

REVIEW

Meta-Analysis of Sulfonylurea Therapy on Long-Term Risk of Mortality and Cardiovascular Events Compared to Other Oral Glucose-Lowering Treatments

W. Ryan Powell · Cindy L. Christiansen · Donald R. Miller

Received: April 11, 2018 / Published online: May 28, 2018
© The Author(s) 2018

ABSTRACT

Introduction: Among the most pressing clinical decisions in type 2 diabetes treatments are which drugs should be used after metformin is no longer sufficient, and whether sulfonylureas (SUs) should remain as a suitable second-line treatment. In this article we summarize current evidence on the long-term safety risks associated with SU therapy relative to other oral glucose-lowering therapies.

Methods: The MEDLINE database and Clinicaltrials.gov were searched for observational

and experimental studies comparing the safety of SUs to that of other diabetes medications in people with type 2 diabetes mellitus through December 15, 2015. Studies with at least 1 year of follow-up, which explicitly examined major cardiovascular events or death in patients who showed no evidence of serious conditions at baseline, were selected for inclusion in meta-analyses.

Results: SU treatment was associated with an elevated risk relative to treatment with metformin (METF), thiazolidinedione (TZD), dipeptidyl peptidase-4 inhibitor (DPP-4), and glucagon-like peptide-1 (GLP-1) agonist classes, either when compared alone (as a monotherapy) or when used in combination with METF. Significant findings were almost entirely derived from nontrial data and not confirmed by smaller, efficacy designed randomized controlled trials whose effects were in the same direction but much more imprecise.

Conclusion: Although much of the evidence is derived and will continue to come from observational studies, the methodological rigor of such studies is questionable. A key challenge for evaluators is the extent to which they should incorporate evidence from study designs that are quasi-experimental.

Enhanced Digital Features To view enhanced digital features for this article go to <https://doi.org/10.6084/m9.figshare.6244904>.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13300-018-0443-z>) contains supplementary material, which is available to authorized users.

W. R. Powell (✉) · C. L. Christiansen · D. R. Miller
Center for Healthcare Organization and Implementation Research, Bedford VA Medical Center, Bedford, MA, USA
e-mail: WilliamRyan.Powell@va.gov

W. R. Powell · C. L. Christiansen · D. R. Miller
Department of Health Law, Policy and Management, Boston University School of Public Health, Boston, MA, USA

D. R. Miller
Department of Dermatology, Boston University School of Medicine, Boston, MA, USA

Keywords: All-cause mortality; Cardiovascular disease; Meta-analysis; Sulfonylurea; Type 2 diabetes mellitus

Abbreviations

CVD	Cardiovascular disease
DPP-4	Dipeptidyl peptidase-4
ES	Effect size
GLP-1	Glucagon-like peptide-1
MEGL	Meglitinide
METF	Metformin
RCT	Randomized controlled trials
RR	Relative risk
SGLT-2	Sodium-glucose co-transporter 2
SU	Sulfonylurea
TZD	Thiazolidinediones

INTRODUCTION

The current clinical consensus is to treat people with type 2 diabetes with metformin (METF) when diet and exercise have failed to control glucose levels. However, to date, the question of whether sulfonylureas (SUs), a class of oral anti-hyperglycemic agents, are a suitable option for second-line therapy remains a focus of contention. While SUs represent a common, inexpensive, and effective treatment to manage glucose levels [1, 2], they have become increasingly controversial because of long-term safety concerns. Emerging evidence links the use of SUs with elevated risks for cardiovascular events and mortality compared to other glucose-lowering drug therapies, but expert opinion remains divided on whether SUs should remain a suitable therapy in the clinical setting [3, 4]. This difference in opinions may be attributed, in part, to the fact that a number of studies reporting elevated risks are observational in nature and thus open to challenge in terms of their methodological rigor. This factor and the lack of safety and efficacy measures in randomized controlled trials (RCTs) designed to evaluate long-term outcomes while also reflecting actual clinical populations have likely contributed to the adoption of different clinical guidelines.

The aim of the study reported here is to pool the existing evidence to summarize the risk of (1) cardiovascular events and (2) mortality (all-cause and cardiovascular) associated with SU use relative to other therapies within a broad range of indicated populations by conducting a series of meta-analyses.

