

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

The analyses and identification of ICD codes for dementias in the research based on NHIRD: A scoping review protocol.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2022-062654
Article Type:	Protocol
Date Submitted by the Author:	09-Mar-2022
Complete List of Authors:	Shih, Ying-Jyun; Asia University College of Medical and Health Science, Department Of Healthcare Administration; Tsaotun Psychiatric Center Ministry of Health and Welfare, Department of Nursing Wang, Jiun-Yi; Asia University, Department of Healthcare Administration Wang, Ya-Hui; Tsaotun Psychiatric Center Ministry of Health and Welfare, Department of Nursing Shih, Rong-Rong; Tsaotun Psychiatric Center Ministry of Health and Welfare, Department of Nursing Yang, Yung-Jen; Institute of Education Social Science Research Unit
Keywords:	Dementia < NEUROLOGY, Old age psychiatry < PSYCHIATRY, EPIDEMIOLOGY, Information management < BIOTECHNOLOGY & BIOINFORMATICS

SCHOLARONE™
Manuscripts

1
2
3 **The analyses and identification of ICD codes for dementias in the researches based on**
4 **NHIRD: A scoping review protocol.**
5
6
7

8 Ying-Jyun Shih¹⁻², Jiun-Yi Wang¹, Ya-Hui Wang², Rong-Rong Shih², Yung-Jen Yang^{3*}
9

10
11 *Corresponding author: Yung-Jen Yang, email: yung-jen.yang.12@ucl.ac.uk
12

13 Affiliation of corresponding author: UCL Social Research Institute, Institute of Education,
14 University College London.
15
16

17
18 ¹Department of Healthcare Administration, College of Medical and Health Science, Asia
19 University, Taichung 413305, Taiwan.
20

21 ²Department of Nursing, Tsaotun Psychiatric Center, Ministry of Health and Welfare, Nantou
22 54249, Taiwan.
23

24 ³UCL Social Research Institute, Institute of Education, University College London, London
25 WC1H 0AL, United Kingdom.
26
27
28
29
30

31 Running title: ICD codes for dementias in database research
32
33

34 Key words: *Dementia, Scoping review, ICD-codes, NHIRD*
35
36

37
38 Word counts: 2816
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Introduction:

Studies based on health claim data (HCD) have been increasingly adopted to in medical research for its strengths in sample size and large amount of information, and Taiwan National Health Insurance Research Database (NHIRD) has been widely used in many disciplines, including dementia. How the diagnostic codes are applied to define the diseases/conditions of interested is pivotal in HCD-related researches, and there is also a lack of consensus on which diagnostic codes can be most appropriately in defining dementias in NHIRD. The objectives of this scoping reviews are (1) to investigate the relevant characteristics in the published reports targeting dementias based on the NHIRD, and (2) to address the diversities by a case study.

Methods and analysis:

This scoping review protocol will follow the methodological framework of the Joanna Briggs Institute Reviewer's Manual and the guidance of the PRISMA extension for scoping reviews. The review will be performed between March and August 2022 in 5 stages, including identifying the relevant studies, developing search strategy, individually screening and selecting evidence, collecting and extracting data, and summarising and reporting the results. We will search the electronic databases of MEDLINE, EMBASE, CENTRAL, CINAHL, and PsychInfo, Airiti Library Academic Database, the National Health Insurance Administration's repository, and Taiwan Government Research Bulletin (GRB). We will perform narrative discussions based on the results to address the mapping and research questions, and will discuss the prevalence across the individual study as a case study.

Ethics and dissemination:

Our scoping review is a review of the published reports and ethical approval is not required. The

1
2
3 results will provide a panorama of the dementia studies based on NHIRD. We will disseminate
4
5
6 our results through peer-reviewed journals and conferences, and share with stakeholders by
7
8 distributing the summaries in social media and emails.
9

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Article summary

Strengths and limitations of this study

- To the best of our knowledge, our study will be the first review of the characteristics and the use of diagnostic codes in the dementia studies based on NHIRD.
- The scoping review methodology allows a broad perspective to depict the heterogeneities across the individual studies while maintaining transparency and accountability through systematic search and data extraction process.
- The results of this scoping review are expected to lay the foundation for future dementia studies based on NHIRD.
- As the review is limited to literature published in English and Chinese, there are still potential publications in other language not considered.
- Due to the lacked consensus on the risk of bias tool, a formal quality and risk of bias assessment of the included literature will not be performed, limiting the understanding of the influence of potential biases in the studies.

Introduction

Health claim data (HCD) have been a major source for the research of epidemiological analyses, health service economics and outcomes research, and disease-specific medical research.[1] Researches based on large population-based database enable researchers to explore medical conditions with low prevalence or interventions with smaller effect sizes.[2] HCDs have been applied in many medical specialities including the field of dementia, and many prestige organisations have advocated that the big data, specifically the HCDs, be applied in the research of dementia.[3, 4] In practice, HCDs have been applied in the research of dementias and their comorbidities and the disease progression trajectory and the interaction between the biological and environmental factors.[3]

However, due to the nature of administrative data or claim data which are usually presumed for reimbursement rather than for research,[5, 6] scholars have proposed important issues that may pose threats and limitations in such studies, including the quality and reliability of the data,[7, 8] the lack of consensus and standardisation across the databases, and the doubt of data accuracy as a result of erroneous linkage.[9, 10]. Furthermore, despite of being a powerful tool, researches on HCDs are still vulnerable to the under-recording of dementia, resulting in under-representativeness of the target group and threats to generalisability of the study results.[11]

As the main components of HCDs, diagnostic codes play an important role in HCD studies, and, by following appropriate algorithms, researchers are able to aim at specific health outcomes of interest or identify populations with specific diagnoses within the HCDs.[12, 13] However, the quality of diagnostic codes fundamentally relies on the avoidance of misdiagnosis, miscoding and misclassification, which will otherwise limit or even flaw the results as mentioned by Stein et al in their systematic review.[14] The issues related with diagnostic codes in dementia database research become more complicated. For example, the diagnostic gap due to under-identification caused by miscoding of dementia[15], misidentification of dementia in HCDs,[16] misclassification of dementia,[17] and the high heterogeneity in selecting the diagnostic codes of dementia[18] is not uncommon and these all probably bring about complexities in dementia studies based on database approach. As differences in selecting diagnostic codes to define dementias in relevant research would result in misidentification of dementia, hence, it is imperative to develop a set of standardised diagnostic codes for dementia to minimise the potential problematic impact and improve the value of HCDs in dementia research.

Scoping review is a relatively young methodology in the family of evidence synthesis, and is

1
2
3 regarded most appropriate to explore, configurate, and aggregate the evidence as well as used as
4 a precursor to a systematic review. Scoping reviews are able to demonstrate the way how
5 researches have been executed in specific topics or fields, and to identify the key concepts,
6 rationales, types of evidence and the research gap. Rather than testing theories or hypotheses,
7 they serve to explore the contents, range, natures, and heterogeneities in the individual studies,
8 summarising the results, and guide the researchers for the future research about the directions
9 and methodologies.[19-21] The methodologies of scoping review have been first proposed by
10 Arksey and O'Malley in 2005,[22] and strengthened by Levac et al by proposing a practical
11 5-step approach.[23] Despite their endeavours, inconsistencies in execution remain. The Joanna
12 Briggs Institute (JBI) developed a comprehensive guideline to standardise the processes first in
13 2014 and updated revisions in both 2017 and 2020.[24] Tricco et al also proposed the Preferred
14 Reporting Items for Systematic Reviews and Meta-Analyses extension for scoping review
15 (PRISMA-ScR) to aid the researchers in reporting their studies.[25] Considering the complex
16 nature in the selections of diagnostic codes to define dementia and the high heterogeneities in
17 the database research, scoping review would be the best appropriate methodology in response to
18 our research questions.

19
20 Since 1995, Taiwan has launched the Taiwan National Health Insurance with an extremely high
21 coverage, 99.9%, of its population and established an HCD cumulating the health-related record
22 of literally the whole 23 million population[26, 27]. Thanks to its abundant information, the
23 National Health Insurance Administration has built up the National Health Insurance Research
24 Database (NHIRD) which has been adopted by the academics and researchers as a resource of
25 secondary database research for medical research with hundreds of publications, including the
26 research for dementia, and as references for healthcare practice guidance and public
27 policymaking. Despite of its strengths, however, NHIRD also bears the same inherent
28 weaknesses of HCDs and there have been heterogeneities in selecting the codes in defining
29 dementia in the researches using NHIRD database. As a result, it is imperative to investigate the
30 characteristics of dementia researches based on NHIRD database, and how the diagnostic codes
31 are selected as well as used in such studies, which will aid in identifying the potential research
32 gaps and reduce the research waste.

33
34 In the present study, by taking the advantages of scoping review methodologies, the research
35 team intends to identify the characteristics, addressing the heterogeneities, and explore the
36 diversities of diagnostic codes used to define dementias in the published studies. With the
37 results of our study, we may lay the foundation for developing a set of standardised codes for
38 defining dementias for future dementia studies based on NHIRD. In this manuscript, we present
39 the protocol for informing the implementation of scoping review.

Aim of the study and research question (study objective)

Based on empirical experience, there are high heterogeneities in NHIRD studies, including the selections of diagnostic codes for defining diseases/conditions, the size of database, time length of the database used, and types of the sub-datasets (i.e., inpatient or outpatient dataset). The main aims of our research are to investigate the relevant characteristics in the published reports targeting dementias based on the NHIRD, and to address the diversities by using the reported prevalence as a case study. We here defined the following research questions for the scoping review.

