J Chin Med Assoc

ORIGINAL ARTICLE

Cancer publications using real-world data from the Taiwan National Health Insurance Research Database: Conceptual framework and bibliometric analysis

Wing Hin Stanford Siu[a,b], Ai Yin Lim[c], Jia-Rou Liu[d], Shu-Hao Chang[d], Wei-Min Chen[d], Pei-Ru Li[d], Lai-Chu See[d,e,f,*]

[a]School of Medicine, College of Medicine, Chang Gung University, Taoyuan, Taiwan, ROC; [b]Department of Medical Education, Chang Gung Memorial Hospital, Taoyuan, Taiwan, ROC; [c]Graduate Institute of Physical Therapy, College of Medicine, Chang Gung University, Taoyuan, Taiwan, ROC; [d]Department of Public Health, College of Medicine, Chang Gung University, Taoyuan, Taiwan, ROC; [e]Biostatistics Core Laboratory, Molecular Medicine Research Center, Chang Gung University, Taoyuan, Taiwan, ROC; [f]Division of Rheumatology, Allergy and Immunology, Chang Gung Memorial Hospital, Taoyuan, Taiwan, ROC

Abstract

Background: Bibliometric analysis often overlooks study-based components such as study aims, design, and statistical methods. In this study, we propose a conceptual framework incorporating these study-based components with disease-based components for the bibliometric analysis of cancer articles using real-world data. This framework is a significant step forward in cancer research. We then investigated the distribution and temporal trends of these components for cancer articles using the Taiwan National Health Insurance Research Database (NHIRD) published from 2006 to 2022.

Methods: Study- and disease-based components were extracted and cross-validated. The distribution and temporal trends of these components were then presented.

Results: We analyzed 1232 articles and found a noticeable increase in the annual publication count from 2011 onward. This upward trend signified the growing momentum in cancer research. Cancer risk factors were the most studied (52%), followed by cancer outcomes (36%) and incidence/prevalence (3%). Among the publications on risk factors and outcomes, most were cohort studies (85%), followed by case–control studies (10.7%). In both study designs, the use of the propensity score method increased steadily from 2.4% in 2011 to 40% in 2022. The most frequently studied cancer site was "all cancers or multiple cancers" (25.6%), followed by breast (9.6%), hepatobiliary (9.2%), and colorectal cancers (8.8%). Among the top 10 cited articles, the first and fourth focused on whether suppressing hepatitis B viral load with nucleoside analogs could reduce hepatocellular carcinoma recurrence and incidence in chronic hepatitis B patients. The remaining eight examined the association between medications and cancer risk.

Conclusion: Beyond citation metrics, our research underscores the importance of considering study-based and disease-based components in bibliometric analysis. These components form the foundation of the real-world data cancer research framework and have practical implications for diseases beyond cancers, providing a broader perspective for researchers and practitioners.

Keywords: Bibliometric analysis; Cancer; Interrupted time series; Taiwan National Health Insurance Research; Target trial emulation

1. INTRODUCTION

1.1. Importance of real-world data

Cancer is the second leading cause of death globally,[1] and has remained the leading cause of death in Taiwan for the past four decades.[2,3] Although randomized controlled trials (RCTs) are considered the gold standard for evaluating the efficacy of interventions under controlled conditions, the effectiveness of interventions can be influenced by various factors. In addition, due to strict eligibility criteria, participants in RCTs may not represent the broader patient population. To overcome these challenges, real-world data (RWD) can be used to investigate how interventions perform in real-world scenarios. RWD commonly refers to health data collected from real-world settings outside the context of RCTs.[4] RWD includes administrative data, claims data, electronic health records, observational studies from prospective data collection, or that derived from personal devices.[5,6] After analysis, the results can be used to evaluate the efficacy of clinical trials, conduct post-marketing safety surveillance for interventions, assess utilization in community practice, examine changes in treatment patterns, generate hypotheses, and develop innovative treatments.[4–6]

1.2. An example of RWD from Taiwan

An example of RWD extensively used for cancer clinical research is the Taiwan National Health Insurance Research Database (NHIRD). The National Health Insurance (NHI) program in

*Address correspondence. Dr. Lai-Chu See, Department of Public Health, Chang Gung University, 259, Wen-Hwa 1st Road, Kweishan, Taoyuan 333, Taiwan, ROC. E-mail address: laichusee.taiwan.hk@gmail.com (L.-C. See).

Conflicts of interest: The authors declare that they have no conflicts of interest related to the subject matter or materials discussed in this article.

Journal of Chinese Medical Association. (2025) 88: 398-409.

Received April 30, 2024; accepted September 6, 2024.

doi: 10.1097/JCMA.0000000000001227

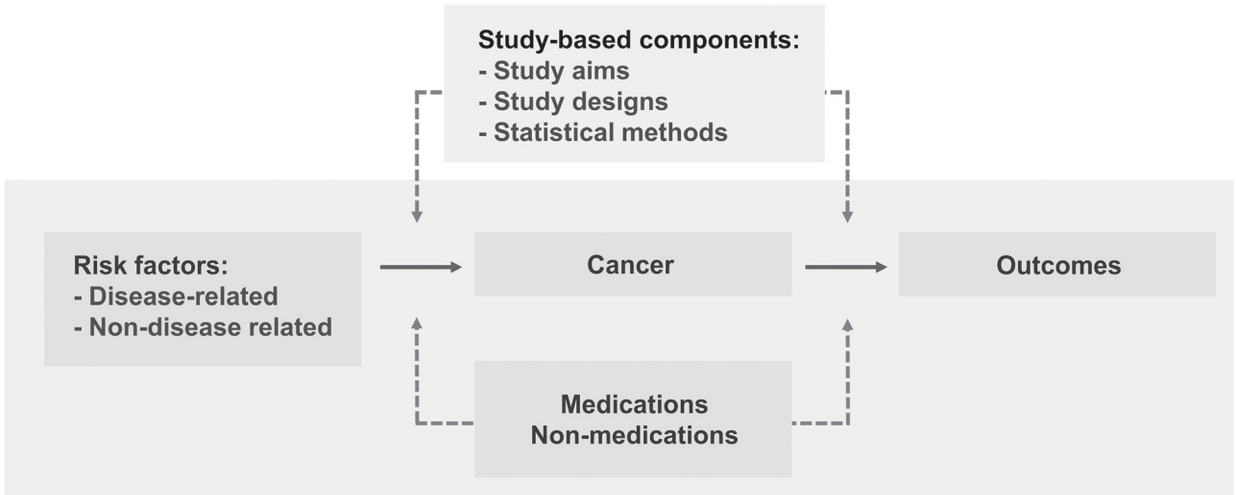

Fig. 1 The conceptual framework of cancer articles for bibliometric analysis.

Taiwan was founded in 1995 to provide medical insurance coverage to the entire population. The coverage rate was 97% in 1995 and has since increased to 99.99%.[7] The NHIRD contains longitudinal patient-level claims data from beneficiaries, including ambulatory care claims, inpatient claims, prescriptions dispensed at pharmacies, and registries for medical facilities and board-certified specialists. However, data on education level, results of laboratory tests, and personal lifestyle habits (such as cigarette smoking, alcohol consumption, diet, and exercise) are not available in the NHIRD.[8] Data from the NHIRD can be linked to other databases via an encrypted personal ID, including the Taiwan Cancer Registry, Cause of Death Data, Cancer Screening Data (Pap smear data, colorectal cancer screening, breast cancer screening, and oral mucosal screening), and Taiwan Biobank.[8] Consequently, the NHIRD provides valuable nationwide longitudinal RWD for healthcare research, which can be used to generate evidence to support clinical decisions and healthcare policymaking. The diagnostic accuracy of major comorbidities, such as cancers, diabetes, and ischemic stroke, in the NHIRD has been validated.[7] In a study investigating the validity of the diagnosis codes used in the NHIRD for cancers compared with the National Cancer Registry, the sensitivity was 91.5%, and the positive predictive value was 93.6%.[9] More than 4000 PubMed-indexed articles based on the NHIRD have been published since its release,[7,10] with an increase from 86 from 2000 to 2005 to 3751 from 2012 to 2017,[10] indicating that the NHIRD is gaining acceptance in population-based research in Taiwan. Notably, cancer was one of the top five medical conditions mentioned in the titles of studies that used the NHIRD in PubMed from 1996 to 2017.[10]

1.3. Framework for bibliometric analysis

Most previous bibliometric analyses of cancer research based on RWD have focused on citation-based components and cancer types. For example, a bibliometric analysis of 589 cancer articles using the NHIRD published from 2002 to 2015 showed that *PLoS ONE* and *Medicine* published the most articles.[11] The top five most studied cancer types were breast, lung, colorectal, liver, and prostate cancers.[11]

In contrast to citation-based components and cancer types, study-based components are often overlooked (Fig. 1). Study-based components refer to the features of methodologies, including study aims, design, and statistical methods. Study- and disease-based components (ie, cancer sites) are interrelated. Studies investigating specific cancer sites are more likely to include cancer staging and risk factors than those investigating all or multiple cancers. In the study aims, researchers may be interested in the incidence or prevalence rates of different cancers, identifying cancer risk factors, and assessing outcomes following a cancer diagnosis. Regarding the risk factors for cancer, disease-related risk factors are diseases associated with cancers, such as cardiovascular disease, diabetes, chronic kidney disease, inflammatory bowel disease, and periodontitis.[12–15] Non–disease-related risk factors include aging, occupational exposure, lifestyle-related risk factors, etc.[13,16] Clinicians are particularly interested in whether medical treatments for one disease increase or reduce a patient's cancer risk. Outcomes may include cause of death, time to recurrence, overall and cancer-specific survival, and second cancer.