METHODS

Data Sources

The MEDLINE database (via PubMed) was searched for studies comparing the safety of SUs (monotherapy or in combination) relative to other oral diabetes medications in patients with type 2 diabetes from 1965 to December 15, 2015, using the search terms reported in Electronic Supplementary Material (ESM) S1. Clinicaltrials.gov, a public database that registers clinical trials, was also searched for unpublished data. In addition, the reference lists of the relevant articles identified by the search of these databases were examined for studies not retrieved from the other search strategies. Finally, references from previous meta-analyses and Cochrane reviews were examined. This article is based on previously conducted studies and does not contain any studies with human participants or animals performed by any of the authors.

A flow chart of the selection process resulting from the MEDLINE and other search strategies is shown in Fig. 1. A total of 1982 articles were extracted from MEDLINE and a further 264 articles were culled from the other search strategies, resulting in a total of 2246 articles extracted for assessment. The abstracts of each of these articles were then reviewed for eligibility, following which 172 of these articles were reviewed in their entirety. Finally, a total of 50 articles met the eligibility requirements to be included in the series of meta-analyses (see ESM S2 for a list of all studies included).

Information on the effect size (e.g., hazard ratio, odds ratio, relative risk [RR]) or the raw information required to calculate it (e.g., number of major cardiovascular events, number of people who died), the standard deviation (or 95% confidence interval [CI]), sample size (number of people in treatment group), and study characteristics relevant to the population, outcome, and exposure were extracted from each study if provided. Adjusted estimates of the effect size were used if provided; otherwise unadjusted estimates were extracted. Authors of individual articles were not contacted to obtain

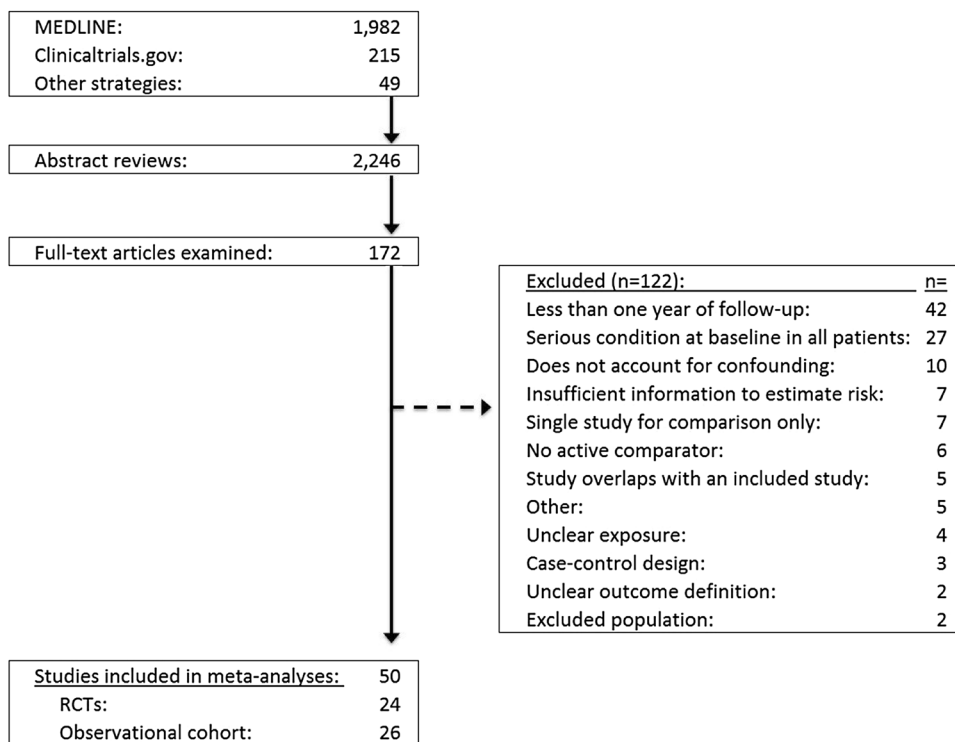


Fig. 1 Flow chart of the study selection process. *RCTs*: Randomized controlled trials

information if missing. For the purposes of the meta-analyses reported here, hazard ratios, odds ratios, and relative risk were treated as equivalent measures when pooling estimates. Article extraction and the culling of information were conducted by one of the authors who is a health services researcher (WRP), with consultation or further reviews by the other two authors (CLC, DRM) who are senior health services researchers.

