

Validity of Meta-analysis in Diabetes: Meta-analysis Is an Indispensable Tool in Evidence Synthesis

To deliver high-quality clinical care to patients with diabetes and other chronic conditions, clinicians must understand the evidence available from studies that have been performed to address important clinical management questions. In an evidence-based approach to clinical care, the evidence from clinical research should be integrated with clinical expertise, pathophysiological knowledge, and an understanding of patient values. As such, in an effort to provide information from many studies, the publication of diabetes meta-analyses has increased markedly in the recent past, using either observational or clinical trial data. In this regard, guidelines have been developed to direct the performance of meta-analysis to provide consistency among contributions. Thus, when done appropriately, meta-analysis can provide estimates from clinically and statistically homogeneous but underpowered studies and is useful in supporting clinical decisions, guidelines, and cost-effectiveness analysis. However, often these conditions are not met, the data considered are unreliable, and the results should not be assumed to be any more valid than the data underlying the included studies. To provide an understanding of both sides of the argument, we provide a discussion of this topic as part of this two-part point-counterpoint narrative. In the point narrative preceding the counterpoint narrative below, Dr. Home provides his opinion and review of the data to date showing that we need to carefully evaluate meta-analysis, and we need to learn what results are reliable. In the counterpoint narrative here, Drs. Golden and Bass emphasize that an effective system exists to guide meta-analysis and that rigorously conducted, high-quality systematic reviews and meta-analyses using established guidelines are an indispensable tool in evidence synthesis despite their limitations.

—WILLIAM T. CEFALU, MD
EDITOR IN CHIEF, *DIABETES CARE*

In an evidence-based approach to clinical care, the evidence from clinical research should be integrated with clinical expertise, pathophysiological knowledge, and an understanding of patient values (1). Some advocates of evidence-based medicine argue that the best evidence comes from systematic reviews of all relevant studies, and they place systematic reviews at the top of the evidence pyramid (Fig. 1). As outlined by Dr. Home in the point narrative (2), we acknowledge that there are limitations and biases associated with the systematic review and meta-analysis process. However, for the purpose of this discussion, we take the view that this process is an indispensable tool in evidence synthesis. In our counterpoint highlighting the critical role of systematic review and meta-analysis in diabetes research and clinical practice, we will 1) review how to ensure that a systematic review follows a scientifically rigorous process for minimizing errors or bias, 2) explain how a systematic review adds to what is known from randomized controlled trials (RCTs), and 3) give examples of how a rigorous systematic review and meta-analysis can and has guided clinical decision making in diabetes care.

THE SCIENCE OF SYSTEMATIC REVIEWS AND META-ANALYSIS

History and current guidelines

According to the definitions used by The Cochrane Collaboration, a systematic review is the process of collecting, reviewing, and presenting all available evidence related to a clearly formulated question that uses systematic and explicit methods (3,4). Meta-analysis is the statistical technique for extracting and combining data to produce a summary result of the included studies (3,4). Depending on the presentation and heterogeneity of the data, a systematic review may or may not include a meta-analysis.

Guidelines have been developed to guide the performance of systematic reviews because of concerns about inconsistent methods and the potential for introducing bias and error in the review process as already discussed by Home. In 1996, an international group developed the QUOROM Statement (Quality Of Reporting Of Meta-analyses) for reporting of meta-analyses of RCTs (5). A similar set of guidelines for reporting meta-analyses of observational studies, MOOSE (Meta-analysis

Of Observational Studies in Epidemiology) guidelines, was published in 2000 (6). In 2009, the QUORUM statement was revised, updated, and renamed the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) statement to address several conceptual and practical advances in the science of systematic reviews (3). The PRISMA statement emphasizes 1) the iterative process of conducting a systematic review, 2) that conduct and reporting of the research are distinct processes, 3) the importance of study level versus outcome level risk of bias assessment, and 4) the importance of reporting bias (3). The goal of the PRISMA statement is to help authors improve reporting of systematic reviews and meta-analyses (3). It includes a 27-item checklist and four-phase diagram (3) (Supplementary Table 1 and Supplementary Fig. 1) that can be applied to studies including RCTs as well as other types of research (3). Most recently, the Institute of Medicine published a report on standards for systematic reviews (7). The report presents specific standards for each of the following aspects of a systematic review: initiating a systematic review, finding and assessing studies, synthesizing evidence, and reporting results (Supplementary Table 2).

