Behavioral Programs for Diabetes Mellitus
Executive Summary

Introduction

The high burden of diabetes necessitates careful attention to factors contributing to optimal diabetes care and self-management, including lifestyle behaviors and medication adherence. Over the past few decades, much of the care and education of people with diabetes in the United States has been transferred from hospitals to outpatient settings, and several guidelines and diabetes management programs have been developed to improve diabetes care in the community.[1] However, an evaluation of initiatives to implement guidelines and processes of care in community health centers did not find improved control of hemoglobin A_{1c} (HbA_{1c}) levels for patients with diabetes.[2]

Approaches for supporting patients with diabetes to change behaviors include interventions such as diabetes self-management education (DSME), with or without an additional support (clinical, behavioral, psychosocial, or educational) phase; lifestyle interventions; and medical nutrition therapy. Interventions vary widely in terms of content, duration, intensity, and delivery methods. The effectiveness of these interventions for patients with type 1 diabetes (T1DM) has not been evaluated in recent years and the few existing reviews have been inconclusive.[3-7] In contrast, there is a diverse evidence base supporting the effectiveness of these approaches for type 2 diabetes (T2DM). However, it is

Evidence-based Practice Program

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

The full report and this summary are available at www.effectivehealthcare. ahrq.gov/reports/final.cfm.

unknown what combination(s) of program components and delivery mechanisms are most effective for success for T2DM.

Epidemiology and Burden of Disease

In 2012, 29.1 million Americans had a form of diabetes (diagnosed and undiagnosed). This represents 9.3 percent of the entire population and 12.3 percent of the adult population 20 years or older.[8] Older adults are disproportionately affected with diabetes; 25.9 percent of people age 65 years or older have diabetes. African Americans, Hispanic Americans, American Indians and Alaska Natives, and some Asian Americans have a higher risk of T2DM than non-Hispanic whites.[8] Although most cases of diabetes are T2DM, T1DM is one of the most common chronic diseases in childhood and adolescence, and its prevalence in the United States (1 of 433 youths <20 years of age) has increased over the past couple of decades.[9] Non-Hispanic white youths are affected with T1DM more often than any other racial or ethnic group.[10]

Diabetes-related care accounts for 11 percent of all U.S. health care expenditures,[11] equating to $245 billion in total costs in 2012.[8] Average medical expenses are more than twice as high for a person with diabetes as they are for someone without diabetes.[12] When considering medical and productivity costs, some calculations provide even more extreme differentials, particularly in relation to T1DM: 2007 national costs per case were $2,864 for undiagnosed diabetes, $9,677 for diagnosed T2DM, and $14,856 for T1DM.[11] Complications from diabetes include cardiovascular disease, retinopathy, neuropathy, nephropathy, and cerebrovascular disease, as well as comorbidities such as depression and other mental health conditions.[13]

Diabetes Care and Self Management

The mainstay of treatment for T1DM is lifelong insulin therapy. In order to achieve optimal glycemic control, people with T1DM (and especially those on multiple-dose insulin or insulin pump therapy) should self-monitor their blood sugar levels frequently during the day and adjust their insulin dose, diet, and/or physical activity accordingly.[14] The benefit of intensive control of glycemia in reducing the incidence and progression of micro- and macrovascular complications was clearly demonstrated in the Diabetes Control and Complications Trial and its related longitudinal study.[15,16] Recently, these findings have extended to demonstrate reduced mortality.[17] Although these findings are promising, a meta-analysis of 12 trials (2,230 participants) of intensive versus conventional glucose control in T1DM confirmed the reduction in development but not progression of microvascular

complications, and stressed that the benefits should be weighed against the risks of severe hypoglycemia.[18]

People with T2DM are often managed progressively, with an initial focus on diet (e.g., medical nutrition therapy) and physical activity, subsequent addition of one or more oral hypoglycemic medications, and in many cases also use of insulin (or sole use of insulin) to obtain optimal blood glucose control. The importance of tight glycemic control for reducing the risk of microvascular complications in T2DM was shown in the United Kingdom Prospective Diabetes Study.[19,20] As with T1DM, though, a meta-analysis pooling results from 28 trials (34,912 participants) of intensive control in T2DM found no significant differences for all-cause mortality or cardiovascular deaths, or for macrovascular complications, including nonfatal myocardial infarction.[21]

Factors other than blood glucose control are important to address. Reducing the risk for diabetes-related complications in T1DM and T2DM often requires lifestyle and/or pharmacological management of body weight, blood pressure, and cholesterol levels.[14,22-24] For instance, intensive lowering of blood pressure in people with diabetes has been shown to reduce major cardiovascular events by 11 percent.[25] Lifestyle interventions targeted at weight loss, diabetes nutrition, and physical activity recommendations have been shown to be associated with weight control and improved glycemic control.[26-29] Additionally, findings from two large cross-national studies—the Diabetes, Attitudes, Wishes, and Needs (DAWN) studies—have demonstrated the need to address other outcomes of importance for patients, such as diabetes-related distress and depression.[30]

A critical element of diabetes care is education and support to enable patients to adopt and adhere to several self-care or self-management and lifestyle behaviors. Because knowledge acquisition alone is insufficient for behavioral changes,[26,31] the focus of many national and international guidelines and recommendations for DSME has shifted from traditional didactic educational services to more patient-centered methodologies incorporating interaction and problem-solving.[32-35] In addition, the national standards for DSME developed by the American Association of Diabetes Educators and the American Diabetes Association have incorporated the provision of ongoing diabetes self-management support "to encourage behavior change, the maintenance of healthy diabetes-related behaviors, and to address psychological concerns."[32] In addition to DSME, a diverse range of interventions and programs have been developed that focus on supporting patients' efforts in changing lifestyle behaviors in order to better manage glycemia and prevent complications.[27]

Despite the availability of new medications and devices (e.g., insulin pumps, continuous glucose monitoring), several standards for care management and DSME programs, and implementation of lifestyle interventions, the National Health and Nutrition Examination Survey found that 45 percent of adults with diabetes in the United States do not achieve glycemic targets.[36]

Rationale for Evidence Review

Health providers working in outpatient and primary care settings in the community struggle with how to best support, educate, and work with patients with diabetes to improve their disease control. To date, it is not clear whether there is (or what constitutes) a set of best practices associated with behavioral programs that can be implemented in the community health setting. For the purpose of this review, community health settings include ambulatory care (i.e., outpatient) clinics, primary care clinics, family physician clinics, and federally qualified health centers (i.e., Community Health Centers and Rural Health Centers).

Self-management and lifestyle interventions have been shown to improve glycemic control for T2DM to a clinically significant extent, at least in the short term;[37-44] the evidence for these programs in T1DM is less conclusive and based on older literature. Many previous systematic reviews on topics relevant to this review for T2DM have included studies evaluating a broad scope of interventions, some of them falling short of meeting current recommendations and others incorporating some enhancement of medical management that may confound the effects of the behavioral program. Many reviews have also included studies evaluating interventions targeted at a single behavior/component (e.g., diet) rather than multiple behaviors, as seems necessary for optimal disease self-management. Moreover, few reviews assessed factors contributing to the success of the interventions,[37,39,43,45,46] and even fewer analyzed the data in a manner that assessed multiple factors simultaneously:[45] the moderating effects of program content and characteristics have therefore not been fully investigated.

Our focus for T1DM was to determine the effectiveness of behavioral programs and for T2DM was to identify factors contributing to the effectiveness of multicomponent programs. We investigated a range of outcomes and conducted a network meta-analysis (enabling simultaneous assessment of multiple variables and a wide variety of comparisons) to analyze potential moderators of effectiveness, such as delivery personnel, effective community linkages, and demographic characteristics.

Because of our focus on moderation of effectiveness for T2DM, we did not examine harms, as we did for T1DM. This review provides information regarding the effectiveness and harms of behavioral programs (T1DM) and the combination of program components and delivery methods that is most effective for implementation of these programs in community health settings (T2DM).

Scope and Key Questions

For the purpose of this review we developed an operational definition of behavioral programs that encompasses DSME (without or with an additional clinical, psychosocial, or behavioral support phase—i.e., "DSME plus support"), as well as other programs incorporating interactive components that target multiple important behavioral changes (e.g., diet and physical activity). A commonality of all programs was that they incorporated one or more behavior change techniques,[47] with or without explicit use of a theory or model of behavior change. Our operational definition of a behavioral program is as follows:

> An organized, multicomponent diabetes-specific program with repeated interactions by one or more trained individuals, with a duration of ≥4 weeks, to improve disease control and/or patient health outcomes, and consisting of at least one of the following: (a) DSME; (b) a structured dietary intervention (related to any of the following: weight loss, glycemic control, or reducing risk for complications) together with one or more additional components; or (c) a structured exercise or physical activity intervention together with one or more additional components. Additional components for (b) and (c) may include interventions related to diet or physical activity; behavioral change (including but not limited to goal-setting, problem-solving, motivational interviewing, coping-skills training, cognitive behavioral therapy strategies); relaxation or stress reduction; blood glucose regulation; medication adherence; or self-monitoring for diabetic complications (foot, eye, and renal tests).

We addressed the following six Key Questions (KQs):

Key Question 1. For patients with T1DM, are behavioral programs implemented in a community health setting effective compared with usual or standard care, or active comparators in—

a.   Improving behavioral, clinical, and health outcomes?

b.   Improving diabetes-related health care utilization?

c. Achieving program acceptability as measured by participant attrition rates?