1. Which diagnostic codes in ICD-9-CM and ICD-10-CM were used to define dementias in the studies based on NHIRD?
2. To what extent the diagnostic codes varied across the studies when being used to define dementias, in relation with the research teams and the size and types of database?
3. What differences in terms of inclusion or exclusion criteria and the time length of database were adopted across the studies?
4. How other important publication characteristics of database studies were reported across the studies?
5. As a case study, how does the prevalence differ across the studies based on NHIRD in relation to the major variables above?

Methods

The present scoping review protocol has been developed based on the methodological framework of the Joanna Briggs Institute (JBI) Reviewer's Manual and has been constructed following the guidance of the PRISMA extension for scoping reviews (PRISMA-ScR).[20, 21, 25] A PRISMA-ScR checklist can be found in the Supplementary file 1. This protocol has been registered with the Open Science Framework (OSF: osf.io/fc65g) on 26 February, 2022, and the study will be implemented between March and August 2022 by following the 5 steps (subsections 2.1 to 2.5) :

2.1 Identification of relevant studies.

2.1.1 Inclusion criteria

(1) Participants

The scoping review will target the published research reports on dementias using NHIRD in the literature, and will include all types of dementia, defined and identified in any way in the reports. There will be no restriction on the type of study designs, age of the participants in the reports and comorbidities so as to maximise the coverage of the type of dementia research, the dementia at any age, specifically the young onset dementia (YOD), and the

1
2
3 comorbidity studies of dementia. However, we will only include the reports written in
4 English or Chinese for the ease of data collection and analyses.
5

6 **(2) Concept**

7
8 The present scoping review intends to focus on dementia researches using NHIRD, and
9 explore as well as configure the elements of the researches, including the ways of selecting
10 diagnostic codes to define dementias, the usage of sub-datasets of NHIRD, the
11 methodological spectra of dementia research based on NHIRD, the heterogeneity, and the
12 potential impact on the outcomes.
13
14
15

16 **(3) Context**

17
18 The scoping review will specifically focus on the dementia studies based on NHIRD which
19 would be influenced by the contextual factors of the pragmatic medical practices and
20 services in Taiwan as well as some of the specific regulations in Taiwan National Health
21 Insurance. Rather than a constraint, however, this is the unique contextual characteristic
22 that this scoping review intends address and can inform.
23
24
25

26 **2.1.2 Exclusion criteria**

27
28 We will exclude the reports when the full texts are not readily available, and, in addition,
29 review articles, study protocols, grey literatures and texts that are not peer-reviewed are
30 excluded. The grey literatures here included letters, editorials or leading articles,
31 commentaries, conference abstracts or presentations, and dissertations or theses.
32
33
34

35 **2.2 Search strategy**

36
37 The research team will develop search strategies that aim to include the study reports published
38 between 2002 and 2022, which ranged from the first year when the NHIRD was available to the
39 academic researchers and the time executing this review. Published studies in English or
40 Chinese with all study designs will be included and the research team will search major
41 electronic databases of MEDLINE-OVID, EMBASE-OVID, Cochrane Central Register of
42 Controlled Trials (CENTRAL), CINAHL, and PsychInfo to identify the reports on the topic. For
43 the Chinese literatures and grey literatures, we will search the Airiti Library Academic
44 Database, the National Health Insurance Administration's repository which collects the articles
45 using Taiwan NHIRD, and Taiwan Government Research Bulletin (GRB). The search will
46 consist of three major steps, and, first, we will develop search terms informing the three main
47 concepts of the review, including dementia, database research, and Taiwan National Health
48 Insurance Research Database. The identified search terms of each concept will be first be
49 organised with the Boolean operator OR, and then the three categories will be combined with
50 the Boolean operator AND. (Table 1) The search strategies will be developed by a trained
51 researcher (YJY) and then reviewed by a librarian in the university library. The second step is to
52
53
54
55
56
57
58
59
60

use all the identified keywords and index terms to undertake search in the titles and abstracts of relevant articles in each electronic database. The final step is to export searched results will be exported to the electronic bibliographic software EndNote X8. The second and third steps will be individual carried out by 2 different researchers (YJS and YJW). The pilot search strategy is shown in the Supplementary file 2. In addition to the electronic database search, we will conduct hand searches for the potential studies in the reference lists of the articles initially identified when performing full-text reviewing.

Table 1 Search terms and concepts

Concept 1: Dementia	“Dementia” “Major neurocognitive disorder” “Alzheimer’s dementia” “Parkinson’s disease dementia” “Dementia with Lewy body” “Frontotemporal dementia” “Vascular Dementia” “Mixed dementia”
Concept 2: Database research	“Claim data” “Insurance claim data” “Insurance data/database” “Administrative (or health) data/database” “Health data” “Electronic health record” “Government database”
Concept 3: Taiwan national health insurance	“Taiwan” “National Health Insurance Research Database (NHIRD)” “National Health Programme/ Service” “National data (or nationwide or population-based) research” “Taiwan Longitudinal (or follow-up) Study”

2.3 Selection process (evidence screening and selection)

Following the search, all identified citations will be collated and uploaded into the bibliographic software EndNote X8 (Clarivate Analytics, PA, USA) where duplicates will be removed. Two reviewers (YJS and JYW) will independently screen potentially eligible studies according to the inclusion criteria in our review by screening titles and abstracts yielded by our comprehensive

1
2
3 search. The selection results by the two individual researchers will then be compared and
4 merged. Any discrepancy in the results will be solved through discussion with involvement of a
5 third reviewer (YJY). Reasons for exclusion of full-text studies that do not meet the inclusion
6 criteria and the results of the search in each step will be reported in detail in the final scoping
7 review and presented in PRISMA flowchart diagram.
8
9
10

11 12 13 **2.4 Data collection process**

14 The research team will develop an electronic data extraction form by Microsoft Excel and will
15 be independently pilot tested by all team members for its applicability, achieving the final
16 version based on the feedback. The targeted data to extract will include the characteristics of the
17 reports such as the diagnostic codes to define dementias, type of dementia reported, study
18 designs, time length, subsets of database, research method used, whether additional methods to
19 increase the likelihood of dementia diagnosis were used, statistical methods and outcomes.
20 Other general profiles of the reports including lead authors, affiliations, main discipline of the
21 research team, year of publication, funding, journals where published and other potential factors.
22 (See Table 2) When extracting the data for the review, a reviewer (YJS) will assess each eligible
23 study and then input the assessed results in the prespecified data extraction form. Another
24 review (JYW) will then randomly examine 80% of the extracted results, and any inconsistency
25 will be resolved through consensus with the involvement of the third reviewer (YJY).
26
27
28
29
30
31
32
33
34
35

36 **Table 2. Data extraction table.**

Article information	Title Authors Affiliations with author(s) Main disciplines of lead author Year of publication (based on the first identifiable date) Journal where published and the JCR impact factor Language of publication Funding Other potential factors
Population	Sample size Inclusion and exclusion criteria of the individual study Age of the participants Sex of participants Diagnostic codes to define dementias

	Type of dementia reported Comorbidities or exposures investigated
Methods	Study designs reported Research method used Time length/ follow-up time Subsets of database Whether and what additional methods to increase the likelihood of dementia diagnosis used Statistical methods
Outcomes	The number of dementias identified Prevalence, if any Main outcomes of the studies and the estimation such as incidence and mortality if reported Other reported health-related outcomes

2.5 Collating, summarising and reporting the results

As the scoping review intends to map the characteristics of the published reports and to address the diversities of the specific outcome for case study, we will not perform critical appraisal of each individual study. The extracted data will be analysed by two researchers who will work at the same time by following a prespecified analytical plan. In responding to the research questions, the variables mentioned above will first be summarised quantitatively, and descriptive statistics including number counts, frequencies and rates will be used to summarise the results. Then the researchers will perform narrative discussions based on the results to address the mapping and research questions that cannot be answered quantitatively. Considering the vast differences in reporting prevalence (i.e., non-reporting, different time frame or ways of calculation) in individual studies, we will not attempt to pool the prevalence with statistical methods by our own, and, instead, the reporting conditions of prevalence in each individual study will be presented with potential reasons discussed. Finally, we will abide by the PRISMA-ScR guideline[25] to ensure the reporting standard of the final manuscript, and submit to a peer-reviewed journal.

Patient and public involvement

Our scoping review will not plan to involve patients or the public, but we will also invite clinicians who have to type in the codes for dementia in their practices in the review process, consulting, discussing and verifying the results.

Ethic and dissemination

Our scoping review is a review of the published reports which do not involve access to individual data, and ethical approval is not required. The results will be disseminated through peer-reviewed journals and conferences focusing on dementia and related topics. We will also share our results with stakeholders, including non-government organisations for dementia and health informatics policymakers, by distributing our results in brief and plain language in social media and emails. If there to be any amendment to the protocol, the revised information will be disclosed on the protocol registry space (OSF: osf.io/fc65g) and be stated in any future publication based on the review.

Discussion

The proposed scoping review aims to investigate and address the diversities in the published reports targeting dementias based on the NHIRD, and to discuss the potential influence through a case study of the reported prevalence. To the best of our knowledge, our study will be the first review of the studies based on NHIRD. We believe that, by the strengths of scoping review methodology, a panorama view of the dementia studies based on NHIRD database will inform the researchers of the *status quo* of the research in its field and thereby improve the quality and value of database research by taking the advantages of NHIRD.