Cancer treatment refers to pharmaceuticals or procedures for diagnosing, preventing, treating, offering supportive care, and relieving symptoms. Non-medical approaches include lifestyle changes, dietary interventions, etc. A text-mining analysis of cancer publications based on the NHIRD revealed a gradual increase in the use of hospice care and end-of-life care from 2012 to 2015.[11] Although terms related to cancer treatments—surgery, chemotherapy, and radiotherapy—were infrequent in most study periods, there was a slight increase in their usage from 2014 to 2015.[11] However, apart from the types of cancer treatment, the other components in our framework (especially study-based components) are often overlooked in the literature.

Taken together, we used these components to form a structured framework of cancer articles for bibliometric analysis (Fig. 1). A structured framework, assembled with many published studies, defines the backbone of a research topic (such as cancer research). In this study, we developed a conceptual framework for the analysis of cancer articles based on RWD, including (1) study-based (ie, study aims, design, and statistical methods); (2) disease-based (ie, cancer site, stage); and (3) citation-based components. We then applied this framework to perform a bibliometric analysis of NHIRD-based cancer articles published from 2006 to 2022 and reported the distribution and temporal trends of these components.

2. METHODS

We searched PubMed (n = 1312) and Scopus (n = 1248) for cancer articles that used the NHIRD and were published from

January 1, 2006, to February 20, 2023 (Fig. 2). The search strings used are summarized in Supplementary Table S1 (http://links.lww.com/JCMA/A320). There were 292 unique and 1134 duplicate records. Duplicate results were defined as those having the same title, abstract, ISSN, and DOI. We then manually screened the remaining records (n = 1334) so that only original articles written in English up to December 2022 were included. We excluded publications that were not based on the NHIRD or cancer-related (n = 87), did not mention study aim(s) (n = 5), or were published in 2023 (n = 10). A total of 1232 publications were included in this bibliometric analysis.

We downloaded citation-based and manually extracted study-based and disease-based components. Citation-based components included titles, publication years, journal categories, impact factors, citation counts, author names, affiliations, and countries. The impact factor was based on the Journal Impact Factor released by the Journal Citation Reports of the Web of Science Group in 2022. Study-based components included the study aims, design, and statistical methods. The study aims were categorized into four groups: (1) incidence or prevalence of reported cancer(s); (2) risk factors for reported cancer(s); (3) outcomes (such as survival rates) of reported cancer(s); and (4) others, such as economic analysis and dataset validation. If an article had multiple aims, it was extracted for each aim. We manually extracted the study design and statistical methods from the methodology section. Study designs included cohort, case–control, and others such as nested case–control, cross-sectional, and quasi-experimental. To ensure accuracy, all authors were involved in cross-verification of the interpretations of study design and statistical methods. The disease-based components included cancer sites, stages, risk factors, and outcomes. We determined the cancer sites based on information provided in the title, study aims, and keywords. Articles that studied all cancer sites were categorized as "all cancers." Descriptive statistics including mean, median,

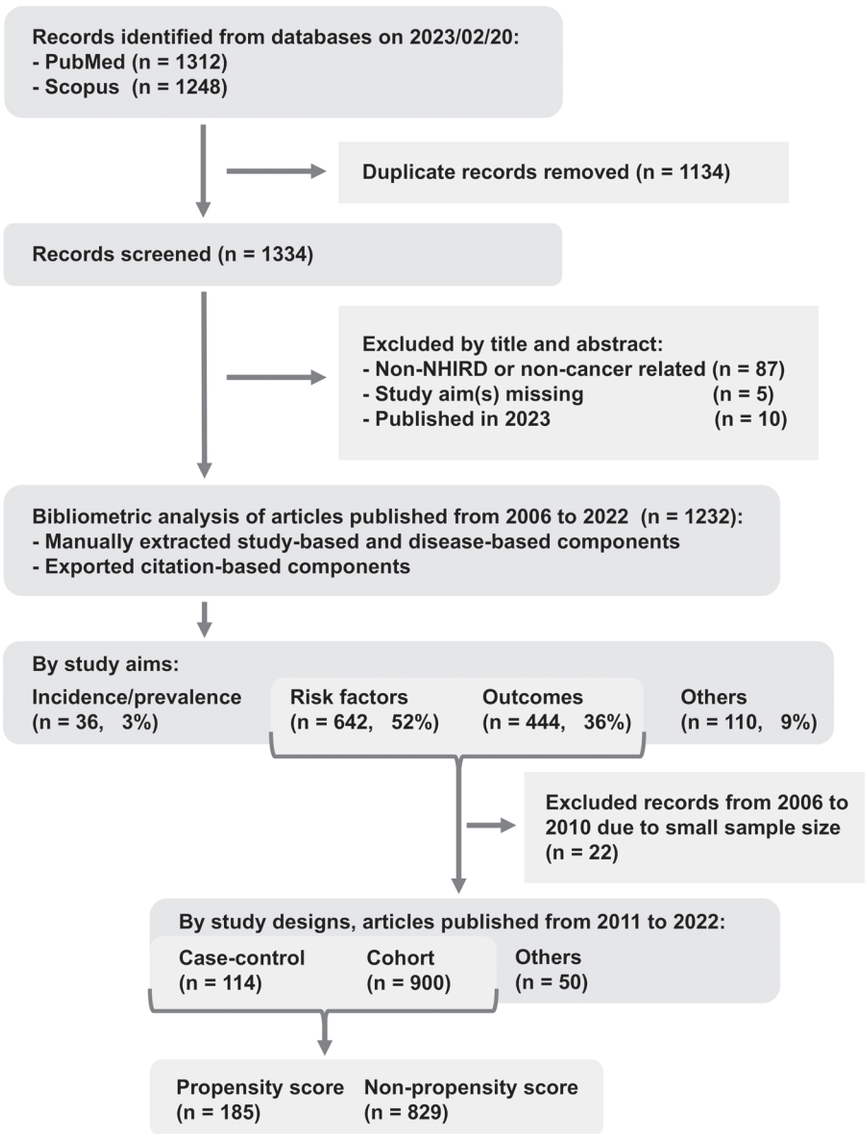

Fig. 2 Flowchart of the literature search for cancer articles using the NHIRD and a brief overview of the results. NHIRD = Taiwan National Health Insurance Research Database.

frequency, and percentage were used for the statistical analysis in this study.

3. RESULTS

3.1. Study-based components

Of the 1232 cancer publications, cancer risk factors were the most studied (n = 642, 52.1%), followed by cancer outcomes (n = 444, 36.0%) and incidence/prevalence (n = 36, 2.9%) (Fig. 2). Between 2006 and 2010, the annual publication count remained at 15 or fewer. It then increased to 48 in 2011 and fluctuated between 100 and 130 from 2014 onward, with a drop to 79 in 2020 due to the COVID-19 pandemic. After 2021, the number of articles increased again, consistently surpassing 120 annually (Fig. 3).

Regarding the temporal trends in cancer publications from 2006 to 2022 by study aims, there was a significant increase in the number of publications on cancer risk factors from 0 in 2006 to 68 (66.0%) in 2014. However, there was a slight decline in the proportion of risk factor publications after 2015, and the trend continued with fluctuations between 39.6% and 53.2%. The trend in publications on cancer outcomes showed a steady increase over the study period. Although the proportion of publications on cancer outcomes fluctuated between 21.7% and 28.1% from 2008 to 2014, a general upward trend emerged. In 2022, the proportion reached its peak at 47.5%. The trend in the number of cancer publications on incidence/prevalence was relatively low and sporadic from 2006 to 2011. There was a noticeable increase starting in 2012, reaching a peak of five publications in 2017. The trend continued with fluctuating proportions ranging from 1.8% to 6.3% in subsequent years (Fig. 3).

After excluding articles published from 2006 to 2010 because of the small percentage (2%), we analyzed the publications on cancer risk factors and outcomes from 2011 to 2022 (n = 1064). Regarding study design, most were cohort studies (84.6%), followed by case–control studies (10.7%) and others (4.7%, including 34 nested case–control studies [3.2%]) (Fig. 4). In addition, 80.8% of the publications on cancer risk factors were cohort studies, and 90.1% were on cancer outcomes. Case–control and nested case–control studies were used in 13.8% and 4.8% of the publications on cancer risk factors and in 6.2% and 0.9% of cancer outcomes, respectively.