Key Question 2. For patients with T1DM, do behavioral programs implemented in the community health setting differ in effectiveness for behavioral, clinical, and health outcomes; their effect on diabetes-related health care utilization; or program acceptability for the following subgroups of patients?

a. Age—children and adolescents (≤18 years) and their families, young adults (19–30 years), adults (31–64 years), older adults (≥65 years)

b. Race or ethnicity

c. Socioeconomic status (e.g., family income, education level, literacy)

d. Time since diagnosis (≤1 year vs. >1 year)

e. Baseline level of glycemic control (HbA$_{1c}$ <7% vs. ≥7%)

Key Question 3. For patients with T1DM, does the effectiveness of behavioral programs differ based on the following factors?

a. Program components

b. Intensity (i.e., program duration, frequency/periodicity of interactions)

c. Delivery personnel (e.g., dietitian, exercise specialist, physician, nurse practitioner, certified diabetes educator, lay health worker)

d. Method of communication (e.g., individual vs. group, face to face, interactive behavior change technology, social media)

e. Degree of tailoring based on needs assessment (e.g., educational/behavioral deficits, age or other demographics, readiness to change)

f. Level and nature of community engagement

Key Question 4. For patients with T1DM, what are the associated harms (i.e., activity-related injury) of behavioral programs implemented in a community health setting compared with usual care, standard care, or active comparators?

Key Question 5. Among behavioral programs targeted at adults with T2DM implemented in a community health setting, what factors contribute to (a) their effectiveness for behavioral, clinical, and health outcomes; (b) their effect on diabetes-related health care utilization; and (c) program acceptability as measured by participant attrition rates? Factors include the following:

a. Program components

b. Program intensity

c. Delivery personnel

d. Methods of delivery and communication

e. Degree of tailoring

f. Community engagement

Key Question 6. Do the factors that contribute to program effectiveness for patients with T2DM vary across the following subpopulations?

a. Age—young adults (19–30 years), adults (31–64 years), older adults (≥65 years)

b. Race or ethnicity

c. Socioeconomic status (e.g., family income, education level, literacy)

d. Time since diagnosis (≤1 year vs. >1 year)

e. Baseline level of glycemic control (HbA$_{1c}$ <7% vs. ≥7%)
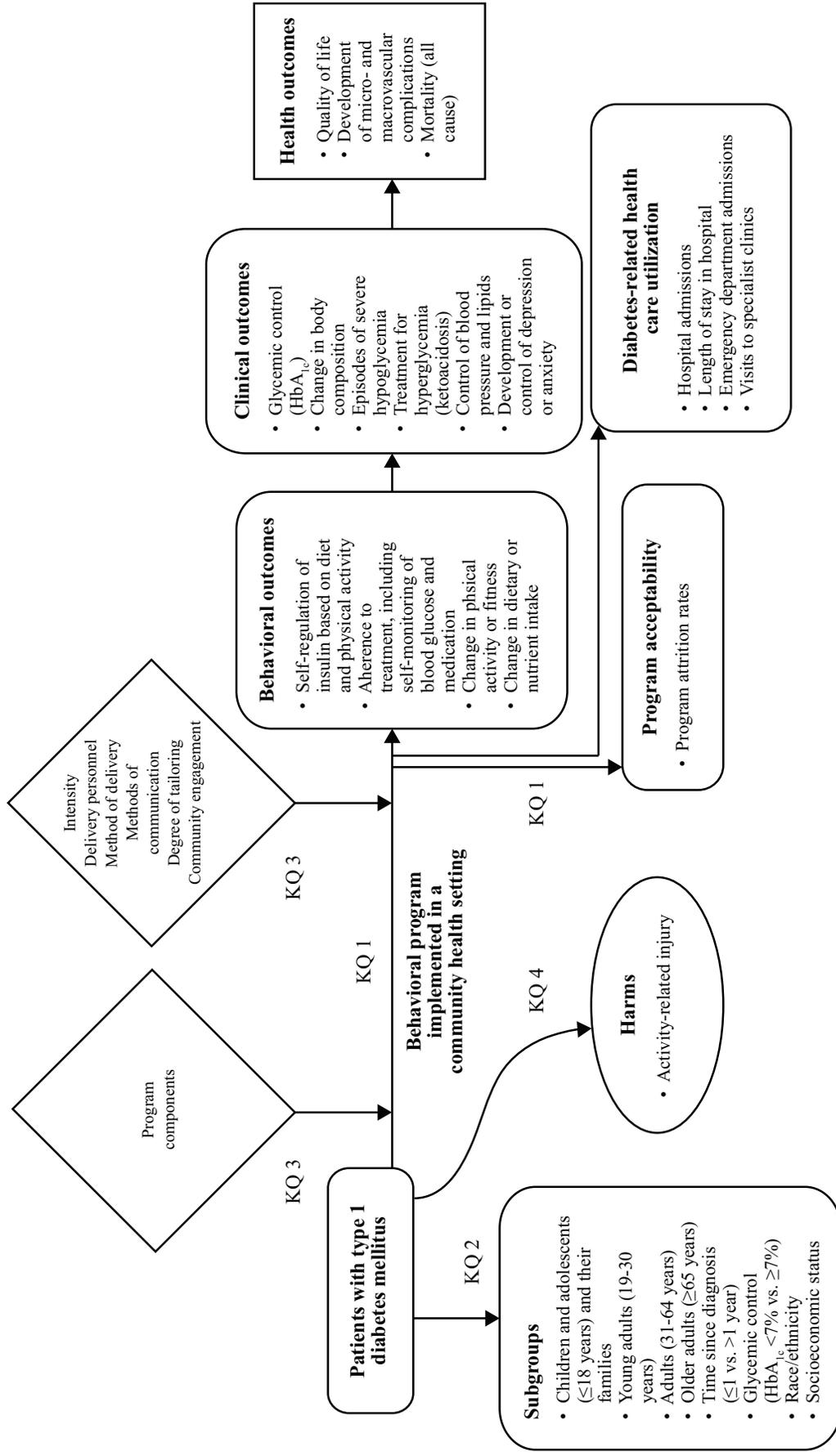
Analytical Frameworks

We developed two analytic frameworks to guide the systematic review process and specific KQs for T1DM and T2DM (Figure A and Figure B, respectively). The figures illustrate the populations of interest and the outcomes that we reviewed.

Methods

Literature Search Strategy

We used the same approach and search strategies for T1DM and T2DM. Our research librarian searched the following bibliographic databases from 1993 to May 2014: Ovid MEDLINE® and Ovid MEDLINE® In-Process & Other Non-Indexed Citations, Cochrane Central Register of Controlled Trials via Cochrane Library, Embase® via Ovid, CINAHL Plus with Full Text via EBSCOhost, PsycINFO® via Ovid, and PubMed® via the National Center for Biotechnology Information Databases. We limited the search to prospective controlled studies published in English. On January 15, 2015, we performed a search update in all databases except Embase, from which none of the previously included studies was exclusively obtained. We reviewed the reference lists of relevant systematic reviews and of all included studies. We searched ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform. We searched the conference proceedings (2011–14) from the American Diabetes Association, American Association of Diabetes

Figure A. Analytic framework for behavioral programs for type 1 diabetes mellitus

Patients with type 1 diabetes mellitus

Program components

Intensity
Delivery personnel
Method of delivery
Methods of communication
Degree of tailoring
Community engagement

KQ 3

KQ 1

KQ 3

Behavioral program implemented in a community health setting

KQ 2

Subgroups
- Children and adolescents (≤18 years) and their families
- Young adults (19-30 years)
- Adults (31-64 years)
- Older adults (≥65 years)
- Time since diagnosis (<1 vs. >1 year)
- Glycemic control (HbA$_{1c}$ <7% vs. ≥7%)
- Race/ethnicity
- Socioeconomic status

KQ 4

Harms
- Activity-related injury

KQ 1

Behavioral outcomes
- Self-regulation of insulin based on diet and physical activity
- Aherence to treatment, including self-monitoring of blood glucose and medication
- Change in phsical activity or fitness
- Change in dietary or nutrient intake

Program acceptability
- Program attrition rates

Clinical outcomes
- Glycemic control (HbA$_{1c}$)
- Change in body composition
- Episodes of severe hypoglycemia
- Treatment for hyperglycemia (ketoacidosis)
- Control of blood pressure and lipids
- Development or control of depression or anxiety

Diabetes-related health care utilization
- Hospital admissions
- Length of stay in hospital
- Emergency department admissions
- Visits to specialist clinics

Health outcomes
- Quality of life
- Development of micro- and macrovascular complications
- Mortality (all cause)

