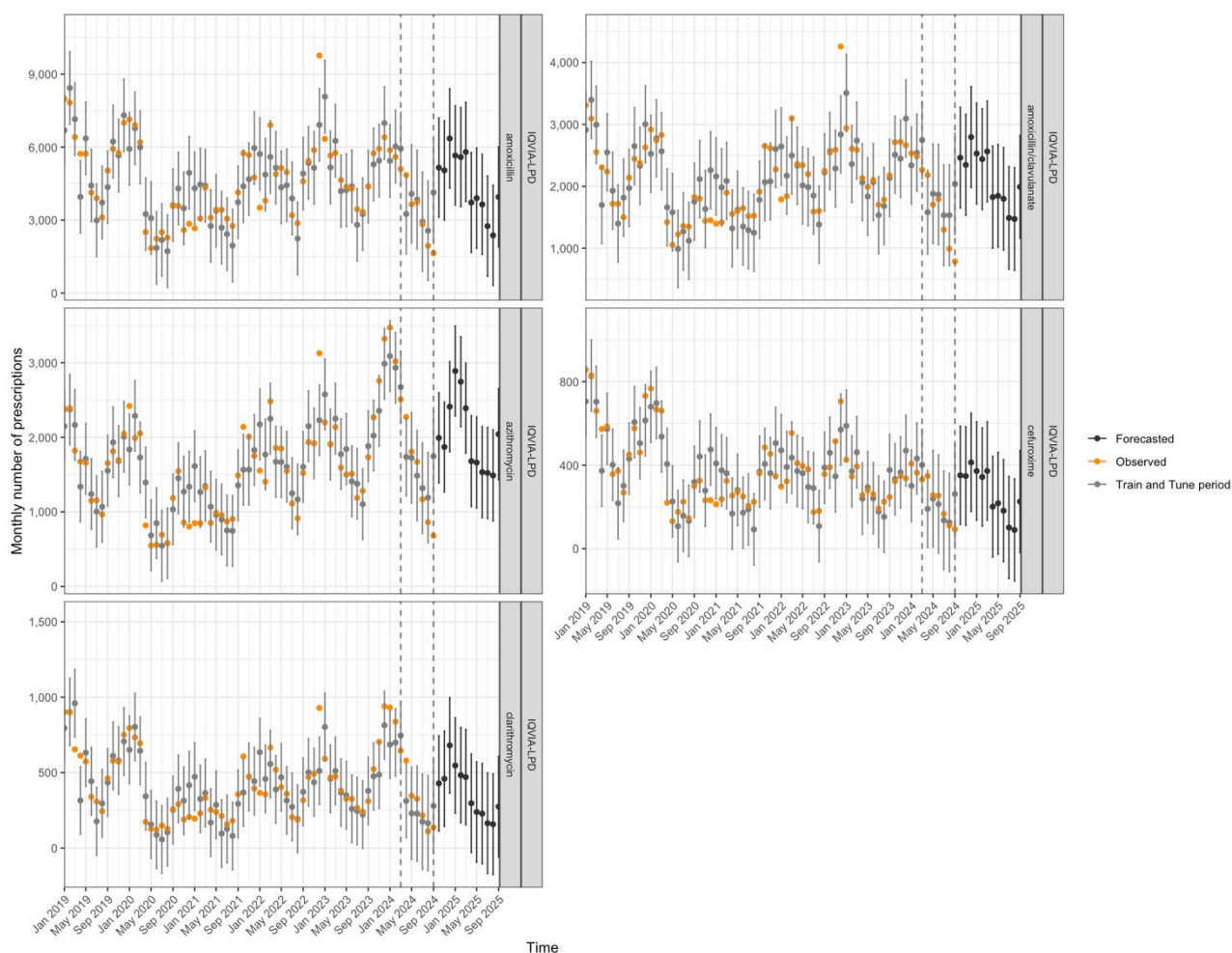


Figure 9. IQVIA LPD Belgium models: forecasted monthly prescriptions of the antibiotic of interest with ARIMAX (in black) vs observed values (in orange). The train and tune values from the ARIMAX are represented in grey. The vertical dotted lines divide time into three sections from left to right: the trained, tuned and forecasted periods. The ARIMAX values are presented with 95% confidence intervals.

14. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

Adverse events/adverse reactions will not be collected or analysed as part of this evaluation. The nature of this non-interventional evaluation, through the use of secondary data, does not fulfil the criteria for reporting adverse events, according to module VI, VI.C.1.2.1.2 of the Good Pharmacovigilance Practices (https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf).

Only in case of prospective data collection, there is a need to describe the procedures for the collection, management, and reporting of individual cases of adverse events/adverse reactions.

15. DISCUSSION

15.1. Key results

Amoxicillin

ARIMAX projected that amoxicillin use was going to maintain similar prescription/dispensation values – i.e., with fluctuations that followed either a non-seasonal trend or a seasonality pattern. Examining the models' fitting, we observed that hospitals (i.e., CDW Bordeaux, IMASIS and H12O) exhibited a linear non-seasonal trend ($d=1$). Conversely, the other four databases, which included primary care data were the ones that exhibited a seasonal trend ($D=1$). Hence, hospitals were more likely to maintain steady trends over the year, whilst primary care were more likely to present seasonal outbreaks.

However, we need to highlight that H12O projected values below 0 after March 2025. This might have been led by the substantial drop in the observed use of amoxicillin between February and March 2024 (from 4876 to 924 prescriptions/dispensations, respectively) likely caused by an incomplete recording of the unstructured sources. In H12O, high 95% predictive intervals between March 2025 and January 2026 ranged 2042-4616 monthly prescriptions/dispensations.

Amoxicillin/clavulanate

Model fitting in six out of the seven databases presented a linear non-seasonal trend ($d=1$), except for IQVIA LPD Belgium that instead presented a seasonal trend ($D=1$). Amoxicillin/clavulanate was projected to increase in InGef and IQVIA DA Germany; to maintain stable values at 522 monthly prescriptions/dispensations in H12O from July 2024 onwards; to display similar prescription/dispensation values, following a linear non-seasonal pattern in IMASIS and IPCI, or following a seasonality pattern in IQVIA LPD Belgium; and to decrease in CDW Bordeaux.

Azithromycin

Azithromycin use was projected to increase in CDW Bordeaux; to maintain values around 280 in H12O from October 2024 onwards; and to decrease in IMASIS. Model fitting in these hospitals had no seasonal autocorrelation (no P, D, Q terms). Additionally, CDW Bordeaux displayed no autoregressive terms ($p=0$).

On the other hand, azithromycin use was projected to maintain similar prescription/dispensation values following a non-seasonality pattern ($d=1$) in IPCI, and following a seasonality pattern ($D=1$) in InGef, IQVIA DA Germany and IQVIA LPD Belgium.

Benzylpenicillin

Benzylpenicillin use was projected to maintain similar prescription/dispensation values, following a non-seasonality pattern ($d=1$) in IMASIS.

Benzylpenicillin use was projected to maintain prescription values as those observed from March 2024 onwards in InGef (around 133) and from May 2023 onwards (around 31-38) in IPCI. Moreover, tune and forecasted values in IQVIA DA Germany became flat from April 2024 onwards, with 53 monthly prescriptions/dispensations. In InGef, IPCI and IQVIA DA Germany, which included primary care data, model fitting displayed no seasonal autocorrelation (no P,D,Q terms), only a linear non-seasonal trend ($d=1$). Moreover, InGef and IQVIA DA Germany models displayed no autoregressive terms ($p=0$).

Cefotaxime

Models for cefotaxime use were only available in IMASIS and H12O databases.

Monthly prescriptions/dispensations of cefotaxime in IMASIS were projected to decrease (from the last observed value of 360 prescriptions/dispensations in June 2024 to 284 expected prescriptions/dispensations in June 2025). Prior to January 2024, observed use seemed to exhibit a seasonal pattern, but the model estimates lost any fluctuations after January 2024. Model fitting of cefotaxime in IMASIS was stationary ($d=0$) and had no seasonal autocorrelation (no P,D,Q terms).

Conversely, projected monthly prescriptions/dispensations of cefotaxime in H12O displayed fluctuations that suggested a similar pattern of use in 2025 to that observed in 2023 and 2024 (with the exception of the observed values during November 2024 and January 2025). Model fitting of cefotaxime in H12O exhibited a linear non-seasonal trend ($d=1$).