Study Selection

Randomized controlled trials and observational cohort studies were included in this meta-analysis. All studies explicitly examining all-cause mortality, cardiovascular-related mortality, or major cardiovascular events were examined. Some heterogeneity in the definition of major cardiovascular composite endpoints used across studies existed; for clarity, each study definition is given in ESM S3. Since the aim was to evaluate long-term cardiovascular and mortality risks, only studies with ≥ 1 year of follow-up from the

date of the first prescription were included for assessment.

Studies were excluded from the meta-analyses if they met any of the following criteria: included only patients with serious conditions at baseline, such as a history of major cardiovascular events or renal failure; had a treatment population of only children (younger than 18 years of age) or only type 1 diabetes patients; did not include an active comparator (e.g., diet/exercise, placebo); had a case-control design; involved research only on animals; written in a language other than English. For studies for which there were more than one publication, the article with the most complete data or which involved the most recent follow-up was selected. For observational studies, an attempt to address confounding factors must have been implemented (matched in the design or model adjustment) by including basic demographic information (i.e., age, sex, and race) and relevant comorbidities at baseline (those adjusting for cardiovascular disease [CVD] risk at a minimum). This resulted in 24 RCTs and 26

observational cohort studies being included in this study.

Data Extraction and Quality Assessment

Details on potential biases in each RCT included in the meta-analyses were assessed using items from the Jadad scale, which assesses the methodological quality of RCTs in terms of study design and its appropriateness (randomization, double blind) as well as whether a description of the dropouts from the study is included [5]. The quality of observational cohort studies was rated using the eight items from the Newcastle–Ottawa Scale [6], which assesses quality in three domains: sample selection, comparability of groups, and outcome assessment. An additional item for both study designs examined whether industry funding explicitly sponsored the study.

Details of the quality assessments are presented in ESM S4 and S5. Results from the Newcastle–Ottawa Scale suggest that all studies met most of the quality assessments in each domain. Regarding the RCTs, all studies were randomized, 20 of the 24 were double blind, and in 23 a description of the participant dropouts was provided. However, industry funding was judged to be high in 64% of all studies (23/24 RCTs; 9/26 observational studies). With the exception of industry-funded studies, most studies were assessed as being at a low risk of bias on the domains assessed, suggesting that the overall quality was fair to good in the selected studies. Total scores from the quality assessments were not used to exclude studies from the meta-analyses.

Data Synthesis and Analysis

Each outcome and comparison required two or more studies. For RCTs and studies with observational designs, both fixed effects and random effect models were conducted and reported. In a fixed effects model, the assumption is that each study provides evidence towards one common effect size; that is, the model assumes the effect size should be the same and that the features of the study (e.g., study design, population)

should not impact the magnitude of the effect size. Therefore, the fixed effect model combines all study information together without taking into account that studies can vary between each other as well as between different study designs. Weights given to each study are determined only by its within-study variance (study weight = $1/\text{within study variance}$). Since variance is a function of sample size, smaller studies will contribute less information to the weighted estimate than larger studies.

In the random effects model, the weights given to each study are determined not only by the within-group variability (as for fixed effects) but also by the between-group variability. The implication is that relatively greater weight tends to be given to smaller studies than it would be in a fixed effect model approach since the weights for each study now account for between-study design variability. In general, since random effects models also include between-study variation, they will tend to have relatively wider confidence intervals compared to fixed effects models [7]. The inverse variance and the DerSimonian–Laird methods were used to estimate the fixed and random effects, respectively, using the METAN command in the Stata version 14.1 data analysis and statistical software [8].

A particular challenge for researchers is how to synthesize results that are produced from two inherently different study designs, namely, RCTs and quasi-experimental observational designs. Therefore, to address this methodological challenge, we used a two-level hierarchical Bayesian design to synthesize result estimates across RCTs and observational designs. This is a random effects model approach and assumes that the effects derived from different study designs will be similar and also different to some extent. The combined effect is the weighted average of these two common effect sizes.

Overall pooled estimates were estimated using the ‘bayesmh’ command with random effect of study design in Stata 14.1 [8]. Thus, the model accounts for heterogeneity from the different study design. This is similar to the approach used by Peters et al. [9] and involved Markov chain Monte Carlo estimation using a

Metropolis–Hastings algorithm and Gibbs sampling with vague conjugate prior distributions specified on unknown parameters. Convergence diagnostics suggested fairly rapid convergence with no trend in trace plots, low autocorrelation, and acceptance rates for the Metropolis–Hastings algorithm of around 75% (well above the 10% rule of thumb) and efficiencies of > 1% for all analyses.

