



EHDEN

EUROPEAN HEALTH DATA & EVIDENCE NETWORK


806968 – EHDEN

European Health Data & Evidence Network

WP1 – Evidence Workflow Development

D1.3 Implementation of Risk Minimisation Measures Effectiveness Testing Methods

Lead contributors	Sara Khalid (UOXF), Anthony Sena (EMC)
Lead contributor email	sara.khalid@ndorms.ox.ac.uk , a.sena@erasmusmc.nl
Other contributors	Daniel Prieto-Alhambra (UOXF), daniel.prieto-alhambra@ndorms.ox.ac.uk UOXF, EMC, NICE, UMC, PFIZER, ABBVIE, SARD, Bayer, AZ, CELGENE
Due date	M26
Delivery date	22/09/2021
Deliverable type	OTHER
Dissemination level	PU
DoA - Version	V1
Date	12/11/2018

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU 2/20



LIST OF ABBREVIATIONS

Abbreviation	Full meaning
CDM	Common Data Model
DDD	Defined Daily Dose
DUS	Drug Utilisation
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EHR	Electronic Health Record
EHDEN	European Health Data and Evidence Network
EMIF	European Medical Information Framework
EU	European Union
ISPE	International Society of Pharmaco-Epidemiology
mMPR	modified Medication Possession Ratio
PDC	Proportion Days Covered
PDD	Prescribed Daily dose
RRE	Remote Research Environment
RMM	Risk minimisation measure
CPD	Change point detection
DXM	Dexamethasone
MTX	Methotrexate
HCQ	Hydroxychloroquine



	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU 3/20




Table of Contents

List of Abbreviations	2
Document History	5
Definitions	6
Publishable Summary	7
1. Introduction	8
2. Aims and Objectives	8
2.1 Primary Objectives.....	8
3. Methods	9
3.1 Study Design	9
3.2 Study Population	9
3.3 Data Sources.....	9
3.4 Variables	9
3.4.1 Diagnosis	9
3.4.2 Calendar time.....	9
3.4.3 Indication of use.....	9
3.5 Sample Size	9
4. Data Management and Resourcing Needs	9
5. Data Analysis	10
5.1 Analysis Pipeline	10
5.1.1 Processing Pipeline	10
5.1.2 Results Sharing Pipeline	11
5.2 Cohort Generation.....	12
5.3 Statistical Analysis	12
5.3.1 Change Point Detection	12
5.3.2 Regression-based methods.....	13
5.3.3 Bayesian Online Change point detection	13
5.3.4 Settings	13
6. Tool Prototype	14
7. Use-cases	15
Example 1: Changes in use of HCQ in RA patients over time.....	16
Example 2: HCQ for covid-19 in 2020.....	16
Example 3: Dexamethasone for covid-19 in 2020.....	17
8. Discussion	18
9. Conclusions and future work	19
10. Quality Control	19
11. Limitations of Research Methods	19
12. Protection of Human Subjects	19
a. Ethical / Data Utilisation Approval and Subject Consent	19
b. Subject Confidentiality	19
c. Report of Adverse Events	20

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development	Version: v3.0 – Final	
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra	Security: PU	4/20




d. EHDEN Data Sharing Agreement	20
13. Plans for Disseminating Study Results	20
14. References	20

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development	Version: v3.0 – Final	
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra	Security: PU	5/20



DOCUMENT HISTORY

Version	Date	Description
V1	15 April 2021	First Draft
V1	15 June 2021	Comments of Formal Review
V2	01 September 2021	Draft for Consortium Review
V3.0	22 September 2021	Final Version To be Submitted to IMI

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU 6/20




DEFINITIONS

Participants of the EHDEN Consortium are referred to herein according to the following codes:

EMC	Erasmus Universitair Medisch Centrum Rotterdam- The Netherlands (Project Coordinator)
Synapse	Synapse Research Management Partners S.L. - Spain
UOXF	The Chancellor, Masters and Scholars of the University of Oxford - United Kingdom
UTARTU	Tartu Ulikool - Estonia
UAVR	Universidade de Aveiro – Portugal
The Hyve	The Hyve BV – the Netherlands
Odysseus	Odysseus Data Services SRO – Czech Republic
EPF	Forum Europeen des Patients (FPE) - Belgium
NICE	National Institute for Health and Care Excellence – United Kingdom
UMC	Stiftelsen WHO Collaborating Centre for International Drug Monitoring - Sweden
ICHOM	International Consortium for Health Outcomes measurement LTD - United Kingdom
Janssen	Janssen Pharmaceutica NV - Belgium (Project Lead)
Pfizer	Pfizer Limited – United Kingdom
Abbvie	AbbVie Inc - United States
IRIS	Institut De Recherches Internationales Servier - France
SARD	Sanofi Aventis Recherche & Developpement - France
Bayer	Bayer Aktiengesellschaft - Germany
Lilly	Eli Lilly and Company Limited – United Kingdom
AZ	AstraZeneca AB - Sweden
Novartis	Novartis Pharma AG - Switzerland
UCB	UCB Biopharma SPRL - Belgium
Celgene	Celgene Management SARL - Switzerland

Grant agreement	The agreement signed between the beneficiaries and the IMI JU for the undertaking of the EHDEN project (806968).
Project	The sum of all activities carried out in the framework of the Grant Agreement.
Consortium	The EHDEN Consortium, comprising the above-mentioned legal entities.
Consortium agreement	Agreement concluded amongst EHDEN participants for the implementation of the Grant Agreement. Such an agreement shall not affect the parties' obligations to the Community and/or to one another arising from the Grant Agreement.

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development	Version: v3.0 – Final	
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra	Security: PU	7/20




PUBLISHABLE SUMMARY

Once a drug has been approved for use, regulatory authorities can introduce risk minimisation measures (RMM) to ensure and improve the safe use of a product. Typical examples of RMM include changes in the conditions of marketing authorisation and/or changes in the conditions of use or label characteristics. Evaluating the impact of such regulatory measures is needed to measure immediate as well as long-term effects on drug uptake and use in the community.

