Optimal Strategies for Determining Washout Period Duration for Identifying Incident Chronic Disease Cases from Administrative Data: A Systematic Review Postprint

Authors: Wenyi Yang, Jingxin Wang, Limei Ai, Xia Wan*

Date: 2023-07-26T00:00:00+00:00

Abstract

Background When utilizing administrative data, establishing a clear and appropriate washout period duration for chronic diseases is fundamental for correctly determining the onset timing in patients with repeated healthcare encounters and for identifying incident cases.

Objective To summarize methods for determining washout period duration through a systematic literature review, thereby providing insights for Chinese researchers on confirming washout period length and correctly identifying incident cases when using administrative data to identify new chronic disease cases.

Methods In October 2021, we systematically searched PubMed, Web of Science, EmBase, CNKI, VIP Chinese Science and Technology Journals Full-text Database, and Wanfang Knowledge Service Platform for literature investigating chronic disease incidence and prevalence using administrative data, with search limits from database inception to October 1, 2022. Two researchers independently screened literature and extracted relevant information. After evaluating methodological quality using the Standards for Reporting Qualitative Research (SRQR), we summarized methods for determining washout period duration using descriptive analysis.

Results A total of 26 articles were included, all with SRQR scores $ $15 points, indicating good methodological quality. The data used primarily originated from countries (regions) with complete and rich administrative data, such as Canada, the United States, and Australia, focusing on various chronic diseases including diabetes, cancer, and schizophrenia. Studies indicated that setting an appropriate washout period duration is essential for accurately identifying incident cases. Currently, methods for determining washout period duration in the literature mainly fall into three categories: direct specification method, consistency test method, and reverse survival function method, with the direct specification method being the most commonly used, while the reverse survival curve method is relatively less utilized.

Conclusion The direct specification method, consistency test method, and reverse survival function method each have corresponding advantages and limitations. The selection criteria, judgment standards, and stability of these methods warrant further investigation.

Full Text

Optimal Strategies for Determining the Duration of Washout Period in the Context of Identifying Chronic Disease Onset Cases Based on Administrative Data: A Systematic Review

YANG Wenyi, WANG Jingxin, AI Limei, ZHAO Jian, WAN Xia*

Institute of Basic Medical Sciences, Chinese Academy of Medical Science & School of Basic Medicine, Peking Union Medical College, Beijing 100005, China

*Corresponding author: Wan Xia, Doctoral supervisor/Professor; E-mail: xiawan@ibms.pumc.edu.cn

Abstract

Background: When using administrative data, establishing a clear and appropriate duration for chronic disease washout periods is fundamental for correctly identifying the onset timing of chronic diseases in patients with recurrent medical visits and for determining new incident cases.

Objective: Through systematic literature review, this study aims to summarize methods for determining washout period duration to provide guidance for Chinese researchers in confirming washout period length and correctly identifying new incident cases when using administrative data to identify chronic disease onset cases.

Methods: In October 2021, we systematically searched PubMed, Web of Science, EmBase, CNKI, CQVIP, and Wanfang Knowledge Service Platform for literature investigating chronic disease incidence and prevalence using administrative data, with the search timeframe from database inception to October 1, 2022. Two researchers independently screened literature and extracted relevant information. After evaluating methodological quality using the Standards for Reporting Qualitative Research (SRQR), we used descriptive analysis to summarize methods for determining washout period duration.

Results: A total of 26 papers were included, all with SRQR scores $ $15, indicating good methodological quality. The data used in these studies primarily came from countries (regions) with complete and abundant administrative data, including Canada, the United States, and Australia. The focused diseases included diabetes, cancer, schizophrenia, and various other chronic diseases. The studies emphasized that setting an appropriate washout period duration is essential for accurate identification of incident cases. Currently, methods for determining washout period duration in the literature mainly fall into three categories: direct restriction method, consistency test method, and retrograde survival function method. The direct restriction method was the most commonly used, while the retrograde survival curve method had relatively low utilization.

