SA104

DETERMINING ADEQUATE LOOKBACK PERIODS FOR BETTER ESTIMATION OF PATIENT HEALTH STATUS AT INCLUSION IN STUDY USING REAL WORLD DATA ELECTRONIC MEDICAL RECORDS FROM THE THIN® DATABASES IN DIFFERENT EUROPEAN COUNTRIES

Costantino De Palma (1), Carlos Iglesias (2), Caroline Eteve-Pitsaer (1), Charlotte Renaudat (1)

1. Cegedim Health Data France, Boulogne-Billancourt, France; 2. Cegedim Health Data Spain, Sant Cugat del Vallès, Barcelona, Spain

BACKGROUND

To calculate key epidemiological indicators like prevalence and incidence using Real World Data (RWD), researchers should carefully determine the lookback period (LP) to ensure well-characterized patient health status at inclusion (1,2). A refined method involves using a LP before including patients with prior records of the disease, thereby reducing the overestimation of incident cases. Variations in LP length affect the accuracy of incidence estimates, with shorter LPs potentially leading to overestimation and longer LPs potentially limiting database utility (3-5).

OBJECTIVES

This study aims to establish appropriate LPs for chronic pathologies (diabetes, arterial hypertension, COPD, stroke, Parkinson's, depression), neoplasia (breast, lung, prostate, colorectal cancer) and acute pathologies (urinary tract infection: UTI), in 7 European countries (Belgium, France, Germany, Italy, Romania, Spain and the United Kingdom). The present work is a first attempt at calibration of a methodology, as one pathology for each group (diabetes, lung cancer and urinary tract infection) will be studied in three selected countries (France, Italy and the United Kingdom) and a further study should be carried out.

METHODS

The design is an observational retrospective multinational study using RWD. Included active patients are used as **denominator** of overall mean rate that approach incidence (‰), referred to as “incidence” in the following, and are those registered in THIN® France, THIN® Italy and THIN® UK databases, when they have at least one visit to a general practitioner during 2023 and at least 24 months of medical data between the first record in the database and the relevant year of study. Patients with prior record of relevant pathology diagnosis in every window defined period are excluded.

Cases described above are included in the **numerator** only when a new diagnosis of the relevant pathology is found in the medical records of 2023. It follows that patients can be counted in the numerator only once, their index date being the earliest recording of the diagnosis in the year of study.

Registration date is calculated based on the date of first electronic record of the patients throughout his complete medical history.

Time from registration (LP) is defined as the differences (in weeks) between the registration date and the earliest date of diagnosis in 2023 (for cases in the numerator) and as the earliest date of contact in 2023 (for all remaining cases).

The **optimal LP window** is chosen across country and pathology based on two criteria : the **absolute differences between the** overall mean **incidence (‰) and the one (‰) at the specific LP window** and the **sign and direction of such difference**. If we assume that the mean incidence generated via the THIN® databases is the closest to the “real” incidence of new diagnosis (with a small overestimation), the chosen LP window should achieving the minimum difference. Similarly, as long as the difference between the overall mean incidence and the incidence at a specific window is of negative sign (-), specifying a LP less then the optimal window will overestimate the number of cases, whereas any LP above the optimal one can lead to underestimation. It follows that the optimal LP is the first window producing both the minimal difference and a reserve of its sign.

RESULTS

The nature of the pathology has the most evident impact in the shape of the curves. The national healthcare system also plays a role (Table 1 and figures 1-3).

Table 2 provides an example for the determination of the optimal LP for Italy. For Diabetes, specifying a LP between 0 and 12 weeks can lead to overestimation of the incidence – due to recording by the GP of prior patient diagnosis upon registration. The optimal LP is of at least 12-24 weeks, yielding the minimal differences from the overall mean incidence and reverting the sign.

For Lung Cancer, similarly, a too small LP can lead to misidentification of cases due to recording practices or ambiguous estimates due to the fact that the primary diagnosis often happens in secondary care setting. The optimal LP is between 36 and 48 weeks, where enough patients have return to their GP and the medical history is transferred.