Steps of the systematic review process

Table 1 summarizes the steps in the systematic review process along with recommendations to avoid bias and error in the review process as articulated by Home. We believe that it is critically important to conduct and report systematic reviews using rigorous, established guidelines (3,6) (see Supplementary Table 1).

CLINICAL QUESTIONS ADDRESSED BY SYSTEMATIC REVIEWS AND META-ANALYSES

—A rigorous review conducted according to the established guidelines summarized above can address important clinical questions that cannot be fully answered by RCTs alone and can be a powerful tool in evidence synthesis. We highlight the types of critical gaps that systematic reviews and meta-analyses can fill, along with examples from diabetes literature, and highlight their value

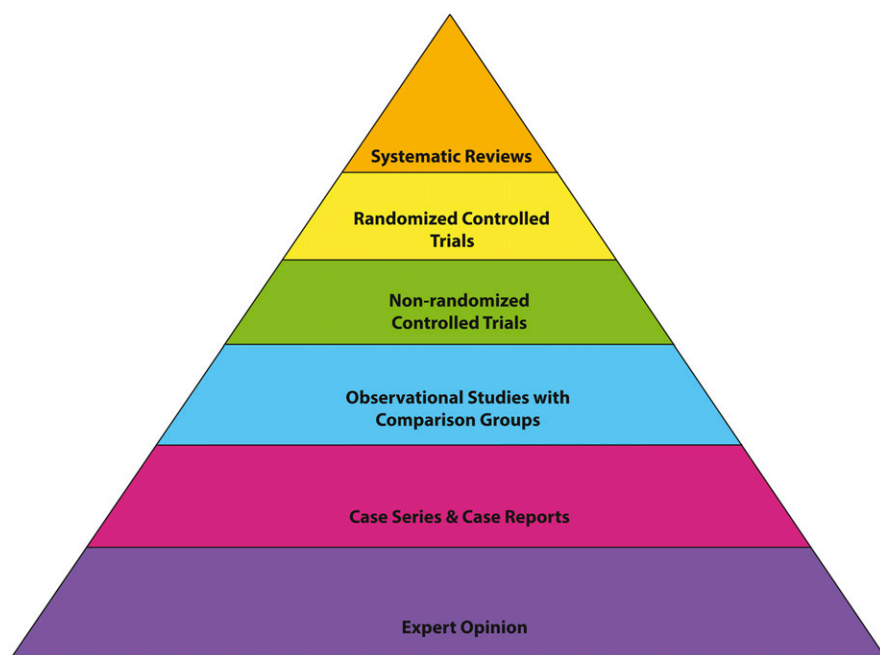


Figure 1—Evidence-based medicine pyramid. The levels of evidence are appropriately represented by a pyramid as each level, from bottom to top, reflects the quality of research designs (increasing) and quantity (decreasing) of each study design in the body of published literature. For example, systematic reviews are higher quality and more labor intensive to conduct, so there is a lower quantity published.

that cannot be replicated by other study designs.

RCTs may be unable to draw definitive conclusions

Well-designed RCTs are considered the gold standard for answering clinical research questions because they have the lowest risk of bias. However, there are many circumstances in which RCTs alone fail to provide definitive conclusions. These include the following: 1) multiple RCTs yield conflicting results making definitive conclusions unclear; 2) certain clinical outcomes, such as cardiovascular end points, require longer term follow-up than the duration of the RCTs; 3) individual RCTs are underpowered to identify significant adverse events; and 4) randomization to certain exposures is unethical. In these circumstances, systematic review and meta-analysis of the existing literature in a specific area, using a rigorous, scientific approach, can be a powerful tool in evidence synthesis to guide future research, clinical guideline development, and health care policy.