Conclusion: The direct restriction method, consistency test method, and retrograde survival function method each have corresponding advantages and limitations. The selection criteria, judgment standards, and stability of these methods require further investigation.

[**Key words**] Chronic disease; Administrative data; Insurance data; Washout period; Incidence; Prevalence

Introduction

To effectively prevent and control various chronic diseases such as diabetes, cancer, and chronic obstructive pulmonary disease, accurate estimation of disease prevalence and incidence is essential. While some researchers have investigated disease prevalence and incidence through large-scale cohort studies and surveys, this approach is time-consuming and costly. Therefore, administrative data such as medical insurance reimbursement records and health monitoring data can be considered for analyzing the epidemiological characteristics of chronic diseases. Unlike lifetime medical records from birth to death, medical insurance data typically cover limited years, and chronic disease patients usually have repeated medical visits. This makes it challenging to effectively identify new incident cases from these truncated, highly repetitive records. To address this issue, researchers generally use a look-back period (washout period) as the basis for identifying new incident cases. The specific method involves identifying suspected new incident cases when patients appear in registration records for a certain chronic disease within the target search years, then looking backward from the target search years. If no disease-related records exist within the defined washout period, the patient can be confirmed as a new incident case. Currently, there is no consensus on the appropriate length of the washout period. A washout period that is too short may lead to overestimation of incidence, while one that is too long results in underutilization of data. Some researchers argue that different diseases have different trajectories and characteristics, and appropriate methods should be used to determine the optimal washout period duration when identifying new incident cases for different diseases using different types of data. This systematic review summarizes methods for determining washout period duration to provide ideas for Chinese researchers to confirm washout period length and correctly identify new incident cases when subsequently using administrative data to identify chronic disease onset cases.

Methods

Literature Search Strategy

In October 2021, we systematically searched three English databases (PubMed, Web of Science, EmBase) and three Chinese databases (CNKI, CQVIP, Wanfang Knowledge Service Platform) for literature investigating chronic disease incidence and prevalence using administrative data, with the search timeframe from database inception to October 1, 2022. Chinese search terms included: medical insurance data, official data, administrative data, hospital data, prevalence, incidence, washout period, look-back period, and window period. English search terms included: administrative data, insurance data, hospital data, Medicare, prevalence, incidence, look-back period, wash-out time, clearance time, disease-free time, and observation time. The specific search strategy for PubMed was: (incidence[Title/Abstract] OR prevalence[Title/Abstract]) AND (insurance data[Title/Abstract] OR administrative data[Title/Abstract] OR hospital data[Title/Abstract] OR Medicare[Title/Abstract]) AND (look-back period[Title/Abstract] OR wash-out time[Title/Abstract] OR clearance time[Title/Abstract] OR disease-free time[Title/Abstract] OR observation time[Title/Abstract]).

Literature Screening and Data Extraction

Two researchers independently screened literature and extracted data, with cross-checking of results. Disagreements were resolved through discussion with a third researcher. During literature screening, titles and abstracts were first read to exclude obviously irrelevant literature, followed by full-text reading to determine inclusion based on predetermined criteria. Extracted data included: (1) basic literature information (authors, publication year, focused disease); (2) basic database information (country/region, database name, database type, data year span, covered population size); and (3) study design (methods and criteria for determining optimal washout period duration, set washout period length).

Inclusion criteria were: (1) studies using administrative data, including medical insurance data, disease registration data, and hospital registration data; (2) studies focusing on chronic diseases with multiple records for the same patient due to the same disease; (3) studies investigating disease prevalence and incidence with a focus on epidemiological characteristics; and (4) studies involving washout periods.

Exclusion criteria were: (1) non-Chinese or non-English literature; (2) studies using non-official data, such as survey data obtained through cohort or case-control studies; (3) non-original research, such as commentaries and systematic reviews; (4) conference papers or research published in abstract form; (5) literature where the full text could not be obtained; (6) duplicate publications; and (7) literature with content not relevant to this study, including studies that only used administrative data to assess disease burden or conduct cohort/case-control research.