Additional Headings

Data statement

Technical appendix, search strategies, electronic data extraction sheets, statistical codes and datasets are available from the OSF repository, DOI: [10.17605/OSF.IO/YAQCM](https://doi.org/10.17605/OSF.IO/YAQCM).

Author Contributions

YJS and YJY conceptualise the research. YJS, JYW, YHW and RRS helped the development of ideas and drafted of the initial protocol. YJY supervised and revised the manuscript. All authors read and agreed the final manuscript.

Author Statement

Acknowledgement

We especially appreciate the help from the staff in the Department of Healthcare Administration, College of Medical and Health Science and the librarians in Asia University.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of interests

We declared no conflict of interests.

Patient consent for publication

Not required.

References

1. Konrad R, Zhang W, Bjarndóttir M, Proaño R. Key considerations when using health insurance claims data in advanced data analyses: an experience report. *Health Systems (Basingstoke, England)* 2019;9(4):317-25.
2. Thygesen LC, Ersbøll AK. When the entire population is the sample: strengths and limitations in register-based epidemiology. *Eur J Epidemiol* 2014;29(8):551-8.
3. OECD. *Dementia Research and Care: Can Big Data Help? Big data approaches to dementia: Opportunities and challenges*. Anderson G, J.Oderkirk, editors. Paris: OECD Publishing; 2015.
4. Pike KJ, Fazio S, Bynum JPW, Trivison TG, Wendler D, Mor V. Resources, methods, and data infrastructure to promote research in dementia care, caregiving, and services. *J Am Geriatr Soc* 2021;69(7):1793-800.
5. Reinke C, Buchmann N, Fink A, Tegeler C, Demuth I, Doblhammer G. Diabetes duration and the risk of dementia: a cohort study based on German health claims data. *Age Ageing* 2021;51(1).
6. Shi L, Steenland K, Li H, Liu P, Zhang Y, Lyles RH, et al. A national cohort study (2000–2018) of long-term air pollution exposure and incident dementia in older adults in the

- 1
2
3 United States. *Nat Commun* 2021;12(1):6754.
- 4
5 7. Feder SL. Data Quality in Electronic Health Records Research: Quality Domains and
6 Assessment Methods. *Western journal of nursing research* 2018;40(5):753-66.
- 7
8 8. Rabinstein AA. Administrative Medical Databases for Clinical Research: The Good, The
9 Bad, and The Ugly. *Neurocrit Care* 2018;29(3):323-5.
- 10
11 9. Bohensky MA, Jolley D, Sundararajan V, Evans S, Pilcher DV, Scott I, et al. Data linkage:
12 a powerful research tool with potential problems. *BMC Health Serv Res* 2010;10:346-.
- 13
14 10. Harron K, Dibben C, Boyd J, Hjern A, Azimae M, Barreto ML, et al. Challenges in
15 administrative data linkage for research. *Big data & society* 2017;4(2):2053951717745678.
- 16
17 11. Davis KAS, Mueller C, Ashworth M, Broadbent M, Jewel A, Molokhia M, et al. What gets
18 recorded, counts: dementia recording in primary care compared with a specialist database.
19 *Age Ageing* 2021;50(6):2206-13.
- 20
21 12. Jaakkimainen RL, Bronskill SE, Tierney MC, Herrmann N, Green D, Young J, et al.
22 Identification of Physician-Diagnosed Alzheimer's Disease and Related Dementias in
23 Population-Based Administrative Data: A Validation Study Using Family Physicians'
24 Electronic Medical Records. *J Alzheimers Dis* 2016;54(1):337-49.
- 25
26 13. Lipscombe LL, Hwee J, Webster L, Shah BR, Booth GL, Tu K. Identifying diabetes cases
27 from administrative data: a population-based validation study. *BMC Health Serv Res*
28 2018;18(1):316.
- 29
30 14. Stein JD, Lum F, Lee PP, Rich WL, 3rd, Coleman AL. Use of health care claims data to
31 study patients with ophthalmologic conditions. *Ophthalmology* 2014;121(5):1134-41.
- 32
33 15. Russell P, Banerjee S, Watt J, Adleman R, Agoe B, Burnie N, et al. Improving the
34 identification of people with dementia in primary care: evaluation of the impact of primary
35 care dementia coding guidance on identified prevalence *BMJ Open*. 2013;3(12):e004023.
- 36
37 16. Zhu CW, Ornstein KA, Cosentino S, Gu Y, Andrews H, Stern Y. Misidentification of
38 Dementia in Medicare Claims and Related Costs. *J Am Geriatr Soc* 2019;67(2):269-76.
- 39
40 17. Health AIO, Welfare. *Predicting early dementia using Medicare claims: a feasibility study*
41 *using the National Integrated Health Services Information Analysis Asset*. Canberra:
42 AIHW; 2021.
- 43
44 18. Wilkinson T, Ly A, Schnier C, Rannikmäe K, Bush K, Brayne C, et al. Identifying
45 dementia cases with routinely collected health data: A systematic review. *Alzheimers*
46 *Dement* 2018;14(8):1038-51.
- 47
48 19. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review
49 or scoping review? Guidance for authors when choosing between a systematic or scoping
50 review approach. *BMC Med Res Methodol* 2018;18(1):143.
- 51
52
53
54
55
56
57
58
59
60

- 1
- 2
- 3 20. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated
- 4 methodological guidance for the conduct of scoping reviews. *JBI Evid Synth*
- 5 2020;18(10):2119-26.
- 6
- 7
- 8 21. Tricco AC, Lillie E, Zarin W, O'Brien K, Colquhoun H, Kastner M, et al. A scoping
- 9 review on the conduct and reporting of scoping reviews. *BMC Med Res Methodol*
- 10 2016;16(1):15.
- 11
- 12
- 13 22. Arksey H, O'Malley L. Scoping studies: towards a methodological framework.
- 14 *International Journal of Social Research Methodology* 2005;8(1):19-32.
- 15
- 16 23. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology.
- 17 *Implement Sci* 2010;5:69.
- 18
- 19 24. Khalil H, Bennett M, Godfrey C, McInerney P, Munn Z, Peters M. Evaluation of the JBI
- 20 scoping reviews methodology by current users. *International journal of evidence-based*
- 21 *healthcare* 2020;18(1):95-100.
- 22
- 23 25. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA
- 24 Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern*
- 25 *Med* 2018;169(7):467-73.
- 26
- 27 26. Lin LY, Warren-Gash C, Smeeth L, Chen PC. Data resource profile: the National Health
- 28 Insurance Research Database (NHIRD). *Epidemiol Health* 2018;40:e2018062.
- 29
- 30 27. Hsieh CY, Su CC, Shao SC, Sung SF, Lin SJ, Kao Yang YH, et al. Taiwan's National
- 31 Health Insurance Research Database: past and future. *Clin Epidemiol* 2019;11:349-58.
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

● **PRISMA-ScR Checklist.**

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	P1
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	P3-P4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	P5-P6
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	P7
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	P7
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	P7-P8
Information sources	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	P8
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	Supplementary table
Selection of sources of evidence	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	P9-P10
Data charting process	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	P10
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	P10-P11, Table 2
Critical appraisal of individual sources of evidence	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	We will not perform critical appraisal due to the nature of the study.
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	P11
RESULTS			
Selection of	14	Give numbers of sources of evidence screened, assessed	N/A, This is a

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
sources of evidence		for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	protocol.
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	N/A, This is a protocol.
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	N/A, This is a protocol.
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	N/A, This is a protocol.
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	N/A, This is a protocol.
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	N/A, This is a protocol.
Limitations	20	Discuss the limitations of the scoping review process.	N/A, This is a protocol.
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	N/A, This is a protocol.
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	We did not receive any grant.

From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA ScR): Checklist and Explanation. *Ann Intern Med.* 2018;169:467–473. doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850).

● **Supplementary material:**

Table: The pilot search strategy in Embase-Ovid.

#	Search strategy
1	exp dementia/
2	dement*.ti,ab.
3	"major neurocognitive disorder".ti,ab.
4	exp Alzheimer Disease/
5	Alzheimer*.ti,ab.
6	(Alzheimer* adj2 dement*).ti,ab.
7	AD.ti,ab.
8	(Parkinson* adj2 dementia).ti,ab.
9	(lewy* adj2 bod*).ti,ab.
10	"frontotemporal lobe degeneration".ti,ab.
11	exp Frontotemporal Lobar Degeneration/
12	FTD.ti,ab.
13	FTLD.ti,ab.
14	"Pick* disease".ti,ab.
15	(Pick* adj2 disease).ti,ab.
16	"vascular dementia".ti,ab.
17	VAD.ti,ab.
18	"muti-infarct".ti,ab.
19	"mixed dementia".ti,ab.
20	exp AIDS Dementia Complex/
21	exp Dementia, Vascular/
22	exp Lewy Body Disease/
23	or 1-22
24	Routinely Collected Health Data/
25	(Claim* adj2 data).ti,ab.
26	(Insurance adj2 claim*).ti,ab.
27	(Insurance adj2 data).ti,ab.
28	(Insurance adj2 reimburse).ti,ab.
29	"administrati* data*".ti,ab.
30	"administrati* health* database*".ti,ab.
31	"Health* data source*".ti,ab.
32	(health* adj3 administrati* adj3 data*).ti,ab.
33	"electr* health* record*".ti,ab.
34	"govern* administrati*".ti,ab.
35	or 24-34
36	"National Health Insurance Research Database".ti,ab.

