

Supplementary Appendices

Appendix 1a: Search strategy

Ovid Interface: MEDLINE

Search Date: 06/21/2024

Search Line	Search Terms	Results
1.	(Map or Mapped or Mapping or Mappings or Maps or Crosswalk* or "cross-walk*" or Convert* or Transcrib* or Conversion*).ti,kf.	218211
2.	("International classification of disease*" or ICD or ICD6 or "ICD-6" or ICD7 or "ICD-7" or ICD8 or "ICD-8" or "ICDA-8" or ICD9 or "ICD-9" or "ICD-9-cm" or ICD10 or "ICD-10" or "ICD-10-cm" or "ICD-10-CA" or "ICD-10-pcs" or "ICD-10-GM" or "ICD-10-AM" or ICD11 or "ICD-11" or "International Statistical Classification of Disease*").ti,ab,kf.	65359
3.	"International Classification of Diseases"/ or Clinical Coding/	11786
4.	2 or 3	70771
5.	1 and 4	472
6.	Limit 5 to English	404

Ovid Interface: EMBASE

Search Date: 06/21/2024

Search Line	Search Terms	Results
1.	(Map or Mapped or Mapping or Mappings or Maps or Crosswalk* or "cross-walk*" or Convert* or Transcrib* or Conversion*).ti,kf.	265746
2.	("International classification of disease*" or ICD or ICD6 or "ICD-6" or ICD7 or "ICD-7" or ICD8 or "ICD-8" or "ICDA-8" or ICD9 or "ICD-9" or "ICD-9-cm" or ICD10 or "ICD-10" or "ICD-10-cm" or "ICD-10-CA" or "ICD-10-pcs" or "ICD-10-GM" or "ICD-10-AM" or ICD11 or "ICD-11" or "International Statistical Classification of Disease*").ti,ab,kf.	145007
3.	"International Classification of Diseases"/ or Clinical Coding/	27283
4.	2 or 3	158388
5.	1 and 4	952
6.	limit 5 to EMBASE and English	544

Ovid Interface: Scopus

Search Date: 06/21/2024

Search Line	Search Terms	Results
1.	TITLE (map OR mapped OR mapping OR mappings OR maps OR crosswalk* OR "cross-walk*" OR convert* OR transcrib* OR conversion*)	758393
2.	KEY (map OR mapped OR mapping OR mappings OR maps OR crosswalk* OR "cross-walk*" OR convert* OR transcrib* OR conversion*)	1701917
3.	TITLE-ABS-KEY ("International classification of disease*" OR icd OR icd6 OR "ICD-6" OR icd7 OR "ICD-7" OR icd8 OR "ICD-8" OR "ICDA-8" OR icd9 OR "ICD-9" OR "ICD-9-cm" OR icd10 OR "ICD-10" OR "ICD-10-cm" OR "ICD-10-CA" OR "ICD-10-pcs" OR "ICD-10-GM" OR "ICD-10-AM" OR icd11 OR "ICD-11" OR "International Statistical Classification of Disease*")	109912
4.	1 OR 2	1998447
5.	3 AND 4 AND LIMIT-TO (LANGUAGE , "English")	1149

Ovid Interface: Web of Science Core Collection

Search Date: 06/21/2024

Search Line	Search Terms	Results
1.	TI= (Map or Mapped or Mapping or Mappings or Maps or Crosswalk* or "cross-walk*" or Convert* or Transcrib* or Conversion*)	664464
2.	AK= (Map or Mapped or Mapping or Mappings or Maps or Crosswalk* or "cross-walk*" or Convert* or Transcrib* or Conversion*)	405955
3.	#1 OR #2	887680
4.	TS= ("International classification of disease*" OR ICD OR ICD6 OR ICD-6 OR ICD7 OR ICD-7 OR ICD8 OR ICD-8 OR ICDA-8 OR ICD9 OR ICD-9 OR ICD-9-cm OR ICD10 OR ICD-10 OR ICD-10-cm OR ICD-10-CA OR ICD-10-pcs OR ICD-10-GM OR ICD-10-AM OR ICD11 OR ICD-11 OR "International Statistical Classification of Disease*")	64487
5.	(#3 AND #4) AND (LA==("ENGLISH"))	381

Supplementary Appendix 1b: Characteristics of the mapping methods and validation methods in the scoping review articles