Observational data are routinely used to assess the effectiveness of such RMMs. However, a recent study by the EMA (1) found that registries that are commonly required for post-marketing purposes frequently experience delays. Making use of existing multinational European data sources through EHDEN has the potential to speed data processing because of the standardization, hence accelerating such studies.

In this work, the prototype of an RMM effectiveness testing tool has been developed. It enables users to apply a range of time-series change detection methods to a) determine the effect of known interventions/changes, and b) to detect plausible changes, in a data-driven manner. The tool is available in the form of a user-friendly, interactive web application. Users, such as regulators, can select the time-series, methods, and interventions of interest and view the results of the RMM test in real-time. A demo is included showing the changes over time in the use of drug for covid-19 treatment.

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU



1. INTRODUCTION

Regulators (e.g. the European Medicines Agency) can require marketing authorisation holders to implement risk minimisation measures (RMM) in order to ensure and improve the safe use of a product. Observational data are routinely used to assess the effectiveness of such RMMs.

However, a recent study by the EMA (1) found that registries that are commonly required for post-marketing purposes frequently experience delays. Making use of existing observational data sources standardized to a common data model can speed up response time to regulatory requests. This is an area where multinational European data are needed, and where EHDEN tools could be of great use to enable rapid, reproducible evidence generation based on real-world data.

In discussion with relevant stakeholders, key proof-of-concept studies in this area were proposed. Feedback suggested the inclusion of a study to develop tools for the formal testing of RMM effectiveness, including interrupted time series methods, to be implemented in EHDEN in collaboration with WP4.

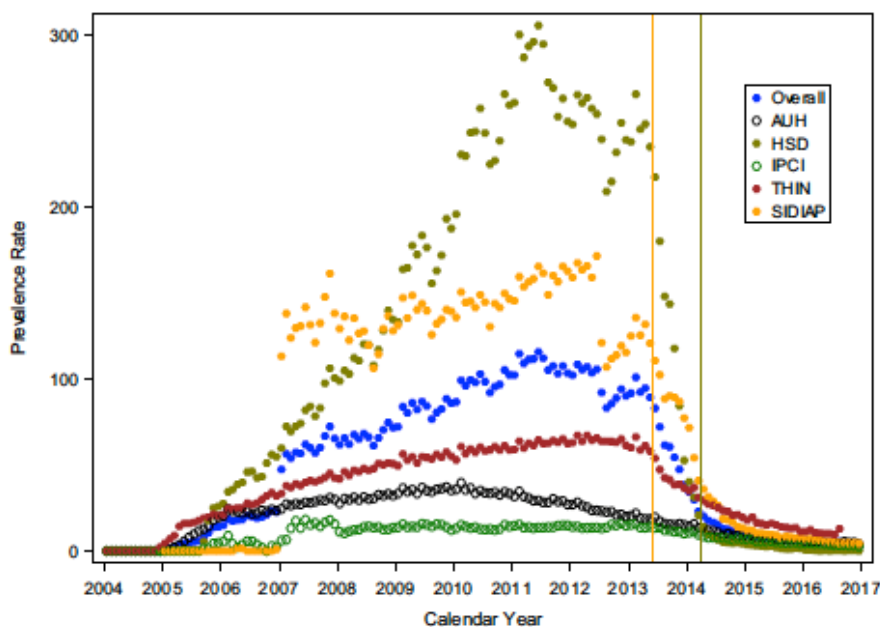



Figure 1. An example of time-series demonstrating the impact of risk minimisation measures: Monthly prevalence rates (/10,000 py) of use of strontium ranelate in 5 European databases, and overall in Europe (2).

2. AIMS AND OBJECTIVES

The overall aim of this work is to develop a tool which enables detection of changes over time in the utilisation of drugs to be studied.

2.1 Primary Objectives

Our specific objective was to develop a software tool that allows users to a) determine the effect of known interventions/changes specified a priori, and b) to detect/discover changes not specified a priori in a data-driven manner.

 EHDEN <small>EUROPEAN HEALTH DATA & EVIDENCE NETWORK</small>	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU



3. METHODS

3.1 Study Design

Cohort study within international data network framework.

3.2 Study Population

For the purposes of prototype development two cohorts were considered. 1) Patients diagnosed with rheumatoid arthritis (RA) and on treatment with hydroxychloroquine (HCQ). 2) Patients diagnosed with covid-19 and on treatment with HCQ and dexamethasone (DXM).

3.3 Data Sources

For this study data mapped to the OMOP Common Data Model (CDM) and available on the EHDEN platform were used, namely IQVIA Open Claims (USA) and IQVIA Hospital (USA), THIN (UK).

3.4 Variables

This section describes covariates that were considered to describe drug utilisation patterns in the use cases included in this report (see section 7 for description of use cases). In principle the tool can be applied to a wider variety of time-series based on various combinations of diagnoses, treatments, and time periods of interest.

3.4.1 Diagnosis

1) RA diagnosis, and 2) covid-19 diagnosis.

3.4.2 Calendar time

Changes over time were described by providing information on drug use (expressed as the number of prescriptions) by calendar time.

3.4.3 Indication of use


Use of hydroxychloroquine and dexamethasone. Drug exposures were identified using the RxNorm Ingredient and all descendants. Multiple exposures to the same drug were combined using a 30-day gap window to create a continuous era of exposure.

3.5 Sample Size

As this is a descriptive drug utilisation study, a sample size calculation to reach significance was not required.