37	NHIRD.ti,ab.
38	"national health insurance".ti,ab.
39	"health insurance".ti,ab.
40	"national health program*".ti,ab.
41	"national health service*".ti,ab.
42	NHI.ti,ab.
43	"national data".ti,ab.
44	nationwide.ti,ab.
45	"insurance, national health".ti,ab.
46	("population-base*" or "population* base*").ti,ab.
47	"Taiwan* longitudinal stud*".ti,ab.
48	("follow-up stud*" or "follow* up stud*").ti,ab.
49	retrospective.ti,ab.
50	or 36-49
51	Taiwan*.ti,ab.
52	23 and 35 and 50 and 51

BMJ Open

The analyses and identification of ICD codes for dementias in the research based on the NHIRD: A scoping review protocol.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2022-062654.R1
Article Type:	Protocol
Date Submitted by the Author:	14-Jun-2022
Complete List of Authors:	Shih, Ying-Jyun; Asia University College of Medical and Health Science, Department Of Healthcare Administration; Tsaotun Psychiatric Center Ministry of Health and Welfare, Department of Nursing Wang, Jiun-Yi; Asia University, Department of Healthcare Administration Wang, Ya-Hui; Tsaotun Psychiatric Center Ministry of Health and Welfare, Department of Nursing Shih, Rong-Rong; Tsaotun Psychiatric Center Ministry of Health and Welfare, Department of Nursing Yang, Yung-Jen; Institute of Education Social Science Research Unit
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Epidemiology, Health informatics, Neurology
Keywords:	Dementia < NEUROLOGY, Old age psychiatry < PSYCHIATRY, EPIDEMIOLOGY, Information management < BIOTECHNOLOGY & BIOINFORMATICS

SCHOLARONE™
Manuscripts

1
2
3 **The analyses and identification of ICD codes for dementias in the research based**
4 **on the NHIRD: A scoping review protocol.**
5
6
7

8 Ying-Jyun Shih¹⁻², Jiun-Yi Wang¹, Ya-Hui Wang², Rong-Rong Shih², Yung-Jen
9 Yang^{3*}
10

11
12 *Corresponding author: Yung-Jen Yang, email: yung-jen.yang.12@ucl.ac.uk
13 Affiliation of corresponding author: UCL Social Research Institute, Institute of
14 Education, University College London.
15
16

17
18 ¹Department of Healthcare Administration, College of Medical and Health Science,
19 Asia University, Taichung 413305, Taiwan.
20

21 ²Department of Nursing, Tsaotun Psychiatric Center, Ministry of Health and Welfare,
22 Nantou 54249, Taiwan.
23

24 ³UCL Social Research Institute, Institute of Education, University College London,
25 London WC1H 0AL, United Kingdom.
26
27

28
29
30 Running title: ICD codes for dementias in database research
31

32
33 Key words: *Dementia, Scoping review, ICD-codes, NHIRD*
34

35
36 Word counts: 2836
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Introduction:

Studies based on health claim data (HCD) have been increasingly adopted in medical research for their strengths in large sample size and abundant information, and the Taiwan National Health Insurance Research Database (NHIRD) has been widely used in medical research across disciplines, including dementia. How the diagnostic codes are applied to define the diseases/conditions of interested is pivotal in HCD-related research, but the consensus on the issue that which diagnostic codes most appropriately define dementias in the NHIRD is lacking. The objectives of this scoping review are (1) to investigate the relevant characteristics in the published reports targeting dementias based on the NHIRD, and (2) to address the diversities by a case study.

Methods and analysis:

This scoping review protocol will follow the methodological framework of the Joanna Briggs Institute (JBI) Reviewer's Manual and the guidance of the PRISMA extension for scoping reviews. The review will be performed between 1st March and 31st December 2022 in 5 stages, including identifying the relevant studies, developing search strategies, individually screening and selecting evidence, collecting and extracting data, and summarising and reporting the results. The electronic databases of MEDLINE, EMBASE, CENTRAL, CINAHL, and PsychInfo, Airiti Library Academic Database, the National Health Insurance Administration's repository, and Taiwan Government Research Bulletin (GRB) will be searched. We will perform narrative syntheses of the results to address research questions, and will analyse the prevalence across the included individual studies as a case study.

Ethics and dissemination:

Our scoping review is a review of the published reports and ethical approval is not

1
2
3
4 required. The results will provide a panorama of the dementia studies based on the
5
6 NHIRD. We will disseminate our findings through peer-reviewed journals and
7
8 conferences, and share with stakeholders by distributing the summaries in social media
9
10 and emails.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Article summary

Strengths and limitations of this study

- To the best of our knowledge, our study will be the first review exploring the characteristics and the utilisation of diagnostic codes in the dementia studies based on the NHIRD.
- The scoping review methodology allows a broad perspective to depict the heterogeneities across the individual studies while maintaining transparency and accountability through systematic search and data extraction process.
- The results of this scoping review are expected to lay the foundation for future dementia studies based on the NHIRD.
- As the review is limited to literature published in English and Chinese, there are still potential publications in other languages that are not considered.
- Due to the lacked consensus on the risk of bias tool for such type of study, a formal quality and risk of bias assessment of the included studies will not be performed, limiting the understanding of how the potential biases in the studies may influence.

Introduction

Health claim data (HCD) have been the major sources for the studies of epidemiological analyses, health service economics and outcomes, and disease-specific medical research.[1] Research based on large population-based databases enables researchers to explore medical conditions with low prevalence or interventions with small effect sizes.[2] HCD have been applied to many medical specialities, and many prestige organisations have advocated that researchers can take good advantages of big data, specifically the HCD, in dementia research.[3, 4] In practice, HCD have been utilised in the research of dementias and their comorbidities, the disease progression trajectory, and the interaction between the biological and environmental factors.[3]

However, due to the nature of administrative data or claim data which are usually presumed for reimbursement rather than for research,[5, 6] scholars have proposed important issues that may pose threats and limitations to such studies, including the quality and reliability of the data,[7, 8] the lack of consensus and standardisation across the databases, and the doubt of data accuracy as a result of erroneous linkage.[9, 10] Furthermore, despite of being powerful, studies based on HCD are still vulnerable to the under-recording of dementia, resulting in under-representativeness of the target group and threats to generalisability of the study results.[11]

Being the main components of HCD, diagnostic codes play important roles in HCD-based studies, and, by following appropriate algorithms, researchers are able to aim at specific health outcomes of interest or identify populations with specific diagnoses within the HCD.[12, 13] However, the quality of diagnostic codes fundamentally relies on the avoidance of misdiagnosis, miscoding and misclassification, which will otherwise limit or even flaw the results as mentioned by Stein et al in their systematic review.[14] The issues related with diagnostic codes in dementia database research become more complicated. For example, the diagnostic gap due to under-identification caused by miscoding of dementia,[15] misidentification of dementia,[16] misclassification of dementia in HCD,[17] and the high heterogeneity in selecting the diagnostic codes of dementia[18] is not uncommon and all the above issues probably bring about complexities in HCD-based dementia studies. As the differences in selecting diagnostic codes to define dementias in relevant research would result in misidentification of dementia, it will be helpful to develop a set of standardised diagnostic codes for dementia to minimise the potential problematic impact and improve the value of HCD in dementia research.

Scoping review is a relatively young methodology in the family of evidence synthesis. It is regarded as an appropriate way to explore, configure, and aggregate the evidence, and can be used as a precursor to a systematic review. Scoping reviews are able to

1
2
3 illustrate the ways how research has been executed in the specific fields or topics, and
4 to identify the key concepts, rationales, types of evidence and the research gap. Rather
5 than testing theories or hypotheses, they can serve to explore the contents, range,
6 natures, and heterogeneities in the individual studies, to summarise the results, and to
7 guide the researchers for the future research about the directions and
8 methodologies.[19-21] The methodologies of scoping review have been first proposed
9 by Arksey and O'Malley in 2005,[22] and later strengthened by Levac et al by
10 proposing a practical 5-step approach.[23] Despite of their endeavours,
11 inconsistencies in execution remain. In this way, the Joanna Briggs Institute (JBI)
12 developed a comprehensive guideline to standardise the processes first in 2014 and
13 updated revisions in both 2017 and 2020.[24] Tricco et al also proposed the Preferred
14 Reporting Items for Systematic Reviews and Meta-Analyses extension for scoping
15 review (PRISMA-ScR) to aid the researchers in reporting their studies.[25]
16 Considering the complex natures in selecting diagnostic codes to define dementias
17 and the high heterogeneities in the database research, scoping review would be the
18 best suitable methodology in response to our research questions.

