



Vjola Hoxhaj¹, **Constanza L Andaur Navarro¹**, Carlos Duran¹, Daniel Weibel¹, Miriam MCJM Sturkenboom¹

GENERATING EVIDENCE FROM RWE: THE CONCEPTION DATA QUALITY PIPELINE

There is a global need for evidence on the use and safety of medicines and vaccines during pregnancy and lactation. The IMI ConcePTION project aims to build a robuts learning ecosystem wherein Real World Data (RWD) can be used to generate Real World Evidence (RWE) (www.imi-conception.eu). Although RWD is promising, the value of RWE to guide evidence-based medicine and regulatory decision-making regarding medicines and vaccines largely depends on the quality and reliability of source.





AIM

To develop a five-level pipeline with 35 quality checks (QC) to verify the overall quality of source data locally at data access providers (DAPs) level.

MATERIALS & METHODS

Design.

- Data formatted according to the ConceP-TION CDM v.2.2 (Thurin et al. 2021)
- Based on Khan et al. (2016) data quality terminology, level checks were designed considered the following quality dimensions: conformance, completeness, and plausibility.

Repository.

- - - https://github.com/UMC-Utrecht-RWE
 - Code license GNU AGLP v3
 - Code versioning system git
 - Software code languages R

Execution.

- Open-source R code runs on CDM tables and generates HTML reports locally.
- 、 天 口 二 口 HTML reports show aggregrated data on counts, rates, and frequencies.
 - Data fitness-for-use can be assessed based on

• Easily shareable and understood.

study-specific requirements.

RESULTS

		2. Missing data ar	nalysis	;						WORKFLOW	
	For easier navigation, in case the	ere is missing data, the table_name, var.	iable_name	and meaning	g will be colored greer	n					
	Missing data overall Missing data analysis is perform	ned over all tables and for each variable.				Counts stratified by meaning (other variables)					
	Show 10 💙 entries			Search:		## [1] "Attention! There is(are) 1 row(s) with a zero value for bot not be displayed in the tables or graphs. This happens when a variab	th count and total. Those will ble is completely missing for a	Counts stratified by meaning	variable_name		
	table_name	variable_name	¢ count	total	percentage	particular meaning category."	10		person id		
	1 VISIT_OCCURRENCE	person_id	0	601	0	Show 10 v entries	Search:	5-	specialty_of_visit		
/el	2 VISIT_OCCURRENCE	visit_occurrence_id	0	601	0	table_name 🔶 variable_name 🔶 meaning 🔶	count 🔶 total 🔶 percentage 单 👸		visit_occurrence_id		
	3 VISIT_OCCURRENCE	visit_start_date	0	601	0	1 VISIT_OCCURRENCE person_id oupatient_specialist_visit 54	545 1078 50.6 ⁵	D-			DAP representative
σ	4 VISIT_OCCURRENCE	visit_end_date	601	601	100	2 VISIT_OCCURRENCE specialty_of_visit oupatient_specialist_visit 1	1047 1047 100 <u>2</u>	5-			
	5 VISIT_OCCURRENCE	specialty_of_visit	22	601	3.7	3 VISIT_OCCURRENCE visit_occurrence_id oupatient_specialist_visit 1	1078 1078 100				
	6 VISIT_OCCURRENCE	specialty_of_visit_vocabulary	0	601	0	Showing 1 to 3 of 3 entries	Previous 1 Next	D			
	7 VISIT_OCCURRENCE	status_at_discharge	601	601	100			. table visit			
	8 VISIT_OCCURRENCE	status_at_discharge_vocabulary	601	601	100			ateh-set			()
	9 VISIT_OCCURRENCE	meaning_of_visit	0	601	0			o ^{sx} Meaning			
	10 VISIT_OCCURRENCE	origin_of_visit	0	601	0	_					
	Showing 1 to 10 of 115 entries	Previous 1	2 3	4 5	12 Next						Data engineer

The following table contains the aggregated issues per table:

	Show 10 🕤 entries		Search:	
	TABLE	lssues 🕴	Total 🔶	Percentage
	EVENTS	493352	18328233	2.7
O	MEDICAL_OBSERVATIONS	87482	2427711	3.6
O	MEDICINES	401330	32315660	1.2
	PROCEDURES	1185150	83717969	1.4
	VACCINES	19348	1296952	1.5
	Showing 1 to 5 of 5 entries			Previous 1 Next

Any warnings/messages generated during the data processing and/or calculations are listed in the table below

Show 5 💿 en	tries	Search:					
TABLE 🔶	d 🔶	Warning					
EVENTS	EVENTS_1_PS.csv	Unable to perform check for 0 out of 1994582 rows.					
EVENTS	EVENTS_1000001_PS.csv	Unable to perform check for 0 out of 1864306 rows.					
EVENTS	EVENTS_2000001_PS.csv	Unable to perform check for 0 out of 1872980 rows.					
EVENTS	EVENTS_3000001_PS.csv	Unable to perform check for 0 out of 1819010 rows.					
EVENTS	EVENTS_4000001_PS.csv	Unable to perform check for 0 out of 1804942 rows.					
Showing 1 to 5 of	257 entries	Previous 1 2 3 4 5 52 Next					





Data Quality Revisor

ш

CONCLUSION

N

As new types of data sources become available, we proposed the ConcePTION QC to demonstrate whether secondary use of existing observational data are reliable and relevant for regulatory decision-making of medicines and vaccines

The QC pipeline has been successfully implemented in the ConcePTION project, the Vaccine monitoring Collaboration for Europe (VAC4EU, https://vac4eu.org/), CONSIGN, EMA and PASS Covid-19 vaccine safety and effectiveness studies

Moral duty to provide evidence to women on reproductive age regarding medicines and vaccine safety



To monitor vaccines and medicines post-licensure, large research network with access to electronic health records needs to be established. This includes global collaboration.



This work has received support from the EU/EFPIA Innovative Medicines Initiative 2 Joing Undertaking ConcePTION grant no.821520 We would like to acknowledged the developers Roel Elbers, Rutger van der Bor, and Ema Alsina for their support on the pipeline construction. We thank the data access providers for their collaboration on testing and evaluation.

EU PAS protocol registration no. 50142

c.l.andaurnavarro@umcutrecht.nl $(\square$

/

2

MPO

¹Department of Data Science & Biostatistics, Julius Center, University Medical Center Utrecht, Utrecht University, The Netherlands