HHS Public Access

Author manuscript

Diabetes Obes Metab. Author manuscript; available in PMC 2019 October 01.

Published in final edited form as:

Diabetes Obes Metab. 2019 September ; 21(9): 2029–2038. doi:10.1111/dom.13766.

Glucose-lowering medications and the risk of cancer: a methodological review of studies based on real-world data.

Katsiaryna Bykov,
Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, USA

Mengdong He,
Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, USA

Jessica M. Franklin,
Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, USA

Elizabeth M. Garry,
Science, Aetion Inc., Boston, USA

John D. Seeger,
Optum Epidemiology, Boston, USA

Elisabetta Patorno
Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, USA

Abstract

Aim: Recent years have witnessed a proliferation of observational research examining the association between glucose-lowering medications and cancer, paired with an increasing concern on the quality of the produced evidence. Our objective was to review published studies in this area to identify the most common methodological challenges and sources of bias.

Methods: We systematically searched PubMed to identify observational studies on glucose-lowering medications and cancer published between January 2000 and January 2016. We assessed the design and analytical methods used in each study, with a focus on their ability to achieve study validity, and further evaluated the prevalence of major methodological choices over time.

Results: Out of 155 studies evaluated, only 26% implemented a new user design, 41% used an active comparator, 33% implemented a lag or latency period, 51% adjusted for diabetes duration. Potential for immortal person-time bias was identified in 63% of the studies; 55% of the studies adjusted for variables measured during the follow-up without appropriate statistical methods. Aside from a decreasing trend in adjusting for variables measured during the follow-up, we observed no trends in methodological choices over time.

Corresponding author: Katsiaryna Bykov, Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, 1620 Tremont Street, Suite 3030, Boston, MA, 02120, USA; Tel: (+1) 617-278-0930; kbykov@partners.org.

Conclusions: The prevalence of well-known design and analysis flaws that may lead to biased results remains high among observational studies on glucose-lowering medications and cancer, limiting the conclusions that can be drawn from these studies. Avoiding known pitfalls could substantially improve the quality and validity of real-world evidence in this field.

Keywords

bias; cancer; epidemiologic methods; glucose-lowering medications; review

Introduction

Cancer is rapidly becoming a major global health burden as longer lifespans lead to more age-related disease, and patients with diabetes, in particular, may be at an increased risk of cancer and cancer-related mortality.[1] Since there are more than 400 million people worldwide with diabetes,[2] the potential role of glucose-lowering therapies in cancer initiation or promotion has attracted substantial attention, further fueled by the rapidly-growing access to large longitudinal databases in the form of healthcare claims data, electronic medical records, or disease registries. More than one hundred observational studies have examined the association between metformin and cancer, with many finding metformin protective, leading to initiation of randomized controlled trials to test the potential for metformin to prevent cancer or improve cancer outcomes.[3] At the same time, higher risk of cancer has been reported in insulin users,[4–8] as well as in patients on sulfonylurea therapies.[7,9,10]

Alongside the proliferation of research on the association between glucose-lowering treatments and cancer, there has been a growing concern about the quality of these studies and their potential susceptibility to bias.[3,11] Many of the studies that reported a protective effect of metformin on cancer have been found to be afflicted with immortal time and other time-related biases that could, at least partially, explain the protective association.[12,13] Arguments have been made that some of the harmful effect of insulin (and other therapies) may represent the natural course of underlying diabetes or prodromal cancer effects.[11] A recent systematic review that evaluated observational studies on the association between long-acting insulin analogs and cancer found important methodological shortcomings in most of them.[14] The Outcomes Reduction with Insulin Glargine Intervention (ORIGIN) randomized controlled trial and subsequent analyses of post-trial effects during an additional 2.7 years reported no effects of insulin glargine on the outcome of any cancer.[15,16]

Nonrandomized studies of drug effects are often not straightforward to conduct and can be affected by methodological challenges that compromise the validity of findings. Medications are not prescribed at random in clinical practice, and the decisions to initiate, monitor, stop or change therapy may be based on factors associated with the outcome.[17–19] Assessing long-term drug effects, such as cancer, and differentiating them from the consequences of the natural progression of the underlying chronic disease, particularly in the context of ever changing therapy, brings in even more complexity. Nevertheless, with randomized controlled trials presenting their own challenges in the evaluation of long-term outcomes, real-world data, i.e., healthcare data from sources outside of traditional clinical trials, represent a

valuable source of information on the safety of medications as they are used in real-world settings.[20–22] A thorough understanding of particular methodological flaws that most commonly occur in nonrandomized studies of glucose-lowering medications and cancer could help improve the quality of evidence produced in this area, its critical appraisal by clinical community and regulators, and, ultimately, treatment decisions.

The objective of this study was to review the methodology of observational studies that examined the association between glucose-lowering therapies and cancer occurrence, with the goal of identifying the most common challenges and sources of bias. We further evaluated whether key methodological choices changed over time.