HbA1c = hemoglobin A$_{1c}$; KQ = Key Question

5

Figure B. Analytic framework for behavioral programs for type 2 diabetes mellitus
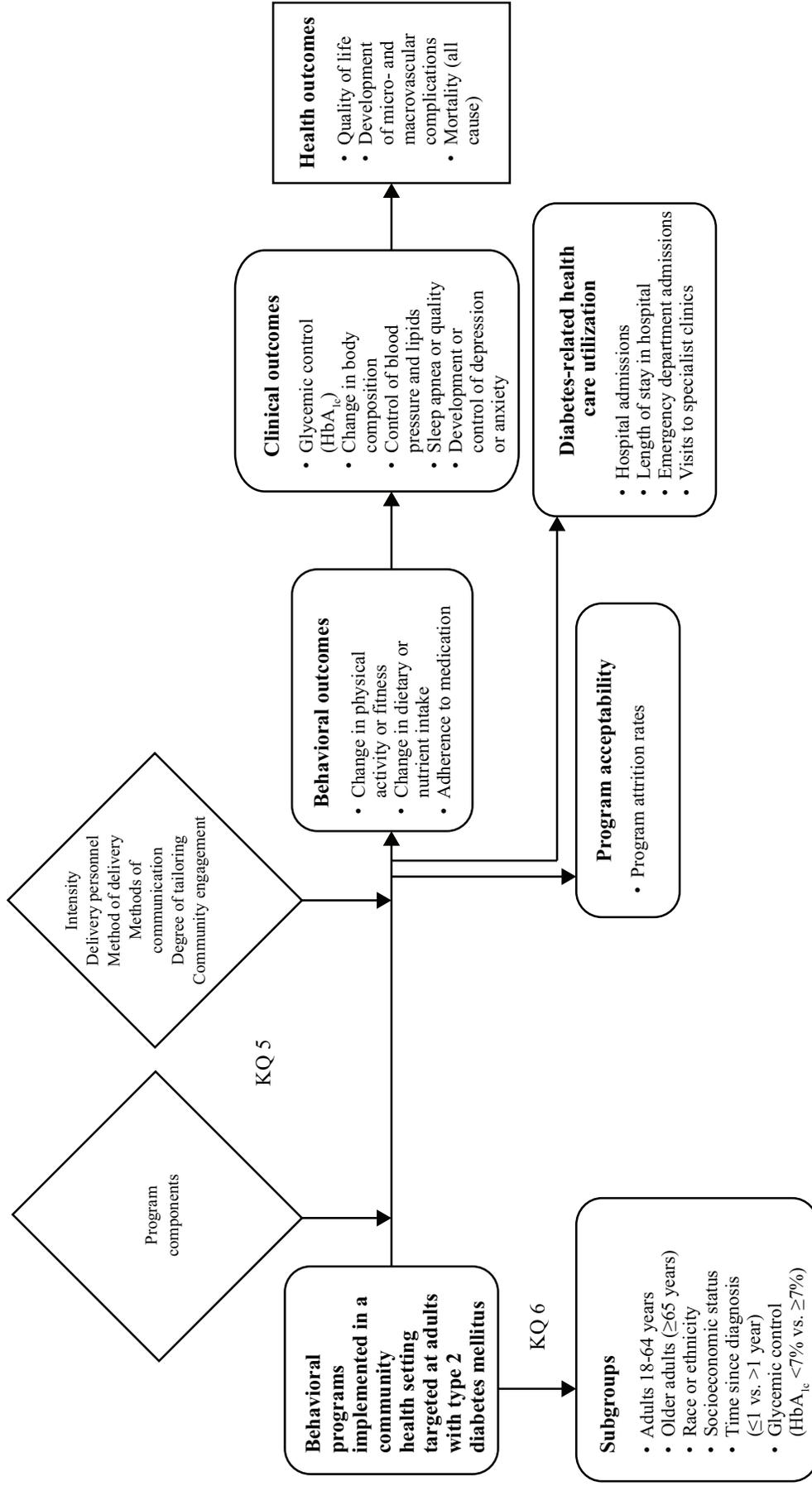
HbA1c = hemoglobin A_{1c}; KQ = Key Question

Educators, National Institute of Diabetes and Digestive and Kidney Diseases, Canadian Diabetes Association, European Association for the Study of Diabetes, International Diabetes Federation, Society of Behavioral Medicine, and International Society for Behavioral Nutrition and Physical Activity.

Eligibility Criteria

The research team developed eligibility criteria with respect to populations, interventions, comparators, outcomes, timing, and setting (PICOTS). For both T1DM and T2DM, we included studies conducted in the United States or other highly developed countries[48] and published in the English language on or after 1993. The publication date limit was chosen because of changes to usual care/medical management (the comparator in most cases in this review) resulting from the findings of landmark trials published from 1993 onward.[15,24,49] For T1DM, we included prospective comparative studies—i.e., randomized controlled trials (RCTs), nonrandomized controlled trials (non-RCTs), prospective cohort studies, and controlled before-after studies. For T2DM, we included RCTs.

For T1DM, we included studies of patients (any age) diagnosed with T1DM who had undergone basic diabetes education. For T2DM, we included studies of adults with T2DM who had undergone basic diabetes education.

For behavioral programs, we included studies of interventions that met the criteria included in our operational definition. The comparators were usual care (i.e., usual medical management provided to all participants), an active comparator (i.e., an intervention not meeting our definition of a behavioral program, such as basic education or a dietary or physical activity intervention), or another behavioral program. When two or more behavioral programs were compared, we considered this an evaluation of comparative effectiveness.

Study Selection

Two reviewers independently screened all titles and abstracts using broad inclusion criteria. We retrieved the full text of any publications marked for inclusion by either reviewer. Two reviewers independently assessed the full texts using a priori inclusion criteria and a standard form. We resolved disagreements by consensus or consulting a third member of the review team.

Risk of Bias

Two reviewers independently assessed the risk of bias of included studies. Discrepancies were resolved through discussion and consensus. We assessed the internal validity of RCTs and non-RCTs using the Cochrane Risk of Bias tool.[50] The tool examines seven domains of potential

bias (sequence generation, concealment of allocation, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting, and "other" sources of data) and is used to categorize the overall risk of bias. Each domain was rated as having low, medium, or high risk of bias.

We assessed the risk of bias for prospective cohort studies and controlled before-after studies using the Newcastle-Ottawa Scale.[51] This tool uses a star system to assess methodological quality across three categories: selection of participants, comparability of study groups, and ascertainment of the outcome of interest. The star rating indicates the quality of a study, with a maximum assessment of nine.

Data Extraction

We used structured data extraction forms to gather pertinent information, including characteristics of study populations, settings, interventions, comparators, and outcomes; study designs; and methods. We extracted data directly into the Systematic Review Data Repository™ (http://srdr.ahrq.gov/).[52] One reviewer extracted data, and a second reviewer checked the data for accuracy and completeness. We resolved disagreements through consensus or by consulting a third member of the review team.

Data Synthesis

We analyzed data separately for T1DM and T2DM, with different approaches for each KQ. For each condition we summarized the characteristics of included studies qualitatively and presented important features of the study populations, interventions, and comparators in summary tables. All outcome data were extracted and reported in figures of meta-analyses (if pooled) or in outcomes tables. We extracted and analyzed data from different postintervention followup timepoints: end of intervention to ≤1 month postintervention, >1 month to ≤6 months, >6 months to ≤12 months, >12 months to ≤24 months, and >24 months.

We focused on the following key outcomes: HbA_{1c}, quality of life, development of micro- and macrovascular complications, all-cause mortality, adherence to diabetes self-management behaviors, change in body composition, change in physical activity or fitness, and change in dietary or nutrient intake. To enable interpretation of the results in terms of clinical significance and the precision of the effect sizes during assessment of the strength of the body of evidence for our key outcomes, discussed later, we defined a threshold for clinical importance when there was literature to provide guidance. For HbA_{1c}, we used a difference of 0.4 percent (e.g., 7.6% vs. 8.0%).[53] For quality-of-life measures and other patient-reported

outcomes represented by continuous data, we used a difference of one-half standard deviation (SD)—i.e., 0.50 standardized mean difference (SMD)—based on the mean SD from the pooled studies, which has been shown to represent a universal conservative estimate of a meaningful difference.[54,55] For adherence to self-management behaviors, we did not apply a threshold for clinical importance because of poor reporting of the scoring and unknown meaning of a threshold for an optimal number of self-monitoring tests (the most common reporting for this outcome).

With input from our Technical Expert Panel, we categorized various components and implementation methods, as outlined in Table A. Many behavioral programs comprised DSME with or without the addition of a support component (i.e., DSME + support); we separated these into two categories to recognize that the support phase was often of a lower intensity (e.g., less frequent contacts) and focused on different content, such as psychosocial support, as compared with the DSME phase. Programs not considered DSME were considered "lifestyle" programs.

Synthesis for T1DM (KQs 1–4)

For each comparison of interest, we conducted a pairwise meta-analysis when two or more eligible trials were sufficiently similar on the basis of study design and clinical homogeneity. We present both pooled and subgroup analysis based on age when there was more than one trial in each age category at any timepoint. We used the Hartung-Knapp-Sidik-Jonkman random-effects model[56,57] for all meta-analyses and used Stata 11.2 and Excel 2010 software. We calculated pooled mean differences (MDs), SMDs, and risk ratios (RRs) with corresponding 95% confidence intervals (CIs), as appropriate, and weighted by sample size and variance. We analyzed outcomes at different postintervention timepoints.

For KQ 2, we searched for subgroup analyses reported by individual trials that focused on whether a particular behavioral program was more or less effective for the outcome reported by the most studies (i.e., HbA_{1c}) based on variables of interest. (See Figure A.) We also compared subgroups of studies—for example, when the mean age of participants fell within one of the age categories.

To assess whether the effectiveness of behavioral programs differed based on various program factors (KQ 3), we performed univariate metaregressions for comparisons between behavioral programs and usual care for HbA_{1c} from each study's longest followup timepoint. Each

behavioral program was coded using the categorization scheme in Table A, and these variables were used in the analysis. For KQ 4, harms (i.e., activity-related injury), we planned to descriptively summarize all outcomes presented in studies.

Synthesis for T2DM (KQs 5 and 6)

Before synthesizing findings to answer KQs 5 and 6, we performed pairwise meta-analyses for all outcomes identified in the PICOTS using the same analytical approach described for KQ 1. To answer KQs 5 and 6, we performed network meta-analyses for key outcomes reported by the most studies (HbA_{1c} and BMI). A network meta-analysis allows for simultaneous evaluation of a suite of comparisons, and considers both direct and indirect evidence while preserving the within-study randomization. A network of different comparisons is constructed (with "nodes" representing groupings of sufficiently similar interventions and comparators). To assess the effectiveness of programs based on different combinations of moderator variables, we grouped the behavioral programs into nodes after coding them in terms of the program components and implementation factors described in Table A. We also formed three categories for the comparator groups: usual care, active "non-DSME education" control (i.e., basic education not meeting our criteria for DSME), and active "other" control (e.g., stand-alone dietary or physical activity interventions). The analysis was conducted using a Bayesian network model. Results are presented as estimates of the treatment effects (MDs) relative to usual care with 95-percent credibility intervals, as well as the rank probabilities for each behavioral program strategy (e.g., probability that a particular combination of components and delivery methods for a behavioral program is the "best program").