Multiple RCTs yield conflicting results

Our scientific approach to medicine generally requires more than one RCT to make a definitive conclusion. This necessary replication can inevitably result in

situations in which multiple trials yield conflicting results, making definitive conclusions unclear. In this circumstance, a well-done meta-analysis can provide a means of data synthesis and reconciliation to improve understanding of conflicting results. For example, an initial RCT of adult intensive care unit (ICU) patients (primarily surgical) demonstrated that tight glycemic control, achieved with intravenous insulin infusion to maintain a glucose target of 80–110 mg/dL, significantly reduced total mortality and mortality secondary to multiorgan failure from a septic focus (8). As a result, professional organizations recommended an ICU glycemic target of 80–110 mg/dL, a noncritical care target of 110 mg/dL preprandially, and a maximum glucose of 180 mg/dL (9). Subsequent trials in medical ICU patients and other settings, however, did not confirm these findings (10,11), and one large clinical trial found increased mortality in the intensive insulin therapy group (11). Similarly, an initial clinical trial of tighter compared with conventional peri- and postoperative glycemic control in myocardial infarction patients showed a significant reduction in mortality (12); however, this was not confirmed in a subsequent follow-up trial (13). Because of conflicting trial data, two meta-analyses were undertaken to synthesize the literature, including these RCTs as well as several

others. While one meta-analysis showed a significant reduction in septicemia with tight glucose control compared with conventional control (14), the other, which was published after 2008 and included the Normoglycemia in Intensive Care Evaluation and Survival Using Glucose Algorithm Regulation (NICE-SUGAR) trial, showed no difference in clinical outcomes (15). Both meta-analyses demonstrated that tight glycemic control was not associated with a reduction in mortality and was associated with a significantly increased risk of hypoglycemia (14,15). Consequently, professional organizations took appropriate steps to alter inpatient glycemic targets to safer and achievable levels and currently recommend an ICU target of 140–180 mg/dL and a preprandial target of less than 140 mg/dL with a maximum glucose of 180 mg/dL in non-ICU settings (16). Thus, well-done meta-analyses have guided appropriate alteration of clinical practice guidelines as new but conflicting data became available.

Informing diabetes diagnosis, monitoring, and clinical treatment

The systematic review process can inform decisions regarding diagnostic strategies for diabetes. A systematic review of the characteristics of postpartum screening tests in women with a prior history of gestational diabetes mellitus showed that a single fasting blood glucose was not a sensitive screening test compared with the standard oral glucose tolerance test for detecting type 2 diabetes, suggesting that it should not replace currently recommended screening test (17). This review also pointed out important limitations of the existing literature in this area (17). Thus, a rigorous systematic review also brings attention to gaps in evidence and the need for more research to address important clinical questions.

Systematic reviews and meta-analysis can also inform treatment practices in the monitoring and clinical management of diabetes. In recent years, many advances have been made in technologies to deliver insulin (continuous subcutaneous insulin infusion [CSII]) and to monitor glucose (real-time continuous glucose monitoring [rt-CGM]); however, their comparative effectiveness as well as the populations most likely to benefit had not been clearly demonstrated (18). Because these technologies are expensive and may be heavily marketed, objective information about their comparative effectiveness with conventional approaches is important so that patients and health care providers can

Table 1—Steps in the systematic review process and guideline recommendations to address potential sources of bias and error