Ceftriaxone

Ceftriaxone use was projected to increase in InGef, IPCI and IQVIA DA Germany; to maintain prescriptions/dispensations values around 944 in IMASIS and around 900 in H12O from March 2024 onwards; and to decrease in CDW Bordeaux. All models' fitting exhibited a non-seasonality trend ($d=1$). Additionally, IMASIS had no autoregressive terms ($p=0$) and no seasonal autocorrelation (no P,D,Q terms).

Cefuroxime

Cefuroxime use was projected to maintain similar prescription/dispensation values to those of the observed monthly prescriptions/dispensations from 2020 onwards in H12O, IMASIS, InGef, IQVIA DA Germany and IQVIA LPD Belgium. IPCI projected a flat number of 12 monthly prescriptions since January 2024.

InGef and IQVIA LPD Belgium models followed a seasonal trend ($D=1$), H12O and IMASIS followed non-seasonal trend ($d=1$), whilst IQVIA DA Germany was stationary ($d=0$). IPCI's model fitting exhibited a linear pattern ($d=1$) but no autoregressive terms ($p=0$). Additionally, IPCI and H12O had no seasonal autocorrelation (no P,D,Q terms).

Clarithromycin

Clarithromycin use was projected to maintain similar prescription/dispensation values to those of the previous year/s, following a linear non-seasonality pattern ($d=1$) in IPCI, and following the seasonality pattern in IQVIA DA Germany and IQVIA LPD Belgium ($D=1$); and to decrease in CDW Bordeaux and InGef. InGef exhibited a seasonality pattern ($D=1$) but the CDW Bordeaux model was stationary ($d=0$).

Meropenem

Meropenem use was projected to display similar fluctuations in the prescription values of CDW Bordeaux to those observed from 2020 onwards, and to maintain monthly prescriptions/dispensations around 315 in H12O. Tune and forecasted values in IMASIS became flat from January 2024 onwards, with 849 monthly prescriptions/dispensations. Similarly, InGef values became flat from March 2024 onwards, with 16 monthly prescriptions/dispensations.

Model fitting of meropenem in these four databases exhibited a linear non-seasonality pattern ($d=1$), and all except H12O had no autoregressive terms ($p=0$). Additionally, IMASIS and InGef had no seasonal autocorrelation (no P,D,Q terms).

Phenoxymethylpenicillin

Models for phenoxymethylpenicillin use were only available in CDW Bordeaux, InGef and IQVIA DA Germany. Phenoxymethylpenicillin tune and forecasted estimates were projected to decrease in CDW Bordeaux from June 2024, and in IQVIA DA Germany from October 2024. Forecasted monthly prescription/dispensation values in InGef fluctuated with values similar to the observed values from 2020 onwards.

Model fitting of phenoxymethylpenicillin in InGef exhibited a linear non-seasonality pattern ($d=1$). Conversely, CDW Bordeaux and IQVIA DA Germany fitting was stationary ($d=0$) and had no seasonal autocorrelation (no P,D,Q terms).

Piperacillin/tazobactam

Models for piperacillin/tazobactam use were only available in CDW Bordeaux and IMASIS. Piperacillin/tazobactam use was projected to maintain similar prescription/dispensation values in IMASIS (between 900 and 1000), and to decrease in CDW Bordeaux. Model fitting exhibited a linear non-seasonality pattern ($d=1$) in both databases.

15.2. Limitations of the research methods

15.2.1. Model diagnostics and limitations:

Based on the model diagnostics, ARIMAX presented a marginally better fit than ARIMA models. ARIMAX models demonstrate adequate levels of goodness of fit, which validates the reliability of their respective coefficients. A number of observations were out of the 95% predictive intervals of the train and tune values but the majority of the ARIMAX models presented a good fitting and captured a linear trend or a seasonality pattern of the antibiotics typically used for bacterial infections frequent in winter, such as amoxicillin or azithromycin.

However, there were seven models out of the 53 executed that had autoregressive terms ($p = 0$) and could not capture seasonal autocorrelations (no P, D, Q). Those were azithromycin in CDW Bordeaux, ceftriaxone and meropenem in IMASIS, benzylpenicillin and meropenem in InGef, cefuroxime in IPCI and benzylpenicillin in IQVIA DA Germany.