Regarding temporal trends, the use of case–control studies in publications on cancer risk factors fluctuated, reaching a peak of 13 studies (22%) in 2019 and declining to three studies (6%) in 2022. Cohort studies were consistently the predominant type of study on cancer risk factors, with proportions ranging from 67.8% in 2019 to a peak of 89.3% in 2021. In publications on cancer outcomes, case–control studies fluctuated between

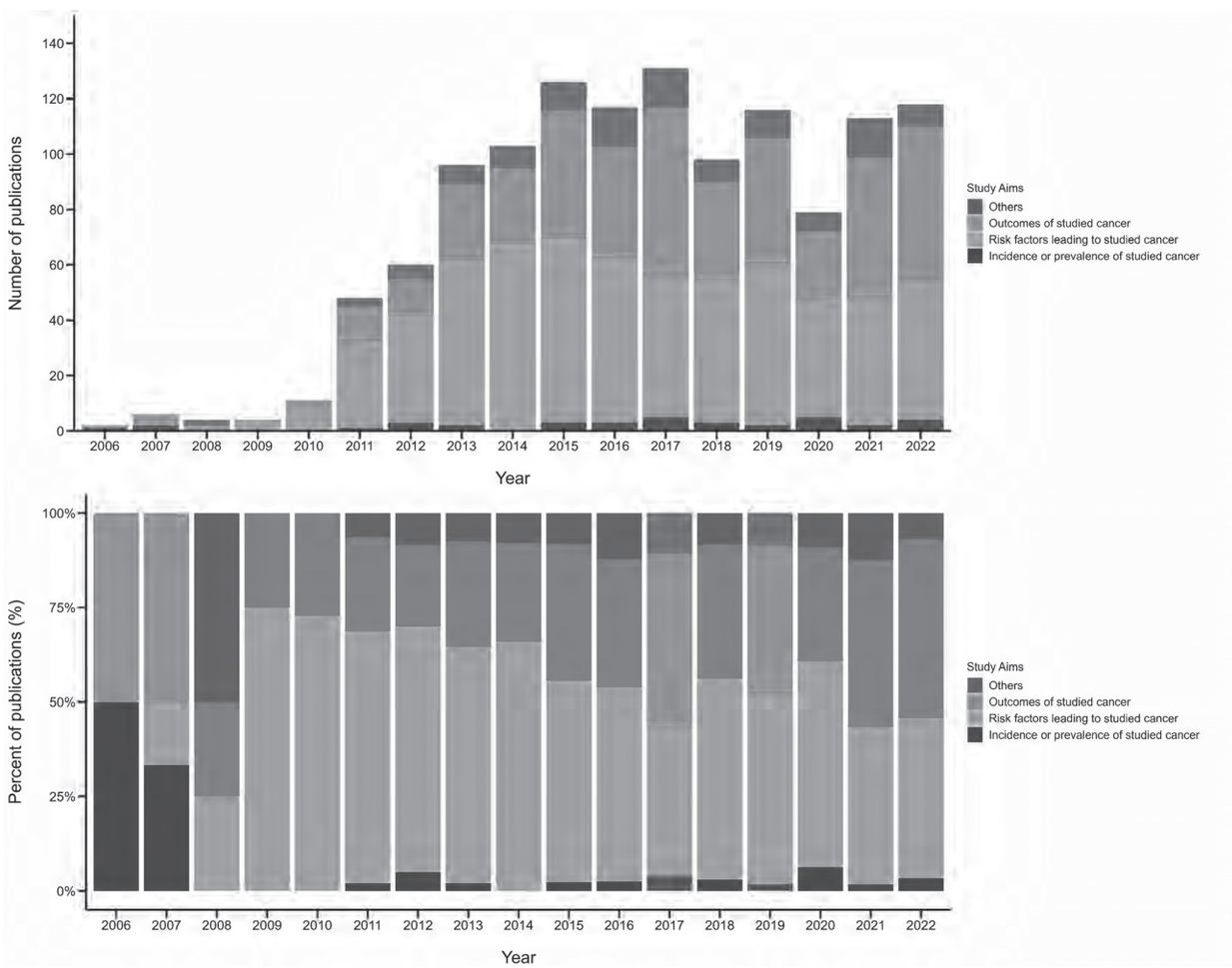

Fig. 3 The annual number and percentage of study aim for cancer research: 1232 articles using the Taiwan National Health Insurance Research Database and published from 2006 to 2022.

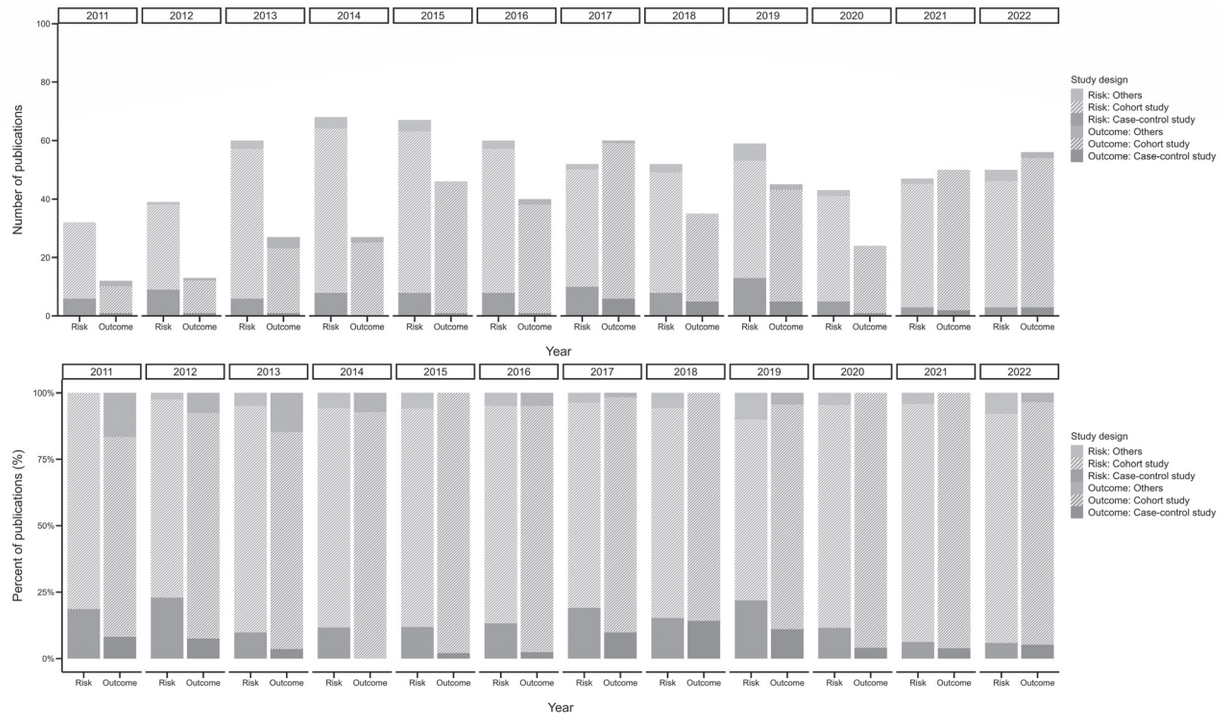

Fig. 4 The annual number and percentage of study designs on cancer risk factors or outcomes: 1064 articles using the Taiwan National Health Insurance Research Database and published from 2011 to 2022. Cancer articles from 2006 to 2010 were excluded because only 2% were published.

0 and 3 publications from 2011 to 2016, reaching a peak of 14.3% in 2018 and then decreasing to a range of 4% to 11.1% in the following years. The proportion of cohort studies on cancer outcomes increased from 75.0% in 2011 to 97.8% in 2015, indicating a consistent preference for a cohort study design in understanding cancer outcomes (Fig. 4).

From 2011 to 2022, the use of the propensity score method in cancer publications with case-control and cohort study designs showed a notable increasing trend. The proportion of studies using the propensity score method increased steadily from 2.4% in 2011 to 6.7% in 2014, followed by a continued increase to 21.9% in 2019. The trend persisted with a further increase ranging between 30.8% and 40% from 2020 to 2022, reaching a total of 185 over the study period (Fig. 5).

3.2. Citation-based components

Publications on risk factors had the highest mean and median impact factors of 5.98 and 4.48, respectively; publications on cancer outcomes had slightly lower mean and median impact factors of 5.27 and 3.82. Publications involving the incidence or prevalence of cancer had mean and median impact factors of 4.75 and 3.75, respectively (Fig. 6).

All of the top 10 cited cancer publications except the first and fourth investigated cancer risk factors. The article with the highest number of citations investigated the association between nucleoside analogs and tumor recurrence in patients with hepatitis B virus (HBV)–related hepatocellular carcinoma (HCC) after curative surgery[17]; the study was published in *JAMA* and had an impact factor of 157.335, with a total of 683 citations. The fourth-ranked article focused on whether suppressing HBV viral load by administering nucleoside analogs could reduce the incidence of HCC among patients with chronic hepatitis B.[18] The remaining eight articles examined whether medical treatments for a disease were associated with an increased/decreased risk of cancer. Of these top 10 cited articles, eight focused on hepatobiliary cancers, and four investigated HCC related to hepatitis B or C virus infections (Table 1).