Finally, for UTI, results are less clear as less often those diagnosis are recorded ex-post. However, a LP between 24 and 48 weeks should avoid over or underestimation of the indicator and produce the most reliable estimates.

Table 1: Optimal LP window by country and pathology

Country	Pathology	Mean Incidence 2023 (‰)	Time since Registration	Incidence (‰)	Distance from Mean
France	Diabetes	3,83	(24,36]	3,77	0,06
France	Lung Cancer	0,06	(48,60]	0,04	0,02
France	UTI	6,82	(12,24]	3,31	3,51
Italy	Diabetes	6,55	(12,24]	5,64	0,90
Italy	Lung Cancer	0,09	(36,48]	0,09	0,01
Italy	UTI	2,76	(36,48]	2,62	0,14
UK	Diabetes	2,22	(60,72]	2,16	0,06
UK	Lung Cancer	0,01	(60,72]	0,03	0,01
UK	UTI	1,89	(72,84]	1,86	0,03

Table 2: LP window 1. Specific LPs window by pathology. THIN® Italy

	Time since registration (Weeks)	Number of cases (numerator)	Incidence (‰)	Difference from mean	Sign of Difference
Diabetes (mean incidence 6.55‰)	[0,12]	17.290	41,09	34,54	-
	(12,24]	2.374	5,64	0,90	+
	(24,36]	1.586	3,77	2,78	+
	(36,48]	1.258	2,99	3,56	+
	(48,60]	1.089	2,59	3,96	+
	(60,72]	911	2,16	4,38	+
	(72,84]	871	2,07	4,48	+
	(84,96]	817	1,94	4,60	+
	(96,108]	734	1,74	4,80	+
	(108,120]	613	1,46	5,09	+
Lung Cancer (mean incidence 0.09‰)	[0,12]	194	0,46	0,37	-
	(12,24]	27	0,07	0,02	+
	(24,36]	25	0,06	0,03	+
	(36,48]	36	0,09	0,01	+
	(48,60]	18	0,04	0,05	+
	(60,72]	12	0,03	0,06	+
	(72,84]	19	0,05	0,05	+
	(84,96]	16	0,04	0,05	+
	(96,108]	20	0,05	0,04	+
	(108,120]	13	0,03	0,06	+
UTI (mean incidence, 2.76‰)	[0,12]	3.131	7,44	4,69	-
	(12,24]	1.393	3,31	0,55	-
	(24,36]	1.203	2,86	0,10	-
	(36,48]	1.101	2,62	0,14	+
	(48,60]	1.020	2,42	0,33	+
	(60,72]	904	2,15	0,61	+
	(72,84]	782	1,86	0,90	+
	(84,96]	718	1,71	1,05	+
	(96,108]	702	1,67	1,09	+
	(108,120]	641	1,52	1,23	+

Figure 1: Diabetes Mellitus. Evolution of specific LPs by time since registration