Methods

Literature search and study selection

We conducted a systematic search in PubMed database for observational studies that examined the association between glucose-lowering medications and cancer, and were published from January 2000 to January 2016. The search strategy included a combination of terms related to glucose-lowering agents, cancer, and key methodological components of observational studies, such as cohort or case-control design (Appendix 1). We excluded studies in a language other than English, studies conducted in patients with secondary, i.e., caused by another disease or drug-induced, diabetes or gestational diabetes, studies in oncological patients, and studies not conducted in humans. Two investigators (KB and EP) screened the titles and abstracts of all identified articles, and further reviewed full texts for eligibility. In addition, we hand-searched relevant reviews and meta-analyses for articles not identified through the systematic search.

Data extraction and review

The included studies were categorized based on their study design, i.e., cohort or case-control (**Glossary**). Relevant information was extracted using data collection forms customized to the specific study design (Appendix 2). Two investigators independently reviewed and extracted information from each article. Disagreements were resolved by consensus between the two investigators, and among all investigators, if necessary.

We extracted basic study characteristics: publication year, data source, geographic region, exposure evaluated, and type of cancer assessed. In addition, the forms contained a set of pre-specified questions assessing study design, analytical methods, transparency of reporting, and potential for bias. While some tools for assessing quality of observational studies have been proposed, most of them are poorly suited to evaluate the variety of design options and possible sources of bias present in pharmacoepidemiologic research.[23] In addition, they are often not specific enough for a particular clinical research area. Thus, we developed a customized list of questions, that was compiled based on the authors' expertise and published literature.[11,12,14,20,24,25] In short, we assessed each study on inclusion of incident or prevalent users, choice of comparator/control group, reporting of how follow-up or exposure risk windows (for cohort and case-control designs, respectively) were defined, exclusion of patients with a prior history of cancer, consideration of lag and latency periods,

assessment of time-varying risks and cumulative duration effects, type of analytical approach (as-treated or intention-to-treat), confounding adjustment, investigation of differential surveillance during follow-up, and outcome identification. For confounding adjustment, we assessed whether the authors adjusted for age and sex, diabetes-related complications, diabetes therapy intensity (in studies where patients had prior treatment with glucose-lowering agents), obesity, HbA_{1c} levels, diabetes duration, cardiovascular disease, smoking, alcohol, and medical surveillance prior to start of follow-up. We further evaluated whether the authors adjusted for any variables that were measured during the follow-up – and if they did, whether appropriate statistical methods, such as inverse probability weighting of marginal structural models or g-estimation, were implemented.[26] These methods do not rely on stratification for confounding control and are recommended when adjustment for variables that may be on the causal pathway between exposure and outcome (e.g., HbA_{1c} levels measured after initiation of treatment) is needed. Conventional adjustment methods, such as including variables in the regression model, may adjust for the effect of treatment on outcome operating through these variables and lead to biased estimation of the overall effect.[26,27]

Since immortal time, defined as period of follow-up during which, by design, outcomes cannot occur (**Glossary**), can result in strong bias in observational studies of medications and has been implicated in a number of studies on glucose-lowering medication and cancer, [12,28] each study was also evaluated for potential for immortal time bias based on whether the authors determined study inclusion or exposure status using information collected during follow-up and whether the follow-up started at the same time for both comparison groups.

Case-control studies were additionally evaluated on selection and treatment of controls: type of control sampling, whether cases were eligible to be controls, whether controls were matched to cases on duration of follow-up (to avoid time-window bias)[29] and whether patients with pre-cancerous conditions (e.g., adenomas or cervical dysplasia) were included.

All analyses (including sensitivity and secondary) performed within each study were considered in the assessment. Since some design and analytical choices may have more impact on the validity of findings than others, and the degree of an impact is hard to predict for a specific study unless a bias analysis is implemented, we assessed each component individually rather than creating an aggregate quality score.

We further plotted the prevalence of methodological characteristics that may have substantial impact on study validity, such as new user design, active comparator, potential for immortal time bias, and adjustment for variables measured during follow-up without use of appropriate statistical methods, over time. In addition, we plotted the prevalence of the cohort design. Since there were few studies published prior to 2009, prevalence over time was evaluated over 2009–2015 period only, using logistic regression model with yearly prevalence as a dependent variable and time (in years) as an independent variable.

Results

Eligible studies

Our search yielded 1613 articles. A total of 155 studies satisfied our inclusion criteria and were included in the review (Appendix figure 1 and Appendix 3).

Study characteristics

Basic study characteristics are presented in Table 1. The most common data sources were administrative health insurance claims (30%) and medical records (28%). Insulin monotherapy was the most frequently evaluated exposure (52%), followed by metformin (50%). More than 10% of studies evaluated exposure to either any glucose-lowering drug or any glucose-lowering drug from a pre-specified group of therapeutic classes. Most of the studies concentrated on a specific cancer; however, a third of the studies evaluated any cancer, either as their primary or secondary outcome of interest.

Methodology assessment

The methodological aspects evaluated in these studies are presented in Table 2 and are described in detail below.