4. DATA MANAGEMENT AND RESOURCING NEEDS

We used the EHDEN platform that allows federated data analyses at scale. Study cohorts, exposure, outcomes and covariates were co-designed by study team. Tool development was based on the available

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU

databases listed in Section 3.3. Once the tool is built, it will be shared with all participating data partners in the EHDEN platform to evaluate on their databases. Data partners will be responsible for obtaining data access approval and perform the analysis using the EHDEN platform and centralised defined codes using their mapped version of data. If the access of unmapped data is required, the same statistical syntax used in the mapped data will be downloaded and run in the unmapped data.

5. DATA ANALYSIS

5.1 Analysis Pipeline

The analysis pipeline for the tool has two parts: the processing pipeline and results sharing pipeline. The pipeline itself is built as an R study package for network research as described in the [Book of OHDSI Chapter 20 \(3\)](#).

5.1.1 Processing Pipeline

Processing Pipeline Workflow

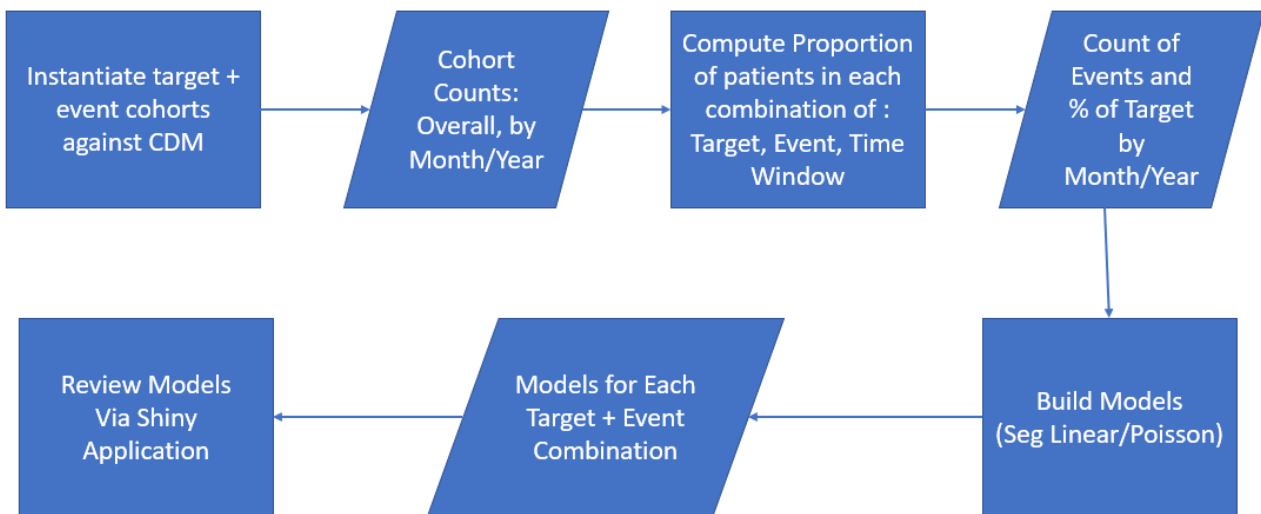



Figure 2 Processing Pipeline Workflow

The pipeline makes extensive use of cohort definitions as described in [Chapter 10 of the Book of OHDSI](#). Cohorts are created using ATLAS or other OHDSI analysis tools to describe populations of interest to use in the RMM study. These cohorts are the main inputs to the pipeline and are detailed below.

Inputs

- **Target Cohorts (T):** one or more populations of interest to use for RMM detection.
 - *Use case Example:* Persons hospitalized with a covid-19 diagnosis record or a SARS-CoV-2 positive test (covid 19 patients)
- **Event Cohorts (E):** events of interest to use to detect changes in the target populations

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU



- *Use case Example:* 1) Prevalent use of Dexamethasone (DXM) and 2) Hydroxychloroquine (HCQ)
- **Time Windows (W):** the start/end dates to evaluate the *event cohorts* in each *target cohort*
 - *Example:* 0 days – 30 days

Using these inputs, we can compute the combinations of people in the **target cohort (T)** who are also in the **event cohorts (E)** for the specified **time window (W)**. For example, we may want to understand the use of DXM (*event cohort*) in the 0-30 days (*time window*) for persons hospitalized with covid-19 (*target cohort*). This operation is computed for every combination of {T x E x W}. Then for each {T x E x W} combination, we compute the number of patients exposed for each month/year. The package allows for censoring of events that fall below a minimum threshold to satisfy various institutional requirements. This results in a time series analytic object describing the changes over time for the full set of {T x E x W}.

Using the time series analytic object, we can fit various change point detection methods (described in the Statistical Analysis). Once the models are fit, an interactive R-Shiny application is available to graphically explore the various change point detection models and results.