KQ 6 focused on whether variability between population groups affected the role of potential factors contributing to effectiveness of behavioral programs for the key outcome with the most data (i.e., HbA_{1c}). We first conducted subgroup analyses of the pairwise meta-analysis results for HbA_{1c} for behavioral programs compared with usual care and active controls at longest followup; subgroup analyses based on between-study baseline glycemic control (HbA_{1c}), age, and ethnicity were performed. For baseline glycemic control and age, we then performed subgroup analysis of the network meta-analysis used for KQ 5 using only studies in which participants had suboptimal baseline glycemic control (>7% HbA_{1c}), or were under 65 years of age. For subgroups based on race/ethnicity (≥75% vs. <75% percent

Table A. Categorization of program components and implementation factors

Program Factors	Categories and Description Variables
Program components[a]	1. DSME 2. DSME + support: DSME plus an added phase to extend program duration and support; often clinically focused but may be psychosocial, educational, or behavioral 3. Lifestyle programs: Behavioral programs focused on diet and/or physical activity rather than on diabetes-specific self-management behaviors; may also include other components as long as program does not meet the criteria for DSME with emphasis on education/training
Duration of program	No categories; duration was used as a continuous variable for the regression analyses for KQs 3 and 6
Intensity[a] (contact hours; where contact hours could not be calculated, we used number of contacts as a proxy)	1. ≤10 hours[b] 2. 11 to 26 hours (e.g., weekly for up to 6 months) 3. ≥27 hours (allowing for monthly followup for 1 year)
Frequency of contacts	No categories; this was a composite variable combining duration and intensity (hours/month); the continuous variable was used for the regression analyses for T1DM
Method of communication[c]	1. In person only 2. Mixture of in person and technology 3. All technology with minimal interaction with providers
Method of delivery[d]	1. Individual 2. Mixed individual and group 3. Group
Delivery personnel[e]	1. Delivered entirely by non–health professional (e.g., lay/community health worker, undergraduate student) after training and under some supervision 2. One health professional for large majority (>75%) of delivery 3. Provision by multidisciplinary team of health professionals
Degree of tailoring[f]	1. None/minimal—no tailoring or only small portion is tailored (e.g., personalized diet prescription in otherwise highly structured lifestyle program or delivery based on flexible hours but same content for all) 2. Moderate/maximum—most of program has content and/or delivery tailoring (e.g., topics are based on needs assessment, and delivery timing/duration/location is based on participant's schedule/needs/location preferences)
Level and nature of community engagement	1. Present—e.g., peer delivery of program or peer support groups for support stage, use of community resources (infrastructure) for delivery or maintenance stages 2. Absent—e.g., nothing reported or, at most, providing written information about community resources
Presence of support person[g]	1. Family or parent involved in >1 session 2. No family or parent involvement in sessions

DSME = diabetes self-management education; KQ = Key Question; T1DM = type 1 diabetes mellitus
[a]In analyses for KQ 5 and 6 only.
[b]Based on the current number of hours billable for patients eligible for public health care administered by the Centers for Medicare & Medicaid Services in the United States (described by Technical Expert Panel as a practical limitation on implementing programs having higher intensity).
[c]2 and 3 were combined for analysis.
[d]1 and 2 were combined for analysis.
[e]2 and 3 were combined for KQs 5 and 6.
[f]Used in summary tables and the analysis for T1DM.
[g]For T1DM only.

nonwhite and/or Hispanic), the number of trials in either subgroup was not sufficient to perform a meaningful network meta-analysis (i.e., the number of studies in each node would be very low, thus limiting the validity of this method), so we conducted a set of univariate metaregressions using the variables in Table A and methods outlined for KQ 2. All of our results for this KQ relied on between-study rather than within-study comparisons, such that the effect of randomization is removed and the results are considered observational and possibly biased through confounding by other study-level characteristics.

Strength of the Body of Evidence

We followed the "Methods Guide for Effectiveness and Comparative Effectiveness Reviews" (Methods Guide)[58] to evaluate the strength of evidence (SOE) for KQ 1 for all health outcomes (i.e., quality of life, development of micro- and macrovascular complications, all-cause mortality) and selected behavioral and clinical outcomes (i.e., glycemic control, adherence to diabetes self-management behaviors, change in body composition, change in physical activity or fitness, and change in dietary or nutrient intake). For KQ 2, we assessed SOE for HbA_{1c}, which was the outcome reported by the most studies and thus the focus of this KQ. SOE assessments were based on evidence from trials. The body of evidence was graded by one reviewer and reviewed by a second reviewer. Disagreements were resolved through discussion or by consulting with a third reviewer, as needed.

For each outcome, we assessed five major domains of most relevance to reviews of RCTs (anticipated to be the large majority of included studies): risk of bias (rated as low, medium, or high), consistency (rated as consistent, inconsistent, or unknown), directness (rated as direct or indirect), precision (rated as precise or imprecise), and reporting bias (rated as suspected or not suspected). A precise estimate is one that allows for a clinically useful conclusion. The overall SOE was graded as high, moderate, low, or insufficient. High, moderate, and low SOE reflect the confidence we have in the effect estimate and the likelihood that the estimate will change with further research. Insufficient SOE implies that we are unable to estimate an effect, that we had no or very little evidence, or that the 95% CI included clinically important effects both for and against behavioral programs.

Applicability

We assessed applicability of the body of evidence following guidance from the Methods Guide.[58] We used the PICOTS framework to explore factors that may affect applicability.

Results

Our database and gray literature searches identified 47,141 citations, and 11 additional records were identified from reference lists of systematic reviews and included studies. For T1DM, we included 34 studies described in 44 publications. For T2DM, we included 132 studies described in 161 publications. Figure C describes the flow of literature through the screening process.

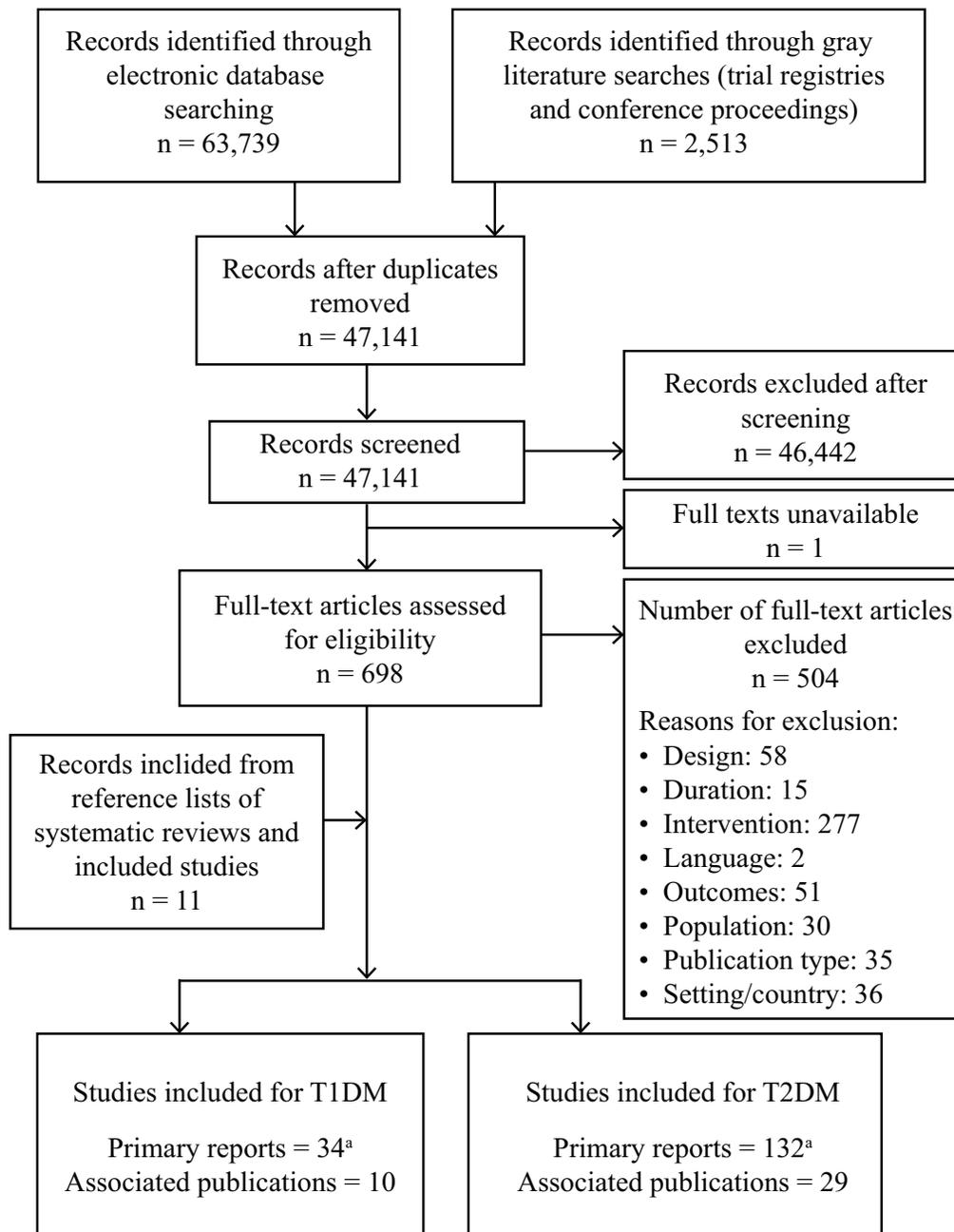## T1DM: Description and Risk of Bias of Studies

Twenty-five studies were conducted in children and adolescents; nine were conducted in adults. Most trials were two-arm trials comparing DSME with usual care. For most studies (70%), the mean HbA_{1c} was 8.5 percent or higher. For studies targeting children and adolescents, the mean age across most studies ranged from 12 to 15 years; because of this, we refer to the included studies as being conducted in "youths." For studies targeting adults, the mean age ranged from 30 to 49 years. No studies specifically targeted older adults (≥65 years). The mean duration of diabetes ranged from 2.7 to 7.3 years among studies that targeted youths and from 2.5 to 23 years for those targeting adults.