Design choices

Of the 155 studies reviewed, 65% used a cohort design while the remaining 35% were case-control studies. In half of the case-control studies, a nested case-control design was implemented (N = 27).

Less than a third of the studies utilized a new user design, which was a more common choice in cohort studies than in case-control studies (36% vs 9%, respectively). Cohort studies were also more likely to have an active comparator (54% of cohort studies vs 16% of case-control studies). Twelve percent of the studies compared patients exposed to the drug(s) of interest to non-diabetic patients, and 7% to non-pharmacologically treated diabetes patients. Case-control studies were more likely to include non-diabetic patients or non-pharmacologically treated diabetic patients in their control groups (78%). The majority of the studies (92%) excluded patients with a prior history of cancer (either any cancer or a specific type of cancer).

In regard to methodological choices specific to case-control design, most case-control studies implemented random control sampling (84%) and risk-set sampling (82%). Only 40% of case-control studies matched cases and controls on duration of follow-up, and in 56% of studies with risk set sampling, cases were eligible to be controls prior to outcome occurrence (Appendix table).

As a consequence of their design choices, more than half of the studies (63%) were found to have potential for immortal time bias. In cohort studies, the potential for immortal time bias was most commonly arising due to assigning exposure status based on prescriptions dispensed during the follow-up (N = 46). For case-control studies, it was assessing study

inclusion based on information collected during the follow-up that was the most common mechanism.

Outcome definition and assessment

In 88 studies (57%) identification of outcome was based on healthcare claims only, while the remaining 43% of the studies utilized medical records or cancer registries either as the only or a supplemental source of outcome information. Among the studies that were relying on claims only, 17% required additional procedure or care (chemotherapy, surgery, radiation, bone marrow transplant, or palliative care), 16% utilized a previously validated algorithm and 13% validated their algorithm within their study population.

Only one-third of studies considered either lag or latency periods in their analyses to allow for development and detection of malignancy. Sixty-eight percent of studies evaluated either cumulative exposure or the effect of the duration of treatment. Most cohort studies (88%) conducted an intention-to-treat analysis, either as the main analysis or in addition to the as-treated analysis.

Confounding adjustment

Almost all studies adjusted for age and sex (99%), but only half adjusted for diabetes duration (51%), diabetes therapy intensity (in studies where patients were not completely naïve to glucose-lowering therapy prior to cohort entry; 55%), smoking (44%), and obesity or BMI (62%). Adjustment for diabetes-related complications or HbA_{1c} was implemented in less than one-third of the studies. Few studies adjusted for medical surveillance prior to drug initiation (19%) or evaluated differential surveillance (to avoid detection bias) during follow-up (21%). Case-control studies were more likely to adjust for variables that were measured during follow-up (87% vs 37% of cohort studies); none of the studies that adjusted for variables measured during follow-up implemented statistical methods that are appropriate for such adjustment (g-methods).

Prevalence over time

Appendix Figure 2 presents the prevalence of studies with 1) a cohort design, 2) an active comparator, 3) potential for immortal time bias, 4) new user design, and 5) adjustment for variables measured during follow-up without proper statistical methods. The prevalence of studies with adjustment for variables measured during follow-up decreased during the study period. We did not observe any noticeable trends in the prevalence of other characteristics. Appendix Figures 3–6 present the prevalence of these characteristics among cohort studies and case-control studies separately.

Discussion

In our review of 155 observational studies that examined the association between glucose-lowering therapy and cancer, we found that more than 60% of the studies had a potential for immortal time bias, at least half of the studies did not control for confounding due to diabetes duration and severity, only a third considered a lag or latency period, most studies grouped patients at various stages of their treatment, and none of the 85 studies that adjusted

for variables measured during follow-up used the appropriate statistical methods. While some of these shortcomings may be driven by limitations of the specific data sources utilized (e.g., lack of information on important confounders in healthcare claims data), many of the methodological flaws observed in this review could have been avoided through careful study design and analysis. Moreover, we found that the prevalence of suboptimal methodological decisions exhibited little change over time.

The persistence of suboptimal design and analytical choices is particularly disconcerting, given that most biases we have focused on in this review have been previously described, including in the context of glucose-lowering medications and cancer.[11,12,24,28,30] There could be various reasons behind the slow penetration of established epidemiological principles into the design and analysis of real-world data. And yet, some of these principles are straightforward and can significantly improve the validity of a study, even when information on important risk factors is missing.

We have provided some recommendations for design and analysis of nonrandomized, observational studies in Table 4. As treatments are not given at random in clinical practice, accounting for confounding through study design and analytical strategies is paramount in observational research. While unexposed comparator groups are possible, comparing treated patients to untreated may introduce bias due to risk factors, often unavailable to researchers, that have led to treatment initiation.[31] In addition, since diabetes itself may be a risk factor for cancer, patients treated with medications may have had diabetes for longer and be at a more advanced stage as compared to untreated patients. For that reason, active comparators are preferred, and in the context of glucose-lowering treatments, it is also important to compare treatments that are given at the same stage of the disease. Comparisons restricted to treated patients at a similar stage of diabetes may not only improve confounding control, but will also reduce potential for immortal time bias by providing a well-defined cohort entry (treatment initiation) for both groups.