In some other cases, the models presented a good fitting during the train period but presented discrepancies between the tune and observed values, with conflicting trends. For instance, amoxicillin/clavulanate observed use during the tune period in InGef and IQVIA DA Germany decreased (down to 25,636 and 7957 prescriptions/dispensations in August 2024, respectively) whilst the models projected an increase (up to 39,292 prescriptions/dispensations in May 2025 and 11,665 prescriptions/dispensations in March 2025, respectively). Similarly, observed prescriptions/dispensations of cefotaxime in H12O during October 2024 and January 2025 were lower than the forecasted estimates.

On the other hand, prescription/dispensation trends during the last six observed months of azithromycin decreased in CDW Bordeaux (down to 130 prescriptions/dispensations in November 2024) and increased in H12O (up to 575 prescriptions/dispensations in January 2025), whilst the model projected a steady increase for CDW Bordeaux (up to 195 prescriptions/dispensations in November 2025) and a flat use for H12O (around 280 monthly prescriptions/dispensations). Similarly, last observed monthly prescriptions/dispensations for piperacillin/tazobactam in IMASIS were higher than expected: observed values ranged around 1148-1312, whereas tune and forecasted values ranged around 900-1000 monthly prescriptions/dispensations.

15.2.2. General limitations:

The study was informed by routinely collected health care data, so data quality issues had to be considered. In particular, a recording of a prescription or dispensation did not mean that the patient actually took the drug.

Additionally, forecasting with ARIMA(X) models had several limitations. ARIMA(X) assumes the homoscedasticity of the data (i.e., variance is constant over time), which might not have held even after integration of the time series. We assessed the residuals from the ARIMA(X) models to ensure the assumption of homoscedasticity was reasonably met. Given that each drug of interest for each country needed its own ARIMA model, model parameters were initially estimated automatically using the function "auto.arima" from the R package forecast and fine-tuned based on diagnostics. Additionally, we might not have been able to capture all the exogenous variables in the models (e.g., having unexpected events leading to abnormal observations), which might have impacted model fitting. For example, the COVID-19 pandemic might have been an example of disruption that negatively impacted estimation and forecasting.

15.2.3. Database-specific limitations:

The observed substantial drop in observed prescriptions/dispensations of amoxicillin and ceftriaxone in H120 between February and March 2024 was likely led by an incomplete recording from unstructured sources such as doctors' clinical notes in the EHR, which are obtained from running NLP algorithms based on internal projects. The monthly counts observed from March 2024 onwards matched the recorded counts from the structured prescription data from the hospital. Amoxicillin and ceftriaxone models in H120 need to be interpreted with caution.

Moreover, an unknown number of individuals were present in both the InGef RDB and IQVIA DA Germany databases. We expected that, since IQVIA DA Germany captured a larger proportion of the German population, it would have provided a broader and more representative coverage of prescriptions and dispensations in the German primary care, whereas InGef RDB was expected to be particularly relevant for medications administered in hospital settings, which were not captured by IQVIA DA Germany. For instance, meropenem use could be modelled in InGef whilst sample size was insufficient to model it in IQVIA DA Germany. Most of the trends of use for the antibiotics of interest modelled in both databases were aligned, with the exception of clarithromycin and phenoxymethylpenicillin use. Moreover, the model fitting exhibited different model parameters for other antibiotics, such as cefuroxime. In other words, the model identified different trends and patterns across the two German databases. This could be interpreted as both databases being representative of the current practice in Germany, or that both databases have a significant overlap that can make them redundant.