3.3. Disease-based components

Overall, the most frequently studied cancer sites were "all cancers or multiple cancers" (n = 316, 25.6%), followed by breast (n = 118, 9.6%), hepatobiliary (n = 113, 9.2%), colorectal (n = 109, 8.8%), prostate (n = 87, 7.1%), lung (n = 86, 7.0%), and hematologic (n = 55, 4.5%) cancers. When categorized by study aims, "all cancers or multiple cancers" remained the most frequently studied category. However, the order differed for other cancer sites. In publications focusing on cancer risk factors, hepatobiliary cancers ranked second (11.3%), followed by colorectal (9.2%), breast (7.0%), lung (6.7%), prostate (5.9%), and hematologic (4.7%) cancers. For publications focusing on cancer outcomes, breast cancer ranked second (14.7%), followed by prostate (11.1%), colorectal (10.4%), lung (8.3%), hepatobiliary (7.3%), and hematologic (4.5%) cancers (Fig. 7).

Among the publications on cancer outcomes, mortality or survival was the most studied (n = 213, 47.9%), followed by second cancer (n = 21, 4.7%), recurrence (n = 17, 3.8%), and metastasis (n = 11, 2.5%).

4. DISCUSSION

4.1. Main findings

In this study, we proposed and applied a conceptual framework (study-based, citation-based, and disease-based components) to analyze NHIRD-based cancer articles published from 2006 to 2022. Publications on cancer risk factors had the highest impact factors and were the most studied. In the publications on cancer risk factors and outcomes, cohort studies were predominant, with the use of propensity scores increasing yearly. "All cancers

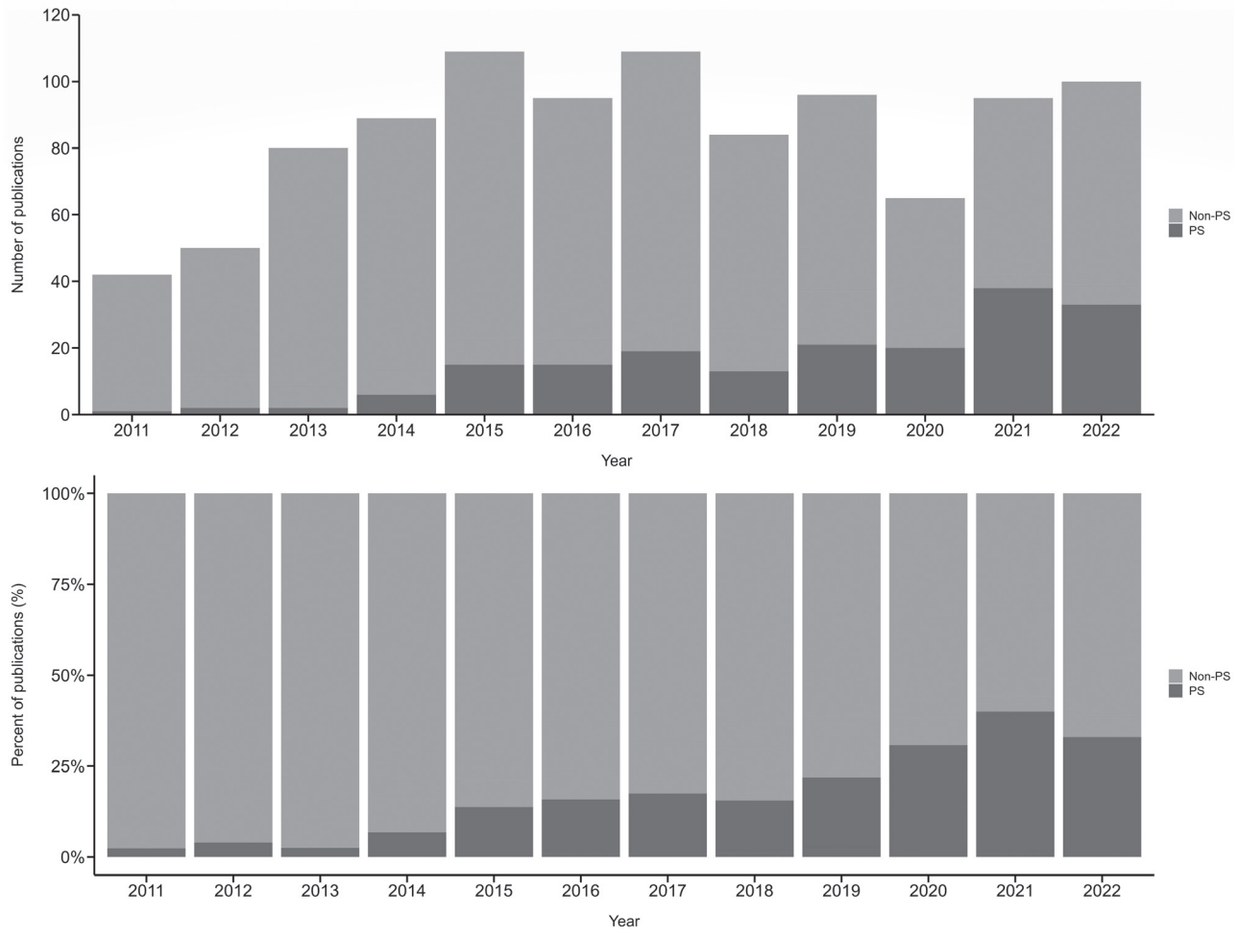

Fig. 5 The annual number and percentage of case–control and cohort studies using the PS method in cancer research: 1014 articles using the Taiwan National Health Insurance Research Database and published from 2011 to 2022. PS = propensity score.

or multiple cancers" were the most frequently studied sites, followed by breast, hepatobiliary, colorectal, prostate, and lung cancers. Among the top 10 cited articles, the first and fourth by the same study team focused on whether suppressing HBV viral load by administering nucleoside analogs could reduce the recurrence[17] and incidence of HCC among patients with chronic hepatitis B,[18] respectively. The remaining eight articles investigated whether medical treatments for a disease were associated with an increased/decreased risk of cancer. Our conceptual framework includes the three components of cancer articles and uses RWD, and we believe that it could be applied to other diseases.

4.2. Overall temporal trends of cancer publications

During the initial 5 years (2006-2010), the yearly publication count was 15 or fewer. This could be due to a lack of awareness regarding the availability of the NHIRD and the time required to establish suitable data analysis methods. In addition, the time required for the peer review and editorial process might have contributed to this finding. As researchers became familiar with the NHIRD and the associated methodologies used for data analysis, the publication count gradually increased to over 100 per year. Despite a temporary drop to about 80 in 2020 due to the COVID-19 pandemic, the number of articles rebounded to over 120 from 2021 onward as the pandemic subsided and restrictions eased. These findings are consistent with the overall trend of a significant decrease in non–COVID-19-related publications and notable increase in COVID–19-related publications, probably because research on COVID-19 was prioritized during this period.[19,20]

4.3. Cancer sites

In the analysis of the distribution of cancer sites by study aims, we found that study aims and disease-based components (ie, cancer sites) were closely related. The most commonly studied sites in publications investigating the incidence or prevalence of cancer were "all cancers" or "multiple cancers." However, in publications investigating cancer risk factors, hepatobiliary, colorectal, breast, and lung cancers were the top four studied types. This pattern could be linked to the epidemiological landscape of cancer in Taiwan. HCC is a consistently prominent topic of cancer research because Taiwan is an endemic region for hepatitis B and C virus infections.[21] Colorectal cancer and lung cancer are the most diagnosed types in Taiwan.[22] Regarding breast cancer, the peak age at diagnosis and age-specific incidence rates differ between Asian and Western countries.[23]

4.4. Cancer risk factors: most studied, highest impact

Regarding study aims, publications on cancer risk factors had the highest impact factor, followed by those on outcomes and incidence/prevalence. This aligns with the finding that cancer risk factor publications comprised half of all NHIRD-based