Quality Assessment of Included Literature

Two researchers independently used the Standards for Reporting Qualitative Research (SRQR) proposed by the University of California to evaluate the methodological quality of the literature, with disagreements resolved through discussion with a third researcher. The SRQR contains 21 items, with each item scored as "yes" (1 point), "no" (0 points), or "unclear" (0 points). The total score ranges from 0 to 21, with literature having SRQR scores $ $15 included in the analysis.

Statistical Methods

We used descriptive analysis to summarize methods for determining washout period duration.

Results

Literature Screening Process and Results

The database search yielded 21 Chinese articles and 1,203 English articles. After initial and secondary screening, 54 articles evaluating chronic disease prevalence/incidence and other epidemiological characteristics using administrative data were obtained. After excluding 28 articles that did not involve washout periods, 26 articles were ultimately included [2-6,8-28]. The literature screening flow chart is shown in Figure 1 [Figure 1: see original paper].

Basic Information of Included Literature and Databases

All 26 included articles were English-language publications. The data used primarily came from countries (regions) with complete and abundant administrative data, including Canada, the United States, and Australia. The focused diseases included diabetes, cancer, schizophrenia, and various other chronic diseases. Eleven articles [2,4,6,8,14,17,19,22-23,25,28] focused on cardiovascular diseases, and five [6,8,10,13,19] focused on diabetes. The main research purpose was to determine the incidence and prevalence of various chronic diseases. Twelve articles [3,6,9-11,13,16,19-20,24,26-27] used medical insurance data, eight [2,4,8,12,22-23,25,28] used hospital patient registration data, and the remaining studies used registered resident health project data [14-15] or registered disease survey project data [5,17-18,21]. The data year spans were substantial, with only two articles [15,19] using data covering less than 5 years, and 14 articles

[2,5,8,10-12,14,16,17,21,24-26,28] using data spanning $ $10 years. The basic information of included literature and databases is shown in Table 1 .

Methodological Quality Assessment Results of Included Literature

The SRQR scores of included literature were all $ $15 (due to space limitations, detailed quality assessment results are not presented in a separate table).

Methods for Determining Optimal Washout Period Duration

Through systematic literature review, we found that current methods for determining washout period duration mainly fall into three categories: direct restriction method, consistency test method, and retrograde survival function method. Among the 26 included articles, 15 [3,13-15,18-28] directly determined washout period duration based on expert opinion, clinical experience, literature review, or data limitations. Ten articles [2,4-6,9-12,16-17] used consistency test indicators such as Kappa values, positive predictive values, and overestimation rates to evaluate the impact of different washout period durations on incidence estimation and determine the optimal duration. Three articles [8-10] used the retrograde survival function method to determine the optimal washout period duration.

2.4.1 Consistency Test Method Ten articles [2,4-6,9-12,16-17] used consistency test indicators including Kappa values, positive predictive values, overestimation rates, negative predictive values, sensitivity, and accuracy. All studies used the longest achievable washout period based on available data as the "gold standard" for washout period duration, quantitatively comparing differences between the number of incident patients defined by different clearance periods (IDD) and the number defined by the gold standard (IDG). Six articles [2,4-6,16-17] used overestimation rate (overestimation rate = (IDD-IDG)/IDG$×$100%) as a consistency evaluation indicator. Three articles [9-11] used Kappa values, three [9,11-12] used positive predictive values, and one [12] used sensitivity, specificity, and negative predictive values. Currently, only judgment standards based on Kappa values have been proposed. According to criteria by BYRT et al. [29]: Kappa values of -1.00 to <0 indicate "no agreement"; 0 to 0.20 indicate "poor agreement"; >0.20 to 0.40 indicate "fair agreement"; >0.40 to 0.60 indicate "moderate agreement"; >0.60 to 0.80 indicate "good agreement"; >0.80 to 0.90 indicate "very good agreement"; and >0.90 to 1.00 indicate "excellent agreement." When the Kappa value is >0.90 to 1.00, indicating excellent agreement between the number of incident cases under the selected washout period and those determined by the gold standard, the selected washout period duration is considered optimal.