15.3. Interpretation

ARIMA(X) models are well-suited to modelling seasonal patterns, such as the peaks in antibiotic use typically observed during winter months, which are associated with the increased incidence of respiratory and other winter infections.(3, 4) In this context, ARIMA(X) models have been widely used to analyse temporal trends and seasonality in antibiotic use, as well as to explore its relationship with antimicrobial resistance.(5-7) Additionally, there have also been attempts to use ARIMA(X) models to forecast short-term antibiotic use (i.e., in the upcoming 4 to 6 months) (8) and long-term trends (9), including our prior Darwin study P2-C1-012 - "Monitoring prescription of medicines for public health emergencies at risk of shortages".(10)

In P2-C1-012, we forecasted the incidence rates of use of the same 11 antibiotics of interest in five databases, including IMASIS and both IQVIA databases. A limitation of this previous study was the presence of artifacts in the denominator population of IMASIS and IQVIA databases that overestimated the incidence rates during the last six months of data availability. To overcome this limitation, we designed the current study to forecast the use of the monthly prescriptions/dispensations, thereby circumventing issues related

to denominator estimation. This decision improved the fitting of the observed and tune estimates but required larger model parameters to achieve an adequate goodness of fit.

In line with the recommendations of *Xie et al.*,⁽⁸⁾ our previous work found that six-month forecasts were more reliable than 12-month forecasts. Consistent with this, we expected the reliability of our current forecasts to diminish with increasing temporal distance from the most recent observed data. This was reflected in the widening of prediction intervals over time in a number of models, such as those for clarithromycin in the CDW Bordeaux, amoxicillin alone or in combination with clavulanate in the H12O, and cefotaxime in IMASIS, among others.

Generally, spikes in time series analysis may be explained by exogenous variables not captured in the model. While ARIMA(X) models could not capture these extreme observations, identification of spikes can still be informative, for instance, by providing estimates of increases in demand that can occur naturally.⁽¹¹⁾ Nonetheless, the forecasted monthly prescriptions/dispensations can be a useful tool to anticipate a minimum stock of the different antibiotics of interest necessary to cover the study population needs during the next 6 to 12 months.

15.4. Generalisability

The study included all individuals at risk of using the antibiotics of interest present in 7 databases from 5 different European countries. These databases covered a varied range of health care settings to observe the medications that, due to their indications, may be administered in primary or secondary care: CDW Bordeaux, H12O and IMASIS as hospital data; IPCI and IQVIA LPD Belgium as primary care databases; InGef containing primary care and hospital data; and IQVIA Germany containing primary care and outpatient care data.

We anticipated that trends in antibiotic use were likely to vary across the databases of the study, reflecting not only different regional differences in infection patterns but also the influence of local health policies and varying access to antibiotics. In other words, the monthly prescriptions for a particular antibiotic could have increased in one database while decreasing in another. For example, we observed different trends in the observed and projected use of azithromycin and cefotaxime among the two Spanish hospitals, H12O and IMASIS. Hence each antibiotic model was trained independently for each database to predict its specific trend behaviour. Therefore, while we consider the results representative for the study population in the respective countries or specific regions, the results should not be generalised to other countries or regions. Instead, they should be understood as reflecting the situation in the areas and settings covered by the respective databases.

16. CONCLUSION

We projected an increase in the use of amoxicillin/clavulanate in InGef and IQVIA DA Germany but a decrease in CDW Bordeaux, an increase in azithromycin use in CDW Bordeaux, a decrease in cefotaxime use in IMASIS, an increase in ceftriaxone use in InGef, IPCI and IQVIA DA Germany but a decrease in CDW Bordeaux, a decrease in clarithromycin use in CDW Bordeaux and InGef, a decrease in phenoxymethylpenicillin use in CDW Bordeaux and in IQVIA DA Germany, and a decrease in piperacillin/tazobactam use in CDW Bordeaux.

Therefore, the French hospital CDW Bordeaux is projected to reduce its use of antibiotics in 6 out of the 8 antibiotics forecasted. Furthermore, databases from the same country did not necessarily exhibit the same trends of use. For instance, the two Spanish hospitals projected different trends of use for cefotaxime, and the two German databases projected different trends of use for clarithromycin and phenoxymethylpenicillin.

Azithromycin use presented a seasonality pattern in most of the primary care databases. Similarly, amoxicillin use presented a seasonality pattern in primary care databases whilst hospitals exhibited a linear

non-seasonal pattern. Conversely, amoxicillin/clavulanate use only presented a seasonality pattern in one of the primary care databases. No apparent patterns were observed when comparing hospital and primary care seasonality trends for the remaining antibiotics. These generally exhibited a linear non-seasonal pattern, and few events exhibited a stationary trend. Additionally, azithromycin in CDW Bordeaux, ceftriaxone and meropenem in IMASIS, benzylpenicillin and meropenem in InGef, cefuroxime in IPCI and benzylpenicillin in IQVIA DA Germany models failed to capture seasonality and autocorrelation.