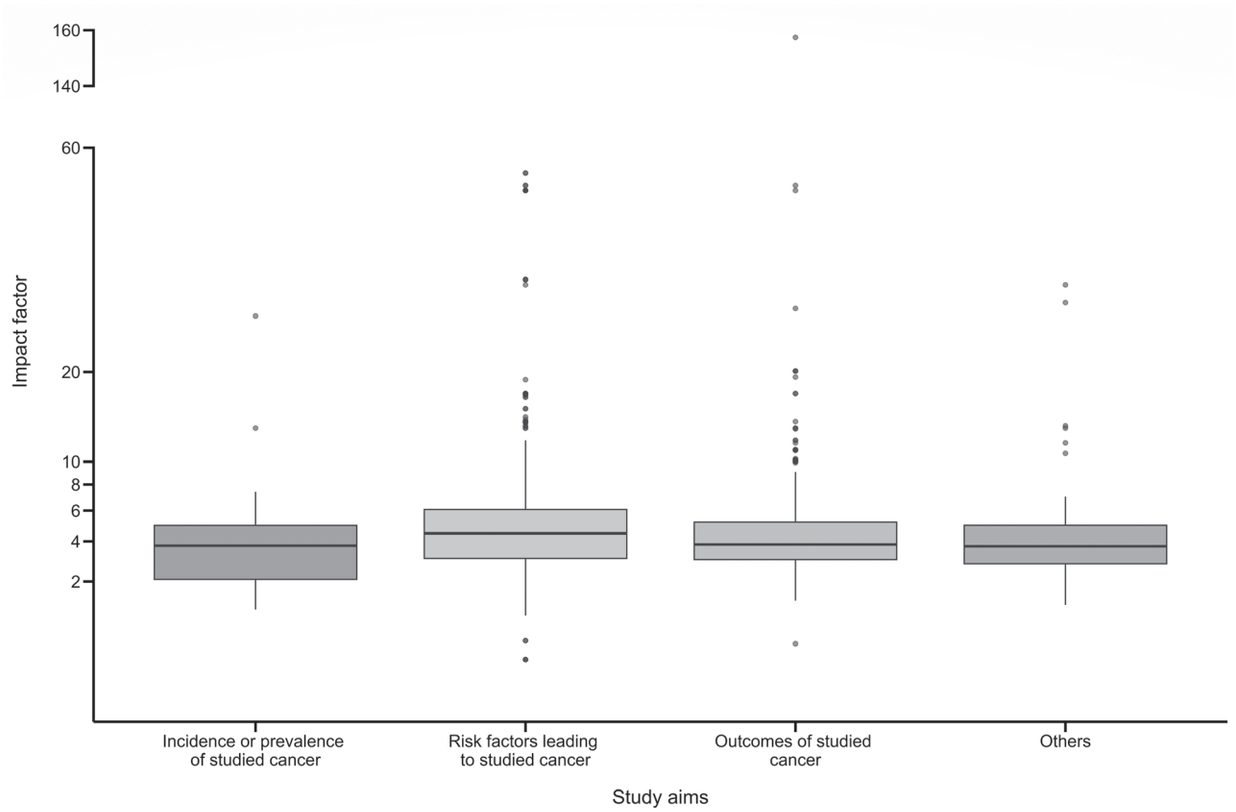

Fig. 6 The impact factor of cancer research: 1232 articles using the Taiwan National Health Insurance Research Database and published from 2006 to 2022.

cancer publications, followed by those on outcomes and incidence/prevalence. These findings suggest a possible relationship between study aims, the number of publications, and their impact factors. In fact, after 2009, there was a consistent trend of more publications focusing on cancer risk factors compared with those addressing cancer outcomes. Other bibliometric analyses of highly cited cancer publications have also reported that cancer risk factors were among the most represented in studies on oral, pharyngeal, and breast cancers.[24,25] "Risk factors" is also one of the most assigned keywords by PubMed indexers in publications on cancer molecular epidemiology.[26] These findings are probably due to the widespread interest in cancer risk factors among researchers from diverse disciplines. In Taiwan, researchers from various fields, not just oncologists, increasingly started to use the NHIRD as it gained recognition. Researchers studying cancer risk factors encompass many clinical specialists, extending beyond oncologists.

Researchers using RWD from the NHIRD directed significant attention and research efforts toward studying risk factors for hepatobiliary, colorectal, breast, and lung cancers. Among cancer risk factors, clinicians were particularly interested in the association between medications for a disease and subsequent cancer risk. In this study, we observed that the eight top-cited articles addressed whether medical treatments for a disease were associated with an increased/decreased risk of cancer (Table 1), aligning with our conceptual framework of cancer research (Fig. 1).

4.5. Cancer outcomes

Regarding cancer outcomes, mortality or survival was most studied (47.9%), and breast and prostate cancers were the second and third most studied sites. Over the past decade, there has been a notable increase in the incidence of breast cancer among Taiwanese females.[22] From 2010 to 2019, prostate cancer had the highest increase in annual percentage change in incidence rate among the top 10 cancers in Taiwanese males.[27] By 2021, female breast cancer ranked fourth, and prostate cancer ranked fifth in terms of mortality in Taiwan. In recent years, an increase in interest regarding the outcomes of these cancers may have led to their ranking as second and third most studied sites.

Few of the included studies investigated recurrence (3.8%). The top-cited article examined whether suppressing HBV viral load by administering nucleoside analogs could reduce the recurrence of HCC.[17] The results supported the association between using nucleoside analogs and reduced risk of HCC recurrence in HBV patients after liver resection. These findings are clinically significant because HBV viral load is one of the most clinically controllable risk factors for HCC.

4.6. Study designs

Since publications on cancer risk factors and outcomes were predominant, we further analyzed the study designs employed in these publications. Cohort studies were the primary design due to the longitudinal nature and large population size of the NHIRD. The advantages of cohort studies include assessing causality, investigating multiple outcomes and rare exposure, and providing disease rates in exposed and unexposed individuals over time (eg, incidence rate and relative risk/rate ratio). However, cohort studies are susceptible to loss to follow-up (ie, attrition bias), selection bias, numerous confounders, and changes in exposure over time.[28,29] Attrition in the NHIRD is negligible because all residents in Taiwan are legally required

Table 1

Top 10 cited cancer articles using the National Health Insurance Research Database of Taiwan published from 2006 to 2022

Authors	Title	Year	Journal	Impact factor	Number of citation
(1) Wu CY, Chen YJ, Ho HJ, Hsu YC, Kuo KN, Wu MS, Lin JT	Association between nucleoside analogues and risk of hepatitis B virus-related hepatocellular carcinoma recurrence following liver resection	2012	*JAMA*	157.335	683
(2) Lee MS, Hsu CC, Wahlqvist ML, Tsai HN, Chang YH, Huang YC	Type 2 diabetes increases and metformin reduces total, colorectal, liver and pancreatic cancer incidences in Taiwanese: A representative population prospective cohort study of 800,000 individuals	2011	*BMC Cancer*	4.638	390
(3) Lai SW, Chen PC, Liao KF, Muo CH, Lin CC, Sung FC	Risk of hepatocellular carcinoma in diabetic patients and risk reduction associated with anti-diabetic therapy: A population-based cohort study	2012	*American Journal of Gastroenterology*	12.045	247
(4) Wu CY, Lin JT, Ho HJ, Su CW, Lee TY, Wang SY, Wu C, Wu JC	Association of nucleos(T)ide analogue therapy with reduced risk of hepatocellular carcinoma in patients with chronic hepatitis B—A nationwide cohort study	2014	*Gastroenterology*	33.883	235
(5) Wu C, Kuo KN, Wu M, Chen Y, Wang C, Lin J	Early helicobacter pylori eradication decreases risk of gastric cancer in patients with peptic ulcer disease	2009	*Gastroenterology*	33.883	232
(6) Tsan YT, Lee CH, Wang JD, Chen PC	Statins and the risk of hepatocellular carcinoma in patients with hepatitis B virus infection	2012	*Journal of Clinical Oncology*	50.717	225
(7) Lai MN, Wang SM, Chen PC, Chen YY, Wang JD	Population-based case-control study of Chinese herbal products containing aristolochic acid and urinary tract cancer risk	2010	*Journal of the National Cancer Institute*	11.816	199
(8) Tsan YT, Lee CH, Ho WC, Lin MH, Wang JD, Chen PC	Statins and the risk of hepatocellular carcinoma in patients with hepatitis C virus infection	2013	*Journal of Clinical Oncology*	50.717	175
(9) Chang CH, Lin JW, Wu LC, Lai MS, Chuang LM, Arnold Chan K	Association of thiazolidinediones with liver cancer and colorectal cancer in type 2 diabetes mellitus	2012	*Hepatology*	17.298	153
(10) Huang WY, Muo CH, Lin CY, Jen YM, Yang MH, Lin JC, Sung FC, Kao CH	Pediatric head CT scan and subsequent risk of malignancy and benign brain tumor: A nationwide population-based cohort study	2014	*British Journal of Cancer*	9.075	150

to be insured by the NHI, which covers more than 99% of the population.[8] Changes in exposure over time, especially with medication use, remain a challenge in studies examining cancer risk factors using the NHIRD.