First Author, Year	Source code	Target code	Research purpose	Direction of mapping	Mapping method	Mapping logic	One-to-many mapping approach	One-to-zero mapping approach	Use of dual-coded data for evaluating mapping quality	Validation method	Loss of information when mapping
Hamad et al. 2021	ICDA-8, ICD-9-CM, ICD-10-CA	ICDA-8, ICD-9-CM, ICD-10-CA	Research, surveillance	Bidirectional	Manual	Rule-based and expert knowledge-based	Handled via expert review	Excluded	No	Expert validation and descriptive analysis	Acknowledged, no steps taken
Columbo et al., 2018	ICD-9, ICD-10	ICD-9, ICD-10	Research, surveillance	Bidirectional	Manual	Rule-based and expert knowledge-based	Included	Reported, no steps taken	No	Expert validation and descriptive analysis	Acknowledged, no steps taken
Xu et al., 2022	ICD-10-CM	ICD-11	Research, clinical decision support	Forward	Manual	Rule-based and ontology-based	Included	Handled by truncating ICD-10-CM codes to broader ICD-10 codes	Yes	Expert validation and quantitative measures	Attempted to minimize by truncation, union of two mapping methods to improve coverage.
Mayhew et al., 2019	ICD-9-CM	ICD-10-CM	Research, surveillance	Forward	Manual	Rule-based and expert knowledge-based	Included	Excluded	No	Trend analysis of prevalence rates across the two coding periods	Acknowledged, no steps taken
Dalton et al., 2023	ICD-9	ICD-10	Research	Forward	Manual	Rule-based and expert knowledge-based	Included	Handled by manual review	No	Expert validation, trend analysis of observed-to-expected (O:E) ratios of admissions	Attempted to minimize by expert review, use of washout period
Turer et al., 2015	ICD-9-CM, ICD-10-CM	ICD-9-CM, ICD-10-CM	Research, billing	Forward, backward, bidirectional	Manual	Rule-based	Included	Included but no steps taken	Yes	Descriptive analysis, % of matched and mismatched.	Acknowledged, no steps taken
Pedersen et al., 2023	ICD-8	ICD-10	Research, clinical decision support	Forward	Manual	Rule-based and ontology-based	Excluded	Excluded	No	Expert validation	Attempted to minimize by expert review, mapping to highest granularity level
Ascensão et al., 2023	ICD-10-CM	ICD-9-CM	Research, surveillance	Backward	Manual	Rule-based	Not reported	Acknowledged	No	Trend analysis of hospital episodes across the two coding periods	Acknowledged, no steps taken
Colls et al., 1980	ICDA-8	ICD-9	Research, surveillance	Forward	Manual	Rule-based and expert knowledge-based	Handled via expert review	Handled via expert review	No	Expert validation	Acknowledged, no steps taken
Tian et al., 2020	ICD-9-CM	ICD-10-CM, ICD-10-PCS	Research, clinical	Bidirectional	Manual	Rule-based and expert knowledge-based	Included	Included via manual review	No	Trend analysis of proportion of the conditions across the two coding periods	Acknowledged, partly addressed by clinical validation