19 Since 1995, Taiwan has launched the Taiwan National Health Insurance (THNI) with
20 the coverage of 99.9% of the whole 23 million population and established an HCD
21 database which cumulates the health-related records of the users in the national health
22 insurance system.[26, 27] In practice, although increasing hospitals have begun to
23 employ the clinical coders to help the task of coding in recent years, physicians across
24 the levels in the health system are still the main persons responsible for coding and
25 inputting the diagnostic codes as well as the interventional codes into the
26 administrative systems for reimbursement in THNI. From 1995 to 2016, the
27 diagnostic codes were based on the International Classification of Diseases, Ninth
28 Revision, Clinical Modification (ICD-9-CM), which was replaced by ICD-10-CM
29 after 2016.[27]

30 Thanks to its abundant information, the National Health Insurance Administration has
31 built up the National Health Insurance Research Database (NHIRD) and released the
32 purchasable datasets which included de-identifiable and encrypted sampling of the
33 health records for the researchers in the academic organisations since 2000. At present
34 time, there are three forms of the NHIRD datasets released at different chronological
35 time, and they are general dataset from 2000, disease-specific dataset, and the latest
36 full population dataset which was released since 2016. These datasets consist both
37 inpatient and outpatient claim data, and the sets include demographic profiles as well
38 as clinical data of the codes for diagnoses, prescriptions and interventions. From
39 2016, the NHIRD data were authorised to link with other government databases at the
40 Data Science Centre, including some national census data, disease registries, health
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 surveys, social service data, death cause data, and welfare registries. The linkage with
4 other large national databases expands the applicability especially in research and
5 health policymaking.[26, 27]
6
7

8 In recent years, the NHIRD has provided the researchers with abundant resource for
9 secondary database medical research, and hundreds of studies have been published,
10 including the research for dementia. Many of them have been used as references for
11 healthcare practice guidance and public policymaking. Despite of its strengths,
12 however, the NHIRD still bears the same inherent weakness of HCD and there have
13 been inconsistencies in selecting the diagnostic codes in defining dementias in the
14 research using the NHIRD. As a result, it is imperative to investigate the
15 characteristics of dementia researches based on the NHIRD and how the diagnostic
16 codes are selected as well as used in such studies. This will aid in identifying the
17 potential research gap and reduce the research waste.
18
19

20 In the present study, by taking the advantages of scoping review methodology, the
21 research team intends to identify the characteristics, address the heterogeneities, and
22 explore the diversities of diagnostic codes used to define dementias in the published
23 studies. With the results of our study, we may lay the foundation for developing a set
24 of standardised codes for defining dementias for future dementia studies based on the
25 NHIRD. In this manuscript, we present the protocol for informing the implementation
26 of scoping review.
27
28
29
30
31
32
33
34

35 **Aim of the study and research question (study objective)**

36 Based on empirical experience, there are heterogeneities in NHIRD studies, including
37 the selections of diagnostic codes for defining diseases/conditions, the size of
38 datasets, the time length of database used, and the types of sub-datasets (i.e., inpatient
39 or outpatient dataset). The main aims of our research are to investigate the relevant
40 characteristics in the published reports targeting dementias based on the NHIRD, and
41 to address the diversities by analysing the reported prevalence as a case study. We
42 here define the following research questions for the scoping review.
43
44
45
46

- 47 1. Which diagnostic codes in ICD-9-CM and ICD-10-CM were used to define
48 dementias in the studies based on the NHIRD?
- 49 2. To what extent the diagnostic codes varied across the studies when being used to
50 define dementias, in relation with the research teams and the size and types of
51 database?
- 52 3. What differences in terms of the additional approaches other than diagnostic codes
53 used in the inclusion or exclusion criteria to identify the individuals with
54 dementias in the databases and the time length of database were adopted across
55 the studies?
56
57
58
59
60

4. How other important publication characteristics of database studies were reported across the studies?
5. As a case study, how does the prevalence differ across the studies based on the NHIRD in relation to the major variables above?

Methods

The present scoping review protocol has been developed based on the methodological framework of the Joanna Briggs Institute (JBI) Reviewer's Manual and has been constructed following the guidance of the PRISMA extension for scoping reviews (PRISMA-ScR).[20, 21, 25] A PRISMA-ScR checklist can be found in the Supplementary file 1. This protocol has been registered with the Open Science Framework (OSF: osf.io/fc65g) on 26 February, 2022, and the study will be implemented between 1st March and 31st December 2022 by following the 5 steps (subsections 2.1 to 2.5):

2.1 Identification of relevant studies.

2.1.1 Inclusion criteria

(1) Participants

The scoping review will aim at the published research reports on dementias using the NHIRD in the literature, and will include all types of dementias, which are defined and identified in any way in the reports. There will be no restriction on the types of study designs, the age of the participants in the reports and the comorbidities so as to maximise the coverage of the types of dementia research. However, we will only include the reports written in English or Chinese for the ease of data collection and analyses.

(2) Concept

The present scoping review intends to focus on dementia research using the NHIRD, and explores as well as configures the elements of the studies. The elements include the ways of selecting diagnostic codes to define dementias, the usage of sub-datasets of the NHIRD, the methodological spectra of dementia research based on the NHIRD, the heterogeneity, and the potential impact on the outcomes.

(3) Context

The scoping review will specifically focus on the dementia studies based on the NHIRD that may be influenced by the contextual factors of the pragmatic medical practices and services in Taiwan as well as some of the specific regulations in Taiwan National Health Insurance. Rather than being a constraint factor, however, this is the unique contextual characteristic that this scoping review intends address and can inform.

2.1.2 Exclusion criteria

We will exclude the reports when the full texts are not readily available, and, in addition, review articles, study protocols, grey literatures and texts that are not peer-reviewed or fail to provide detailed information that is in line with our study are excluded. The grey literatures here included letters, editorials or leading articles, commentaries, conference abstracts or presentations, and dissertations or theses.

2.2 Search strategy

The research team will develop search strategies that aim to include the study reports published between 2000 and 2022, ranging from the first year when the NHIRD was available to the academic researchers and the time to execute this review. Published studies in English or Chinese with all study designs will be included and the research team will search the major electronic databases of MEDLINE-OVID, EMBASE-OVID, Cochrane Central Register of Controlled Trials (CENTRAL), CINAHL, and PsychInfo to identify the reports on the topic. For the Chinese literatures and grey literatures, we will search the Airiti Library Academic Database, the National Health Insurance Administration's repository which collects the articles using Taiwan NHIRD, and Taiwan Government Research Bulletin (GRB). The search will consist of three major steps, and, the first step is to identify the search terms informing the three main concepts of the review, including dementia, database research and National Health Insurance Research Database, and Taiwan. The identified search terms of each concept will be first be organised with the Boolean operator OR into each category, and then the three categories will be combined with the Boolean operator AND. (Table 1) The search strategies will be developed by a trained researcher (YJY) and then reviewed by a librarian in the university library. The second step is to use all the identified keywords and index terms to undertake search in the titles and abstracts of relevant articles in each electronic database. The final step is to export searched results into the electronic bibliographic software EndNote X8. The second and third steps will be individual carried out by 2 different researchers (YJS and YJW). The pilot search strategies are shown in the Supplementary file 2. In addition to the electronic database search, we will conduct hand search for the potential studies in the reference lists of the articles initially identified when performing full-text reviewing.

Table 1 Search terms and concepts

Concept 1: Dementia	"Dementia" "Major neurocognitive disorder" "Alzheimer's dementia"
----------------------------	---

	“Parkinson’s disease dementia” “Dementia with Lewy body” “Frontotemporal dementia” “Vascular Dementia” “AIDS Dementia Complex” “Mixed dementia”
Concept 2: Database research & National health insurance	“Claim data” “Insurance claim data” “Insurance data/database” “Administrative (or health) data/database” “Health data” “Electronic health record” “Government database” “National Health Insurance Research Database (NHIRD)” “National Health Programme/ Service” “National data (or nationwide or population-based) research” “Longitudinal (or follow-up) Study”
Concept 3: Taiwan	“Taiwan”

2.3 Selection process (evidence screening and selection)

Following the search, all identified citations will be collated and exported into the bibliographic software EndNote X8 (Clarivate Analytics, PA, USA) where duplicates will be removed. Two researchers (YJS and JYW) will independently screen potentially eligible studies according to the inclusion criteria in our review by screening titles and abstracts yielded by our comprehensive search. The selection results by the two individual researchers will then be compared and merged. Any discrepancy in the results will be solved through discussion by involving a third reviewer (YJY). Reasons for exclusion of full-text studies that do not meet the inclusion criteria and the results of the search in each step will be reported in detail in the final scoping review and presented in PRISMA flowchart diagram.

2.4 Data collection process

The research team will develop an electronic data extraction form. This Microsoft Excel data extraction form will be independently pilot tested by all team members for its applicability and the research team will achieve its final version based on the feedback. The targeted data to extract will include the characteristics of the reports such as the diagnostic codes to define dementias, the type of dementia reported, the

study designs, the time length, the subsets of database, whether additional methods to increase the likelihood of dementia diagnosis were used, the statistical methods and the outcomes. Other general profiles of the reports including lead authors, affiliations, main discipline of the research team, year of publication, funding, journals where published and other potential factors. (See Table 2) When extracting the data for the review, a reviewer (YJS) will assess each eligible study and then input the assessed results in the prespecified data extraction form. Another review (JYW) will then randomly examine 80% of the extracted results, and any inconsistency will be resolved through consensus with the involvement of the third reviewer (YJY).

Table 2. Data extraction table.

