Coded Diagnoses Versus Proxy-Based Diagnoses in Primary Care: Registration By General Practitioners (GP) in the PHARMO GP Data

Jetty A. Overbeek[1], Brenda N. Baak[1], Ron M.C. Herings[1], Karin M.A. Swart-Polinder[1]

[1]PHARMO Institute for Drug Outcomes Research, Utrecht, Netherlands

Background

The quality of real-world data (RWD) is highly dependent on the nature of routine clinical practice. General practitioners (GPs) in the Netherlands are not obliged to enter a diagnosis in the electronic health records (EHR) by a coded diagnosis, which can consequently result in missed cases when conducting observational studies using RWD.

Objective

To quantify the extent to which diagnoses are registered by means of a diagnosis codes or a proxy using three use cases.

Methods

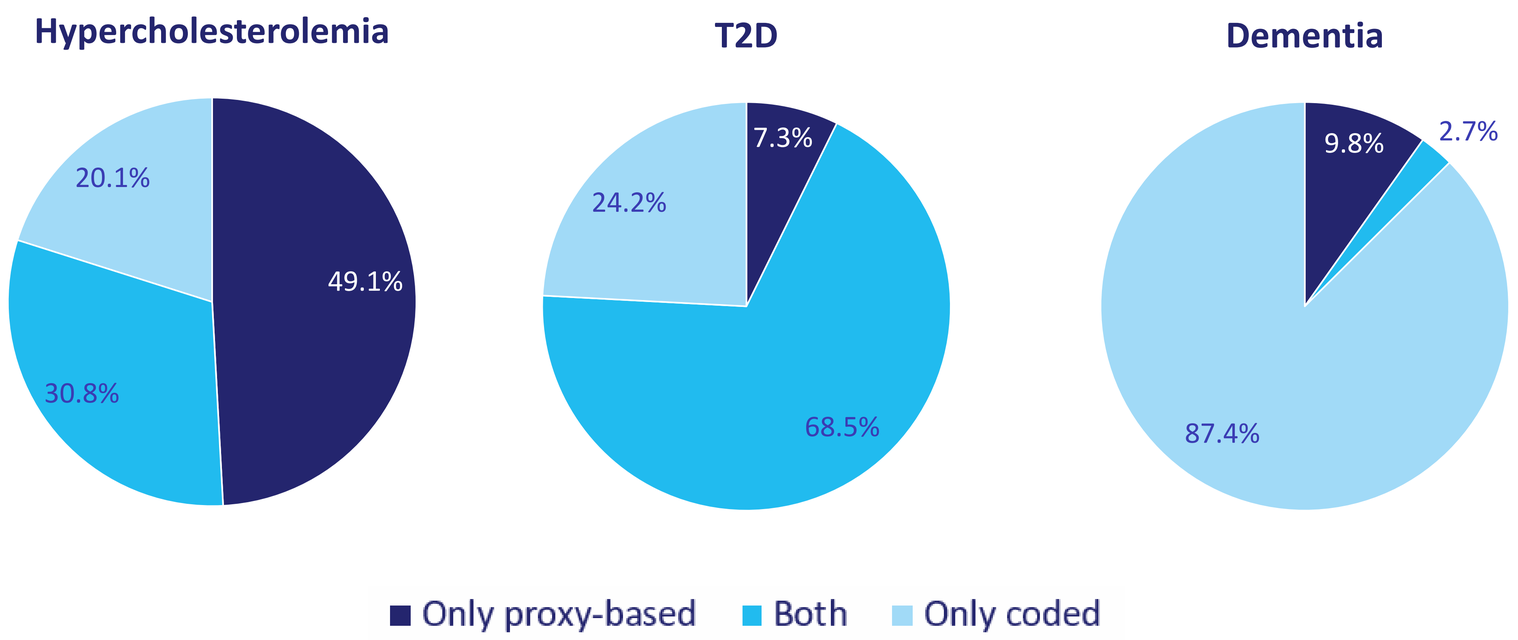
- A retrospective cohort design was used.
- The PHARMO GP Data was used to select all patients with a coded or proxy-based diagnosis of hypercholesterolemia, type 2 diabetes (T2D) and dementia.
- The PHARMO GP Data is part of the PHARMO Data Network and comprise EHRs registered by GPs. The records include information on diagnoses and symptoms, laboratory test results, referrals to specialists and healthcare product/drug prescriptions.
- International Classification of Primary Care (ICPC) coded diagnoses were compared with a proxy-based diagnosis (Table 1).

Table 1. Definition of coded and proxy-based diagnosis

	ICPC coded diagnosis	Proxy-based diagnosis
Hypercholesterolemia	T93	TC measurement ≥6.5 mmol/L
T2D	T90.02	≥2 non-insulin prescriptions used in diabetes for T2D (ATC code: A10B)
Dementia	P70	Unstructured free text notes for dementia

Results

- From all available GP data, the following numbers of patients were included:
  - 750,509 patients with hypercholesterolemia (150,897 with an ICPC coded diagnosis, 368,821 with a proxy-based diagnosis and 230,791 with both)
  - 318,491 patients with T2D (76,953 with an ICPC coded diagnosis, 23,261 with a proxy and 218,277 with both)
  - 69,036 patients with dementia (60,366 with an ICPC coded diagnosis, 6,781 with a proxy and 1,889 with both).
- A proxy-based diagnosis resulted in 49.1%, 7.3%, 9.8% more patients with hypercholesterolemia, T2D and dementia, respectively.



Conclusion

The results of this study suggest that additional information available in GP data, such as measurements, treatment and unstructured free text notes, should be considered to complement coded diagnoses when selecting diseases/symptoms. The added value depends on the diagnosis of interest. In the future, automation of structuring this additional information can improve data preparation enhance the quality of the research.

Funding: no funding was received