A new user design, which starts following patients as they initiate treatment, is similar to randomized clinical trials that randomize patients to treatment initiation. The new user design “anchors” the study timeline so that study inclusion criteria and covariates are assessed before treatment initiation, and exposure is assessed prior to the start of follow-up. In our review, the most common reason for immortal time bias potential was treatment group assignment based on prescriptions dispensed during follow-up, making all the time from the cohort entry (start of follow-up) until that first prescription immortal. The second most common reason was assessing study inclusion based on information collected during follow-up. Applying the new-user design with a clear temporal sequencing of study inclusion, covariates evaluation, treatment group assignment and outcomes assessment will safeguard not only against immortal-time bias, but also numerous other biases, such as bias due to confounding by differential baseline risk and selection bias due to depletion of patients susceptible to the outcome.[20,32,33] Thus, while requiring patients to be new users may restrict the size of a cohort, as well as the total follow-up and the number of outcomes, the threats to internal validity caused by immortal time bias, confounding, and selection bias outweigh considerations about sample size and reduced power. In addition, the new-user

design allows evaluation of drug effects that vary over time or depend on duration of treatment.

While a new-user design with active comparator may generate cohorts of patients who are very similar in regard to baseline characteristics, adjusting for potential confounders in the analysis is still essential. Many techniques to adjust for variables exist, however, propensity score methods are particularly beneficial in high-dimensional data as they allow balancing many variables even when the number of outcomes is low, which may be the case with many cancers.[34] A propensity score is an estimated probability of receiving one treatment over the other given the complete set of all selected covariates.[35] The score can be further used via matching, stratification, regression or weighting.[36,37] A propensity score approach that incorporates many clinical and healthcare utilization variables may substantially reduce confounding and could even balance important, but unmeasured, confounders by proxy.[38] It will balance the unmeasured confounders, however, only to the extent that they are correlated with the measured variables included in the propensity score. Thus, confounding can never be completely ruled out in observation studies, although it can be reduced to minimal when an appropriate active comparator, along with propensity score methods, is utilized. Researchers who need to adjust for variables measured after the start of follow-up (either post-initiation confounders or adherence) should consult the appropriate resources on g-methods,[26,27,39] since adjustment for variables on the causal pathway between exposure and outcome without the appropriate statistical modeling may lead to biased estimation.

Some additional points to consider when designing a study of glucose-lowering medications and cancer include differential surveillance, bias due to reverse causation, when either treatment initiation or discontinuation is a consequence of prodromal cancer symptoms, the reliability of outcome ascertainment (high specificity is preferred) in the data, and exposure or covariate misclassification. Only a third of the reviewed studies considered a lag period in their analyses. And yet most cancers require time to develop and manifest, which should be reflected in exposure risk window definition. The use of a lag period also reduces bias due to reverse causation, when worsening of a patient's condition due to yet undiagnosed cancer triggers a change in therapy. While exact length of a lag or a latency period may be unknown, sensitivity analyses can be implemented to address the uncertainty.

Analyzing data using intention-to-treat (ITT) approach, when initial exposure status is carried forward regardless any treatment changes that occur during the follow-up, is another safeguard against reverse causation bias or informative censoring, particularly when compared to as-treated analysis, when patients are censored upon treatment change or discontinuation. In addition, ITT approach is more appropriate when effects are suspected to persist long after treatment discontinuation or to be irreversible, which could be the case with some cancers. It is worth noting, however, that ITT analysis is still open to bias due to differential loss to follow-up, and may make unsafe therapies appear safe when the rate of treatment changes and non-adherence is high.[40] Thus, we would recommend both ITT and as-treated analyses when evaluating the effects of glucose-lowering medications on cancer. If data on reasons for non-adherence and treatment changes are available (in as-treated analysis) or when adjustments for differential loss to follow-up are needed (in either ITT or

as-treated analyses), adjustments should be done using methods that can handle post-treatment-initiation confounders (g-methods).[40]

Since a third of the studies reviewed were case-control studies, we would like to point out that although a case-control study can provide unbiased estimates when rigorously conducted,[25] ensuring appropriate temporal sequence of study inclusion, covariate assessment, and treatment group assignment is more challenging with case-control design as compared to cohort design. Almost 90% of the case-control studies we reviewed assessed exposure and covariates during the same period preceding an outcome. Some of these covariates could have been impacted by the exposure at the time of measurement and many represent variables on the causal pathway from treatment to outcome. Some additional issues we identified in case-control studies included potential for differential length of follow-up between cases and controls, which can lead to time-window bias,[29] and potential for selection bias when choosing controls. Thus, unless restricted by data available, a cohort design may present a more intuitive and less error-prone approach.