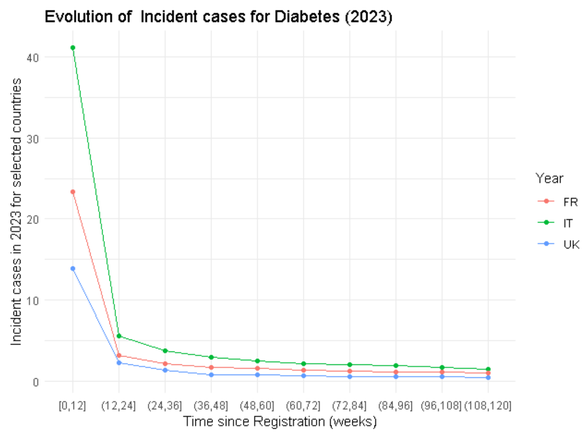


Table 2: Lung cancer. Evolution of specific LPs by time since registration

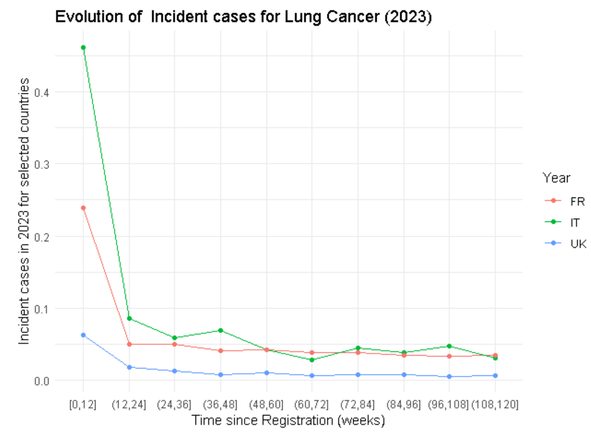
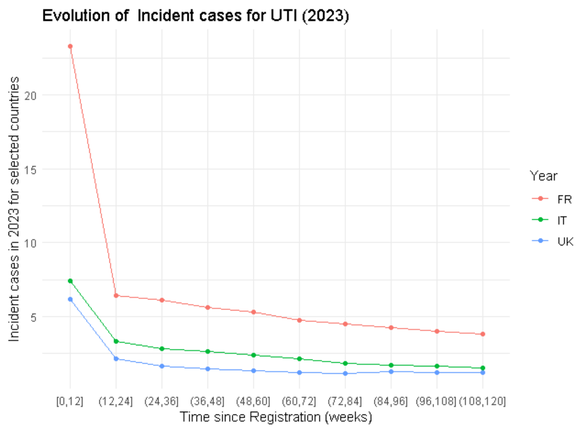


Table 3: UTI. Evolution of specific LPs by time since registration



CONCLUSION

These results show that to develop a standardized methodology across countries and pathologies for epidemiological indicators using RWD requires a deep knowledge of LPs. These depend on the national health system - registration based systems (UK) or GP-based (France)- and differs across pathologies. Although UTI, as an episode, can be defined as an acute pathology, being an eminently recurrent disease, its behaviour is similar to that of chronic pathology with this methodology. The knowledge of the best LP in every database is crucial and its must be analysed previous to perform an epidemiologic study.

REFERENCES

Bagley SC, Altman RB. Computing disease incidence, prevalence and comorbidity from electronic medical records. J Biomed Inform. 2016 Oct;63:108-111. doi: 10.1016/j.jbi.2016.08.005. Epub 2016 Aug 4. PMID: 27498067; PMCID: PMC6642638.
Lewis JD, Bilker WB, Weinstein RB, Strom BL. The relationship between time since registration and measured incidence rates in the General Practice Research Database. Pharmacoepidemiol Drug Saf. 2005 Jul;14(7):443-51. doi: 10.1002/pds.1115. PMID: 15898131.
Ødegaard KM, Lirhus SS, Melberg HO, Hallén J, Halvorsen S. A nationwide registry study on heart failure in Norway from 2008 to 2018: variations in lookback period affect incidence estimates. BMC Cardiovasc Disord. 2022 Mar 5;22(1):88. doi: 10.1186/s12872-022-02522-y. PMID: 35247979; PMCID: PMC8898410.
Rassen JA, Bartels DB, Schneeweiss S, Patrick AR, Murk W. Measuring prevalence and incidence of chronic conditions in claims and electronic health record databases. Clin Epidemiol. 2018 Dec 17;11:1-15. doi: 10.2147/CLEP.S181242. PMID: 30588119; PMCID: PMC6301730.
Sulo G, Igland J, Vollset SE, Nygård O, Egeland GM, Ebbing M, Sulo E, Tell GS. Effect of the Lookback Period's Length Used to Identify Incident Acute Myocardial Infarction on the Observed Trends on Incidence Rates and Survival: Cardiovascular Disease in Norway Project. Circ Cardiovasc Qual Outcomes. 2015

Presented at ISPOR Europe 2024; 17-20 November, 2024 Corresponding author: Carlos.IGLESIASGARCIA@cegedim.com