Systematic review step	Potential causes of bias	Potential causes of error	Guideline recommendations to avoid bias and error
Defining the question	Conflict of interest of analyst	Ambiguous formulation of question	Establish team with appropriate expertise
Determine priorities	Conflict of interest of stakeholders	Insufficient time and resources	Ensure user and stakeholder input
Create analytic framework	Selective input from stakeholders		Manage stakeholder and research team bias and conflict of interest
Engage potential stakeholders (e.g., clinicians, researchers policy makers, insurance plans, patients, community advocates)			Formulate topic
Formulate question in specific measurable terms (PICOTS)			
<u>P</u> opulation			
<u>I</u> ntervention			
<u>C</u> omparisons			
<u>O</u> tcomes			
<u>T</u> iming			
<u>S</u> etting			
Set eligibility criteria and exclusion criteria a priori	Exclusions influenced by conflict of interest	Inadequate specification of inclusion/exclusion criteria	Handling of excluded studies
Review PICOTS	Exclusions influenced by knowledge of study results		Enumerate and explain
Choose study design(s)			Include in sensitivity analysis
Set other criteria (study size, publication year, peer-reviewed, language)			Comment in discussion section
Search literature and review citations	Publication bias	Inadequate monitoring of search process	Conduct comprehensive search and document search for reproducibility
Select highest yield sources (PubMed, EMBASE, Cochrane database)	Redundant reporting	Imperfect reliability of reviewers	Use independent dual reviewers
Consider other sources	Exclusion of important sources		Use key articles to validate search
Use hand searches	Incomplete development of search strategy		Assess samples early
			Discuss discrepancies
			Watch for duplication
Extract data	Reporting bias	Careless or rushed extraction	Manage data collection
Set priorities for extraction	Data dredging	Sloppy or inconsistent reporting of results	Dual review to reduce errors (independent or sequential) (37)
Choose data management method (systematic review software vs. database software vs. tables)	Extractor influenced by conflict of interest	Imperfect fit between study report and extraction form	Use early random auditing to estimate error rates
		Inadequate training and auditing	Adjudication process to discuss discrepancies
Assess study quality and applicability	Evaluator influenced by conflict of interest or by another reviewer	Inadequate specification of assessment criteria	Address potentially biased reporting of results
		Limited reliability of assessment instruments	Critically appraise each study
		Inadequate training and auditing	Use standard instruments for RCTs (38) and observational studies (39)
			Use prespecified method to evaluate evidence
			Independent dual review
			Adjudication process to reconcile discrepancies

Continued on p. 3371

Table 1—Continued

Systematic review step	Potential causes of bias	Potential causes of error	Guideline recommendations to avoid bias and error
Summarize and synthesize evidence			
Prepare summary of evidence on each question			
Assemble evidence tables (qualitative summary)	Improper selection of outcome measures (risk difference vs. relative risk or odds ratio)	Mistakes in transferring or pooling data	Conduct qualitative synthesis
Perform meta-analysis (quantitative summary)	Improper choice of pooling technique (random vs. fixed effects)	Poor understanding of GRADE/evidence grading approach	Consider quantitative synthesis
Assess SOE using GRADE or similar approach, accounting for risk of bias, consistency, directness, and precision (40,41)	SOE grading influence by conflict of interest		Use structured reporting format (3,5)

SOE, strength of evidence.

make informed decisions. A recent systematic review and meta-analysis of RCTs showed that multiple daily injections (MDIs) and analog-based CSII had similar effects on hemoglobin A_{1c} (HbA_{1c}) levels (18) and severe hypoglycemia in children and adults with type 1 diabetes and in adults with type 2 diabetes, indicating that glycemic goals can be achieved with either method of intensive insulin delivery and that CSII is not superior (18). In contrast, two recent meta-analyses demonstrated that compared with self-monitoring of blood glucose (SMBG), rt-CGM achieved lower HbA_{1c} without differences in severe hypoglycemia in individuals with type 1 diabetes (18,19). In addition, sensor-augmented insulin pumps, which combine CSII with rt-CGM, significantly decreased HbA_{1c} more than MDI with SMBG in patients with type 1 diabetes (18). These studies suggest that using rt-GCM may have a favorable impact on glycemic control regardless of the methods of insulin delivery; however, as pointed out as a limitation of the review, the current literature does not allow a comparison of rt-CGM versus SMBG in patients only using CSII or only using MDI because the modes of insulin delivery were mixed (18). Therefore, the systematic reviews and meta-analyses identified a need for additional clinical trials in the area of glucose monitoring.