In summary, while our review may have brought into question the quality of most observational studies on the cancer risk associated with glucose-lowering medications, it should not be interpreted as questioning the value of observational studies in clinical research and decision making. Quite to the contrary, since randomized clinical trials are often underpowered for long-term or rare safety endpoints and may suffer from their own challenges, such as differential dropout and short follow-up, real-world data may be our best source of relatively quick and inexpensive information on the link between medications and cancer. In addition, real-world data often reflect experiences of a wide spectrum of patients, many of whom would not be included in highly-selective clinical trials. With long-term data accumulating and additional data sources, such as electronic medical records and genetic data, becoming available and linkable between each other, and with our understanding on how to conduct observational studies of medications constantly growing, the value of real-world data for evidence generation is only expected to increase. With growing access to these data, however, it is vital to apply rigorous study design and analysis principles to maximize study validity.

A few limitations of our review should be noted. Our review period ended in January 2016; thus, we did not evaluate whether the prevalence of methodological choices has changed in more recent years. While we evaluated the design and analytic choices in each study, we did not evaluate the amount of bias or confounding that resulted from these choices, which could range from minimal to substantial. Our review was based on our current understanding of biases and challenges in observational studies on medications and long-term outcomes, and as that understanding and knowledge is constantly evolving, we may have missed some important considerations. Finally, as with any systematic review, there is potential for publication bias. If studies with positive findings are more likely to be published, and positive findings are associated with biases and suboptimal study design, it is possible that the high prevalence of suboptimal design choices in published literature has been, at least partially, driven by publication bias.

In conclusion, our review indicates that most observational studies on the association between glucose-lowering medications and cancer published prior to 2016 suffered from major methodological flaws that could have been avoided through the use of rigorous design and analysis. The clinical research community should strive to increase awareness of methodological missteps and addressable biases in observational studies of medications so that real-world data are effectively leveraged to provide the needed and valid information on the safety of treatments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

This study was funded by the Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital and Harvard Medical School. KB was supported by a training grant from the National Institute of Child Health and Human Development (T32 HD40128-14). EP was supported by a career development grant K08AG055670 from the National Institute on Aging.

References

1. Giovannucci E, Harlan DM, Archer MC, et al. Diabetes and cancer: a consensus report. Diabetes Care 2010;33:1674–1685. [PubMed: 20587728]
2. International Diabetes Federation diabetes atlas 2017 (Accessed 5 November, 2018, at http://www.diabetesatlas.org/.)
3. Suissa S Metformin to Treat Cancer: Misstep in Translational Research from Observational Studies. Epidemiology 2017;28:455–458. [PubMed: 28166103]
4. Suissa S, Azoulay L, Dell'Aniello S, Evans M, Vora J, Pollak M. Long-term effects of insulin glargine on the risk of breast cancer. Diabetologia 2011;54:2254–2262. [PubMed: 21614572]
5. Mannucci E, Monami M, Balzi D, et al. Doses of insulin and its analogues and cancer occurrence in insulin-treated type 2 diabetic patients. Diabetes Care 2010;33:1997–2003. [PubMed: 20551014]
6. Hemkens LG, Grouven U, Bender R, et al. Risk of malignancies in patients with diabetes treated with human insulin or insulin analogues: a cohort study. Diabetologia 2009;52:1732–1744. [PubMed: 19565214]
7. Bowker SL, Majumdar SR, Veugelers P, Johnson JA. Increased cancer-related mortality for patients with type 2 diabetes who use sulfonylureas or insulin. Diabetes Care 2006;29:254–258. [PubMed: 16443869]
8. Yang YX, Hennessy S, Lewis JD. Insulin therapy and colorectal cancer risk among type 2 diabetes mellitus patients. Gastroenterology 2004;127:1044–1050. [PubMed: 15480982]
9. Currie CJ, Poole CD, Gale EA. The influence of glucose-lowering therapies on cancer risk in type 2 diabetes. Diabetologia 2009;52:1766–1777. [PubMed: 19572116]
10. Bodmer M, Becker C, Meier C, Jick SS, Meier CR. Use of antidiabetic agents and the risk of pancreatic cancer: a case-control analysis. Am J Gastroenterol 2012;107:620–626. [PubMed: 22290402]
11. Hernandez-Diaz S, Adami HO. Diabetes therapy and cancer risk: causal effects and other plausible explanations. Diabetologia 2010;53:802–808. [PubMed: 20177658]
12. Suissa S, Azoulay L. Metformin and the risk of cancer: time-related biases in observational studies. Diabetes Care 2012;35:2665–2673. [PubMed: 23173135]
13. Suissa S, Azoulay L. Metformin and cancer: mounting evidence against an association. Diabetes Care 2014;37:1786–1788. [PubMed: 24963109]

