

ASSESSING THE EXTERNAL VALIDITY OF THIN® DATABASE IN SPAIN AS A RELIABLE SOURCE OF REAL-WORLD DATA FOR OBSERVATIONAL RETROSPECTIVE RESEARCH

Valente AJC¹, Morros M², Artés M², Iglesias C¹

¹Cegedim Health Data, Barcelona, Spain - ²Adelphi Targis, Barcelona, Spain

BACKGROUND

Reliable real-world databases (RWD) are essential across Europe to generate real-world evidence studies. The Health Improvement Network (THIN®), a Cegedim database, is a comprehensive RWD containing longitudinal data from approximately 1.8 million anonymised outpatients from 2014 to the present.

OBJECTIVES

To evaluate the external validity of THIN® Spain as a secondary data source for research.

METHODS

We compared data on demographic distribution, prevalence of major chronic diseases and drug prescriptions from THIN® Spain, a large research database of primary care records, with those from official sources.

We obtained demographic distribution data for 2020 from Spain's National Institute of Statistics (INE)¹. We investigated the prevalence of major chronic diseases using the 2020 European Health Survey in Spain² and the 2017 Spanish National Health Survey³, using the data from 2017 to address prevalence biases caused by COVID-19.

We also analysed prescribed drug dispensations⁴ by retail pharmacies in 2021 based on the Anatomical Therapeutic Chemical (ATC) classification.

We used descriptive statistics and Pearson correlations to assess dataset concordance.

RESULTS

THIN® exhibited a strong correlation ($r=0.99$) with INE's demographic distribution by age and gender, with a mean age of 45 years, of whom 51% were female and 49% were male.

The age groups where we observed the greatest discrepancies were 2 to 7 years and 28 to 36 years (**Figure 1**).

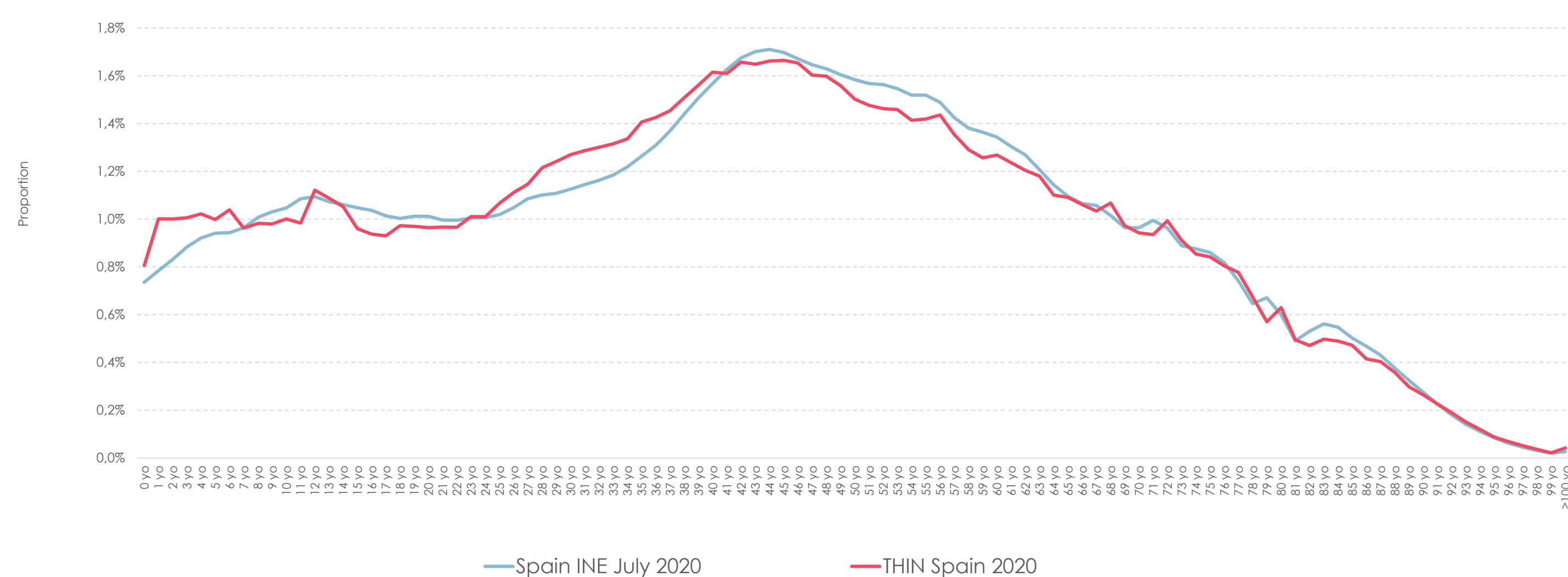


Figure 1:

Age demographics of THIN® Spain's active patients (men + women) in 2020 compared with those of Spain's National Statistics Institute for July 2020.