Detecting adverse events

Most RCTs have inadequate power to identify serious adverse events, but systematic reviews and meta-analyses of RCTs, supplemented with observational studies, can provide more power to detect such events, particularly when multiple, small studies have been conducted. Meta-analyses initially raised concern about the increased risk of myocardial infarction in individuals taking rosiglitazone compared with placebo, other agents, or pioglitazone (20,21). This ultimately led the Food and Drug Administration to issue a black box warning for rosiglitazone. In addition, meta-analyses have shown that thiazolidinediones are associated with increased risk for congestive heart failure and bone fracture compared with other agents (22,23). These signals may not have become apparent without well-conducted meta-analyses.

Estimating effect sizes for clinical treatments and clinical trial design

Well done meta-analyses provide reliable information about anticipated effect sizes for clinical treatment decisions. This is

particularly valuable when the optimal choice of treatment is sensitive to a patient's views about the seriousness of potential complications or outcomes. Information about expected treatment effect sizes can also be helpful to investigators when they are planning RCTs. For example, treatment effect sizes from a meta-analysis of observational studies can be used to estimate 1) anticipated treatment effect sizes in an RCT comparing two interventions and 2) the number of participants necessary to have adequate power to detect that effect. A recent meta-analysis of the comparative effectiveness of various two-drug combinations for treatment of type 2 diabetes found that the combinations similarly reduced HbA_{1c} by 1 percentage point, although with different side effect profiles (23). Which drug combinations are most effective for various patients remains an important clinical question. Building on this meta-analysis, an RCT, the Glycemic Reduction Approaches in Diabetes: A Comparative Effectiveness Study (GRADE), is currently underway to compare various combination treatment approaches with metformin in patients with type 2 diabetes (24). This trial will confirm or refute the effect sizes for various outcomes suggested by the prior meta-analysis.

Answering clinical questions not amenable to RCTs

Though considered the gold standard, RCTs are unable to answer all questions because it is not feasible and/or ethical to randomize patients to certain exposures, necessitating observational study designs that can be incorporated into systematic reviews to identify novel risk factors for disease outcomes. For example, individuals cannot be randomly assigned to having a depressive disorder versus not, but observational studies can be used to define the experience and outcomes of individuals who do and do not have depressive disorders. Well-done meta-analyses published in *Diabetes Care* have identified depression as a risk factor for insulin resistance (25), metabolic syndrome (26), and type 2 diabetes (27). Identification of such novel risk factors for adverse metabolic outcomes set the stage for future preventive intervention trials to determine if treatment for depression improves insulin resistance and/or prevents metabolic syndrome and diabetes.

In some instances, it is unethical to randomize individuals to certain exposures. The Thiazolidinedione Intervention with Vitamin D Evaluation (TIDE) trial was

designed as a postmarketing study to compare the cardiovascular safety of rosiglitazone and pioglitazone (28). However, as concerns accumulated about the adverse events of rosiglitazone, the TIDE trial was halted because it was deemed unethical to continue exposure to rosiglitazone (28). Hence, meta-analyses of prior RCTs and observational studies may provide the best way to examine the association of rosiglitazone with important clinical outcomes. Finally, certain clinical outcomes, such as cardiovascular end points, require longer term follow-up than the duration of the RCTs and necessitate longitudinal, observational studies that can subsequently be included in systematic reviews and meta-analyses.

Enhancing the generalizability/ applicability of RCTs

Many RCTs have very restrictive entry criteria, often resulting in homogenous populations or populations that are not truly reflective of the race/ethnic or sex distribution of the individuals with the disorder of interest, limiting their generalizability. A systematic review of observational studies that includes more diverse populations can address these concerns. For example, RCTs comparing bariatric surgery to medical therapy for weight loss in patients with type 2 diabetes (29,30) or for weight loss in the prevention of type 2 diabetes (31) have included primarily Caucasian populations. Two studies were conducted exclusively in European Caucasian populations (30,31), and the participants in one U.S. study were 74% Caucasian (29), but it is well established that ethnic minority populations have the highest prevalence and risk of type 2 diabetes (32). A recently published meta-analysis attempted to initially address the important issue of effectiveness of bariatric surgery in individuals of African and Caucasian descent and showed that the percent estimated weight loss was greater in Caucasians compared with African Americans (33). These data, which could not be derived from existing RCTs, point to the need for future bariatric surgery trials to reflect the race/ethnic composition of the population that has diabetes and to determine the reasons for these disparities.