14. Wu JW, Filion KB, Azoulay L, Doll MK, Suissa S. Effect of Long-Acting Insulin Analogs on the Risk of Cancer: A Systematic Review of Observational Studies. Diabetes Care 2016;39:486–494. [PubMed: 26740633]
15. Bordeleau L, Yakubovich N, Dagenais GR, et al. The association of basal insulin glargine and/or n-3 fatty acids with incident cancers in patients with dysglycemia. Diabetes Care 2014;37:1360–1366. [PubMed: 24574355]
16. Investigators OT. Cardiovascular and Other Outcomes Postintervention With Insulin Glargine and Omega-3 Fatty Acids (ORIGINALE). Diabetes Care 2016;39:709–716. [PubMed: 26681720]
17. Walker AM. Confounding by indication. Epidemiology 1996;7:335–336. [PubMed: 8793355]
18. Glynn RJ, Knight EL, Levin R, Avorn J. Paradoxical relations of drug treatment with mortality in older persons. Epidemiology 2001;12:682–689. [PubMed: 11679797]
19. Brookhart MA, Patrick AR, Dormuth C, et al. Adherence to lipid-lowering therapy and the use of preventive health services: an investigation of the healthy user effect. Am J Epidemiol 2007;166:348–354. [PubMed: 17504779]
20. Franklin JM, Schneeweiss S. When and How Can Real World Data Analyses Substitute for Randomized Controlled Trials? Clin Pharmacol Ther 2017;102:924–933. [PubMed: 28836267]
21. Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. Br J Cancer 2014;110:551–555. [PubMed: 24495873]
22. Corrigan-Curay J, Sacks L, Woodcock J. Real-World Evidence and Real-World Data for Evaluating Drug Safety and Effectiveness. JAMA 2018;320:867–868. [PubMed: 30105359]
23. Losilla JM, Oliveras I, Marin-Garcia JA, Vives J. Three risk of bias tools lead to opposite conclusions in observational research synthesis. J Clin Epidemiol 2018;101:61–72. [PubMed: 29864541]
24. Patorno E, Patrick AR, Garry EM, et al. Observational studies of the association between glucose-lowering medications and cardiovascular outcomes: addressing methodological limitations. Diabetologia 2014;57:2237–2250. [PubMed: 25212258]
25. Rothman KJ, Greenland S, Lash TL. Modern epidemiology 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
26. Robins JM, Hernan MA. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, eds. Longitudinal data analysis Boca Raton: Chapman & Hall/CRC; 2009:553–599.
27. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11:550–560. [PubMed: 10955408]
28. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ 2010;340:b5087. [PubMed: 20228141]
29. Suissa S, Dell'aniello S, Vahey S, Renoux C. Time-window bias in case-control studies: statins and lung cancer. Epidemiology 2011;22:228–231. [PubMed: 21228697]
30. Suissa S Immortal time bias in observational studies of drug effects. Pharmacoepidemiol Drug Saf 2007;16:241–249. [PubMed: 17252614]
31. Glynn RJ, Schneeweiss S, Wang PS, Levin R, Avorn J. Selective prescribing led to overestimation of the benefits of lipid-lowering drugs. J Clin Epidemiol 2006;59:819–828. [PubMed: 16828675]
32. Emilsson L, Garcia-Albeniz X, Logan RW, Caniglia EC, Kalager M, Hernan MA. Examining Bias in Studies of Statin Treatment and Survival in Patients With Cancer. JAMA Oncol 2018;4:63–70. [PubMed: 28822996]
33. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol 2003;158:915–920. [PubMed: 14585769]
34. Franklin JM, Eddings W, Austin PC, Stuart EA, Schneeweiss S. Comparing the performance of propensity score methods in healthcare database studies with rare outcomes. Stat Med 2017;36:1946–1963. [PubMed: 28208229]
35. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41–55.

36. Glynn RJ, Schneeweiss S, Sturmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. Basic Clin Pharmacol Toxicol 2006;98:253–259. [PubMed: 16611199]

37. Brookhart MA, Wyss R, Layton JB, Sturmer T. Propensity score methods for confounding control in nonexperimental research. Circ Cardiovasc Qual Outcomes 2013;6:604–611. [PubMed: 24021692]

38. Patorno E, Gopalakrishnan C, Franklin JM, et al. Claims-based studies of oral glucose-lowering medications can achieve balance in critical clinical variables only observed in electronic health records. Diabetes Obes Metab 2018;20:974–984. [PubMed: 29206336]

39. Hernan MA, Robins JM. Causal inference Boca Raton: Chapman & Hall/CRC 2019.

40. Hernan MA, Robins JM. Per-Protocol Analyses of Pragmatic Trials. N Engl J Med 2017;377:1391–1398. [PubMed: 28976864]

Table 1.