Article information	Title Authors Affiliations with author(s) Main disciplines of lead author Year of publication (based on the first identifiable date) Journal where published and the JCR impact factor Language of publication Funding Type of the literature Other potential factors
Population	Sample size Inclusion and exclusion criteria of the individual study Age of the participants Sex of participants Diagnostic codes to define dementias Type of dementia reported Comorbidities or exposures investigated
Methods/Results	Study designs reported Time length/follow-up time Subsets of database Whether and what methods used to validate the selection of diagnostic codes to define dementia Whether and what additional methods to increase the likelihood of dementia diagnosis used Statistical methods The number of people with dementia identified Prevalence, if any

	<p>Main outcomes of the studies and the estimation such as incidence and mortality if reported</p> <p>Other reported health-related outcomes</p>
--	--

2.5 Collating, summarising and reporting the results

As the scoping review intends to map the characteristics of the published reports and to address the diversities of the specific outcome for case study, we will not perform critical appraisal of each individual study. The extracted data will be analysed by two researchers who will work at the same time by following a pre-specified analytical plan. In responding to the research questions, the variables mentioned above will first be summarised quantitatively, and descriptive statistics including number counts, frequencies and rates will be used to summarise the results. Then the researchers will perform narrative discussions based on the results to address the mapping and research questions that cannot be answered quantitatively. Considering the vast differences in reporting prevalence (i.e., non-reporting, different time frame or ways of calculation) in individual studies, we will not attempt to pool the prevalence with statistical methods by our own, and, instead, the reporting conditions of prevalence in each individual study will be presented with potential reasons discussed. We will also perform a subgroup analysis based on the type of literature to compare whether there are differences between the formal and grey literatures. Finally, we will abide by the PRISMA-ScR guideline [25] to ensure the reporting standard of the final manuscript, and submit to a peer-reviewed journal.

Patient and public involvement

Our scoping review will not plan to involve patients with dementia or the public, but we will also invite clinicians who have to type in the codes for dementia in their practices in the review process and when consulting, discussing and verifying the results.

Ethic and dissemination

Our scoping review is a review of the published reports which does not involve access to individual data, and ethical approval is not required. The results will be disseminated through peer-reviewed journals and conferences focusing on dementia and related topics. We will also share our results with stakeholders, including the non-government organisations for dementia and the policymakers in the field of health informatics, by distributing our results in brief and plain language in the social media and emails. If there to be any amendment to the protocol, the revised information will be disclosed on the protocol registry space (OSF: osf.io/fc65g) and will be stated in

any future publication based on the review.

Discussion

The proposed scoping review aims to investigate and address the diversities in the published reports targeting dementias based on the NHIRD, and to discuss the potential influence through a case study of the reported prevalence. To the best of our knowledge, our study will be the first review of the studies based on the NHIRD. We believe that, by the strengths of scoping review methodology, a panorama view of the dementia studies based on the NHIRD database will help to inform the researchers of the *status quo* of the research in its field and thereby improve the quality and value of database research by taking the advantages of the NHIRD.

Additional Headings

Data statement

Technical appendix, search strategies, electronic data extraction sheets, statistical codes and datasets are available from the OSF repository, DOI: 10.17605/OSF.IO/YAQCM.

Author Contributions

YJS and YJY conceptualised the research. YJS, JYW, YHW and RRS helped the development of ideas and drafted of the initial protocol. YJY supervised and revised the manuscript. All authors read and agreed the final manuscript.

Author Statement

Acknowledgement

We especially appreciated the help from the staff in the Department of Healthcare Administration, College of Medical and Health Science and the librarians in Asia University.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of interests

We declared no conflict of interests.

Patient consent for publication

Not required.

References

1. Konrad R, Zhang W, Bjarndóttir M, Proaño R. Key considerations when using health insurance claims data in advanced data analyses: an experience report. *Health Systems (Basingstoke, England)* 2019;9(4):317-25.
2. Thygesen LC, Ersbøll AK. When the entire population is the sample: strengths and limitations in register-based epidemiology. *Eur J Epidemiol* 2014;29(8):551-8.
3. OECD. *Dementia Research and Care: Can Big Data Help? Big data approaches to dementia: Opportunities and challenges*. Anderson G, J.Oderkirk, editors. Paris: OECD Publishing; 2015.
4. Pike KJ, Fazio S, Bynum JPW, Trivison TG, Wendler D, Mor V. Resources, methods, and data infrastructure to promote research in dementia care, caregiving, and services. *J Am Geriatr Soc* 2021;69(7):1793-800.
5. Reinke C, Buchmann N, Fink A, Tegeler C, Demuth I, Doblhammer G. Diabetes duration and the risk of dementia: a cohort study based on German health claims data. *Age Ageing* 2021;51(1):afab231.
6. Shi L, Steenland K, Li H, Liu P, Zhang Y, Lyles RH, et al. A national cohort study (2000–2018) of long-term air pollution exposure and incident dementia in older adults in the United States. *Nat Commun* 2021;12(1):6754.
7. Feder SL. Data Quality in Electronic Health Records Research: Quality Domains and Assessment Methods. *Western journal of nursing research* 2018;40(5):753-66.
8. Rabinstein AA. Administrative Medical Databases for Clinical Research: The Good, The Bad, and The Ugly. *Neurocrit Care* 2018;29(3):323-5.
9. Bohensky MA, Jolley D, Sundararajan V, Evans S, Pilcher DV, Scott I, et al. Data linkage: a powerful research tool with potential problems. *BMC Health Serv Res* 2010;10:346-.
10. Harron K, Dibben C, Boyd J, Hjern A, Azimaee M, Barreto ML, et al. Challenges in administrative data linkage for research. *Big data & society* 2017;4(2):2053951717745678.
11. Davis KAS, Mueller C, Ashworth M, Broadbent M, Jewel A, Molokhia M, et al. What gets recorded, counts: dementia recording in primary care compared with a specialist database. *Age Ageing* 2021;50(6):2206-13.
12. Jaakkimainen RL, Bronskill SE, Tierney MC, Herrmann N, Green D, Young J, et al. Identification of Physician-Diagnosed Alzheimer's Disease and Related

- 1
2
3
4 Dementias in Population-Based Administrative Data: A Validation Study Using
5 Family Physicians' Electronic Medical Records. *J Alzheimers Dis*
6 2016;54(1):337-49.
7
- 8 13. Lipscombe LL, Hwee J, Webster L, Shah BR, Booth GL, Tu K. Identifying
9 diabetes cases from administrative data: a population-based validation study.
10 *BMC Health Serv Res* 2018;18(1):316.
11
- 12 14. Stein JD, Lum F, Lee PP, Rich WL, 3rd, Coleman AL. Use of health care claims
13 data to study patients with ophthalmologic conditions. *Ophthalmology*
14 2014;121(5):1134-41.
15
- 16 15. Russell P, Banerjee S, Watt J, Adleman R, Agoe B, Burnie N, et al. Improving
17 the identification of people with dementia in primary care: evaluation of the
18 impact of primary care dementia coding guidance on identified prevalence *BMJ*
19 *Open*. 2013;3(12):e004023.
20
- 21 16. Zhu CW, Ornstein KA, Cosentino S, Gu Y, Andrews H, Stern Y.
22 Misidentification of Dementia in Medicare Claims and Related Costs. *J Am*
23 *Geriatr Soc* 2019;67(2):269-76.
24
- 25 17. Health AIo, Welfare. *Predicting early dementia using Medicare claims: a*
26 *feasibility study using the National Integrated Health Services Information*
27 *Analysis Asset*. Canberra: AIHW; 2021.
28
- 29 18. Wilkinson T, Ly A, Schnier C, Rannikmäe K, Bush K, Brayne C, et al.
30 Identifying dementia cases with routinely collected health data: A systematic
31 review. *Alzheimers Dement* 2018;14(8):1038-51.
32
- 33 19. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E.
34 Systematic review or scoping review? Guidance for authors when choosing
35 between a systematic or scoping review approach. *BMC Med Res Methodol*
36 2018;18(1):143.
37
- 38 20. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al.
39 Updated methodological guidance for the conduct of scoping reviews. *JBIM Evid*
40 *Synth* 2020;18(10):2119-26.
41
- 42 21. Tricco AC, Lillie E, Zarin W, O'Brien K, Colquhoun H, Kastner M, et al. A
43 scoping review on the conduct and reporting of scoping reviews. *BMC Med Res*
44 *Methodol* 2016;16(1):15.
45
- 46 22. Arksey H, O'Malley L. Scoping studies: towards a methodological framework.
47 *International Journal of Social Research Methodology* 2005;8(1):19-32.
48
- 49 23. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the
50 methodology. *Implement Sci* 2010;5:69.
51
- 52 24. Khalil H, Bennett M, Godfrey C, McInerney P, Munn Z, Peters M. Evaluation of
53 the JBI scoping reviews methodology by current users. *International journal of*
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

evidence-based healthcare 2020;18(1):95-100.

25. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med* 2018;169(7):467-73.
26. Lin LY, Warren-Gash C, Smeeth L, Chen PC. Data resource profile: the National Health Insurance Research Database (NHIRD). *Epidemiol Health* 2018;40:e2018062.
27. Hsieh CY, Su CC, Shao SC, Sung SF, Lin SJ, Kao Yang YH, et al. Taiwan's National Health Insurance Research Database: past and future. *Clin Epidemiol* 2019;11:349-58.