5.1.2 Results Sharing Pipeline

Results Sharing Workflow

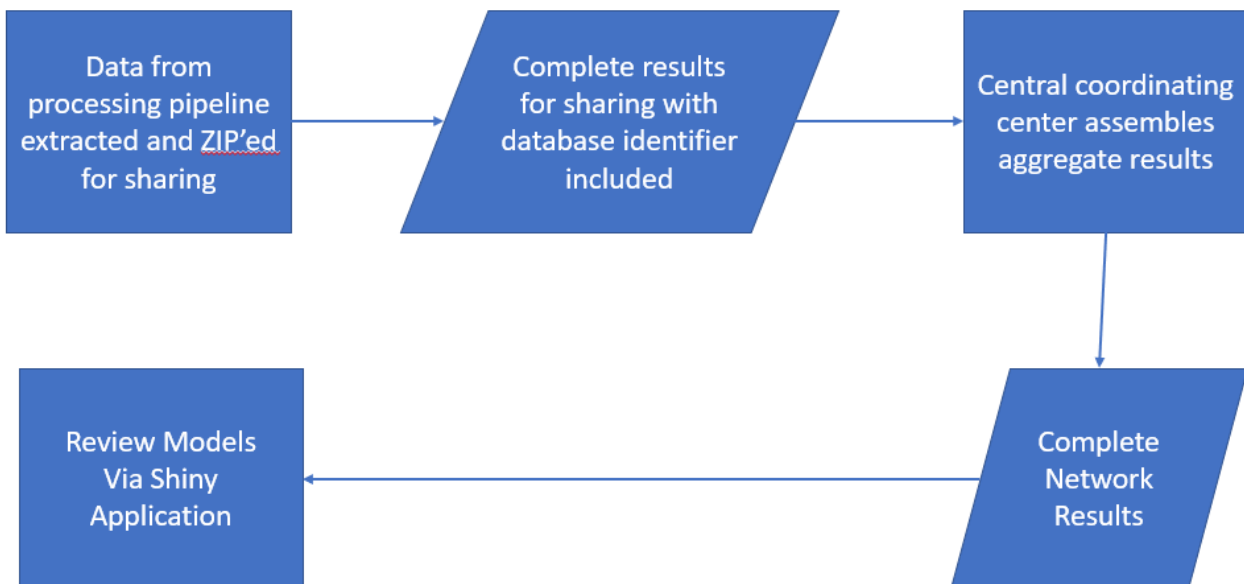



Figure 3 Results sharing pipeline workflow

Aggregate results are stored in comma-separated value (CSV) files for full transparency of results for inspection outside of R. The data generated in the pipeline along with log files are then added to a ZIP file for sharing with a central study coordinator. The central coordination center can assemble the full set of

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU

results from the different data partners and use the same Shiny application to review the full set of network results.

5.2 Cohort Generation

The OHDSI definition of a cohort is: ‘a cohort is a set of persons who satisfy one or more inclusion criteria for a duration of time.’ This results in a table that describes episodes for a given person as follows:

cohort_definition_id	subject_id	cohort_start_date	cohort_end_date
1	1000	01/02/2020	30/04/2020
1	1001	15/02/2020	01/03/2020
2	1000	03/02/2020	14/02/2020

The table above contains 4 columns: a cohort_definition_id that uniquely identifies the cohort, the subject_id that identifies the patient and the cohort_start_date and cohort_end_date period of time when the person satisfied the inclusion criteria for the cohort. In our pipeline example, let’s assume covid-19 patients are cohort_definition_id = 1 (**Target cohort**) while DXM is cohort_definition_id = 2 (**Event cohort**). If we’d like to evaluate if a covid-19 patient is provided with DXM in the 0 – 30 d (**Time Window**) of hospitalization, we can compare the relative start dates for each patient that appear in both the Target and Event cohort and if they are within 30 d of each other, they count as having that event. In the example above subject_id 1000 is provided with DXM 2 days after satisfying the inclusion criteria for the covid-19 cohort therefore this patient counts as having the event.

5.3 Statistical Analysis

Statistical analyses were performed for the two objectives re-stated below.

Objective 1) Determine the effect of known interventions/changes specified a priori, and

Objective 2) to detect/discover changes not specified a priori in a data-driven manner.


All analyses were conducted in R version 4.0.0.

5.3.1 Change Point Detection

Change point detection refers to the study of known or unknown abrupt changes or disturbances in the behaviour of a series of data points over time (hereafter referred to as a time-series). Change point detection methods can be applied to determine the effect of a known intervention i.e. when the estimated change-point(s) is located at known/pre-specified time-point(s), or when it is of interest to discover any naturally-occurring change points in the time-series.

Most change point detection methods can be classified based on criteria such as ability to detect single or multiple change points, capacity to model univariate or multivariate time-series, use of parametric or non-parametric approaches, and online or offline design (4).

For the purposes of development of our prototype which is designed to detect one or more change-points in a single time-series at a time, this work is restricted to univariate methods, and our design permits extension to (multivariate)methods for multiple time-series if required.

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU



Following the latest comprehensive review and validation of change point detection methods (4) we included the following methods based on suitability to our use-case, with a possibility to extend to further change point detection methods, depending on user feedback.

5.3.2 Regression-based methods

Segmented regression is one of the most common approaches for change point detection. The analysis has two stages (5): 1) fitting a regression equation to the time-series to model its underlying distribution, and 2) to search for the optimal change point(s) in that distribution, using an optimisation procedure such as maximum likelihood. The time-series is split into segments and the optimisation problem seeks the segments which result in the smallest root mean squared error. The number and initial location of segments can be pre-specified by user or determined by the model (i.e. data-driven). This analysis was implemented using R package *segmented* version 1.3.1 (6).

5.3.3 Bayesian Online Change point detection


This approach applies the principles of Bayesian inference to the problem of change point detection (7). Beginning with an initial guess (prior), the posterior probability of there being a change at each time point in the time-series is computed based on the observed data. Hence at each time-point the probability of change is based on the data observed up to that point (online detection). The output is an estimate of the probability of change at each timepoint, along with a confidence interval to assess the uncertainty around the estimated probability. This analysis was implemented using R package *ocp* version 0.1.1 (8).

5.3.4 Settings

Table 1 shows the change point detection methods included in the tool. For each method, a number of options can be provided to the user. For segmented regression, the regression model can be chosen between linear or Poisson. The initial estimate of the location of the change point can be pre-specified, e.g. if the time of the intervention is known, after which a final estimate is determined by the model. Alternatively, the location of the change point can be estimated by the model based solely on the data (data-driven). Similarly, the number of desired change points can also be set to one or more or can be determined automatically (data-driven). Finally, it is possible to fix the location of the change-point to the initial estimate (fixed). The outputs of the segmented regression method include estimated change points and the slope and intercept values of each of the segments, along with 95% CI and p-values. The Bayesian method is purely data-driven and provides the estimated probability of change at each time point, along with 95% CI.

Table 1 Change point detection methods included in the tool along with the settings that can be selected by the user.