Sylvestre et al., 2020	French ICD-10, ICD-10-CM	French ICD-10, ICD-10-CM	Research	Bidirectional	Hybrid	Rule-based NLP and lexical similarity based	Included	Included but no steps taken	No	Expert validation, descriptive analysis	Acknowledged, allowed fallback to parent code and expert review
Panozzo et al., 2018	ICD-9-CM, ICD-10-CM	ICD-9-CM, ICD-10-CM	Research	Forward, backward, bidirectional	Manual	Rule-based	Included	Excluded	No	Trend analysis of the incidence rates of the health conditions of interest	Acknowledged, no steps taken
Noorbakhsh et al., 2022	ICD-9-CM	ICD-10-CM	Research, surveillance	Forward	Manual	Rule-based and ontology-based	Included	Acknow-ledged	No	Trend analysis of proportions of diagnosis codes identified in datasets across the two coding periods Expert validation	Acknowledged, no steps taken
Simeone et al., 2021	ICD-9-CM	ICD-10-CM	Research, billing	Forward	Manual	Rule-based and expert knowledge-based	Included	Not reported	No		Acknowledged, no steps taken
Fenton et al., 2014	ICD-9-CM	ICD-10-CM	Research, surveillance	Bidirectional	Manual	Rule-based and expert knowledge-based	Included	Included and manually recoded	Yes	Comparability factors and ratios	Acknowledged, no steps taken
Lee et al., 2021	KCD-7 (ICD-10-KM)	ICD-11	Research, health care planning	Bidirectional	Manual	Rule-based and lexical similarity based	Included, used post-coordination	Excluded	No	Expert and descriptive analysis	Attempted to minimize by using post-coordination system, expert review
Utter et al., 2018	ICD-9-CM	ICD-10-CM	Research	Forward, backward, reverse forward and reverse backward	Manual	Rule-based	Included, used expert review	Excluded after expert review	No	Expert validation	Attempted to minimize by multi-stage expert review and categorization of mapping complexity.
Schulz et al., 1998	ICD-9, ICD-10	ICD-9, ICD-10	Research, clinical use and billing	Bidirectional	Hybrid	Rule, lexical and vector-space text similarity based	Included, reduced to 1:1 by expert review	Included by labelling as "undefined"	No	Expert validation and descriptive analysis	Attempted to minimize by using pragmatic mappings and residual categories like "other conditions" or "unspecified"
Nam et al., 2021	ICD-9-CM, ICD-10-CM	ICD-9-CM, ICD-10-CM	Research, surveillance	Bidirectional	Manual	Rule-based and expert knowledge-based	Not reported	Not reported	No	Expert validation, trend analysis of the incidence and prevalence of health outcomes of interest across the two coding periods Expert validation	Acknowledged, no steps taken
Nikiema et al., 2017	ICD-10	ICD-O3	Research, oncology surveillance	Bidirectional	Hybrid	Ontology-based (OWL-EL), rule-based, morphosyntactic and logical reasoning.	Used post-coordination method	Excluded	No		Attempted to minimize by a combination of filtering, expert review, semantic integration leveraging local definitions

He et al., 2020	ICD-9	ICD-10	Surveillance , Research	Bidirectional	Manual	Rule-based	Included	Not reported	No	Expert validation and trend analysis	Acknowledged, no steps taken
Hernandez-Ibarburu et al., 2019	ICD-9-PCS	ICD-10-PCS	Clinical research	Bidirectional	Manual	Rule-based	Included, used expert review to reduce one-to-many mapping	Excluded	No	Expert validation	Acknowledged, partly addressed by expert review
Hernandez-Ibarburu et al., 2019	ICD-9	ICD-10	Clinical research	Forward	Manual	Rule-based and hierarchy-based	Included by grouping	Included via manual review	No	Expert validation and descriptive analysis	Attempted to minimize by inserting ICD-9 codes into ICD-10 hierarchy and adjusting mappings
Venepalli et al., 2014	ICD-9-CM	ICD10-CM	Billing	Forward	Manual	Rule-based and network-based	Included	Included but no steps taken	No	Expert validation	Acknowledged, partly addressed by identification and categorization
Colliver et al., 1984	ICD-9	ICDA-8	Research, surveillance	Backward	Manual	Rule-based	Expert review to reduce one-to-many mapping	Not reported	No	Expert review ad trend analysis	Acknowledged, partly addressed by expert review

Supplementary Appendix 1c: Summary of mapping challenges and strategies to handle ambiguity and loss of information in reviewed articles

Ambiguous Mappings	Strategies to handle ambiguous mappings (frequency)	Strategies to handle loss of information when mapping (frequency)
One-to-many mappings	- Retained without modification (n=11) - Grouping (n=6) - Reduced to single code mapping via expert review (n=4) - Not reported (n=2) - Handled using post-co-ordination method (n=1) - Excluded from analysis (n=1)	- None (n=12) - Expert review only (n=6) - Expert review in combination with grouping, allowing to fallback to parent code, and post-coordination (n=4) - Expert review and categorization of mapping complexity (n=2) - Expert review and mapping to highest granularity level (n=1)
One-to-zero mappings	- Excluded from analysis (n=8) - Retained without modification (n=6) - Handled via expert review (n=5) - Not reported (n=4) - Labelled as undefined (n=1) - Truncation to parent code (n=1)	- None (n=21) - Expert review and/or adjusted mappings (n=4)

The summary shows that 84% (n = 21) of the articles did not adopt any strategy to account for the loss of information arising from one-to-zero mappings. On the other hand, the loss of information arising from one-to-many mappings was *acknowledged* in all the included articles, but 48% (n = 12) of them *did not attempt to mitigate the loss of information*.