Heterogeneity across the studies was assessed using the I^2 statistic, with values of > 50% benchmarked as indicating substantial heterogeneity [10]. This statistic represents the percentage of variance in the effect size attributable to heterogeneity, with larger values indicating less overlap in confidence intervals across studies. A benefit of the statistic is that the number of studies involved in each meta-analysis has little influence on the I^2 statistic, unlike other estimates.

In drug comparisons that included ≥ 10 studies, publication bias was assessed by testing for asymmetry in funnel plots (scatterplot for the log effect size by the log standard error) using Egger’s tests [11] via the METABIAS Stata command [12]. Tests for funnel plot asymmetry are not recommended in comparisons with < 10 studies since power may be too low to detect moderate asymmetry [13].

RESULTS

A total of 24 randomized clinical trials and 26 observational cohort studies were included in the series of meta-analyses. Meta-analysis summaries of the effect size (and 95% CIs or credible intervals) for each comparison and outcome are presented in Figs. 2, 3, and 4. Further information, including both fixed and random effect models for each analysis, is presented in ESM S6.

Pooled Effects by Design

Observational Cohort Design

Sixteen meta-analyses (from eight drug-to-drug comparisons) of only observational cohort studies suggest that treatment with SUs poses a greater risk than other therapies. Three of these comparisons involved SU monotherapy against

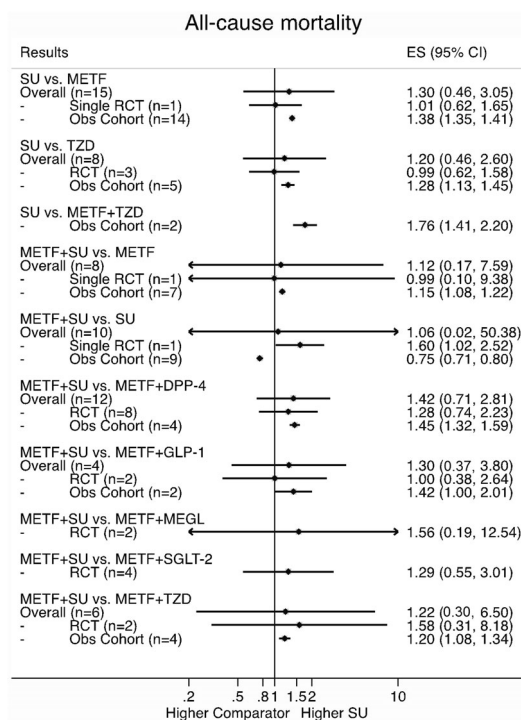


Fig. 2 Pooled relative risks (RR) for all-cause mortality. Inverse variance fixed effect estimates are shown for pooled estimates by study design, and two-level hierarchical Bayesian estimates are shown for overall pooled estimates. RR and the 95% confidence interval (CI) are presented for results by study design, and RR and 95% credible intervals are presented for overall pooled estimates. *DPP-4* Dipeptidyl peptidase-4, *ES* effect size, *GLP-1* glucagon-like peptide-1, *MEGL* meglitinide, *METF* metformin, *Obs* observational, *SGLT-2* sodium–glucose co-transporter 2, *SU* sulfonylurea, *TZD* thiazolidinedione

METF (all-cause mortality: RR 1.38, 95% CI 1.35, 1.41; cardiovascular mortality: 1.21 95% CI 1.16, 1.27; cardiovascular composite RR 1.18, 95% CI 1.15, 1.22), thiazolidinedione (TZD) (all-cause mortality: RR 1.28, 95% CI 1.13, 1.45), and combination METF + TZD (all-cause mortality: RR 1.76, 95% CI 1.41, 2.20; cardiovascular composite: RR 1.99, 95% CI 1.47, 2.69).