Identification of low/insufficient strength of evidence and future research needs

Modern systematic reviews require assessment of the strength of evidence for the key research questions examined (3).

Associations and comparisons for which the strength of evidence is low or insufficient can point to important areas for future research and where to allocate research resources. In the area of diabetes treatment, strength of evidence was moderate for race/ethnic differences in percentage estimated weight loss in patients with diabetes undergoing bariatric surgery; however, strength of evidence was poor for race/ethnic differences in remission of type 2 diabetes with bariatric surgery (34). In comparing various type 2 diabetes medications on long-term clinical outcomes, strength of evidence was low or insufficient for all-cause mortality, cardiovascular disease morbidity and mortality, and the microvascular outcomes of neuropathy and retinopathy (23). Finally, while strength of evidence was moderate for the comparative effectiveness of CSII versus MDI on HbA_{1c} in patients with diabetes, it was low or insufficient for their effects on hyperglycemia, hypoglycemia, weight gain, quality of life, and long-term clinical outcomes (18). All of these data point to critical areas for additional research and the need for rigorous and appropriate study designs to assess the most relevant outcomes. For example, it may not be feasible to conduct a large, long-term study of the comparative effectiveness of CSII versus MDI on cardiovascular outcomes, and future studies will have to consider the appropriateness of using surrogate or intermediate outcomes as alternatives.

Contributing to conflict-free guideline development

Finally, using rigorously conducted systematic reviews and meta-analyses to inform guideline development permits them to be compiled in an objective and conflict-free manner (34). The preference is to use systematic reviews of RCTs when available, but in the absence of RCTs, systematic reviews of observational data can provide another, albeit slightly lower, level of evidence. A study recently examined whether guidelines on oral medications for treatment of type 2 diabetes were consistent with systematic review of the current evidence (35). Eleven guidelines were identified through their systematic review process from several different professional organizations and institutions. They found that 7 out of 11 guidelines agreed that metformin is favored as first-line therapy, 10 out of 11 agreed that thiazolidinediones were associated with higher rates of edema and congestive heart failure compared with other agents, and only 5 of 11 guidelines agreed with all 7 conclusions from the prior

systematic review (35). The guidelines with recommendations that were consistent with current evidence were the highest quality. A recent study found that low-quality systematic reviews were cited in 24 endocrinology clinical practice guidelines and used as the main evidence for five recommendations, with only one recommendation acknowledging the systematic review quality (36). Thus, the quality of the current guideline development process that informs clinical practice is quite variable, and high-quality systematic reviews and meta-analyses, using the approaches we have outlined, can contribute to evidence-based recommendations.

CONCLUSIONS—Systematic reviews and meta-analyses are indispensable tools in evidence synthesis when they are performed in a rigorous manner following established guidelines for minimizing bias and error. They are indispensable in the context of conflicting RCT data, detecting adverse events, informing disease monitoring and treatment, addressing clinical questions not amenable to RCTs, enhancing the generalizability/ applicability of RCTs, and identifying areas of low and/or insufficient strength of evidence for future research. Systematic reviews and meta-analyses will continue to play a critical role in developing unbiased guidelines that take into consideration an assessment of the risk of bias of individual studies as well as the overall strength of available evidence.

SHERITA HILL GOLDEN, MD, MHS^{1,2}
ERIC B. BASS, MD, MPH^{1,2,3,4}

From the ¹Department of Medicine, Division of Endocrinology, Diabetes, and Metabolism, and Division of General Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland; the ²Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland; the ³Department of Health Policy and Management, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland; and the ⁴Johns Hopkins University Evidence-Based Practice Center, Baltimore, Maryland.