The total duration of the behavioral programs for youths ranged from 1.2 to 25 months (median = 5.6 months). The number of contact hours ranged from 1 to 48 hours (median = 9.5 hours). Five trials delivered the programs to youths only; 16 delivered the programs to both youths and their parents or family members. There was a mixture of delivery to individuals and to groups, and programs were delivered by a variety of personnel, with seven trials not using health care professionals.

In studies on adults, the total duration of the behavioral programs ranged from 1.5 to 12 months (median = 6 months), and the number of contact hours ranged from 9 to 52 hours (median = 16 hours). There was a mixture of individual and group formats. All trials were provided by health care professionals; one used a peer who served as coleader.

All trials were assessed as having either moderate or high overall risk of bias. For objective outcomes (i.e., HbA_{1c}), 58 percent of trials had a medium risk of bias and 42 percent had a high risk. The assessment of high risk was largely driven by incomplete outcome data (i.e., loss to followup). For trials reporting subjective outcomes of interest to this review (e.g., health-related quality of life [HRQL], patient-reported self-management behaviors), all but one trial had a high risk of bias (95%), primarily because of lack of blinding of participants, study personnel, and outcome assessors.

Figure C. Flow diagram of study retrieval and selection

```
┌─────────────────────────────┐        ┌─────────────────────────────┐
│ Records identified through  │        │ Records identified through  │
│   electronic database       │        │ gray literature searches    │
│      searching              │        │ (trial registries and       │
│     n = 63,739              │        │ conference proceedings)     │
│                             │        │     n = 2,513               │
└─────────────────────────────┘        └─────────────────────────────┘
                 │                                   │
                 └──────────────┬────────────────────┘
                                ▼
                 ┌─────────────────────────────┐
                 │ Records after duplicates    │
                 │        removed              │
                 │      n = 47,141             │
                 └─────────────────────────────┘
                                │
                                ▼
                 ┌─────────────────────────────┐      ┌─────────────────────────────┐
                 │    Records screened         │─────▶│ Records excluded after       │
                 │      n = 47,141             │      │       screening              │
                 └─────────────────────────────┘      │     n = 46,442               │
                                │                      └─────────────────────────────┘
                                │                      ┌─────────────────────────────┐
                                ├─────────────────────▶│ Full texts unavailable       │
                                │                      │         n = 1                │
                                ▼                      └─────────────────────────────┘
                 ┌─────────────────────────────┐      ┌─────────────────────────────┐
                 │ Full-text articles assessed │─────▶│ Number of full-text articles │
                 │      for eligibility        │      │        excluded              │
                 │        n = 698              │      │        n = 504               │
                 └─────────────────────────────┘      └─────────────────────────────┘
```

Reasons for exclusion:
- Design: 58
- Duration: 15
- Intervention: 277
- Language: 2
- Outcomes: 51
- Population: 30
- Publication type: 35
- Setting/country: 36

Records included from reference lists of systematic reviews and included studies
n = 11

Studies included for T1DM

Primary reports = 34[a]
Associated publications = 10

Studies included for T2DM

Primary reports = 132[a]
Associated publications = 29

T1DM = type 1 diabetes mellitus; T2DM = type 2 diabetes mellitus
[a]One study was included for both T1DM and T2DM.

T1DM: Results for KQs 1–4

A summary of the key findings and SOE assessments for behavioral programs compared with usual care and active controls are presented in Tables B and C, respectively.

When comparing behavioral programs with usual care, there was moderate SOE showing reduction in HbA_{1c} at 6-month postintervention followup, with percent HbA_{1c} reduced by 0.31. This result failed to reach our threshold of clinical significance of a change by 0.4 percent HbA_{1c}.

For all other timepoints, there was no significant difference in HbA_{1c}; the SOE was low because of risk of bias and imprecise effect estimates. For followup timepoints of 12 months or longer, the 95% CIs included our threshold for clinical importance such that we cannot rule out benefit for behavioral programs based on the available evidence. For individuals who were enrolled in behavioral programs compared with those receiving an active control, there was moderate SOE showing a statistically significant and clinically important reduction in percent HbA_{1c} of

Table B. Type 1 diabetes: summary of key findings and strength of evidence for behavioral programs compared with usual care

Outcome	Outcome Timing	# Trials (# Subjects); Tool if Applicable	Mean Difference or Standardized Mean Difference	Strength of Evidence
HbA_{1c}	EOI	16 (1,155)	MD, -0.11; 95% CI, -0.33 to 0.11[a]	Low for no significant difference
HbA_{1c}	6m followup	12 (1,463)	MD, -0.31; 95% CI, -0.47 to -0.15	Moderate for benefit[b]
HbA_{1c}	12m followup	7 (1,333)	MD, -0.22; 95% CI, -0.49 to 0.05	Low for no significant difference
HbA_{1c}	≥12m followup	4 (1,138)	MD, -0.40; 95% CI, -0.92 to 0.12 (>12m to <24m) MD, -0.08; 95% CI, -1.96 to 1.8 (≥24m)	Low for no significant difference
Adherence to diabetes self-management	EOI	4 (282); SMBG 1 (74); SDSCA 1 (54); DSMP 1 (74); DSCI	MD, 0.15; 95% CI, -0.54 to 0.84 MD, 1.4 days; 95% CI, 0.35 to 2.43 MD, 5.00; 95% CI, 0.60 to 9.40 MD, 0.22; 95% CI, -0.60 to 1.04	Low for no significant difference
Adherence to diabetes self-management	6m followup	5 (252); SMBG 1 (244); SDSCA 2 (471); DSMP	MD, 0.40; 95% CI, -0.36 to 1.16 MD, -0.06; 95% CI, -0.60 to 0.48 No difference (different measures)	Low for no significant difference
Adherence to diabetes self-management	12m followup	1 (54); DSMP 1 (180); skipping 1 or more doses in past month	MD, 4.00; 95% CI, -1.69 to 9.69 OR, 0.82; 95% CI, 0.48 to 0.1.38	Insufficient
Adherence to diabetes self-management	>12m followup	1 (390); SMBG 1 (190); skipping 1 or more doses in past month	MD, -0.36; 95% CI, -0.69 to -0.03 (≥24m) OR, 1.30; 95% CI, 0.78 to 2.17 (24m)	Insufficient
Change in body composition (BMI [kg.m^{-2}])	EOI	1 (60)	MD, 0.08; 95% CI, -0.35 to 0.51	Insufficient
Change in body composition (BMI [kg.m^{-2}])	6m followup	1 (227)	MD, -0.21; 95% CI, -0.62 to 0.20	Insufficient
Change in body composition (kg)	EOI	1 (61)	MD, -0.50; 95% CI, -5.69 to 4.69	Insufficient
Change in physical activity (fitness, VO$_2$ max)	EOI	1 (43)	MD, 0.59; 95% CI, 0.22 to 0.96	Insufficient

Table B. Type 1 diabetes: summary of key findings and strength of evidence for behavioral programs compared with usual care (continued)

Outcome	Outcome Timing	# Trials (# Subjects); Tool if Applicable	Mean Difference or Standardized Mean Difference	Strength of Evidence
Change in physical activity (intensity/ duration)	EOI	2 (91)	SMD, 0.16; 95% CI, -0.25 to 0.57	Insufficient
Change in physical activity (intensity/ duration)	6m followup	2 (272)	SMD, -0.26; 95% CI, -1.00 to 0.49	Insufficient
Change in dietary or nutrient intake (energy [kcal/ day])	EOI	1 (61)	MD, -247.10; 95% CI, -281.7 to -212.5	Insufficient
Change in dietary or nutrient intake (% saturated fat)	EOI	1 (61)	MD, -1.80; 95% CI, -3.53 to -0.07	Insufficient
Generic HRQL	EOI	7 (474)	SMD, 0.10; 95% CI, -0.18 to 0.38	Moderate for no difference
Generic HRQL	6m followup	1 (53)	SMD, -0.29; 95% CI, -0.83 to 0.26	Insufficient
Generic HRQL	12m followup	2 (405)	SMD, 0.02; 95% CI, -0.11 to 0.15	Insufficient
Generic HRQL	≥12m followup	1 (291)	SMD, -0.04; 95% CI, -0.27 to 0.19	Insufficient
Diabetes-specific quality of life	EOI	3 (212)	SMD, 0.08; 95% CI, -1.44 to 1.60	Insufficient
Diabetes distress	EOI	4 (209)	SMD, -0.31; 95% CI, -0.83 to 0.21	Low for no significant difference
Diabetes distress	6m followup	4 (236)	SMD, -0.28; 95% CI, -0.94 to 0.38	Low for no significant difference

BMI = body mass index; CI = confidence interval; DSCI = Diabetes Self-Care Inventory (scale not reported; higher scores better); DSMP = Diabetes Self-Management Profile (scale not reported; higher scores better); EOI = end of intervention to ≤1 month postintervention followup (interventions 1.5–25 months); HbA1c = hemoglobin A1c; HRQL = health-related quality of life; m = month; MD = mean difference; OR = odds ratio; SDSCA = Summary of Diabetes Self-Care Activities (days per week adhering to self-management behaviors); SMBG = self-monitoring of blood glucose (frequency; tests per day); SMD = standardized mean difference; VO_2max = maximal oxygen uptake

[a]Negative values for MDs or SMDs are favorable for HbA1c, change in body composition, change in dietary intake, and diabetes distress.