There were also differential risks when SU combination therapy was evaluated against SU and METF monotherapy, respectively. A lower risk was associated with METF + SU combination therapy when compared to SU monotherapy (all-cause mortality: RR 0.75, 95% CI 0.71,

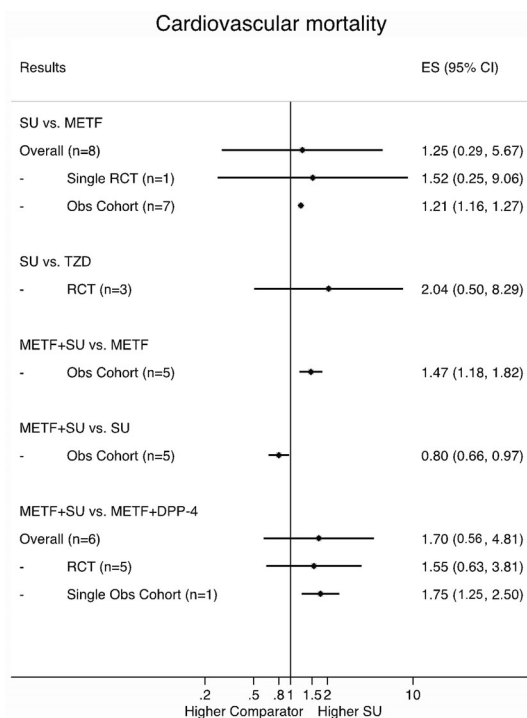


Fig. 3 Pooled RR for cardiovascular mortality. Inverse variance fixed effect estimates are shown for pooled estimates by study design, and two-level hierarchical Bayesian estimates are shown for overall pooled estimates. RR and the 95% CI are presented for results by study design, and RR and 95% credible intervals are presented for overall pooled estimates

0.80; cardiovascular mortality: RR 0.80, 95% CI 0.66, 0.97; cardiovascular composite: RR 0.84, 95% CI 0.77, 0.93), and a higher risk was associated with SU + METF combination therapy compared against METF monotherapy (all-cause mortality: RR 1.15, 95% CI 1.08, 1.22; cardiovascular mortality: RR 1.47, 95% CI 1.18, 1.82).

The remaining analyses found elevated effects for SU + METF combination therapy relative to other METF combinations, such as METF + TZD (all-cause mortality: RR 1.20, 95% CI 1.08, 1.34; cardiovascular composite: RR 1.12, 95% CI 1.03, 1.23), METF + dipeptidyl peptidase-4 (DPP-4) (all-cause mortality: RR 1.45, 95% CI 1.32, 1.59; cardiovascular composite: RR 1.46, 95% CI 1.28, 1.68), and METF + glucagon-like peptide-1 (GLP-1) (all-cause mortality: RR 1.42, 95% CI 1.00, 2.01).

In addition, pooled results were statistically inconsistent in four analyses between the fixed

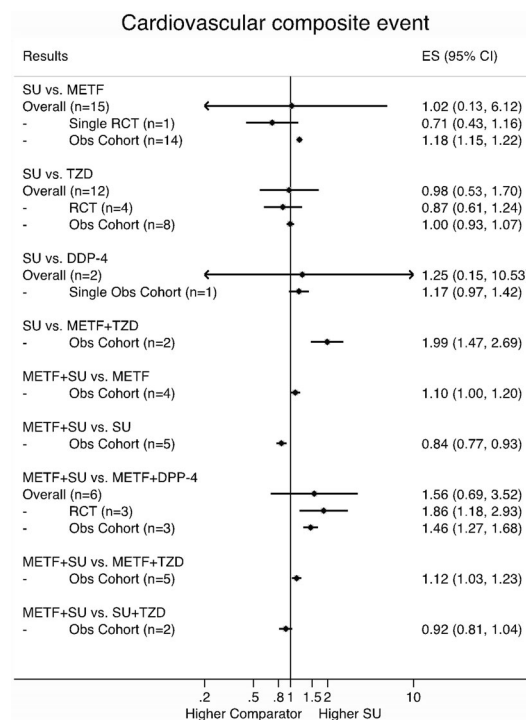


Fig. 4 Pooled RR for cardiovascular composite events. Inverse variance fixed effect estimates are shown for pooled estimates by study design, and two-level hierarchical Bayesian estimates are shown for overall pooled estimates. RR and the 95% CI are presented for results by study design, and RR and 95% credible intervals are presented for overall pooled estimates

inverse variance method and the DerSimonian–Laird random effect method, such that the added between-study variance included in the random effects estimates produced wider confidence intervals for the pooled effect in all cases, giving statistically non-significant estimates. Thus, substantial heterogeneity existed within each of these analyses, with the I^2 statistic ranging from 74 to 93%. All of these analyses involved METF + SU combination therapy compared to monotherapies, and they found a lower risk when compared to SU alone (all-cause, cardiovascular composite) and a higher risk when compared to METF monotherapy (on all-cause mortality, cardiovascular death). With the exception of this last drug comparison, all of the inconsistent comparisons had a similar magnitude and directions of the estimated pooled effects between random effects and fixed effects estimates (see ESM S6).