Corresponding author: Sherita Hill Golden, sahill@jhmi.edu.

DOI: 10.2337/dc13-1196

© 2013 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.

Acknowledgments—No potential conflicts of interest relevant to this article were reported.

References

1. Tonelli MR. The philosophical limits of evidence-based medicine. *Acad Med* 1998;73:1234–1240
2. Home PD. Validity of meta-analysis in diabetes: we need to be aware of its limitations. *Diabetes Care* 2013;36:3361–3367
3. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264–269, W64
4. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0*. The Cochrane Collaboration, 2011
5. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet* 1999;354:1896–1900
6. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283:2008–2012
7. Institute of Medicine. *Finding What Works in Health Care: Standards for Systematic Reviews*. Eden J, Levit L, Berg A, Morton S, Eds. Washington, DC, The National Academies Press, 2011
8. van den Berghe G, Wouters P, Weekers F, et al. Intensive insulin therapy in critically ill patients. *N Engl J Med* 2001;345:1359–1367
9. Clement S, Braithwaite SS, Magee MF, et al.; American Diabetes Association Diabetes in Hospitals Writing Committee. Management of diabetes and hyperglycemia in hospitals. *Diabetes Care* 2004;27:553–591
10. Van den Berghe G, Wilmer A, Hermans G, et al. Intensive insulin therapy in the medical ICU. *N Engl J Med* 2006;354:449–461
11. Finfer S, Chittock DR, Su SY, et al.; NICE-SUGAR Study Investigators. Intensive versus conventional glucose control in critically ill patients. *N Engl J Med* 2009;360:1283–1297
12. Malmberg K. Prospective randomised study of intensive insulin treatment on long term survival after acute myocardial infarction in patients with diabetes mellitus. DIGAMI (Diabetes Mellitus, Insulin Glucose Infusion in Acute Myocardial Infarction) Study Group. *BMJ* 1997;314:1512–1515
13. Malmberg K, Rydén L, Wedel H, et al.; DIGAMI 2 Investigators. Intense metabolic control by means of insulin in patients with diabetes mellitus and acute myocardial infarction (DIGAMI 2): effects on mortality and morbidity. *Eur Heart J* 2005;26:650–661
14. Wiener RS, Wiener DC, Larson RJ. Benefits and risks of tight glucose control in critically ill adults: a meta-analysis. *JAMA* 2008;300:933–944
15. Kansagara D, Fu R, Freeman M, Wolf F, Helfand M. Intensive insulin therapy in hospitalized patients: a systematic review. *Ann Intern Med* 2011;154:268–282
16. Moghissi ES, Korytkowski MT, DiNardo M, et al.; American Association of Clinical Endocrinologists; American Diabetes Association. American Association of Clinical Endocrinologists and American Diabetes Association consensus statement on inpatient glycemic control. *Diabetes Care* 2009;32:1119–1131
17. Bennett WL, Bolen S, Wilson LM, Bass EB, Nicholson WK. Performance characteristics of postpartum screening tests for type 2 diabetes mellitus in women with a history of gestational diabetes mellitus: a systematic review. *J Womens Health (Larchmt)* 2009;18:979–987
18. Yeh HC, Brown TT, Maruthur N, et al. Comparative effectiveness and safety of methods of insulin delivery and glucose monitoring for diabetes mellitus: a systematic review and meta-analysis. *Ann Intern Med* 2012;157:336–347
19. Pickup JC, Freeman SC, Sutton AJ. Glycaemic control in type 1 diabetes during real time continuous glucose monitoring compared with self monitoring of blood glucose: meta-analysis of randomised controlled trials using individual patient data. *BMJ* 2011;343:d3805
20. Singh S, Loke YK, Furberg CD. Long-term risk of cardiovascular events with rosiglitazone: a meta-analysis. *JAMA* 2007;298:1189–1195