● **PRISMA-ScR Checklist.**

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	P1
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	P3-P4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	P5-P7
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	P7-P8
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	P8
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	P8-P9
Information sources	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	P9
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	Supplementary file 2.
Selection of sources of evidence	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	P10
Data charting process	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	P10-P11
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	P11-P12, Table 2
Critical appraisal of individual sources of evidence	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	We will not perform critical appraisal due to the nature of the study.
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	P12
RESULTS			
Selection of	14	Give numbers of sources of evidence screened, assessed	N/A, This is a

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
sources of evidence		for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	protocol.
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	N/A, This is a protocol.
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	N/A, This is a protocol.
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	N/A, This is a protocol.
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	N/A, This is a protocol.
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	N/A, This is a protocol.
Limitations	20	Discuss the limitations of the scoping review process.	N/A, This is a protocol.
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	N/A, This is a protocol.
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	We did not receive any grant.

From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med.* 2018;169:467–473. doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850).

● **Supplementary table:**

Search strategy

Source	#	Search strategy
EMBASE 1980 to 2022 (Ovid SP)	1	exp dementia/
	2	dement*.ti,ab.
	3	"major neurocognitive disorder".ti,ab.
	4	exp Alzheimer Disease/
	5	Alzheimer*.ti,ab.
	6	(Alzheimer* adj2 dement*).ti,ab.
	7	AD.ti,ab.
	8	(Parkinson* adj2 dementia).ti,ab.
	9	(lewy* adj2 bod*).ti,ab.
	10	"frontotemporal lobe degeneration".ti,ab.
	11	exp Frontotemporal Lobar Degeneration/
	12	FTD.ti,ab.
	13	FTLD.ti,ab.
	14	"Pick* disease".ti,ab.
	15	(Pick* adj2 disease).ti,ab.
	16	"vascular dementia".ti,ab.
	17	VAD.ti,ab.
	18	"multi-infarct".ti,ab.
	19	"mixed dementia".ti,ab.
	20	exp AIDS Dementia Complex/
	21	exp Dementia, Vascular/
	22	exp Lewy Body Disease/
	23	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22
	24	Routinely Collected Health Data/
	25	(Claim* adj2 data).ti,ab.
	26	(Insurance adj2 claim*).ti,ab.

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
	27 (Insurance adj2 data).ti,ab.
	28 (Insurance adj2 reimburse).ti,ab.
	29 "administrati* data*".ti,ab.
	30 "administrati* health* database*".ti,ab.
	31 "Health* data source*".ti,ab.
	32 (health* adj3 administrati* adj3 data*).ti,ab.
	33 "electr* health* record*".ti,ab.
	34 "govern* administrati*".ti,ab.
	35 "National Health Insurance Research Database".ti,ab.
	36 NHIRD.ti,ab.
	37 "national health insurance".ti,ab.
	38 "health insurance".ti,ab.
	39 "national health program*".ti,ab.
	40 "national health service*".ti,ab.
	41 NHI.ti,ab.
	42 "national data*".ti,ab.
	43 nationwide.ti,ab.
	44 "insurance, national health".ti,ab.
	45 ("population-base*" or "population* base*").ti,ab.
	46 "Taiwan* longitudinal stud*".ti,ab.
	47 ("follow-up stud*" or "follow* up stud*").ti,ab.
	48 retrospective.ti,ab.
	49 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48
	50 Taiwan*.ti,ab.
	51 23 and 49 and 50
	52 limit 51 to yr="2000 -Current"

Source	#	Search strategy
MEDLINE 1946 to 2022 (Ovid)	1	exp dementia/
	2	dement*.ti,ab.
	3	"major neurocognitive disorder".ti,ab.
	4	exp Alzheimer Disease/
	5	Alzheimer*.ti,ab.
	6	(Alzheimer* adj2 dement*).ti,ab.
	7	AD.ti,ab.
	8	(Parkinson* adj2 dementia).ti,ab.
	9	(lewy* adj2 bod*).ti,ab.
	10	"frontotemporal lobe degeneration".ti,ab.
	11	exp Frontotemporal Lobar Degeneration/
	12	FTD.ti,ab.
	13	FTLD.ti,ab.
	14	"Pick* disease".ti,ab.
	15	(Pick* adj2 disease).ti,ab.
	16	"vascular dementia".ti,ab.
	17	VAD.ti,ab.
	18	"multi-infarct".ti,ab.
	19	"mixed dementia".ti,ab.
	20	exp AIDS Dementia Complex/
	21	exp Dementia, Vascular/
	22	exp Lewy Body Disease/
	23	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22
	24	Routinely Collected Health Data/
	25	(Claim* adj2 data).ti,ab.
	26	(Insurance adj2 claim*).ti,ab.
	27	(Insurance adj2 data).ti,ab.

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
28	(Insurance adj2 reimburse).ti,ab.
29	"administrati* data*".ti,ab.
30	"administrati* health* database*".ti,ab.
31	"Health* data source*".ti,ab.
32	(health* adj3 administrati* adj3 data*).ti,ab.
33	"electr* health* record*".ti,ab.
34	"govern* administrati*".ti,ab.
35	"National Health Insurance Research Database".ti,ab.
36	NHIRD.ti,ab.
37	"national health insurance".ti,ab.
38	"health insurance".ti,ab.
39	"national health program*".ti,ab.
40	"national health service*".ti,ab.
41	NHI.ti,ab.
42	"national data*".ti,ab.
43	nationwide.ti,ab.
44	"insurance, national health".ti,ab.
45	("population-base*" or "population* base*").ti,ab.
46	"Taiwan* longitudinal stud*".ti,ab.
47	("follow-up stud*" or "follow* up stud*").ti,ab.
48	retrospective.ti,ab.
49	24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48
50	Taiwan*.ti,ab.
51	23 and 49 and 50
52	limit 51 to yr="2000 -Current"

Source	#	Search strategy
APA PsycInfo 1806 to 2022 (Ovid)	1	exp dementia/
	2	dement*.ti,ab.
	3	"major neurocognitive disorder".ti,ab.
	4	exp Alzheimer Disease/
	5	Alzheimer*.ti,ab.
	6	(Alzheimer* adj2 dement*).ti,ab.
	7	AD.ti,ab.
	8	(Parkinson* adj2 dementia).ti,ab.
	9	(lewy* adj2 bod*).ti,ab.
	10	"frontotemporal lobe degeneration".ti,ab.
	11	"frontotemporal lob* degeneration".ti,ab.
	12	FTD.ti,ab.
	13	FTLD.ti,ab.
	14	"Pick* disease".ti,ab.
	15	(Pick* adj2 disease).ti,ab.
	16	"vascular dementia".ti,ab.
	17	VAD.ti,ab.
	18	"multi-infarct".ti,ab.
	19	"mixed dementia".ti,ab.
	20	exp AIDS Dementia Complex/
	21	"dementia vascular".ti,ab.
	22	exp Lewy Body Disease/
	23	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22
	24	Routinely Collected Health Data/
	25	(Claim* adj2 data).ti,ab.
	26	(Insurance adj2 claim*).ti,ab.
	27	(Insurance adj2 data).ti,ab.

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
28	(Insurance adj2 reimburse).ti,ab.
29	"administrati* data*".ti,ab.
30	"administrati* health* database*".ti,ab.
31	"Health* data source*".ti,ab.
32	(health* adj3 administrati* adj3 data*).ti,ab.
33	"electr* health* record*".ti,ab.
34	"govern* administrati*".ti,ab.
35	"National Health Insurance Research Database".ti,ab.
36	NHIRD.ti,ab.
37	"national health insurance".ti,ab.
38	"health insurance".ti,ab.
39	"national health program*".ti,ab.
40	"national health service*".ti,ab.
41	NHI.ti,ab.
42	"national data".ti,ab.
43	nationwide.ti,ab.
44	"insurance, national health".ti,ab.
45	("population-base*" or "population* base*").ti,ab.
46	"Taiwan* longitudinal stud*".ti,ab.
47	("follow-up stud*" or "follow* up stud*").ti,ab.
48	retrospective.ti,ab.
49	24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48
50	Taiwan*.ti,ab.
51	23 and 49 and 50
52	limit 51 to yr="2000 -Current"

Source	#	Search strategy
CENTRAL Cochrane Library	1	MeSH descriptor: [Dementia] explode all trees
	2	major neurocognitive disorder
	3	Alzheimer Disease
	4	dement*
	5	Alzheimer*
	6	Alzheimer* adj3 dement*
	7	AD
	8	Parkinson* adj3 dementia
	9	lewy* adj3 bod*
	10	Lewy Body Disease
	11	MeSH descriptor: [Lewy Body Disease] explode all trees
	12	DLB
	13	Dementia with Lewy bodies
	14	Dementia with Lewy body
	15	Frontotemporal Lobar Degeneration
	16	MeSH descriptor: [Frontotemporal Lobar Degeneration] explode all trees
	17	Frontotemporal Loba* Degeneration
	18	FTD
	19	FTLD
	20	Pick* disease
	21	Pick* adj3 disease
	22	MeSH descriptor: [Dementia, Vascular] explode all trees
	23	VaD
	24	mixed dementia
	25	MeSH descriptor: [AIDS Dementia Complex] explode all trees
	26	AIDS Dementia Complex
	27	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18