Randomized Controlled Trials

One significant elevated effect was found in the series of analyses using only RCTs. People randomized to receive the combination METF + SU had an 86% increased risk of a cardiovascular composite event than those assigned combined therapy with METF + DPP-4 (pooled RR 1.86, 95% CI 1.18, 2.93). All other pooled estimates of RCT design studies failed to detect a difference in risk between SU therapy and other regimens for all outcomes. While most comparisons had the same direction in the effect as pooled observational cohort estimates, precision was often worse in the RCT than in its pooled observational cohort counterpart.

Overall Combined Across Study Design

None of the analyses suggested an elevated effect for SUs when results were combined across RCT and observational cohort study designs according to all two-level hierarchical Bayesian models. While the overall direction and magnitude of the effect estimates are similar to that of the pooled estimates from observational cohort designed studies, overall pooled estimates have considerably wider credible intervals. This is most likely a result of the added variation existing between study designs.

Publication Bias

Assessing publication bias was limited since most analyses were excluded if there were < 10 studies included. There was no significant test result suggesting publication bias according to Egger's test.

DISCUSSION

Cardiovascular disease is the main cause of death in people with diabetes, yet evidence on whether particular drug therapies contribute to an increase in cardiovascular events and mortality has been unclear and insufficient. Early evidence for concerns over SU use came from the UK Prospective Diabetes Study [2] and from studies showing that their use is associated with weight gain, fluid retention, and hypoglycemia,

all of which are known risk factors for CVD. Certain SUs affect vascular ATP-sensitive potassium channels (KATP channels); this results in interference with ischemic preconditioning and the KATP channels possibly not being selective for pancreatic β -cells and rather binding to receptors in other tissues, such as cardiomyocytes and vascular smooth muscle cells [14]. These findings, together with mounting evidence from epidemiologic studies, have further raised concerns over the use of SUs.

The pooled results of the series of meta-analyses reported here suggest that SU therapy is associated with an elevated health risk relative to METF, TZD, GLP-1 agonists, and DPP-4 inhibitors when either compared as a monotherapy or when used in combination with METF. These findings are almost entirely derived from observational data (with one exception).

While most RCT-derived estimates were in the same direction as and had a similar magnitude to those for their observational cohort counterpart, the uncertainty surrounding each effect was much larger for the former. Therefore, when evidence was pooled using both types of study design, there was high variability around the effect estimates (wide credible intervals) as a result of the imprecise estimates reported from prior RCT studies. Across all RCTs in this study, the majority that evaluated long-term safety outcomes had small sample sizes with relatively few or no events in a given drug group occurring during the follow-up period. As a result, existing RCTs were not sufficiently powered to evaluate long-term safety outcomes.

Pooled estimates from the observational studies suggest worse outcomes for SUs versus older type 2 diabetes drug classes. For the monotherapy regimens, a higher pooled relative risk was reported for SU monotherapy in comparison to METF on all three safety outcomes, and for TZD on all-cause mortality. The results also suggest a higher risk for both SU monotherapy and METF + SU combination therapy than with METF + TZD combination therapy for all-cause mortality and cardiovascular composite events.

Beginning in 2008, all novel type 2 diabetes medications have to undergo a trial focused on

cardiovascular outcomes. These studies have typically involved the enrollment of patients with high cardiovascular risks (those with numerous CVD risk factors or with existing CVD). Evidence from most studies indicate that novel agents do not pose an increased cardiovascular risk compared to placebo (exception being saxagliptin, for which an increased risk of hospitalization for heart failure has been shown [15]). However, there are several shortcomings to these studies, with criticism focused on their lack of a clear interpretation of the cardiovascular risk among the broader indicated population, an insufficient study period duration which does not allow understanding of the cardiovascular safety profile (there is no mandatory minimum duration set for these studies), and the fact that placebo-controlled trial designs do not provide insight into clinically relevant questions [16].