Case–control (10.7%) and nested case–control (3.2%) study designs were less frequently used. The case–control design is particularly advantageous when studying rare outcomes because there is no need to enroll all controls. The core of a case–control study is that cases with the condition under study are derived from a source population, and controls are a representative sample of the same source population. Hence, whether the cases and controls are representative samples from the same source population is uncertain.[30] This drawback is not an issue in nested case–control studies because cases and controls are nested within well-defined cohorts.[31] In fact, many case–control studies examining cancer risk factors using the NHIRD should be nested case–control studies. For example, the seventh most cited article identified all patients newly diagnosed with urinary tract cancer (case subjects) from 2001 to 2002. The control subjects were randomly selected from a 200,000-person random sample of the entire insured population from 1997 to 2002.[32] Unfortunately, this article did not specify whether the controls were obtained by random selection with or without replacement from non-cases. Lubin and Gail[33] stated that bias arises when (1) requiring controls to remain completely disease-free for a fixed time interval, and (2) excluding all incident cases during observation as controls. They suggested that controls should be sampled randomly with replacement from the entire risk set or without replacement from the noncases. The exclusion of future cases (ie, a noncase member of the risk set that later becomes a case) as controls should be avoided.[34]

In observational studies addressing cancer risk factors or outcomes, the problem of confounding can seriously impact the validity of the analysis. Therefore, we investigated temporal trends using the propensity score method in the case–control and cohort studies. We found a gradual increase in the use of propensity scores from 2010. Unlike the random assignment in clinical trials to balance the differences between treatment and nontreatment groups, observational studies have no randomization, such as those using the NHIRD. The propensity score method gained popularity after 2010 to resolve this limitation and reduce confounding in observational studies. This approach uses propensity score weighting and matching to estimate the average treatment effect for the population and the average treatment effect for the treated, respectively.[35] Our findings are consistent with the increasing popularity of the propensity score method in cancer research literature based on large databases of RWD.[36]

4.7. Discrepancies in the causal effect of certain medications on cancer between observational studies and clinical trials

According to our conceptual framework of cancer research (Fig. 1), the clinicians were particularly interested in whether a medication treatment for one disease increased or reduced cancer risk. Notable examples included metformin (the first-line diabetic drug) and statins (the first-line medication for hypercholesterolemia). Among four of the top 10 cited articles, two examined the association of metformin with gastrointestinal cancer, and another two examined the association of statins with HCC. The findings from these four observational studies and clinical trials are conflicting. Observational studies[37,38] and healthcare claims data[39–41] have reported a protective effect of metformin against cancer incidence. However, RCTs have concluded that metformin provides no benefit for breast cancer[42,43] or worsens the prognosis for gastrointestinal system cancer.[44] For statins, observational studies have suggested that their use in patients

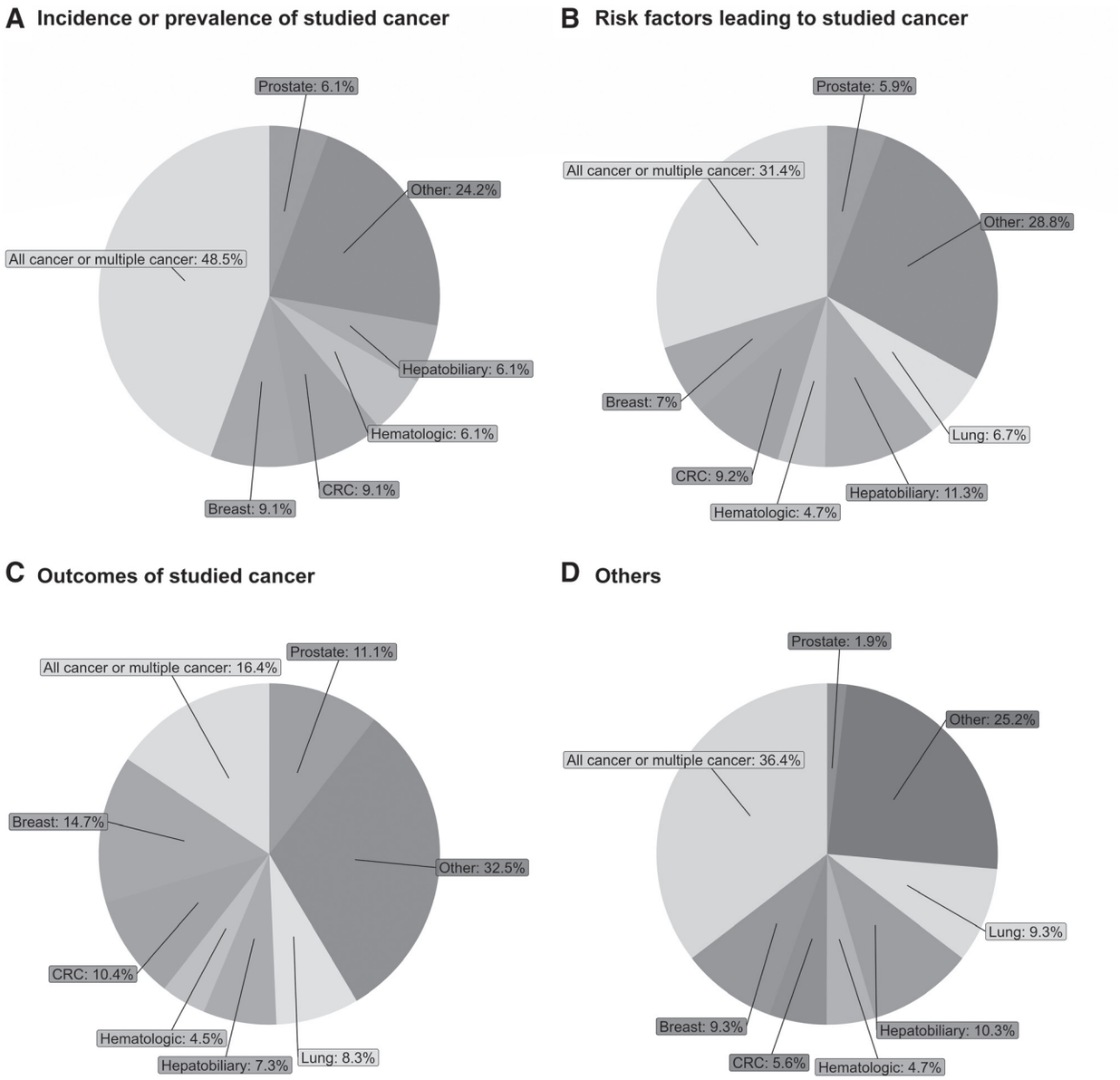

Fig. 7 Distribution of cancer sites by study aims in cancer research: 1232 articles using the Taiwan National Health Insurance Research Database and published from 2006 to 2022. CRC = colorectal cancer.

with cancer is associated with reduced cancer-related mortality.[45,46] However, the evidence from several trials was insufficient to confirm or refute the efficacy and safety of statins in patients with solid malignant tumors,[47] or whether they improved overall survival or progression-free survival in patients with advanced cancer and a prognosis of <2 years.[48] Selection bias, immortal time bias, residual/unmeasured confounding, and reporting/detection bias may explain the discrepancies in the causal effect of medications on the risk of cancer between observational studies and clinical trials. Selection bias occurs when cancer patients may die from the disease before they can derive cardiovascular benefits from statins.[48] Immortal time bias is likely when participants must remain event-free until they start taking the medication of interest. The outcome rate is systematically underestimated in the medication-exposed group but overestimated in the unexposed group, thereby falsely suggesting that exposure to the medication of interest prevents the outcome under study.[49] Residual or unmeasured confounding in observational studies persists even after statistical adjustment. These biases may be further amplified by concomitant changes in lifestyle in patients who took the medication of interest.[50] Reporting/detection bias of health outcomes occurs between patients who took and did not take the medication of interest because those who took the drug knew the side effects and were closely monitored.[50] Reporting bias also occurs due to exposure and poor adherence to the study medication.[48]

Without conducting a real trial, Hernán and Robins[51] first proposed target trial emulation (TTE) to assess the real-world effectiveness and safety of medical treatments using observational big data. The framework of TTE has seven key components: (1) eligibility criteria, (2) treatment strategies, (3) assignment procedures, (4) follow-up period, (5) outcomes, (6)

causal contrasts of interest, and (7) analysis plan. In TTE, these seven components of the target trial protocol are compared with an observational study to identify potential sources of bias and adjust the results accordingly.[51] (1) Regarding the participants' eligibility criteria, strict inclusion criteria are necessary. Patients should not be included based on future information. Prevalent users, those exposed to the medication of interest before the study begins, should also be avoided. Confounding may also be introduced when the medication influences covariates for participants exposed to the medication of interest.[52] (2) Regarding treatment strategies, tight monitoring and enforcing adherence to the study protocol is not recommended when emulating a placebo-controlled trial.[51] A new-user design, which restricts the analysis to persons under observation at the start of the current course of their treatment, is suggested to avoid selection bias. Because individuals meet eligibility criteria defined after initiating a treatment strategy, they may be influenced by the strategy itself. (3) Regarding assignment procedures, randomization and blinding are impossible in observational studies. In other words, the two study groups have very different baseline characteristics, and the patients and their healthcare workers know the treatments they receive. It is suggested that baseline confounders should be adjusted by matching, stratification or regression, inverse probability weighting, etc.[51] Unfortunately, laboratory data or personal lifestyle habits are usually unavailable in healthcare claim datasets, as with the NHIRD. It is also recommended to compare an active treatment with another active treatment rather than with no treatment (or usual care).[51] (4) Regarding the follow-up period, Matthews et al[53] stated that starting follow-up should coincide with three conditions: when eligibility criteria are met, treatment strategies are assigned, and study outcomes begin to be counted. However, these three conditions usually differ in an observational study when selecting prevalent users.[53] (5) Regarding outcomes, it is recommended to emulate a target trial with systematic and blinded outcome ascertainment to mitigate the influence of treatment status on clinicians' decisions to pursue desired outcomes.[51] For example, a death registry is a better ascertainment source than a medical chart when the outcome is death. (6) Regarding causal contrasts of interest, analogs of intention-to-treat and per-protocol effects from observational data should be provided. The intention-to-treat effect is the comparative effect of being assigned to the treatment strategy at baseline, regardless of whether the individual continues to follow the strategy after baseline. The per-protocol effect is the comparative effect of following the treatment strategy specified in the study protocol. (7) Regarding the analysis plan, it is recommended to examine different treatment strategies based on therapy prescriptions at baseline and to conduct an intention-to-treat analysis.[51] It is also recommended that the results of the per-protocol analysis be provided if possible.[51] Since introducing the TTE concept, many studies have been published, including three studies using the NHIRD.[54–56] None of these three articles was cancer-related.