Method	Model	Location	Number of change points	Output
Segmented Regression (reference method)	<ul style="list-style-type: none"> ○ Linear piece-wise regression ○ Poisson piece-wise regression 	<ul style="list-style-type: none"> ○ Pre-specified ○ Data-driven 	<ul style="list-style-type: none"> ○ 1 ○ 2 ○ Data-driven ○ Fixed (1 changepoint) 	<ul style="list-style-type: none"> ○ Slope and intercept of each segment ○ Estimated change point(s)

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU

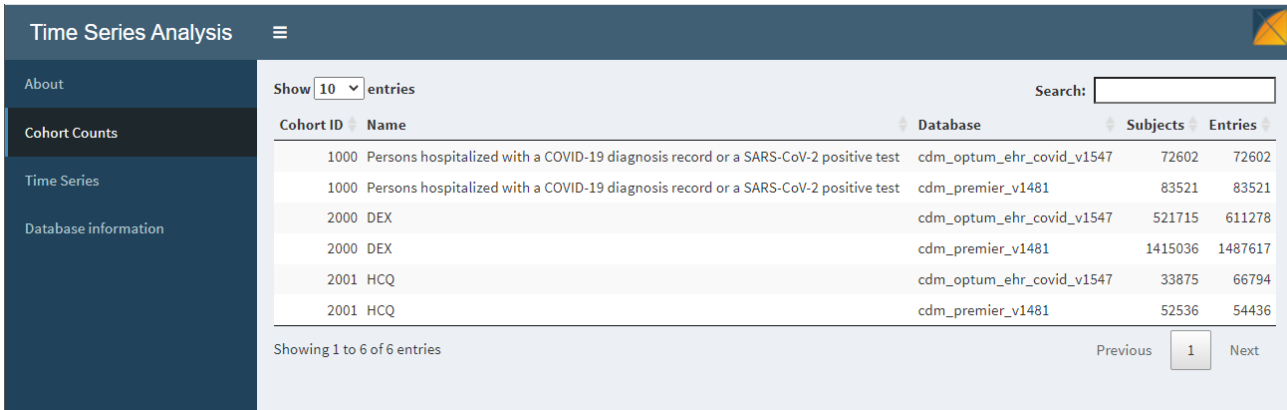
Bayesian Online	Bayesian inference	Data-driven	Data-driven	<ul style="list-style-type: none"> ○ Probability of change at each time point ○ Estimated change point(s)
-----------------	--------------------	-------------	-------------	---

6. TOOL PROTOTYPE

The prototype application for this tool is available at <https://data.ohdsi.org/EhdenTimeSeriesAnalysis/> while the source code is available at <https://github.com/EHDEN/TimeSeriesAnalysis>. The application allows for the exploration of change point detection methods along with providing a high-level summary of the counts of the cohorts used in the study.

Cohort Counts

Using the left-hand menu, the application provides a summary of cohort counts used in the analysis for each database, as shown in the following figure:




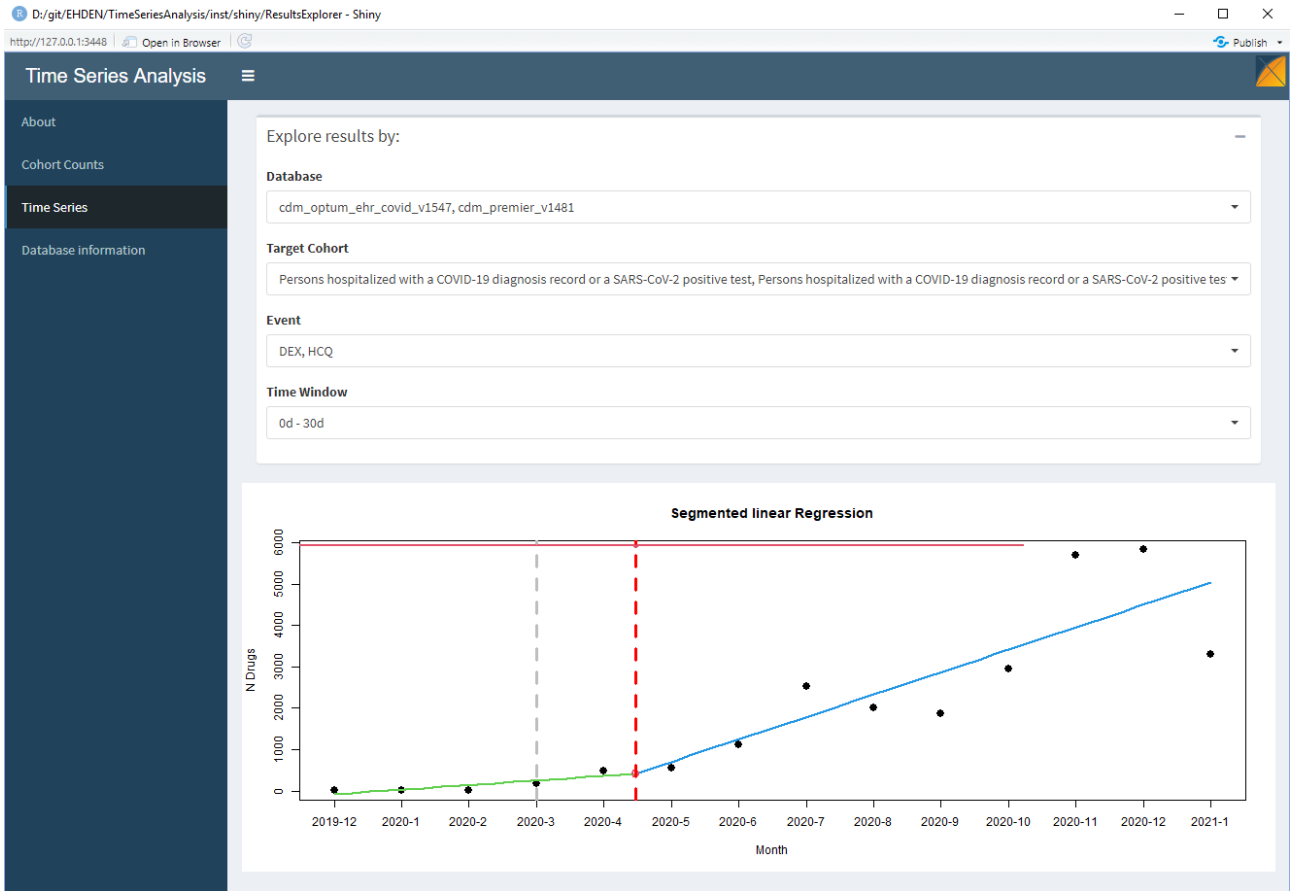
The screenshot shows the 'Time Series Analysis' application interface. On the left is a navigation menu with options: 'About', 'Cohort Counts' (selected), 'Time Series', and 'Database information'. The main content area displays a table of cohort counts with columns for Cohort ID, Name, Database, Subjects, and Entries. A search bar and a 'Show 10 entries' dropdown are at the top. The table contains 6 rows of data. At the bottom, it shows 'Showing 1 to 6 of 6 entries' and pagination controls for 'Previous', '1', and 'Next'.