Our comparisons of the prevalence of chronic diseases showed correlations of $r=0.90$ (2017) and $r=0.92$ (2020).

According to THIN® data, in 2020 the most prevalent major chronic diseases were arterial hypertension (24.86%), hypercholesterolaemia (15.81%) and chronic (low) back pain (14.03%). These findings match those of the 2020 European Health Survey in Spain (**Table 1**).

CONCLUSIONS

The THIN® database exhibits good external validity for real-world evidence research in Spain, as demonstrated by its alignment with age and gender demographics, disease prevalence, and prescribed medications. Lower prevalence data in 2020 may be due to the pandemic situation which biased prevalence and lowered figures, and to different data collection methods (self-reported in surveys vs. diagnosed prevalence in THIN®).

	Patients ≥15 years old	Circulatory system			Metabolism		Respiratory		Mental health		Other			
		Arterial Hypertension (1)	MI	Angina pectoris, coron. heart dis.	Diabetes (2)	Hyperchol. (3)	Asthma (4)	COPD (5)	Depression (6)	Chronic anxiety (6)	Migraine or freq. Headache	Osteopor.	Chronic back pain (lumbar)	Incontinence
2020	Results of Spain's official 2020 health survey	19.34%	0.71%	0.72%	7.54%	15.47%	4.11%	2.75%	5.7%	6.4%	7.64%	3.38%	14.73%	3.61%
	Findings based on THIN Spain's data for 2020 (active patients from 15 July 2019 to 24 July 2020)	24.86%	0.96%	0.76%	8.70%	15.81%	4.49%	5.33%	12.45%	12.04%	6.30%	2.35%	14.03%	3.03%
2017	Results of Spain's official 2017 health survey	20.13%	0.7%	0.97%	7.82%	18.03%	4.73%	3.23%	7.17%	7.36%	9.47%	3.84%	19.76%	4.72%
	Findings based on THIN Spain's data for 2017 (active patients from 16 October 2016 to 15 October 2017)	18.63%	0.89%	0.77%	8.72%	15.64%	3.70%	5.63%	12.0%	10.69%	5.07%	2.12%	13.94%	2.51%

Table 1: Prevalence of chronic diseases in study patients: Arterial hypertension diagnosis or a blood pressure reading of 140/90 or higher; Diabetes diagnosis or antidiabetic treatment (ATC code A10) or an HbA1c level of 6.5% or more; hypercholesterolaemia diagnosis or total cholesterol level of over 200; allergic asthma; chronic bronchitis; COPD; and emphysema. Declared prevalence may be substantially lower than diagnosed prevalence.

Average prevalence differences were 1.5% (2020) and 0.6% (2017) with depression displaying the highest difference (5.7% in the 2020 survey and 7.17% in the 2017 survey vs. 12.45% according to 2020 THIN® data), followed by chronic anxiety (6.4% in 2020 and 7.36% in 2017 vs. 11.91% according to 2020 THIN® data). The fieldwork for the 2020 European Health Survey in Spain started in July 2019 and ended in July 2020, so the official prevalence figures probably do not yet incorporate the increase in people with self-declared anxiety.

Lower prevalence data were observed in all diseases in the 2020 Health Survey compared with 2017, with an average difference of -1.2%. The average prevalence difference between 2017 and 2020 observed in all diseases according to THIN® Spain was 0.6%.

Regarding drug distribution (ATC classification), THIN® Spain data had correlations of $r=0.997$ (ATC1), $r=0.992$ (ATC2), and $r=0.986$ (ATC3) with official sources (2021).