2.4.2 Retrograde Survival Function Method This method is generally implemented using Kaplan-Meier or life table methods. In this approach, the endpoint event is a patient being recorded during the washout period, while survival is defined as no record of the patient being captured during the given washout period (in which case the patient can be considered an incident case). Specifically, to calculate the number of incident cases for chronic disease i in year j: let a represent the date when a patient was first recorded for disease i in year j, b represent the most recent record date for disease i by that patient during the washout period before a, and c represent the start date of the washout period. If the patient was recorded during the given washout period, the survival time is lx = a - b; if not recorded, the patient is defined as "censored" (i.e., "survived"), with censoring time lc = a - c. Based on this, a survival function is established, and Kaplan-Meier or life table methods are used to calculate the survival probability [S(t)] of study subjects under a given washout period T, with survival curves plotted. Additionally, BRAMELD et al. [8] and BEAUDET et al. [9] used hazard functions [h(t)] to calculate the instantaneous risk of endpoint events in the next Δt time period for individuals who have already survived to time t, and plotted hazard function curves. The studies suggest that if there exists a time point tf where h(tf) approaches zero and S(tf) approaches a constant value, that time point represents the optimal washout period duration. In practice, BRAMELD et al. [8] considered $h(tf) < 0.00001$ as approaching zero, while BEAUDET et al. [9] did not provide specific criteria. ASGHARI et al. [10] did not calculate instantaneous risk through hazard functions, instead suggesting that when the survival probability of study subjects no longer changes at a certain time point tf (i.e., becomes stable), that time point represents the optimal washout period duration.

Washout Period Durations

The washout period durations set in included literature ranged from 0.5 to 15.0 years. When different literature focused on different diseases, there were significant differences in set washout period durations. However, even when different literature focused on the same disease, substantial differences existed. For example, among five articles [6,8,10,13,19] focusing on diabetes, the longest washout period was 13 years and the shortest was 1 year. The methods for determining washout period duration and the set durations in included literature are shown in Table 2 .

Discussion

This study systematically reviewed methods for determining optimal washout period duration when identifying chronic disease onset cases using administrative data. Currently, approximately half [48.1% (26/54)] of literature evaluating chronic disease prevalence/incidence and other epidemiological characteristics using administrative data set washout periods when identifying chronic disease onset cases, but no consensus exists regarding washout period duration. The washout period durations set in the 26 included articles ranged from 0.5 to 15.0 years. Methods for determining washout period duration mainly include three approaches: direct restriction method, consistency test method, and retrograde survival function method, with the direct restriction method being most commonly used.

This study found that more than half [57.7% (15/26)] of included literature used the direct restriction method to set washout period duration. Under this method, researchers directly set washout period duration based on expert opinion, literature review results, etc. While the direct restriction method is convenient and easy to implement, it lacks quantitative data support, making the set washout period duration less convincing to some extent. Some literature noted that directly setting washout period duration may lead to inaccurate incidence estimates [18,26]. Two articles [14,23] chose to use (maximum available data year - minimum available data year) as the washout period duration, but this approach only allows researchers to investigate disease incidence in the maximum available data year, resulting in inefficient data utilization.

Some literature [38.5% (10/26)] determined optimal washout period duration based on consistency test indicators [2,4-6,9-12,16-17], but different studies used different consistency test indicators. Currently, the most commonly used consistency test indicators are Kappa values and overestimation rates. The Kappa value criteria were proposed by BYRT et al. [29] in 1993. Most researchers chose Kappa values >0.80 ("good agreement" or "excellent agreement") as the standard for determining washout period duration [9-11]. However, some studies suggest that when identifying incident cases of high-prevalence diseases (such as diabetes and cardiovascular diseases), high Kappa values may be related to the large number of prevalent cases, and the stability of Kappa values requires further investigation [10]. Currently, there are no established standards for determining washout period duration based on consistency evaluation indicators other than Kappa values. For example, ABBAS et al. [6] used an overestimation rate of 10% as the critical standard, while ROSENLUND et al. [4] set this indicator at 20%. For positive predictive value, some researchers chose 80% as the threshold [11], while BENCHIMOL et al. [12] chose accuracy reaching 90% as the evaluation standard.