Cohort ID	Name	Database	Subjects	Entries
1000	Persons hospitalized with a COVID-19 diagnosis record or a SARS-CoV-2 positive test	cdm_optum_ehr_covid_v1547	72602	72602
1000	Persons hospitalized with a COVID-19 diagnosis record or a SARS-CoV-2 positive test	cdm_premier_v1481	83521	83521
2000	DEX	cdm_optum_ehr_covid_v1547	521715	611278
2000	DEX	cdm_premier_v1481	1415036	1487617
2001	HCQ	cdm_optum_ehr_covid_v1547	33875	66794
2001	HCQ	cdm_premier_v1481	52536	54436

Time Series Exploration

Additionally, you can visually inspect the results of the various time series analyses using the “Time Series” menu option:

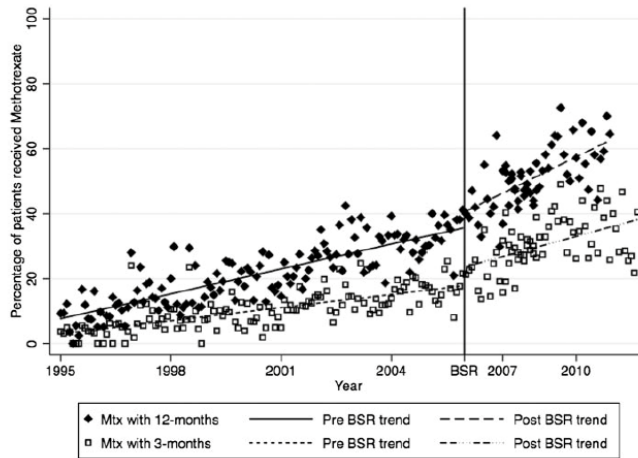
	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 - Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU 15/20



The menu items allow you to filter based on the database(s) of interest, the target & event cohorts and the time window as described in section 5.


7. USE-CASES

Fig. 1 Trends in prescribing of MTX before and after publication of BSR guidance



BSR: British Society for Rheumatology.

Figure 4 Methotrexate prescription timeseries. Figure from (9).

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU 16/20



Example 1: Changes in use of HCQ in RA patients over time

- **Background:** RA guidelines recommend using Methotrexate as first line therapy from 2008, possibly ‘displacing’ hydroxychloroquine (HCQ), Figure 4.
- **Data:**
 - Initiation of first line HCQ for RA
 - GP prescriptions in THIN (UK)
- **Period:** 2000 – 2018
- **Change Point Detection Results:**

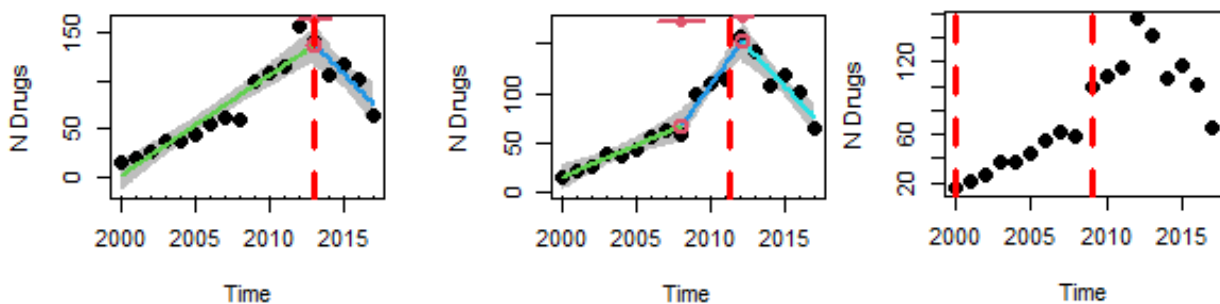



Figure 5. The number of annual HCQ prescriptions in UK (y-axis) over time from the year 2000 to 2018 (x-axis). Left: Segmented linear regression with one change point detected. The initial location of the change point pre-specified by user is indicated by a red circle, along with a horizontal line marking the confidence interval around that change point. Regression lines representing the regression equations fitted to each segment are also shown in green and blue colours. The (data-driven) location of the change point estimated by the model is marked by a vertical dashed line in red on the x-axis. Centre: Segmented linear regression with two initial change points pre-specified by user (red circles) and the (data-driven) location of the single changepoint estimated by the model (vertical dashed line). Right: Bayesian online change point detection, where no location is pre-specified by user, and the method can estimate one or more change points. Here the correct change point is at the year 2009 (the first timepoint shown here as a changepoint was a feature of the original model that has now been corrected in the prototype). The Bayesian method provides a probability of there being a change point for each given time point (not shown here), whereas the segmented regression provides a confidence interval around each estimated change point.