[b]This point estimate did not meet the threshold for clinical significance, although the 95% CI included a clinically important difference.

Table C. Type 1 diabetes: summary of key findings and strength of evidence for behavioral programs compared with an active control

Outcome	Outcome Timing	# Trials (# Subjects); Tool if Applicable	Mean Difference	Strength of Evidence
HbA$_{1c}$	EOI	4 (566)	MD, -0.32; 95% CI, -0.78 to 0.14[a]	Low for no significant difference
HbA$_{1c}$	6m followup	4 (504)	MD, -0.43; 95% CI, -0.62 to -0.24	Moderate for benefit
HbA$_{1c}$	12m followup	3 (342)	MD, -0.34; 95% CI, -0.71 to 0.03	Low for no significant difference
Adherence to diabetes self-management	EOI	1 (54); DSMP 1 (149); DBRS	MD, 2.40; 95% CI, -2.46 to 7.26 No data reported; those in behavioral program did more poorly	Insufficient
Adherence to diabetes self-management	6m followup	1 (149); SMBG 1 (149); DBRS	MD, -0.20; 95% CI, -0.76 to 0.36 No data reported; those in behavioral program did more poorly	Insufficient
Adherence to diabetes self-management	12m followup	1 (54); DSMP 1 (149); DBRS	MD, 2.00; 95% CI, -3.78 to 7.78 No data reported; those in behavioral program did more poorly	Insufficient

CI = confidence interval; DBRS = Diabetes Behavior Rating Scale (scale not reported; higher scores better); DSMP = Diabetes Self-Management Profile (scale not reported; higher scores better); EOI = end of intervention; HbA1c = hemoglobin A1c; m = month; MD = mean difference; SMBG = self-monitoring of blood glucose (frequency; tests per day)
[a]Negative values for MDs are favorable for HbA$_{1c}$.

0.44 at 6-month postintervention followup. There was no difference in HbA$_{1c}$ at other timepoints; however, the SOE was low and we cannot rule out a benefit for behavioral programs.

There was low SOE showing no difference in adherence to diabetes self-management behaviors (i.e., frequency of blood glucose checks or overall self-management) at end of intervention and 6-month followup for comparisons with usual care; for comparisons with active controls, there was insufficient SOE for this outcome at any followup timepoint. For participants receiving behavioral programs compared with usual care, there was no difference in generic HRQL at the end of intervention (moderate SOE). Few trials reported on generic HRQL at longer followup timepoints. In comparisons with usual care, there was insufficient SOE to assess whether there was any effect on diabetes-specific HRQL at any timepoint, and low SOE of no difference for diabetes distress at end of intervention and 6-month followup. The 95% CIs for diabetes distress included our threshold for clinical importance such that we cannot rule out a favorable effect for behavioral programs. There were no data on HRQL for comparisons

of behavioral programs with active controls. Few trials reported on symptoms of depression or on episodes of severe hypo- or hyperglycemia. No trials reported on micro- and macrovascular complications or on all-cause mortality.

Few trials reported on the number of diabetes-related hospital admissions or emergency department admissions. Behavioral programs appear to be acceptable to patients with T1DM; our meta-analysis found a 21-percent higher risk of attrition for individuals receiving usual care compared with those receiving the behavioral program.

For KQ 2, we examined the differential effect of patient characteristics on the effectiveness of behavioral programs for T1DM. In comparisons with usual care, results for the subgroups of studies in adults and in youths were consistent with the results when looking at all studies combined for KQ 1. At 6 months, behavioral programs reduced HbA$_{1c}$ in studies of youths by a statistically significant 0.28 percent and in studies of adults by a non–statistically significant 0.38 percent. At end of intervention, the point estimates indicated greater benefit in the adult subgroup (0.28) than in the youth subgroup

(0.00), although neither of these values reached statistical significance. None of the point estimates exceeded the clinically important difference of 0.4 percent HbA_{1c}, which was established a priori. In the comparisons with active controls based on age of study participants, the small number of studies in most subgroups provided insufficient SOE.

One trial reported results separately for youths with baseline $HbA_{1c} \geq 8$ percent and found favorable results for this subgroup; no other subgroup analysis was conducted because the majority of trials enrolled participants with poor control ($HbA_{1c} >8.5\%$). No trials reported on HbA_{1c} by race or ethnicity, socioeconomic status, or time since diagnosis.

For KQ 3, our univariate metaregressions did not find any statistically significant differences for moderation by any program factor. Examining the coefficients (e.g., change in HbA_{1c} from switching from one category to another or adding an increment in a continuous variable such as program hours) and their 95% CIs suggested that program intensity (duration, contact hours, frequency of contacts) did not influence effectiveness, and that individual (vs. group) delivery was beneficial. No studies reported on the associated harms (i.e., activity-related injury) of behavioral programs (KQ 4).

T2DM: Description and Risk of Bias of Studies

The majority of RCTs were two-arm trials, with many comparing DSME with usual care (55 trials) or an active control (7 trials); 16 three- or four-arm trials were included, as were several trials comparing two different behavioral programs (21 trials). Trials were conducted in 16 countries, but the majority (63%) were undertaken in the United States. Several trials evaluated more than one behavioral program; there were 166 intervention arms in total. The mean age of the participants ranged from 45 to 72 years (median = 58). Baseline HbA_{1c} ranged from 6.3 to 12.3 percent (median = 8%). Median duration of diabetes was 8.1 years (range, 1–18 years). The proportion of nonwhite and/or Hispanic participants was between 0 and 100 percent; the majority ($\geq 75\%$) of participants in 32 trials reported nonwhite and/or Hispanic race/ethnicity.

Overall, median program duration was 6 months (range, 1–96) and median number of contact hours was 12 (range, 1–208). Sixty-four programs were delivered to individuals only, 56 were delivered to groups only, and 44 had some mixture of individual and group delivery. A small majority of programs were delivered by one health care professional, with or without the assistance of a non–health care professional; other programs were delivered by a multidisciplinary team or solely by non–health care

professionals. Technology was the primary method of communication for 17 programs studied in 16 trials and was used in combination with in-person communication in 25 programs; based on our inclusion criteria, all programs were delivered with some form of communication with delivery personnel.

All trials were assessed as having a medium or high overall risk of bias. For objective outcomes (e.g., HbA_{1c}, weight, blood pressure), 42 percent of trials had a medium risk of bias and 58 percent had a high risk. The assessment of high risk was largely driven by incomplete outcome data (i.e., loss to followup). Of trials (n = 92) reporting on subjective outcomes of interest for this review (e.g., HRQL, depression), 13 percent had a medium risk of bias; the remainder (87%) had a high risk of bias. This was primarily because of lack of blinding of participants, study personnel, and outcome assessors. See the Supplementary File: Full Text Screening Form, Risk of Bias Tools, and Results of Meta-Analyses for T2DM Across Outcomes (available at http://srdr.ahrq.gov) for a description of decision rules for these assessments.

T2DM: Overall Effectiveness of Behavioral Programs and Results for KQs 5 and 6
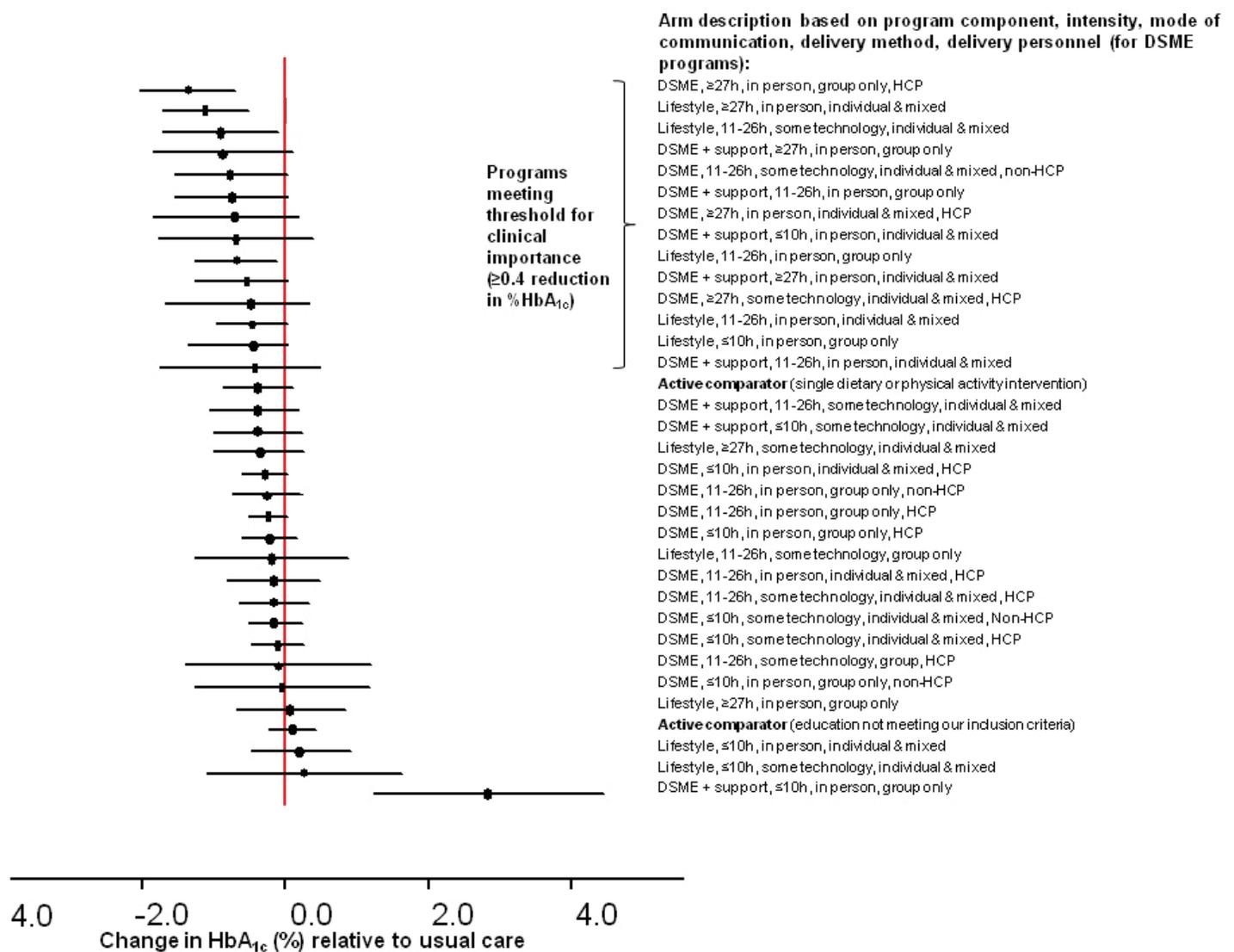
Effectiveness of Behavioral Programs Across Outcomes