Only a small proportion of studies [11.5% (3/26)] used the retrograde survival function method to determine optimal washout period duration [8-10]. After BRAMELD et al. [8] first proposed this method in 2003, it has been subsequently used [9-10]. The retrograde survival function method emphasizes plotting survival and hazard function curves to determine optimal washout period duration based on survival and hazard probabilities. Researchers believe this method maximizes data utilization and ensures full use of all available records to estimate survival functions. Simultaneously, survival and hazard function curves allow intuitive observation of central trends in survival and hazard probabilities, enabling effective and quantitative determination of optimal washout period duration. However, this method also has limitations, as judgment criteria have not been fully standardized. For example, BRAMELD et al. [8] considered that if a time point tf exists where $h(tf) < 0.00001$, that time point represents the optimal washout period duration. According to this standard, the washout period in their study (focusing on diabetes) was as long as 13 years. ASGHARI et al. [10] (also focusing on diabetes) did not use this criterion, instead confirming a washout period duration of 5 years based on the standard that survival probability of study subjects becomes stable. Notably, due to lack of consensus on methods and judgment standards for determining washout period duration, many researchers chose to calculate multiple consistency indicators and combine them with results from the retrograde survival function method for comprehensive judgment [8-10].

Administrative data represent valuable resources with wide coverage, long time spans, and convenient accessibility, offering certain advantages over traditional survey data. Since the medical insurance system reform in 2012, China has vigorously promoted medical insurance policies and expanded coverage [30-31]. With gradual deepening of reforms, the volume of effective data continues to increase. However, few studies in Chinese databases (Wanfang, CNKI, CQVIP) have analyzed chronic disease epidemiological characteristics using medical insurance data or hospital patient registration data (only 21 articles as of October 1, 2022), far fewer than in other countries/regions with complete administrative data and mature research paradigms. This indicates that China is still in the exploratory stage in this field, with huge future development potential. Meanwhile, mainland Chinese researchers have not paid sufficient attention to washout periods when analyzing medical insurance data (none of the literature retrieved from Chinese databases involved washout periods), and no mainland researchers have explored how to determine washout period duration when analyzing disease epidemiological characteristics using Chinese population medical insurance data. Therefore, the three methods systematically summarized in this study can provide relevant ideas and methodological support for subsequent researchers to accurately identify chronic disease onset cases and explore disease epidemiological characteristics based on Chinese medical insurance data and other administrative data.

This study has certain limitations. First, due to the special nature of administrative data, some data may be internal or confidential, and some unpublished literature could not be retrieved during the search process, which may affect the research results. Second, the search was limited to Chinese and English databases, and relevant literature published in other languages was not retrieved, which may result in incomplete comprehensiveness of the identified methods.

References

[1] MENG Q, XU L, ZHANG Y, et al. Trends in access to health services and financial protection in China between 2003 and 2011: a cross-sectional study [J]. Lancet, 2012, 379(9818): 805-814. DOI: 10.1016/s0140-6736(12)60278-5.

[2] ØDEGAARD K M, LIRHUS S S, MELBERG H O, et al. A nationwide registry study on heart failure in Norway from 2008 to 2018: variations in lookback period affect incidence estimates [J]. BMC Cardiovasc Disord, 2022, 22(1): 88. DOI: 10.1186/s12872-022-02522-y.

[3] CZWIKLA J, JOBSKI K, SCHINK T. The impact of the lookback period and definition of confirmatory events on the identification of incident cancer cases in administrative data [J]. BMC Med Res Methodol, 2017, 17(1): 122. DOI: 10.1186/s12874-017-0388-1.