General characteristics of the reviewed studies

Characteristic	All (N = 155)	Cohort studies (N = 100)	Case-control studies (N = 55)
	N (%)	N (%)	N (%)
Publication year			
2000–2008	11 (7%)	4 (4%)	7 (13%)
2009	10 (6%)	5 (5%)	5 (9%)
2010	11 (7%)	5 (5%)	6 (11%)
2011	18 (12%)	12 (12%)	6 (11%)
2012	34 (22%)	27 (27%)	7 (13%)
2013	33 (21%)	18 (18%)	15 (27%)
2014	26 (17%)	21 (21%)	5 (9%)
2015	12 (8%)	8 (8%)	4 (7%)
Geographic region of study population			
North America	36 (23%)	21 (21%)	15 (27%)
Europe	65 (42%)	38 (38%)	27 (49%)
Southeast Asia	51 (33%)	39 (39%)	12 (22%)
Middle east	1 (1%)	1 (1%)	0 (0%)
Multiple regions	2 (1%)	1 (1%)	1 (2%)
Data source			
Administrative insurance claims	47 (30%)	38 (38%)	9 (16%)
Medical records	44 (28%)	21 (21%)	23 (42%)
Disease registry	6 (4%)	5 (5%)	1 (2%)
Longitudinal survey/Secondary data analysis[†]	8 (5%)	7 (7%)	1 (2%)
Primary data collection[‡]	10 (6%)	0 (0%)	10 (18%)
Multiple linked data sources	40 (26%)	29 (29%)	11 (20%)
Exposure[§]			
Alpha-glucosidase inhibitors	12 (8%)	8 (8%)	4 (7%)
Dipeptidyl peptidase-4 inhibitors	2 (1%)	1 (1%)	1 (2%)
Injectable glucagon-like peptide-1 receptor agonists	2 (1%)	1 (1%)	1 (2%)
Insulin	80 (52%)	49 (49%)	31 (56%)
Meglitinides	8 (5%)	4 (4%)	4 (7%)
Metformin	77 (50%)	41 (41%)	36 (65%)
Sulfonylureas	48 (31%)	25 (25%)	23 (42%)
Thiazolidinediones	55 (35%)	34 (34%)	21 (38%)
Any glucose-lowering medication	11 (7%)	8 (8%)	3 (5%)
Any agent from a pre-specified group of therapeutic classes	8 (5%)	6 (6%)	2 (4%)
Polytherapy with multiple therapeutic agents	11 (7%)	8 (8%)	3 (5%)

Characteristic	All (N = 155)	Cohort studies (N = 100)	Case-control studies (N = 55)
	N (%)	N (%)	N (%)
Outcomes[§]			
Lip, oral cavity and pharynx cancer	3 (2%)	3 (3%)	0 (0%)
Digestive organs cancer	60 (39%)	31 (31%)	29 (53%)
Respiratory and intrathoracic organs cancer	15 (10%)	11 (11%)	4 (7%)
Skin cancer	4 (3%)	4 (4%)	0 (0%)
Breast cancer	27 (17%)	22 (22%)	5 (9%)
Female genital organs cancer	10 (6%)	8 (8%)	2 (4%)
Male genital organs cancer	26 (17%)	19 (19%)	7 (13%)
Urinary tract cancer	28 (18%)	22 (22%)	6 (11%)
Eye, brain and other parts of central nervous system cancer	1 (1%)	0 (0%)	1 (2%)
Thyroid and other endocrine glands cancer	3 (2%)	3 (3%)	0 (0%)
Lymphoid, hematopoietic and related tissue cancer	4 (3%)	4 (4%)	0 (0%)
Any cancer[¶]	49 (32%)	42 (42%)	7 (13%)
Any cancer from a pre-specified group of cancer types	2 (1%)	2 (2%)	0 (0%)

[†]Secondary data analyses: analyses of data collected from clinical trials, prospective cohort studies or case-control studies conducted for a different research question

[‡]Primary data collection: prospective cohort studies, case-control studies with recruitment, interview, questionnaire or follow-up

[§]Categories are not mutually exclusive

[¶]Including studies that evaluated any cancer except nonmelanoma skin cancer (NMSC) (7 cohort and 0 case-control studies)

Table 2.