Meta-analysis of the results of observational cohort studies suggests that SUs have higher long-term risks than do the newer potential second-line drug classes on one or more outcomes. Compared to the combination METF + DDP-4, our results suggest that the combination METF + SU poses an increased risk for cardiovascular composite events (which is in agreement with the RCT pooled results) as well as for all-cause mortality, and compared to the combination METF + GLP-1, there was an elevated risk for all-cause mortality.

While new evidence from trials on second-line medications after METF with a long follow-up would ideally be a welcome addition, such studies are typically neither feasible nor timely. This is particularly true for any study investigating comparative safety among the older drug classes, such as SUs and TZDs. For example, the one trial of second-line TZD use that had a follow-up of > 1 year (TOSCA.IT) was underpowered, with only about one-third of the actual events necessary to detect a 20% reduction in the cardiovascular composite outcome with 80% power [17]. In addition, the forthcoming GRADE study does not have a TZD treatment arm [18].

While ongoing trials such as the GRADE and CAROLINA trials may provide evidence on newer classes of drugs with a longer follow-up

than previously reported [18, 19], there is increasing pressure to include evidence derived from non-randomized designs [20]. With this increasing demand, a methodological challenge for researchers is how/whether evidence from observational cohort and RCTs can be combined to inform key treatment decisions. In our study, we used a two-level Bayesian model to explore how results can be synthesized across study designs. Since there were fewer RCTs than observational studies in our meta-analyses, this strategy tended to give more weight to RCTs than otherwise would occur if the results were simply combined without any consideration of the study design. However, this strategy also included additional variance in the form of between-study design variance in the Bayesian models.

Future studies should explore whether there are other suitable methods to account for uncertainty and pooling estimates across study designs for the purpose of advancing empirical knowledge and informing evidence-based medicine practice. In particular, Bayesian multilevel models that use informed prior distributions that are formally specified to reflect the relative strength of RCT designs compared to observational designs would be most beneficial. Such an approach would assign less weight to study design types that are more susceptible to bias (e.g., observational designs) relative to RCT designs. Empirically, these weights might be developed via meta-regression examining how effect estimates vary by study design, as has been suggested previously [21]. Additionally, expert judgments may be elicited via survey or using a Delphi or group consensus approach, where this information may be quantified in the form of a prior probability distribution.

Finally, it is important to note that there are several shortcomings in existing comparative safety analyses that need to be explored in future research. While SU therapy is commonly compared to metformin and TZD, there is limited comparative safety research on how the newer classes of medications compare against SU therapy (e.g., sodium–glucose co-transporter 2 inhibitors). There are even fewer comparative safety analyses that parse out the different sequencing possibilities involving SU

combination therapy, such as whether existing therapy (e.g., often METF monotherapy) is discontinued or augmented when a second-line therapy is introduced.

Also, there were few comparisons that included ≥ 10 studies to examine publication bias, and so this factor cannot be ruled out. In addition, other biases beyond the types assessed in this study could influence study effect sizes. In future work, meta-regression is one way to explore the influence that various study characteristics as well as other effect-modifying factors have on these estimates.

CONCLUSIONS

While the results from previous studies suggest that type 2 diabetes medications other than SUs appear to have equal glucose-lowering efficacy both alone and when combined with METF [1], further research is needed to determine whether they also provide greater long-term safety. In this study, meta-analyses using only observational cohort evidence suggest that SUs pose an elevated risk when compared to other drug classes. RCTs to date have been poorly designed to evaluate long-term outcomes with type 2 diabetes medications, resulting in few events and providing little evidence. The focus of many of these trials has been to make direct head-to-head comparisons to assess which medications work best at managing glucose levels, and they were not designed to examine long-term risks. These trials have typically been small in size with relatively short follow-up periods, thereby limiting the ability to obtain precise estimates of risk.

While much of the evidence is derived and will continue to come from observational database studies, the methodological rigor of such studies is questionable (e.g., internal threats to validity such as selection bias and unmeasured confounding are possible). Since evidence from RCTs on the long-term risks is typically not feasible or underpowered, a greater emphasis on designing frameworks for comparative safety research that incorporate evidence from well-designed, rigorous observational studies is needed.

ACKNOWLEDGEMENTS

The authors would like to thank Lewis Kazis, Manuel Cifuentes, and Varsha Vimalananda for their helpful feedback.