	OR #19 OR #20 OR #21 OR #22 OR# 23 OR #24 OR #25 OR #26
28	MeSH descriptor: [Routinely Collected Health Data] explode all trees
29	Claim* adj2 data
30	Insurance adj2 claim*
31	Insurance adj2 data
32	Insurance adj2 reimburse
33	administrati* data*
34	administrati* adj3 health* adj3 database*
35	Health* data source*
36	health* adj3 administrati* adj3 data*
37	electr* health* record*
38	govern* administrati*
39	National Health Insurance Research Database
40	NHIRD
41	national health insurance
42	health insurance
43	national health program*
44	national health service*
45	NHI
46	national data*
47	nationwide
48	insurance adj3 national adj3 health
49	insurance national health
50	population-base*
51	population* base*
52	Taiwan* longitudinal stud*
53	follow-up stud*
54	follow* up stud*
55	retrospective

56	#28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51 OR #52 OR #53 OR #54 OR #55
57	Taiwan*
58	#27 AND #56 AND #57

Source	S	Search strategy
CINAHL (EBSCOhost)	1	TI dementia OR AB dementia
	2	MH dementia+ OR TX dementia+
	3	MH "dement*" OR TX "dement*"
	4	TI dement* OR AB dement*
	5	TI "Alzheimer* Disease" OR AB "Alzheimer* Disease"
	6	TX "Alzheimer Disease"
	7	TI Alzheimer* OR AB Alzheimer*
	8	TX Alzheimer*
	9	MH Alzheimer*
	10	TX "Alzheimer*" N3 "dement*"
	11	TX "AD" OR TI "AD" OR AB "AD"
	12	TX "major neurocognitive disorder"
	13	TI "major neurocognitive disorder" OR AB "major neurocognitive disorder"
	14	TI "Parkinson* N3 dement*" OR AB "Parkinson* N3 dement*"
	15	TX "Parkinson* N3 dement*"
	16	TX "lewy* N3 bod*"
	17	TI "lewy* N3 bod*" OR AB "lewy* N3 bod*"
	18	MH Lewy Body Disease+
	19	TX "Lewy* Bod* Disease" OR TI "Lewy* Bod* Disease" OR AB "Lewy* Bod* Disease"
	20	TX "DLB" OR TI "DLB" OR AB "DLB"
	21	TX "frontotemporal lobe* degeneration"

22	MH frontotemporal lobar degeneration
23	TI "frontotemporal lobar degeneration" OR AB "frontotemporal lobar degeneration"
24	MH frontotemporal dement*
25	TX "frontotemporal dementia" OR TI "frontotemporal dementia" OR AB "frontotemporal dementia"
26	TX "FTD" OR TI "FTD" OR AB "FTD"
27	TX "FTLD" OR TI "FTLD" OR AB "FTLD"
28	TX "Pick* disease" OR TI "Pick* disease" OR AB "Pick* disease"
29	TX "vascular dementia" OR TI "vascular dementia" OR AB "vascular dementia"
30	TX "multi-infarct" OR TI "multi-infarct" OR AB "multi-infarct"
31	TX "mixed dementia" OR TI "mixed dementia" OR AB "mixed dementia"
32	TX "mixed N3 dementia" OR TI "mixed N3 dementia" OR AB "mixed N3 dementia"
33	TX "mixed N3 dementia" OR TI "mixed N3 dementia" OR AB "mixed N3 dementia"
34	MH AIDS Dementia Complex+
35	TX "AIDS Dement* Complex" OR TI "AIDS Dement* Complex" OR AB "AIDS Dement* Complex"
36	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35
37	MH Routinely Collected Health Data+
38	TX "Routinely Collected Health Data+" OR TI "Routinely Collected Health Data" OR AB "Routinely Collected Health Data"
39	TX "Claim* N3 data" OR TI "Claim* N3 data" OR AB "Claim* N3 data"
40	TX "Insurance N3 claim*" OR TI "Insurance N3 claim*" OR AB "Insurance N3 claim*"
41	TX "Insurance N3 data*" OR TI "Insurance N3 data*" OR AB "Insurance N3 data*"

42	TX "Insurance N3 reimburse" OR TI "Insurance N3 reimburse" OR AB "Insurance N3 reimburse"
43	TX "administrati* data*" OR TI "administrati* data*" OR AB "administrati* data*"
44	TX "Health* data* source*" OR TI "Health* data* source*" OR AB "Health* data* source*"
45	TX "administrati* health* data*" OR TI "administrati* health* data*" OR AB "administrati* health* data*"
46	TX "health* administrati* data*" OR TI "health* administrati* data*" OR AB "health* administrati* data*"
47	TX "electr* health* record*" OR TI "electr* health* record*" OR AB "electr* health* record*"
48	TX "govern* administrati*" OR TI "govern* administrati*" OR AB "govern* administrati*"
49	TX "National Health Insurance Research Database" OR TI "National Health Insurance Research Database" OR AB "National Health Insurance Research Database"
50	TX "NHIRD" OR TI "NHIRD" OR AB "NHIRD"
51	TX "national health insurance" OR TI "national health insurance" OR AB "national health insurance"
52	TX "national health program" OR TI "national health program" OR AB "national health program"
53	TX "national health service*" OR TI "national health service*" OR AB "national health service*"
54	TX "NHI" OR TI "NHI" OR AB "NHI"
55	TX "national data*" OR TI "national data*" OR AB "national data*"
56	TX "nationwide insurance" OR TI "nationwide insurance" OR AB "nationwide insurance"
57	TX nationwide OR TI nationwide OR AB nationwide
58	TX "insurance N3 national health" OR TI "insurance N3 national health" OR AB "insurance N3 national health"
59	TX "insurance national health*" OR TI "insurance national health*" OR AB "insurance national health*"
60	TX "population-base*" OR TI "population-base*" OR AB "population-base*"

61	TX "population* base*" OR TI "population* base*" OR AB "population* base*"
62	TX "Taiwan* longitudinal stud*" OR TI "Taiwan* longitudinal stud*" OR AB "Taiwan* longitudinal stud*"
63	TX "follow-up stud*" OR TI "follow-up stud*" OR AB "follow-up stud*"
64	TX "follow* up stud*" OR TI "follow* up stud*" OR AB "follow* up stud*"
65	TX "retrospective" OR TI "retrospective" OR AB "retrospective"
66	MH Taiwan*
67	TX "Taiwan*" OR TI "Taiwan*" OR AB "Taiwan*"
68	S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR S62 OR S63 OR S64 OR S65
69	S66 OR S67
70	S36 AND S68 AND S69

websites Source	#	Search strategy
Airiti library (https://www.airitilibrary.com/)	1	"Dementia" OR "dement*" OR "major neurocognitive disorder" OR "Alzheimer Disease" OR "Alzheimer*" OR "Alzheimer Dementia" OR "Parkinson dementia" OR "Dementia with lewy body" OR "DLB" OR "Lewy Body Disease" OR "frontotemporal lobe degeneration" OR "Frontotemporal Lobar Degeneration" OR "Frontotemporal dementia" OR "FTD" OR "FTLD" OR "Pick's disease" OR "vascular dementia" OR "multi-infarct" OR "mixed dementia" OR "AIDS Dementia Complex"
	2	"Routinely Collected Health Data" OR "Claim data" OR "Insurance claim" OR "Insurance reimburse" OR "Administration data" OR "Administration health database" OR "Health data source" OR "National Health Insurance Research Database" OR "NHIRD" OR "National health insurance" OR "National health program" OR "National health service" OR "NHI" OR "National data" OR "Nationwide" OR "Population-base study" OR "Taiwan's longitudinal" OR "Follow-up study" OR "Cohort database" OR "Big

		data" OR "Nationwide population-based" OR "Epidemiologic study"
	3	Taiwan
	4	1 AND 2 AND 3
	5	失智症 OR 阿茲海默氏症 OR 阿茲海默型失智症 OR 路易氏體失智症 OR 血管性失智症 OR 額顳葉型失智症 OR 混合型失智症 OR 愛滋感染失智
	6	全民健康保險資料庫 OR 健保資料庫 OR 資料庫研究 OR 行政申報資料庫 OR 台灣健保申報資料 OR 全民健康保險學術研究資料庫 OR 健保大數據 OR 世代研究 OR 全國性研究 OR 健康保險資料庫
	7	臺灣
	8	5 AND 6 AND 7
	9	4 AND 8

website Source	#	Search strategy
Taiwan Government Research Bulletin (GRB) (https://www.grb.gov.tw/)	1	失智症 OR 阿茲海默氏症 OR 阿茲海默型失智症 OR 路易氏體失智症 OR 血管性失智症 OR 額顳葉型失智症 OR 混合型失智症 OR 愛滋感染失智
	2	全民健康保險資料庫 OR 健保資料庫 OR 資料庫研究 OR 行政申報資料庫 OR 台灣健保申報資料 OR 全民健康保險學術研究資料庫 OR 健保大數據 OR 世代研究 OR 全國性研究 OR 健康保險資料庫
	3	臺灣
	4	1 AND 2 AND 3

website Source	#	Search strategy
The National Health Insurance Administration's repository of Journal articles	1	Dementia OR major neurocognitive disorder OR Alzheimer Disease OR Alzheimer OR Alzheimer Dementia OR Parkinson dementia OR dementia with lewy body OR Lewy Body Disease OR frontotemporal lobe degeneration OR Frontotemporal Lobar Degeneration OR Frontotemporal dementia OR Pick's disease OR vascular dementia OR multi-infarct OR mixed dementia OR AIDS

<p>using Taiwan's National Health Insurance research data. (https://www.nhi.gov.tw/Query/Query_AcademicResearch.aspx)</p>	Dementia Complex
---	------------------

For peer review only