Methodological characteristics of the reviewed studies

Characteristic	Total (N=155)	Cohort studies (N=100)	Case-control studies (N=55)
New user design	41 (26%)	36 (36%)	5 (9%)
Comparator			
Active comparator	63 (41%)	54 (54%)	9 (16%)
Non-diabetic patients	19 (12%)	13 (13%)	6 (11%)
Non-pharmacologically treated diabetic patients	11 (7%)	7 (7%)	4 (7%)
Comparator group that includes either non-diabetic patients or non-pharmacologically treated diabetic patients, in addition to pharmacologically treated patients with diabetes	88 (57%)	45 (45%)	43 (78%)
Exclusion of patients with any or specific cancer history	142 (92%)	95 (95%)	47 (85%)
Potential for immortal person-time bias	98 (63%)	66 (66%)	32 (58%)
Clear definition of follow-up period (cohort) or exposure assessment window (case-control) provided in the manuscript	124 (80%)	91 (91%)	33 (60%)
Inclusion of lag and/or latency period	51 (33%)	33 (33%)	18 (33%)
Duration/cumulative exposure analysis	105 (68%)	68 (68%)	37 (67%)
Intention-to-treat analysis	88 (57%)	88 (88%)	0 (0%)
Adjustment for confounders			
Age and sex	154 (99%)	99 (99%)	55 (100%)
Diabetes-related complications	43 (28%)	37 (37%)	6 (11%)
Diabetes therapy intensity †	84 (55%)[a]	55 (56%)[a]	29 (53%)
BMI/obesity	96 (62%)	61 (61%)	35 (64%)
HbA_{1c} level	42 (27%)	29 (29%)	13 (24%)
Diabetes duration	79 (51%)	47 (47%)	32 (58%)
Cardiovascular disease	47 (30%)	39 (39%)	8 (15%)
Smoking	68 (44%)	43 (43%)	25 (45%)
Alcohol	42 (27%)	20 (20%)	22 (40%)
Adjustment for medical surveillance prior to drug initiation	29 (19%)	21 (21%)	8 (15%)
Investigation of differential surveillance during follow-up	32 (21%)	12 (12%)	20 (36%)
Adjustment for variables measured during follow-up without appropriate statistical models	85 (55%)	37 (37%)	48 (87%)
Source of information of cancer outcome			
Healthcare claims	88 (57%)	62 (62%)	26 (47%)
Medical records and/or cancer registry, with or without claims	67 (43%)	38 (38%)	29 (53%)
Cancer outcome definition in claims-based studies (N=62 for cohort studies, N=26 for case-control studies)			
1 code of diagnosis	82 (93%)	57 (92%)	25 (96%)
At least 2 diagnosis codes (within a specified time-period)	6 (7%)	5 (8%)	1 (4%)
Additional information of surgery, chemotherapy, radiation, bone marrow transplantation, or palliative care	15 (17%)	4 (6%)	11 (42%)
Validation of outcomes in claims-based studies (N=62 for cohort studies, N=26 for case-control studies)			

Characteristic	Total (N=155)	Cohort studies (N=100)	Case-control studies (N=55)
Previously validated algorithm	14 (16%)	9 (15%)	5 (19%)
Validated within the study	11 (13%)	4 (6%)	7 (27%)

† Among cohort studies that allowed treatment with glucose-lowering agents prior to start of follow-up (N=98).

Table 3.

Glossary of terms

Term	Definition
Active comparator design	A study design that compares the effect of the drug of interest to another drug used in clinical practice instead of non-use.
As-treated analysis	An analytical approach that terminates exposure to a medication and follow-up when a patient discontinues that medication
Case-control design	A study design in which cases (patients with the disease) are identified and compared with controls (patients without the disease) with respect to the exposure of interest
Cohort design	A study design in which a group of patients (a cohort) is identified and followed to ascertain the occurrence of an outcome
Confounding	A mixing of effects that arises when patients with different baseline risks are compared – the resulting effect measure is a mix of drug effects and risk factor effects
Detection bias	Bias that occurs when the degree of outcome surveillance is related to exposure and is differential among the exposure groups
Exposure risk window	The time period during which a drug of interest puts a patient at risk of a harmful or beneficial effect with regard to a specific outcome
Immortal time bias	Bias that derives from including a period of follow-up during which, by design, outcomes cannot occur
Intention-to-treat analysis	An analytical approach that carries forward the initial exposure status and disregards changes in treatment status over time
Lag period	A time period following drug initiation during which a specific outcome cannot be attributed to the initiated drug
Latency period	A time period after drug discontinuation during which a specific outcome can still be attributed to the discontinued drug
New user design	A study design that starts following patients at the time they initiate a new drug (also known as *incident user design*)
Selection bias	Bias that occurs when patient inclusion is related to both treatment and outcome, e.g., when prevalent users of a drug are compared to nonusers or incident users

Table 4.

Recommendations for design and analysis of observational studies on glucose-lowering medications and incidence of cancer

- Include patients who initiate the drug(s) of interest (new user design) which increases the chances of identifying more comparable patients with respect to the underlying risk of cancer, and is particularly suited to detect and evaluate medication effects that vary over time
- Identify comparison groups of new users of medications with similar medical indications and disease stage (active comparator)
- Follow a well-defined temporal sequence of study inclusion, covariate assessment, exposure definition, and start of follow-up, to reduce the chances of immortal time bias
- Consider specific cancers and specific glucose-lowering medications; consider validating the outcomes if previously validating algorithms are not available
- Exclude individuals with prior cancer history
- Consider lag and latency times based on cancer biology and use sensitivity analyses to explore and identify the optimal exposure risk window with regard to the lag time between exposure and start of follow-up and the grace period after drug discontinuation
- Adjust adequately for all baseline factors associated with the choice of treatment and the risk of specific cancers. Consider implementing techniques that guarantee a high-dimensional confounding adjustment, such as propensity score methodology
- Investigate the possibility of differential surveillance and detection bias prior to and following diabetes therapy initiation and account for it in the analysis or in the interpretation of the results
- Include intention-to-treat analyses that do not censor patients when a treatment is stopped or changes, but also consider analyses accounting for adherence to the initial treatment over follow-up. If data on the drivers of adherence and treatment changes are reliably collected, adjustments should be done using methods that allow for control for time-varying confounders (marginal structural models and other g-estimation methods).
- Consider multiple sensitivity analyses to assess robustness of findings