Example 2: HCQ for covid-19 in 2020

- **Background:**
 - HCQ received emergency approval by FDA in March 2020
 - Regulatory action followed in late April due to cardiovascular safety issues
- **Data:** HCQ for covid-19 treatment
- **Database:** IQVIA Open Claims (USA)
- **Period:** Jan – Oct 2020

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU 17/20



- **Change point detection Results:**

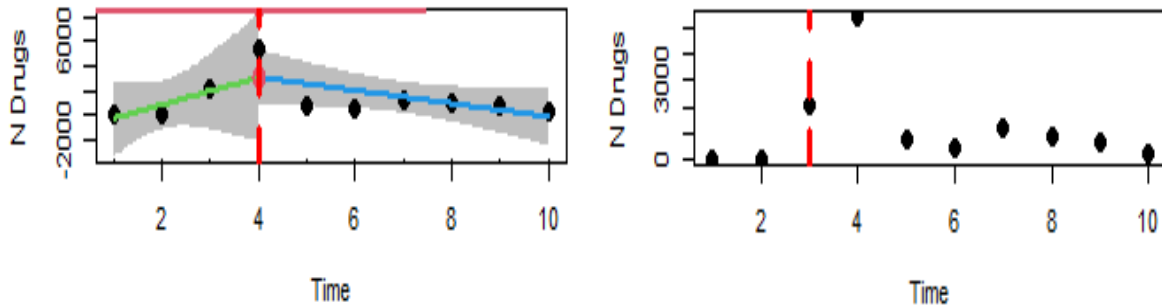
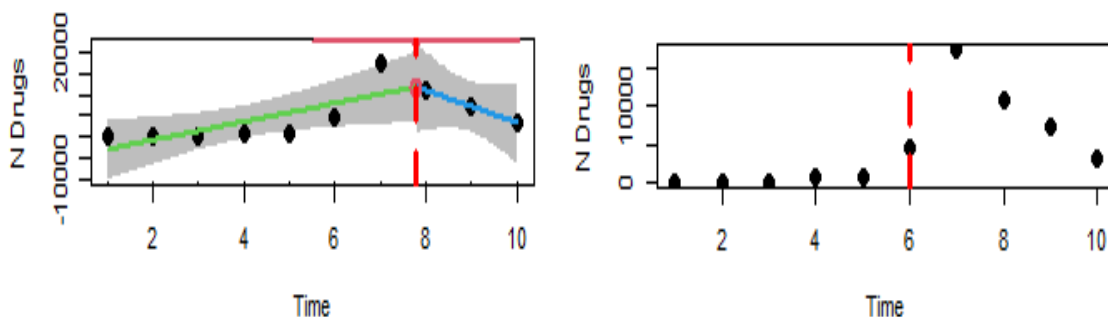



Figure 6 The number of monthly HCQ prescriptions for covid-19 patients (y-axis) over time from January (month 1) to October 2020 (month 10) (x-axis), available in the IQVIA Open Claims (USA) database. Left: Segmented linear regression. The location of a changepoint is marked by a vertical dashed line in red on the x-axis, along with a horizontal line marking the confidence Interval around the estimated change point. Regression lines representing the regression equations fitted to each segment are also shown. Right: Bayesian online change point detection, with the location of a changepoint marked by a vertical dashed line in red on the x-axis. Note: same data used in both plots, however Y-axis of the left plot wider to allow plotting of regression model including error bar in grey.

Example 3: Dexamethasone for covid-19 in 2020

- **Background:** RECOVERY RCT report 30% reduction in mortality in late June 2020 with dexamethasone (DXM)
- **Data:** DXM for covid-19 treatment
- **Databases:** IQVIA Open Claims (USA), IQVIA Hospital Records (USA)
- **Period:** Jan – Oct 2020
- **Change point detection Results:**



	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU 18/20

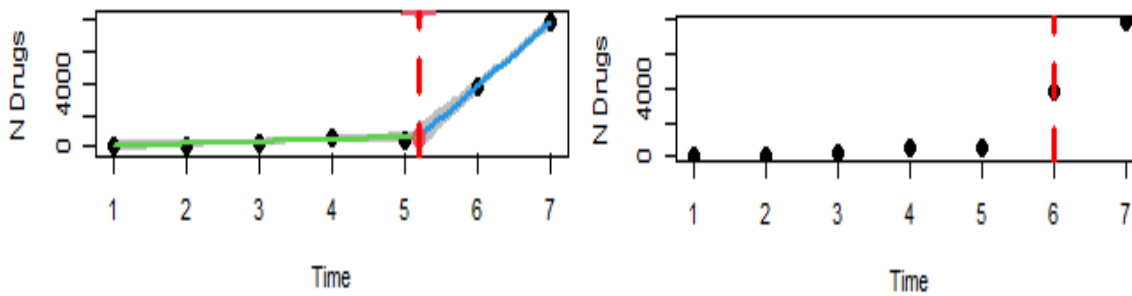


Figure 7 The number of monthly Dexamethasone prescriptions for covid-19 patients (y-axis) available in the IQVIA Open Claims (USA) database. Top: IQVIA Open Claims (USA), over time from January (month 1) to October 2020 (month 10) (x-axis). Bottom: IQVIA Hospital (USA), over time from January (month 1) to July 2020 (month 7) (x-axis). Left: Segmented linear regression. The location of a changepoint is marked by a vertical dashed line in red on the x-axis, along with a horizontal line marking the confidence Interval around the estimated change point. Regression lines representing the regression equations fitted to each segment are also shown. Right: Bayesian online change point detection, with the location of a changepoint marked by a vertical dashed line in red on the x-axis.