Funding. No funding or sponsorship was received for this study or publication of this article. The article processing charges were funded by the authors.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Authorship Contributions. W. Ryan Powell designed the study, performed the data analysis, interpreted the results, wrote the manuscript, and approved the final version. Donald R. Miller and Cindy L. Christiansen contributed to the design, reviewed/edited the manuscript, and gave final approval. W. Ryan Powell is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Disclosures. The authors (W. Ryan Powell, Cindy L. Christiansen, and Donald R. Miller) have nothing to disclose.

Compliance with Ethics Guidelines. This article is based on previously conducted studies and does not contain any studies with human participants or animals performed by any of the authors.

Data Availability. All study-specific effect sizes analyzed in this series of meta-analyses are included in ESM S6.

Open Access. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any

medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

1. Bolen S, Tseng E, Hutfless S, et al. Diabetes medications for adults with Type 2 diabetes: an update. Rockville: Agency for Healthcare Research and Quality (US); 2016.
2. United Kingdom Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet* 1998;352(9131):837–853.
3. Abrahamson MJ. Should sulfonylureas remain an acceptable first-line add-on to metformin therapy in patients with type 2 diabetes? Yes, they continue to serve us well! *Diabetes Care*. 2015;38(1):166–9.
4. Genuth S. Should sulfonylureas remain an acceptable first-line add-on to metformin therapy in patients with type 2 diabetes? No, it's time to move on! *Diabetes Care*. 2015;38(1):170–5.
5. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17:1–12.
6. Wells G, Shea B, O'Connell D, et al. The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in metaanalyses [Internet]; 2001. <http://www.medicine.mcgill.ca/rtamblyn/Readings%5CThe%20Newcastle%20-%20Scale%20for%20assessing%20the%20quality%20of%20nonrandomised%20studies%20in%20meta-analyses.pdf>.
7. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Chichester: Wiley; 2009.
8. StataCorp LP. Stata statistical software: release 14. College Station: StataCorp LP; 2015.
9. Peters JL, Rushton L, Sutton AJ, Jones DR, Abrams KR, Mugglestone MA. Bayesian methods for the cross-design synthesis of epidemiological and toxicological evidence. *J R Stat Soc Ser C Appl Stat*. 2005;54(1):159–72.
10. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–60.
11. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–34.
12. Harbord RM, Harris RJ, Sterne JAC. Updated tests for small-study effects in meta-analyses. *Stata J*. 2009;9(2):197–210.
13. Higgins, JPT, Green S. *Cochrane handbook for systematic reviews of interventions* (5.1.0 [updated March 2011]). Hoboken: John Wiley & Sons. <http://www.handbook.cochrane.org>.
14. Gribble FM, Tucker SJ, Seino S, Ashcroft FM. Tissue specificity of sulfonylureas: studies on cloned cardiac and β -cell K(ATP) channels. *Diabetes*. 1998;47:1412–8.
15. Scirica BM, Braunwald E, Raz I, et al. Heart failure, saxagliptin and diabetes mellitus: observations from the SAVOR-TIMI 53 randomized trial. *Circulation* 2014;130(18):1579–88.
16. Regier EE, Venkat MV, Close KL. More than 7 years of hindsight: revisiting the FDA's 2008 guidance on cardiovascular outcomes trials for Type 2 diabetes medications. *Clin Diabetes*. 2016;34(4):173–80.
17. Vaccaro O, Masulli M, Nicolucci A, et al. Effects on the incidence of cardiovascular events of the addition of pioglitazone versus sulfonylureas in patients with type 2 diabetes inadequately controlled with metformin (TOSCA.IT): a randomised, multicentre trial. *Lancet Diabetes Endocrinol*. 2017;5(11):887–97.
18. Nathan DM, Buse JB, Kahn SE, et al. Rationale and design of the glycemia reduction approaches in diabetes: a comparative effectiveness study (GRADE). *Diabetes Care*. 2013;36(8):2254–61.
19. Marx N, Rosenstock J, Kahn SE, et al. Design and baseline characteristics of the CARDiovascular Outcome Trial of LINagliptin Versus Glimepiride in Type 2 Diabetes (CAROLINA®). *Diab Vasc Dis Res*. 2015;12(3):164–174.
20. 114th U.S. Congress. 21st Century Cures Act; H.R. 34 (114-255 Public Law 114–255); 2015.
21. Goodman SN. Bayesian methods for evidence evaluation. *Circulation*. 2013;127(24):2367–9.