Apart from TTE, we suggest using interrupted time-series (ITS) analysis to assess the effects of specific medical treatments on cancer when the medical treatment of interest was introduced at a specific time. ITS analysis includes one time series (single ITS analysis) or more than one time series (controlled interrupted time-series analysis, CITS) of observations on the same outcome (mainly at an aggregate level), which is interrupted by an intervention at a known point in time. Hence, the intervention effect can be easily estimated by comparing the change in the level and slope between the pre- and postintervention periods using segmented regression analysis.[57] The strengths of these statistical methods include (1) using observational data in an aggregate format rather than at an individual level, (2) providing clear and easy-to-interpret graphical results, (3) utilizing statistical methods readily available, and (4) minimizing selection bias if every individual is included when calculating the outcomes.[58] In this study, four cancer articles used ITS. The first analyzed temporal trend changes in the incidence and mortality due to HCC before and after the National Antiviral Treatment Program was implemented in 2003, which reimbursed patients for antiviral drugs and interferon for chronic hepatitis B and C.[59] The second article explored whether changes in reimbursement policies had a significant impact on the accessibility of targeted therapies for metastatic non–small-cell lung cancer treatment.[60] The third article assessed the real-world impact of adjuvant oxaliplatin treatment on the survival of patients with stage III colon cancer,[61] and the fourth examined changes in cancer incidence rates in three phases of disease-modifying antirheumatic drugs among patients with rheumatoid arthritis.[62]

We suggest a collaborative approach, encouraging researchers with diverse clinical, epidemiologic, and biostatistics expertise to work together. RWD from the NHIRD can be analyzed using TTE and ITS to evaluate the benefit–risk of clinical interventions, fostering a sense of shared responsibility and collective progress in the field.

This study has several limitations. In citation-based components, we limited our literature search to cancer articles published in English, potentially leading to an underestimation of the publication count. However, in Taiwan, the assessment of research performance relies heavily on the number of articles published in the journals indexed by the Science Citation Index. Therefore, we believe most cancer publications based on the NHIRD can be found in PubMed and Scopus. In addition, although impact factors are commonly used in research evaluation, they are not without limitations and are vulnerable to manipulation.[63] A holistic assessment of journal-level bibliometrics is recommended instead of relying solely on the impact factor.[64]

In study-based components, we did not further identify the specific statistical tests employed or assess their appropriateness of use. Notably, propensity score analysis may not be conducted properly in cancer publications using RWD.[36] For example, the variables used for propensity score estimation may not be specified clearly, and non-baseline variables may be incorrectly included in the analysis.[36] Further research is needed to assess the validity and appropriateness of statistical tests used in cancer publications based on the NHIRD.

In disease-based components, we did not further categorize the risk factors into disease-related and non–disease-related risk factors, as illustrated in our proposed conceptual framework (Fig. 1). These areas can be further investigated in future studies. In addition, we did not analyze the interventions (ie, medications and nonmedications) reported in the publications because they have been extensively discussed elsewhere.[11]

In conclusion, this study proposed a conceptual framework for cancer research based on RWD for bibliometric analysis. In addition to the citation-based components (such as impact factor and citation count), it is also essential to consider study-based components, including the study design, aims, and statistical tests. Moreover, a holistic evaluation of cancer publications should include disease-based components, medical or nonmedical interventions, and outcomes. In our proposed framework, all of these components not only define the backbone of cancer research, but could also be extended to encompass other diseases.

ACKNOWLEDGMENTS

This work was partially supported by the Ministry of Science and Technology (Taiwan) under grant number 111-2813-C-182-033-B and the Chang Gung Medical Foundation under CMRPD1M0211.

We thank the Ministry of Science and Technology (Taiwan) for grant 111-2813-C-182-033-B and the Chang Gung Medical Foundation for grant CMRPD1N0451.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data related to this article can be found at http://links.lww.com/JCMA/A320.