The forecasted monthly prescriptions/dispensations can be a useful tool to anticipate a minimum stock of the different antibiotics of interest necessary to cover the study population needs during the next 6 to 12 months.

17. REFERENCES

1. Aronson JK, Heneghan C, Ferner RE. Drug shortages. Part 1. Definitions and harms. *Br J Clin Pharmacol*. 2023;89(10):2950-6.
2. Reyes C, Pistillo A, Fernandez-Bertolin S, Recalde M, Roel E, Puente D, et al. Characteristics and outcomes of patients with COVID-19 with and without prevalent hypertension: a multinational cohort study. *BMJ Open*. 2021;11(12):e057632.
3. Elseviers MM, Ferech M, Vander Stichele RH, Goossens H, group Ep. Antibiotic use in ambulatory care in Europe (ESAC data 1997-2002): trends, regional differences and seasonal fluctuations. *Pharmacoepidemiol Drug Saf*. 2007;16(1):115-23.
4. Goossens H, Ferech M, Vander Stichele R, Elseviers M, Group EP. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. *Lancet*. 2005;365(9459):579-87.
5. Martinez EP, van Rosmalen J, Bustillos R, Natsch S, Mouton JW, Verbon A. Trends, seasonality and the association between outpatient antibiotic use and antimicrobial resistance among urinary bacteria in the Netherlands. *J Antimicrob Chemother*. 2020;75(8):2314-25.
6. Trinh NTH, Chalumeau M, Bruckner TA, Levy C, Bessou A, Milic D, et al. Monitoring outpatient antibiotic utilization using reimbursement and retail sales data: a population-based comparison in France, 2012-17. *J Antimicrob Chemother*. 2021;76(9):2446-52.
7. Zeng S, Xu Z, Wang X, Liu W, Qian L, Chen X, et al. Time series analysis of antibacterial usage and bacterial resistance in China: observations from a tertiary hospital from 2014 to 2018. *Infect Drug Resist*. 2019;12:2683-91.
8. Xie J, Kanghuai Z, Cai Y, Wang N, Wang H, Zhao Y, et al. Application of ARIMA model in monitoring the use rate of antibiotics in outpatients in 2010 - 2018 . *Int J Clin Pharmacol Ther*. 2020;58(5):282-8.
9. Bindel LJ, Seifert R. Long-term forecast for antibacterial drug consumption in Germany using ARIMA models. *Naunyn Schmiedeberg's Arch Pharmacol*. 2025.
10. EMA | RWD Catalogues. DARWIN EU® – Monitoring prescription of medicines for public health emergencies at risk of shortages (EUPAS1000000473) 2025 [Available from: <https://catalogues.ema.europa.eu/node/4374/administrative-details>].
11. Goin DE, Ahern J. Identification of Spikes in Time Series. *Epidemiologic Methods*. 2019;8(1).

18. ANNEXES

Appendix I: Preliminary code lists for the antibiotics identified as potentially critical in PHE

Antibiotic	Class	ATC code	Ingredient Concept ID
Amoxicillin	Penicillins with extended spectrum	J01CA04	1713332
Amoxicillin/clavulanate (Amoxicillin + Clavulanic acid)	Combinations of penicillins	J01CR02	1713332 + 1759842
Azithromycin	Macrolides	J01FA10	1734104
Benzylpenicillin (Penicillin G)	Beta-lactamase sensitive penicillins	J01CE01	1728416
Clarithromycin	Macrolides	J01FA09	1750500
Cefotaxime	Third-generation cephalosporins	J01DD01	1774470
Ceftriaxone	Third-generation cephalosporins	J01DD04	1777806
Cefuroxime	Second-generation cephalosporins	J01DC02	1778162
Meropenem	Carbapenems	J01DH02	1709170
Phenoxymethylpenicillin (Penicillin V)	Beta-lactamase sensitive penicillins	J01CE02	1729720
Piperacillin + Tazobactam	Combinations of penicillins	J01CR05	1746114 + 1741122

ATC = Anatomical Therapeutic Chemical