There is evidence showing a beneficial effect of behavioral programs compared with both usual care and active interventions at end of intervention for glycemic control; however, for followup timepoints of 6 and 12 months, only the results at 6 months for comparisons with active controls were statistically significant. None of the results were considered to be clinically important based on our prespecified threshold of a 0.4 change in percent HbA_{1c}. There was substantial statistical heterogeneity in these pairwise meta-analyses, supporting our subsequent analysis for KQs 5 and 6 to determine which program factors and population characteristics mediate (and optimize) the effects.

Compared with usual care but not active controls, behavioral programs showed some benefits in terms of reducing BMI (0.2–0.9 kg/m^2) up to 12-month followup. There were reductions in weight (1.3–1.7 kg) and waist circumference (3.2 cm) at end of intervention, and (vs. usual care) in daily energy intake (65–150 kcal per day at 6 months). Few studies reported on outcomes related to changes in physical activity and medication adherence, and findings were consistently of no difference.

HRQOL was reported by fewer studies than anticipated, and the results mostly showed no difference. Results for diabetes distress favored behavioral programs compared with usual care at end of intervention (MD, -1.8; not

Figure D. Plot of network meta-analysis results for HbA$_{1c}$

Arm description based on program component, intensity, mode of communication, delivery method, delivery personnel (for DSME programs):

DSME, ≥27h, in person, group only, HCP
Lifestyle, ≥27h, in person, individual & mixed
Lifestyle, 11-26h, some technology, individual & mixed
DSME + support, ≥27h, in person, group only
DSME, 11-26h, some technology, individual & mixed, non-HCP
DSME + support, 11-26h, in person, group only
DSME, ≥27h, in person, individual & mixed, HCP
DSME + support, ≤10h, in person, individual & mixed
Lifestyle, 11-26h, in person, group only
DSME + support, ≥27h, in person, individual & mixed
DSME, ≥27h, some technology, individual & mixed, HCP
Lifestyle, 11-26h, in person, individual & mixed
Lifestyle, ≤10h, in person, group only
DSME + support, 11-26h, in person, individual & mixed
Active comparator (single dietary or physical activity intervention)
DSME + support, 11-26h, some technology, individual & mixed
DSME + support, ≤10h, some technology, individual & mixed
Lifestyle, ≥27h, some technology, individual & mixed
DSME, ≤10h, in person, individual & mixed, HCP
DSME, 11-26h, in person, group only, non-HCP
DSME, 11-26h, in person, group only, HCP
DSME, ≤10h, in person, group only, HCP
Lifestyle, 11-26h, some technology, group only
DSME, 11-26h, in person, individual & mixed, HCP
DSME, 11-26h, some technology, individual & mixed, HCP
DSME, ≤10h, some technology, individual & mixed, Non-HCP
DSME, ≤10h, some technology, individual & mixed, HCP
DSME, 11-26h, some technology, group, HCP
DSME, ≤10h, in person, group only, non-HCP
Lifestyle, ≥27h, in person, group only
Active comparator (education not meeting our inclusion criteria)
Lifestyle, ≤10h, in person, individual & mixed
Lifestyle, ≤10h, some technology, individual & mixed
DSME + support, ≤10h, in person, group only

Programs meeting threshold for clinical importance (≥0.4 reduction in %HbA$_{1c}$)

DSME = diabetes self-management education; HbA1c = hemoglobin A1c; HCP = health care practitioner

This plot depicts the results from our network meta-analysis for the outcome of HbA$_{1c}$ (negative values favorable) when comparing groups ("nodes") of interventions, with each group differing by at least 1 level in the categories of program component, intensity, mode of communication, delivery method, and (for DSME programs only) delivery personnel. (See Table A for categorization schema.) The factors of program duration, program tailoring, and community engagement were not used for the analysis because of overlap in meaning with other factors (e.g., community engagement often attained through use of non–health care providers) and ability to categorize based on reporting (e.g., tailoring). The dots and lines represent the mean difference (MDs) and 95% credibility intervals for the represented programs relative to usual care; the figure indicates which MDs meet or exceed our predetermined threshold for clinical importance (change of ≥0.4% HbA).

clinically important based on prespecified threshold of 0.5 SD from the pooled studies), but not at longer followup. Diabetic retinopathy was reduced by 14 percent and very high–risk chronic kidney disease was reduced by 31 percent in participants receiving an intensive lifestyle program lasting 8 years or longer compared with didactic education and support in the largest trial, conducted by the LookAHEAD research group. All-cause mortality was 14 percent lower for those receiving behavioral programs than active control groups (RR, 0.86).

KQ 5. Potential Mediators of Effectiveness for T2DM

When interpreting the results for potential modifiers of effectiveness (components, intensity, delivery personnel, method of communication, degree of tailoring, and level of community engagement), we relied primarily on the relative ranking of the nodes that represented grouped factors and looked for trends in the findings based on program variables that appeared to determine whether the effects would offer clinical benefit. Some nodes had very

few studies, small sample sizes, and/or wide credibility intervals. Thus we did not make any firm conclusions for a single node or for differences in 561 potential comparisons, but rather from looking across nodes with similar features.

In a network meta-analysis with usual care serving as the main reference, programs demonstrating relative effect sizes for HbA_{1c} above our threshold for clinical importance (i.e., 0.4%) represented all three major program component categories of DSME, DSME plus support, and lifestyle. The effect sizes of minimally intensive DSME programs (≤10 contact hours) were all less than our threshold for clinical importance but were all higher than the effect sizes of active controls of educational interventions not meeting our criteria for a behavioral program (e.g., didactic education programs). Programs having higher effect sizes were more often delivered in person rather than including technology; the effective programs incorporating technology were all of moderate or high intensity (>10 contact hours). Figure D summarizes the results of the network meta-analysis for HbA_{1c}.

For the network meta-analysis of BMI, we created nodes using four variables (i.e., program component, program intensity, method of communication, and method of delivery). Lifestyle programs resulted in the highest effect sizes for BMI. Program intensity appeared to be less important than method of delivery; providing some in-person delivery appears to be beneficial.

KQ 6. Subgroups for Factors Mediating Effectiveness in T2DM

In terms of overall effectiveness at longest followup for HbA_{1c}, participants with suboptimal glycemic control (≥7% HbA_{1c}) appear to benefit more than those with good control (<7%) from behavioral programs when compared with usual care and active controls. The effect sizes were not clinically important for either group. Few differences were evident when a network meta-analysis was used to evaluate potential mediation by program factors in a subgroup of studies having participants with suboptimal baseline glycemic control.

At longest followup, older adults (≥65 years) did not benefit in terms of reduction in HbA_{1c} from behavioral programs compared with usual care or active controls. In adults <65 years, the effect size for behavioral programs compared with active controls at longest followup (up to 12 months) was clinically important. When using the studies of only participants <65 in the network analysis, the active "other" control group (e.g., dietary or physical activity intervention) showed clinically important benefit for glycemic control (MD, -0.55).

Programs offered to predominantly minority participants (≥75% nonwhite and/or Hispanic) appear to provide more benefit than those offered to populations with a lower proportion (<75%) of minority participants. The effect size for minority participants reached clinical importance. None of the program implementation factors (e.g., intensity, delivery personnel) reached statistical significance for influencing the effectiveness of behavioral programs compared with usual care on HbA_{1c}. Lifestyle programs appeared to be favorable over DSME or DSME plus support for the group of studies (n = 24) with predominantly white non-Hispanic individuals (p = 0.07); the difference in reduction in HbA_{1c} between these two categories approached our threshold for clinical importance. Our results for ethnicity need to be interpreted with caution because of the apparent worse baseline glycemic control in studies of minority versus white non-Hispanic participants (8.8% vs. 7.6% HbA_{1c}); because behavioral programs seem to preferentially benefit those with higher baseline HbA_{1c}, this factor may account for much of the increase in benefit.