REFERENCES

1. Kocarnik JM, Compton K, Dean FE, Fu W, Gaw BL, Harvey JD, et al; Global Burden of Disease 2019 Cancer Collaboration. Cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life years for 29 cancer groups from 2010 to 2019: a systematic analysis for the global burden of disease study 2019. *JAMA Oncol* 2022;8:420–44.
2. Department of Statistics, Ministry of Health and Welfare, Taiwan. Cause of Death Statistics, 2020. Available at https://www.mohw.gov.tw/cp-5256-63399-2.html. Accessed January 26, 2023.
3. Tai SY, Cheon S, Yamaoka Y, Chien YW, Lu TH. Changes in the rankings of leading causes of death in Japan, Korea, and Taiwan from 1998to 2018: a comparison of three ranking lists. *BMC Public Health* 2022;22:926.
4. Saesen R, Van Hemelrijck M, Bogaerts J, Booth CM, Cornelissen JJ, Dekker A, et al. Defining the role of real-world data in cancer clinical research: the position of the European Organisation for Research and Treatment of Cancer. *Eur J Cancer* 2023;186:52–61.
5. Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology. *J Natl Cancer Inst* 2017;109. Doi: 10.1093/jnci/djx187.
6. Booth CM, Karim S, Mackillop WJ. Real-world data: towards achieving the achievable in cancer care. *Nat Rev Clin Oncol* 2019;16:312–25.
7. 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.
8. Lin LY, Warren-Gash C, Smeeth L, Chen PC. Data resource profile: the National Health Insurance Research Database (NHIRD). *Epidemiol Health* 2018;40:e2018062.
9. Kao WH, Hong JH, See LC, Yu HP, Hsu JT, Chou IJ, et al. Validity of cancer diagnosis in the National Health Insurance database compared with the linked National Cancer Registry in Taiwan. *Pharmacoepidemiol Drug Saf* 2018;27:1060–6.
10. Sung SF, Hsieh CY, Hu YH. Two decades of research using Taiwan's National Health Insurance claims data: bibliometric and text mining analysis on PubMed. *J Med Internet Res* 2020;22:e18457.
11. Chiang JK, Lin CW, Wang CL, Koo M, Kao YH. Cancer studies based on secondary data analysis of the Taiwan's National Health Insurance Research Database: a computational text analysis and visualization study. *Medicine (Baltim)* 2017;96:e6704.
12. Gordon H, Biancone L, Fiorino G, Katsanos KH, Kopylov U, Al Sulais E, et al. ECCO Guidelines on inflammatory bowel disease and malignancies. *J Crohns Colitis* 2023;17:827–54.
13. Tran KB, Lang JJ, Compton K, Xu R, Acheson AR, Henrikson HJ, et al. The global burden of cancer attributable to risk factors, 2010-19: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2022;400:563–91.
14. Tu H, Wen CP, Tsai SP, Chow WH, Wen C, Ye Y, et al. Cancer risk associated with chronic diseases and disease markers: prospective cohort study. *BMJ* 2018;360:k134.
15. Bui FQ, Almeida-da-Silva CLC, Huynh B, Trinh A, Liu J, Woodward J, et al. Association between periodontal pathogens and systemic disease. *Biomed J* 2019;42:27–35.
16. Collatuzzo G, Boffetta P. Cancers attributable to modifiable risk factors: a road map for prevention. *Annu Rev Public Health* 2023;44:279–300.
17. Wu CY, Chen YJ, Ho HJ, Hsu YC, Kuo KN, Wu MS, et al. Association between nucleoside analogues and risk of hepatitis B virus–related hepatocellular carcinoma recurrence following liver resection. *JAMA* 2012;308:1906–14.
18. Wu CY, Lin JT, Ho HJ, Su CW, Lee TY, Wang SY, et al. Association of nucleos (t) ide analogue therapy with reduced risk of hepatocellular carcinoma in patients with chronic hepatitis B—a nationwide cohort study. *Gastroenterology* 2014;147:143–51.e5.
19. Aviv-Reuven S, Rosenfeld A. Publication patterns' changes due to the COVID-19 pandemic: a longitudinal and short-term scientometric analysis. *Scientometrics* 2021;126:6761–84.
20. Raynaud M, Goutaudier V, Louis K, Al-Awadhi S, Dubourg Q, Truchot A, et al. Impact of the COVID-19 pandemic on publication dynamics and non-COVID-19 research production. *BMC Med Res Methodol* 2021;21:255.
21. Liao SH, Chen CL, Hsu CY, Chien KL, Kao JH, Chen PJ, et al. Long-term effectiveness of population-wide multifaceted interventions for hepatocellular carcinoma in Taiwan. *J Hepatol* 2021;75:132–41.
22. Kuo CN, Liao YM, Kuo LN, Tsai HJ, Chang WC, Yen Y. Cancers in Taiwan: practical insight from epidemiology, treatments, biomarkers, and cost. *J Formos Med Assoc* 2020;119:1731–41.
23. Lin CH, Yap YS, Lee KH, Im SA, Naito Y, Yeo W, et al; Asian Breast Cancer Cooperative Group. Contrasting epidemiology and clinicopathology of female breast cancer in Asians vs the US population. *J Natl Cancer Inst* 2019;111:1298–306.
24. Hassona Y, Qutachi T. A bibliometric analysis of the most cited articles about squamous cell carcinoma of the mouth, lips, and oropharynx. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2019;128:25–32.e6.
25. Bayar B, Peksoz R. Analysis of trending topics in breast cancer articles from an altmetric perspective. *Cureus* 2022;14:e26565.
26. Ugolini D, Puntoni R, Perera FP, Schulte PA, Bonassi S. A bibliometric analysis of scientific production in cancer molecular epidemiology. *Carcinogenesis* 2007;28:1774–9.
27. Health Promotion Administration, Ministry of Health and Welfare, Taiwan. Health Promotion Administration Annual Report, 2022. Available at https://www.hpa.gov.tw/EngPages/Detail.aspx?nodeid=1070&pid=16384. Accessed November 2022.
28. Song JW, Chung KC. Observational studies: cohort and case-control studies. *Plast Reconstr Surg* 2010;126:2234–42.
29. Morrow B. An overview of cohort study designs and their advantages and disadvantages. *Int J Ther Rehabil* 2010;17:518–23.
30. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959;22:719–48.
31. Biesheuvel CJ, Vergouwe Y, Oudega R, Hoes AW, Grobbee DE, Moons KG. Advantages of the nested case-control design in diagnostic research. *BMC Med Res Methodol* 2008;8:1–7.
32. Lai MN, Wang SM, Chen PC, Chen YY, Wang JD. Population-based case–control study of Chinese herbal products containing aristolochic acid and urinary tract cancer risk. *J Natl Cancer Inst* 2010;102:179–86.
33. Lubin JH, Gail MH. Biased selection of controls for case-control analyses of cohort studies. *Biometrics* 1984;40:63–75.
34. Robins JM, Gail MH, Lubin JH. More on" Biased selection of controls for case-control analyses of cohort studies". *Biometrics* 1986;42:293–9.
35. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399–424.
36. Yao XI, Wang X, Speicher PJ, Hwang ES, Cheng P, Harpole DH, et al. Reporting and guidelines in propensity score analysis: a systematic review of cancer and cancer surgical studies. *J Natl Cancer Inst* 2017;109:djw323.
37. Evans JM, Donnelly LA, Emslie-Smith AM, Alessi DR, Morris AD. Metformin and reduced risk of cancer in diabetic patients. *BMJ* 2005;330:1304–5.
38. Gong TT, Wu QJ, Lin B, Ruan SK, Kushima M, Takimoto M. Observational studies on the association between post-diagnostic metformin use and survival in ovarian cancer: a systematic review and meta-analysis. *Front Oncol* 2019;9:458.
39. Lee MS, Hsu CC, Wahlqvist ML, Tsai HN, Chang YH, Huang YC. Type 2 diabetes increases and metformin reduces total, colorectal, liver and pancreatic cancer incidences in Taiwanese: a representative population prospective cohort study of 800,000 individuals. *BMC Cancer* 2011;11:1–10.
40. Hsieh MC, Lee TC, Cheng SM, Tu ST, Yen MH, Tseng CH. The influence of type 2 diabetes and glucose-lowering therapies on cancer risk in the Taiwanese. *J Diabetes Res* 2012;2012:413782.
41. Lai SW, Chen PC, Liao KF, Muo CH, Lin CC, Sung FC. Risk of hepatocellular carcinoma in diabetic patients and risk reduction associated with anti-diabetic therapy: a population-based cohort study. *Am J Gastroenterol* 2012;107:46–52.
42. Goodwin PJ, Chen BE, Gelmon KA, Whelan TJ, Ennis M, Lemieux J, et al. Effect of metformin vs placebo on invasive disease–free survival in

patients with breast cancer: the MA. 32 randomized clinical trial. *JAMA* 2022;327:1963–73.
43. Zhang ZJ, Yuan J, Bi Y, Wang C, Liu Y. The effect of metformin on biomarkers and survivals for breast cancer—a systematic review and meta-analysis of randomized clinical trials. *Pharmacol Res* 2019;141:551–5.
44. Wen J, Yi Z, Chen Y, Huang J, Mao X, Zhang L, et al. Efficacy of metformin therapy in patients with cancer: a meta-analysis of 22 randomised controlled trials. *BMC Med* 2022;20:402.
45. Nielsen SF, Nordestgaard BG, Bojesen SE. Statin use and reduced cancer-related mortality. *N Engl J Med* 2012;367:1792–802.
46. Mei Z, Liang M, Li L, Zhang Y, Wang Q, Yang W. Effects of statins on cancer mortality and progression: a systematic review and meta-analysis of 95 cohorts including 1,111,407 individuals. *Int J Cancer* 2017;140:1068–81.
47. Thomas JP, Loke YK, Alexandre L. Efficacy and safety profile of statins in patients with cancer: a systematic review of randomised controlled trials. *Eur J Clin Pharmacol* 2020;76:1639–51.
48. Farooqi MA, Malhotra N, Mukherjee SD, Sanger S, Dhesy-Thind SK, Ellis P, et al. Statin therapy in the treatment of active cancer: a systematic review and meta-analysis of randomized controlled trials. *PLoS One* 2018;13:e0209486.
49. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol* 2008;167:492–9.
50. Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet* 2016;388:2532–61.
51. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol* 2016;183:758–64.
52. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol* 2003;158:915–20.
53. Matthews AA, Danaei G, Islam N, Kurth T. Target trial emulation: applying principles of randomised trials to observational studies. *BMJ* 2022;378:e071108.
54. Tsai HR, Lin YJ, Yeh JI, Lin SM, Liu PPS, Chang YC, et al. Use of sodium-glucose co-transporter 2 inhibitors in patients with type 2 diabetes and the incidence of retinal vein occlusion in Taiwan. *Invest Ophthalmol Vis Sci* 2024;65:19.
55. Wu CS, Yang AC, Chang SS, Chang CM, Liu YH, Liao SC, et al. Validation of machine learning-based individualized treatment for depressive disorder using target trial emulation. *J Pers Med* 2021;11:1316.
56. Yen FS, Hwu CM, Liu JS, Wu YL, Chong K, Hsu CC. Sodium–glucose cotransporter-2 inhibitors and the risk for dialysis and cardiovascular disease in patients with stage 5 chronic kidney disease. *Ann Intern Med* 2024;177:693–700.
57. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. *J Clin Pharm Ther* 2002;27:299–309.
58. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int J Epidemiol* 2017;46:348–55.
59. Chang CH, Lin JW, Wu LC, Liu CH, Lai MS. National antiviral treatment program and the incidence of hepatocellular carcinoma and associated mortality in Taiwan: a preliminary report. *Med Care* 2013;51:908–13.
60. Hsu JC, Wei CF, Yang SC. Effects of removing reimbursement restrictions on targeted therapy accessibility for non-small cell lung cancer treatment in Taiwan: an interrupted time series study. *BMJ Open* 2019;9:e022293.
61. Huang WK, Hsu HC, Chang SH, Chou WC, Chang PH, Chiang SF, et al. Real-world effectiveness of adjuvant oxaliplatin chemotherapy in stage III colon cancer: a controlled interrupted time series analysis. *Front Pharmacol* 2021;12:693009.
62. Fang YF, Liu JR, Chang SH, Kuo CF, See LC. Trends of adverse events and mortality after DMARDs in patients with rheumatoid arthritis: interrupted time-series analysis. *Immun Inflamm Dis* 2022;10:e630.
63. Schmid SL. Five years post-DORA: promoting best practices for research assessment. *Mol Biol Cell* 2017;28:2941–4.
64. Roldan-Valadez E, Salazar-Ruiz SY, Ibarra-Contreras R, Rios C. Current concepts on bibliometrics: a brief review about impact factor, Eigenfactor score, CiteScore, SCImago Journal Rank, Source-Normalised Impact per Paper, H-index, and alternative metrics. *Ir J Med Sci* 2019;188:939–51.