The Alimentary Tract and Metabolism (A) category showed the highest difference (13.87% according to THIN® in 2021 vs. 15.87% according to the Spanish Ministry of Health in 2021), followed by the Nervous System (N) category (28.18% according to THIN® in 2021 vs. 29.24% according to the Spanish Ministry of Health in 2021) (**Table 2**).

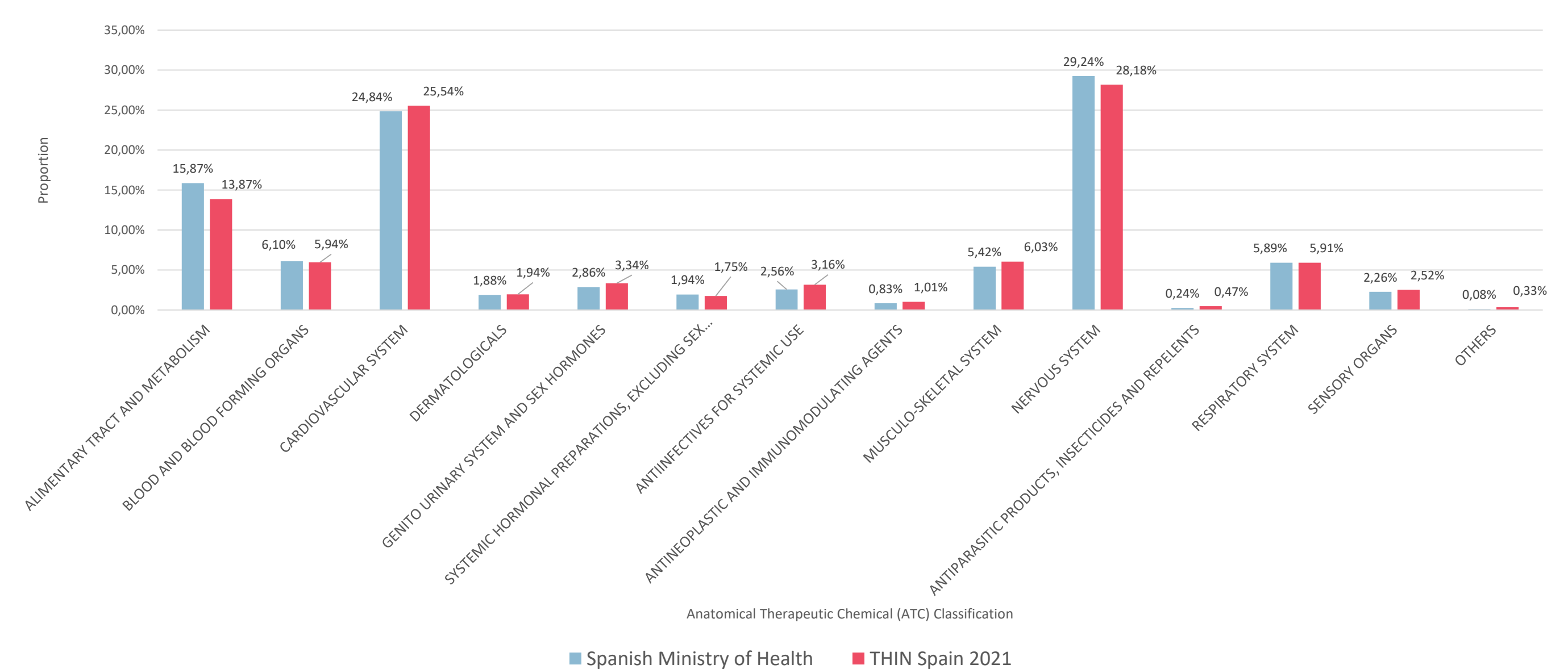


Table 2: Proportion of prescribed drugs dispensed by retail pharmacies based on the Anatomical Therapeutic Chemical 1 (ATC1) Classification - a comparison between the Spanish Ministry of Health's data and THIN® Spain's data for 2021.

The ATC classes of prescribed drugs most frequently dispensed by retail pharmacies are: Nervous system (N class) drugs, cardiovascular system (C class) drugs and Alimentary Tract and Metabolism (A class) drugs. These findings confirm the prevalence data for the major chronic diseases shown above.

The average proportion difference for ATC class drugs between THIN® and the Spanish Ministry of Health in 2021 was 0.49%.