[4] ROSENLUND M, EKSTROM N, TORNBLOM M, et al. Impact of variable look-back periods on the incidence rates of chronic diseases using real world data [J]. Pharmacoepidemiology and Drug Safety, 2020, 29(9): 1086-1092. DOI: 10.1002/pds.5066.

[5] SCHMEDT N, KHIL L, BERGER K, et al. Incidence of multiple sclerosis in Germany: a cohort study applying different case definitions based on claims data [J]. Neuroepidemiology, 2017, 49(3/4): 91-98. DOI: 10.1159/000481990.

[6] ABBAS S, IHLE P, KOESTER I, et al. Estimation of disease incidence in claims data dependent on the length of follow-up: a methodological approach [J]. Health Services Research, 2012, 47(2): 746-755. DOI: 10.1111/j.1475-6773.2011.01325.x.

[7] O'BRIEN B, HARRIS T, BECKMAN J, et al. Standards for reporting qualitative research: a synthesis of recommendations [J]. Acad Med, 2014, 89(9): 1245-1251. DOI: 10.1097/acm.0000000000000388.

[8] BRAMELD K J, HOLMAN C D, LAWRENCE D M, et al. Improved methods for estimating incidence from linked hospital morbidity data [J]. Int J Epidemiol, 2003, 32(4): 617-624. DOI: 10.1093/ije/dyg191.

[9] BEAUDET N, COURTEAU J, SARRET P, et al. Improving the selection of true incident cases of low back pain by screening retrospective administrative data [J]. Eur J Pain, 2014, 18(7): 923-931. DOI: 10.1002/j.1532-2149.2013.00437.x.

[10] ASGHARI S, COURTEAU J, CARPENTIER A C, et al. Optimal strategy to identify incidence of diagnostic of diabetes using administrative data [J]. BMC Med Res Methodol, 2009, 9: 62. DOI: 10.1186/1471-2288-9-62.

[11] VANASSE A, COURTEAU J, FLEURY M J, et al. Treatment prevalence and incidence of schizophrenia in Quebec using a population health services perspective: different algorithms, different estimates [J]. Soc Psychiatry Psychiatr Epidemiol, 2012, 47(4): 533-543. DOI: 10.1007/s00127-011-0371-y.

[12] BENCHIMOL E I, GUTTMANN A, MACK D R, et al. Validation of international algorithms to identify adults with inflammatory bowel disease in health administrative data from Ontario, Canada [J]. J Clin Epidemiol, 2014, 67(8): 887-896. DOI: 10.1016/j.jclinepi.2014.02.019.

[13] LIN J, SHAU W, LAI M. Sex- and age-specific prevalence and incidence rates of sight-threatening diabetic retinopathy in Taiwan [J]. JAMA Ophthalmology, 2014, 132(8): 922-928. DOI: 10.1001/jamaophthalmol.2014.859.

[14] SULO G, IGLAND J, VOLLSET S, et al. Effect of the lookback period's length used to identify incident acute myocardial infarction on the observed trends on incidence rates and survival: cardiovascular disease in Norway project [J]. BMC Cardiovasc Disord, 2022, 22(1): 88. DOI: 10.1186/s12872-022-02522-y.

[15] RASSEN J, BARTELS D, SCHNEEWEISS S, et al. Measuring prevalence and incidence of chronic conditions in claims and electronic health record databases [J]. Clin Epidemiol, 2019, 11: 1-15. DOI: 10.2147/clep.S181242.

[16] STEIN J, BLACHLEY T, MUSCH D. Identification of persons with incident ocular diseases using health care claims databases [J]. Am J Ophthalmol, 2013, 156(6): 1169-1175. DOI: 10.1016/j.ajo.2013.06.035.

[17] CAMPLAIN R, KUCHARSKA-NEWTON A, CUTHBERTSON J, et al. Misclassification of incident hospitalized and outpatient heart failure in administrative claims data: the Atherosclerosis Risk in Communities (ARIC) study [J]. Pharmacoepidemiol Drug Saf, 2017, 26(4): 421-428. DOI: 10.1002/pds.4162.