8. DISCUSSION

The use cases demonstrated the potential application of the tool on real-world data under real-world contexts.


Example 1 showed that both the reference method (segmented linear regression) and the Bayesian approach suggested changes in the trends of HCQ use in the years following the introduction of MTX as the preferred treatment (intervention) in 2008. Segmented regression provides confidence interval for each estimated change point, whereas the Bayesian approach provides an estimate of the probability of change at each time-point, therefore the two approaches can be used complementarily.

In Example 2, where the FDA warning was issued in March 2020, the methods estimated the change point in the use of HCQ to be March 2020 (Bayesian) and April 2020 (segmented linear regression).

In Example3, the methods were similarly able to detect the change point in DXM use following the June 2020 findings of the Recovery trial favouring the use of DXM for covid-19 treatment, with the estimated change points being June 2020 and July 2020 by the Bayesian and segmented linear regression methods, respectively. Interestingly, the results were similar even when Example 3 was repeated in a database with fewer samples (IQVIA Hospital, Figure 7 bottom plots), suggesting the potential for the methods to capture change points well as long as the data sufficiently captured underlying trends.

The focus of this work was on tool development. In future work it would be of interest to explore in detail comparative application scenarios for the different approaches considered, for instance scenarios where the segmented regression and Bayesian techniques can be used complementarily or where one has particularly advantages over the other. Further methods can also be incorporated based on the feedback on individual data partners and their respective use cases.

There are a number of extensions to this work that are underway as part of future work beyond the scope of this deliverable.

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU



- a) Extension to further change detection methods (e.g. non-parametric change point detection)
- b) Replication in other databases mapped to the OMOP CDM in EHDEN
- c) Extension to clinical applications beyond drug utilisation, e.g. changes in outcomes e.g. mortality, and healthcare provision and utilisation before, during, and after the current pandemic to name one.
- d) Reporting the number of drugs used as proportion of persons in the target population.

9. CONCLUSIONS AND FUTURE WORK

A tool for studying changes in the use of a drug over time was developed. Use of the tool was demonstrated through three real-world examples, illustrating how a number of methods implemented in the tool detect the real-world changes in a data-driven manner. Validation through further use-cases e.g. an approach for quantification of the accuracy and precision of the suggested change-points, while beyond the scope of this deliverable (D 1.3), is a natural next step and is recommended as future work. The tool is available as an interactive web application where data custodians can select their database, drug, cohort, and time period of interest, and view the estimated changes in these data over time. The design of the workflow supports extension of this work to include further methods for change point detection, and application to a wide variety of clinical/healthcare data series over time, beyond the context of drug utilisation.

10. QUALITY CONTROL

The study was conducted according to the guidelines for Good Pharmaco-epidemiology Practice (GPP) (ISPE 2008) and according to the ENCePP code of conduct (EMA 2010).

11. LIMITATIONS OF RESEARCH METHODS

The proposed study design is observational in nature. Despite all the efforts made to minimise confounding, residual (unobserved) confounders are always possible, and will remain a limitation for the assessment of causality in the proposed studies, although in this work. Although this work uses only drug utilisation data, confounding is a consideration during the cohort definition process. Indeed, the mapping to OMOP could potentially lead to information loss that could result in misclassification/information bias in some scenarios, e.g. of drugs and diseases.


12. PROTECTION OF HUMAN SUBJECTS

a. Ethical / Data Utilisation Approval and Subject Consent

Local data partners are responsible for obtaining ethical/data approval prior to the start of the analysis (also see table 16.1)

b. Subject Confidentiality

No direct subject contact or collection of additional subject data will occur.

	D1.3 - Implementation of RMM Effectiveness testing methods		
	WP1 -Evidence Workflow Development		Version: v3.0 – Final
	Author(s): Sara Khalid, Anthony Sena, Daniel Prieto-Alhambra		Security: PU



c. Report of Adverse Events

N/A

d. EHDEN Data Sharing Agreement

No local data can be shared in this project.

13. PLANS FOR DISSEMINATING STUDY RESULTS

The code, results, and all study documentation are deposited in the EHDEN Timeseries GitHub:

<https://github.com/EHDEN/TimeSeriesAnalysis>

14. REFERENCES

1. Bouvy JC, Blake K, Slattery J, De Bruin ML, Arlett P, Kurz X. Registries in European post-marketing surveillance: a retrospective analysis of centrally approved products, 2005-2013. *Pharmacoepidemiology and drug safety*. 2017;26(12):1442-50.
2. Berencsi K, Sami A, Ali MS, Marinier K, Deltour N, Perez-Gutthann S, et al. Impact of risk minimisation measures on the use of strontium ranelate in Europe: a multi-national cohort study in 5 EU countries by the EU-ADR Alliance. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA*. 2020;31(4):721-55.
3. The Book of OHDSI [Internet]. 2019. Available from: <https://ohdsi.github.io/TheBookOfOhdsi/>.
4. van den Burg GJ, Williams CK. An evaluation of change point detection algorithms. arXiv preprint arXiv:200306222. 2020.
5. Muggeo VM. Estimating regression models with unknown break-points. *Statistics in medicine*. 2003;22(19):3055-71.
6. segmented: Regression Models with Break-Points / Change-Points Estimation [Internet]. Available from: <https://cran.r-project.org/package=segmented>.
7. Adams RP, MacKay DJ. Bayesian online changepoint detection. arXiv preprint arXiv:07103742. 2007.
8. ocp: Bayesian Online Changepoint Detection [Internet]. Available from: <https://cran.r-project.org/package=ocp>.
9. Judge A, Wallace G, Prieto-Alhambra D, Arden NK, Edwards CJ. Can the publication of guidelines change the management of early rheumatoid arthritis? An interrupted time series analysis from the United Kingdom. *Rheumatology (Oxford, England)*. 2015;54(12):2244-8.