21. Nissen SE, Wolski K. Rosiglitazone revisited: an updated meta-analysis of risk for myocardial infarction and cardiovascular mortality. *Arch Intern Med* 2010;170:1191–1201
22. Loke YK, Singh S, Furberg CD. Long-term use of thiazolidinediones and fractures in type 2 diabetes: a meta-analysis. *CMAJ* 2009;180:32–39
23. Bennett WL, Maruthur NM, Singh S, et al. Comparative effectiveness and safety of medications for type 2 diabetes: an update including new drugs and 2-drug combinations. *Ann Intern Med* 2011;154:602–613
24. Nathan DM, Buse JB, Kahn SE, et al.; the GRADE Study Research Group. Rationale and design of the Glycemia Reduction Approaches in diabetes: A Comparative Effectiveness Study (GRADE). *Diabetes Care* 2013;36:2254–2261
25. Kan C, Silva N, Golden SH, et al. A systematic review and meta-analysis of the association between depression and insulin resistance. *Diabetes Care* 2013;36:480–489
26. Pan A, Keum N, Okereke OI, et al. Bidirectional association between depression and metabolic syndrome: a systematic review and meta-analysis of epidemiological studies. *Diabetes Care* 2012;35:1171–1180
27. Mezuk B, Eaton WW, Albrecht S, Golden SH. Depression and type 2 diabetes over the lifespan: a meta-analysis. *Diabetes Care* 2008;31:2383–2390
28. Mello MM, Goodman SN, Faden RR. Ethical considerations in studying drug safety—the Institute of Medicine report. *N Engl J Med* 2012;367:959–964
29. Schauer PR, Kashyap SR, Wolski K, et al. Bariatric surgery versus intensive medical therapy in obese patients with diabetes. *N Engl J Med* 2012;366:1567–1576
30. Mingrone G, Panunzi S, De Gaetano A, et al. Bariatric surgery versus conventional medical therapy for type 2 diabetes. *N Engl J Med* 2012;366:1577–1585
31. Carlsson LM, Peltonen M, Ahlin S, et al. Bariatric surgery and prevention of type 2 diabetes in Swedish obese subjects. *N Engl J Med* 2012;367:695–704
32. Golden SH, Brown A, Cauley JA, et al. Health disparities in endocrine disorders: biological, clinical, and nonclinical factors—an Endocrine Society scientific statement. *J Clin Endocrinol Metab* 2012;97:E1579–E1639
33. Admiraal WM, Celik F, Gerdes VE, Dallal RM, Hoekstra JB, Holleman F. Ethnic differences in weight loss and diabetes remission after bariatric surgery: a meta-analysis. *Diabetes Care* 2012;35:1951–1958
34. Institute of Medicine. *Clinical Practice Guidelines We Can Trust*. Graham R, Mancher M, and Wolman DM, et al., Eds. Washington, DC, National Academies Press, 2011
35. Bennett WL, Odelola OA, Wilson LM, et al. Evaluation of guideline recommendations on oral medications for type 2 diabetes mellitus: a systematic review. *Ann Intern Med* 2012;156:27–36
36. Brito JP, Tsapas A, Griebeler ML, et al. Systematic reviews supporting practice guideline recommendations lack protection against bias. *J Clin Epidemiol* 2013;66:633–638
37. Buscemi N, Hartling L, Vandermeer B, Tjosvold L, Klassen TP. Single data extraction generated more errors than double data extraction in systematic reviews. *J Clin Epidemiol* 2006;59:697–703
38. Oremus M, Wolfson C, Perrault A, Demers L, Momoli F, Moride Y. Interrater reliability of the modified Jadad quality scale for systematic reviews of Alzheimer's disease drug trials. *Dement Geriatr Cogn Disord* 2001;12:232–236
39. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52:377–384
40. GRADE Working Group. GRADE guidelines—best practices using the GRADE framework, 2013. Available from http://www.gradeworkinggroup.org/publications/JCE_series.htm
41. Atkins D, Best D, Briss PA, et al.; GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490