[18] OHLMEIER C, SAUM K, GALETZKA W, et al. Epidemiology and health care utilization of patients suffering from Huntington's disease in Germany: real world evidence based on German claims data [J]. Bmc Neurology, 2019, 19(1): 318. DOI: 10.1186/s12883-019-1556-3.

[19] KOO B, LEE C, YANG B, et al. The incidence and prevalence of diabetes mellitus and related atherosclerotic complications in Korea: a National Health Insurance Database Study [J]. PLoS One, 2014, 9(10): e110650. DOI: 10.1371/journal.pone.0110650.

[20] JÜRISSON M, VOROBJOV S, KALLIKORM R, et al. The incidence of hip fractures in Estonia, 2005—2012 [J]. Osteoporosis International, 2015, 26(1): 77-84. DOI: 10.1007/s00198-014-2820-4.

[21] BUKHBINDER A, LING Y, HASAN O, et al. Risk of Alzheimer's disease following influenza vaccination: a claims-based cohort study using propensity score matching [J]. J Alzheimers Dis, 2022, 88(3): 1061-1074. DOI: 10.3233/jad-220361.

[22] ZHOU L, NGO L, FIELD T. Population incidence of cerebral venous thrombosis (CVT) from a large Canadian cohort 2001—2017. Stroke, 2022, 53(s1): 61. DOI: 10.1161/str.53.suppl-1.61.

[23] KATZENELLENBOGEN J, BOND-SMITH D, CUNNEEN R, et al. Contemporary Australian profile of rheumatic fever: results from multiple jurisdictions [J]. Heart Lung and Circulation, 2019, 28(s2): 42-43. DOI: 10.1016/j.hlc.2019.05.111.

[24] OU Y, GROSSMAN D, LEE P, et al. Glaucoma, Alzheimer disease and other dementia: a longitudinal analysis [J]. Ophthalmic Epidemiology, 2012, 19(5): 285-292. DOI: 10.3109/09286586.2011.649228.

[25] CLISSOLD B, SUNDARARAJAN V, CAMERON P, et al. Stroke incidence in Victoria, Australia-emerging improvements [J]. Front Neurol, 2017, 8: 180. DOI: 10.3389/fneur.2017.00180.

[26] AHN J, HWANG J, SEO G. Incidence and risk of developing rheumatic diseases in 19,724 patients with palindromic rheumatism in South Korea: a nationwide population-based study [J]. Joint Bone Spine, 2021, 88(3): 105128. DOI: 10.1016/j.jbspin.2020.105128.

[27] ANDERSON K, FUHRER R, ABRAHAMOWICZ M, et al. The incidence of first-episode schizophrenia-spectrum psychosis in adolescents and young adults in montreal: an estimate from an administrative claims database [J]. Can J Psychiatry, 2012, 57(10): 626-633. DOI: 10.1177/070674371205701007.

[28] KATZENELLENBOGEN J, SANFILIPPO F, HOBBS M, et al. Variable effects of prevalence correction of population denominators on differentials in myocardial infarction incidence: a record linkage study in Aboriginal and non-Aboriginal Western Australians [J]. J Clin Epidemiol, 2011, 64(6): 658-666. DOI: 10.1016/j.jclinepi.2010.08.008.

[29] BYRT T, BISHOP J, CARLIN J B. Bias, prevalence and kappa [J]. J Clin Epidemiol, 1993, 46(5): 423-429. DOI: 10.1016/0895-4356(93)90018-v.

[30] XIE L, QIN P, GAO X, et al. Development process, challenges and coping strategies of basic medical insurance system for urban and rural residents in China [J]. Chinese Journal of Public Health, 2020, 36(12): 1673-1676. DOI: 10.11847//zgggws1126456.

[31] WU R, LI N, ERCIA A. The effects of private health insurance on universal health coverage objectives in China: a systematic literature review [J]. Int J Environ Res Public Health, 2020, 17(6): 1660-4601. DOI: 10.3390/ijerph17062049.

Note: Figure translations are in progress. See original paper for figures.