Discussion

Type 1 Diabetes Mellitus

Overall, behavioral programs appear to have benefit in T1DM for reducing HbA_{1c} when followup extends beyond the immediate postintervention period up to 6 months. The delay in benefit may in part reflect the time required for this marker of glycemic control, indicating control over the past 2 to 3 months, to demonstrate change. Notable, though, is the large diversity in program duration, whereby end of intervention was anywhere between 1.5 and 25 months from the beginning of the program. Another contributor to the delay in benefit may be that a period of time is needed to integrate newly learned self-management behaviors into one's life; however, the largely insufficient level of evidence for the behavioral outcomes does not allow us to determine this with any certainty. These beneficial findings for HbA_{1c} at 6 months appear to be tempered by those of no difference at longer followup timepoints (≥12 months), although we are unable to confidently rule out benefit at long-term followup because of low SOE. Our findings may underestimate the effect of these programs should they be implemented in routine practice. The usual care group in several studies received some form of attention from the investigators (e.g., periodic telephone calls to maintain contact and encourage study participation), which may have resulted in improved glycemic control for the comparator group and reduced the relative effects observed for the behavioral program. Participants, or their providers, in the usual care or active control groups (not being blinded to group assignment in most studies) may have become more motivated to practice better self-management (including blood glucose regulation using insulin

titrations), which could also attenuate differences between groups. Differences in the "usual care" provided may have also played a role, although this effect may be minimal considering recent evidence that variations in standard care in studies of behavioral interventions for youths with T1DM did not significantly impact study results.[58]

The positive findings for behavioral programs compared with active controls are notable. By offering an intervention to both study arms, these studies may also have introduced less potential bias from lack of allocation concealment and blinding. Our finding of a statistically significant and clinically important reduction (by 0.44%) in HbA_{1c} at 6-month followup for these comparisons is promising.

Self-management of T1DM during adolescence is complex, often characterized by personal challenges and uncertainty, transitions to adult care, less frequent health care visits, and diminished parental involvement; consequently, glycemic control deteriorates over the course of childhood and adolescence for many youths with T1DM.[59-62] For these reasons, many of the studies included in this review aimed to prevent deterioration of glycemic control rather than to improve it. The statistically significant reductions in HbA_{1c} at 6-month followup (vs. usual care) and the clinically important reductions in HbA_{1c} at 6- and 12-month followup (0.60% and 0.52%, respectively) in comparisons with active controls in youths lend substantial support for these programs. Likewise, incorporating more demanding self-management behaviors may negatively impact social and emotional functioning, such that our findings of no difference in generic HRQL at end of intervention may be viewed positively.

For T1DM, there was the suggestion that effectiveness was not moderated by program intensity (i.e., duration, contact hours, or frequency of contacts) and that individual versus group delivery may be beneficial. Because of insufficient data, we were unable to examine the difference between educational and lifestyle programs, or the benefit from addition of a support component to DSME programs.

Type 2 Diabetes Mellitus

Moderate- and high-intensity (≥11 hours contact time) programs appear to be necessary to provide individuals with clinically important effects on glycemic control. This outcome may also benefit from in-person delivery rather than incorporating technology. For BMI, providing some individual delivery, rather than solely relying on group formats, appears to be beneficial.

Lifestyle programs, focusing more on weight reduction and increases in physical activity than diabetes self-care, may provide similar or more benefit than DSME programs

for improving glycemic control for individuals with T2DM. Our review also confirms previous suggestions that programs that have an interactive nature and employ behavioral change techniques are beneficial when compared with didactic educational interventions. While some of our findings may not result in clinically important changes at an individual level, the burgeoning growth of this disease means that even small gains in glycemic control from behavioral programs may serve as a substantial benefit for public health.

Our network meta-analysis results suggest that both individual and group delivery of programs is beneficial. Delivery format may be highly dependent on the population served and program content. Studies having clinically important effect sizes that offered programs in groups tended to be those offered to minorities, in which support from peers was incorporated as a key program feature.

We were unable to draw any conclusions about the choice of delivery personnel from the network meta-analysis. Drawing from the pairwise meta-analysis of five RCTs (647 subjects) comparing two or more interventions, there may be no difference between program delivery conducted by health care professionals or by lay providers (e.g., peers with diabetes, community health workers). One reason that programs delivered by health care professionals were not superior may be that physicians, nurses, and dietitians receive little or no training in behavioral techniques as part of their formal education.

Our findings suggest that people with suboptimal, or poor, baseline glycemic control (≥7% HbA_{1c}), younger age (<65 years), and racial/ethnic minority status may benefit the most from behavioral programs. Because there were apparent differences in baseline glycemic control between subgroups of race/ethnicity (i.e., 8.8% HbA_{1c} in the ≥75% minority group vs. 7.6% HbA_{1c} in the <75% minority group), it is hard to distinguish if ethnicity or glycemic control is more likely to have the greater influence in moderating program effectiveness. There are likely several other factors to also consider. Many investigators enrolling a large proportion of ethnic minorities in the trials included in this review also adapted programs in ways to make them more culturally and linguistically acceptable, often including peers in the delivery or social support groups, which appeared to enhance their effectiveness. Our reliance on study-level data to create subgroups (i.e., the entire study was delivered to minorities) may have limited our ability to capture differences in effects from programs delivered to a wider population base, which may reflect routine practice in many community health settings.

Applicability

Type 1 Diabetes Mellitus

The results of this report may be most applicable to individuals with suboptimal and poor glycemic control. Nevertheless, clinicians may view the results as highly relevant to their patient population, of whom many—particularly in their pubertal years—are struggling to achieve optimal control. The results should be generally applicable to older children and adolescents (youth studies), and middle-aged adults.

It is unclear whether the results are applicable to youths or adults with recently diagnosed T1DM. We did not find evidence to confirm or refute whether behavioral programs are more or less efficacious for other subgroups, including males or females, or racial or ethnic minorities.

All of the studies targeting adults were conducted in the United Kingdom, Europe, or New Zealand. It is unclear whether the results from these studies are applicable to community health settings in the United States. For youths, most studies (73%) were conducted in the United States; the remaining studies were conducted in Europe and Australia. Despite potential differences in settings and health systems, results were similar across the studies. The studies were conducted primarily in outpatient diabetes clinics affiliated with a secondary or tertiary care hospital. Our findings are generally applicable to these settings in the United States.

Type 2 Diabetes Mellitus

Our results appear to be applicable to the majority of people enrolling in behavioral programs. There were few studies of older (≥65 years) adults or for those with good glycemic control. Our exclusion criteria related to duration of diabetes (mean <1 year)—implemented in order to capture programs providing training in ongoing self-management and lifestyle behaviors—limit the relevance of this review for newly diagnosed patients. The results appear to be applicable to both men and women, and for people on a variety of diabetes treatment regimens (19.2% were on insulin). Overall, there was fairly good representation of individuals reporting a minority racial/ethnic background.

 The results seem to be applicable to community health settings in the United States. The majority (63%) of trials were conducted in the United States, and based on our inclusion criteria related to the Human Development Index,[48] all studies were performed in countries of similar development status. Although reported inconsistently, health systems differences (i.e., usual care) may vary widely between study populations and could potentially influence the results from behavioral programs. The effect from this difference should be minimal for this review, since we limited our results to changes from baseline between groups randomly assigned and judged to receive similar medical care.

Limitations of the Comparative Effectiveness Review Process

This review followed rigorous methodological standards, which were detailed a priori. Nevertheless, several limitations are inherent within systematic reviews in general.

First, there is a possibility of selective reporting bias (e.g., reporting only positive outcomes) and publication bias, whereby unexpectedly strong results from large trials are selectively reported. In terms of selective outcome reporting, we were able to locate several trial registries and protocols to compare planned and published outcome reporting; most studies included in this review were judged as having low bias in this respect. Our prespecified tests for publication bias provided no significant indication of bias. Selected studies were confined to the English language because we felt that these reports would be most applicable to the end-users of this review, who create recommendations or implement programs for people with diabetes within the United States. Moreover, effect sizes in language-restricted reviews have shown to not differ significantly (overestimating effect sizes by 2%) from those not having restrictions.[63] Study selection bias was limited by having two independent reviewers perform screening and selection; we feel confident that study exclusion was based on explicit and appropriate reasoning, which was clearly understood by reviewers.

The interventions evaluated in the included trials were highly diverse in their content, delivery, and setting; accordingly, some of our statistical analyses indicated substantial heterogeneity. Our analyses for KQs 3, 5, and 6 were designed to determine some of the factors leading to variability in success for behavioral programs. Variability may still exist in terms of several factors. An example is length of followup; our analyses for these KQs were based on longest followup to maximize study inclusion and capture outcome durability. Another example, applicable to T2DM, is within-program intensity; DSME plus support and lifestyle programs often had lower intensity maintenance phases of varying durations.

The effects of programs delivered solely through technology (i.e., no interaction with personnel) were not assessed. Cost analysis of implementing differing behavioral programs was not addressed in this review.

Limitations of the Evidence Base

The evidence base was inadequate to fully answer the KQs, particularly with respect to the limited number of outcomes evaluated in several studies. We were unable to fully evaluate all outcomes of interest for several KQs. For KQ 1, for T1DM, limited data were available to assess the SOE for many outcomes, including behavioral outcomes related to changes in dietary intake or physical activity, and clinical and health outcomes apart from HbA_{1c}. No studies contributed data for our assessment of harms (KQ 4). Our assessment of factors contributing to effectiveness of behavioral programs for T1DM (KQ 3) was limited to the outcome of HbA_{1c} and to univariate metaregressions.

For KQs 5 and 6, related to T2DM, our network meta-analysis allowed for multiple comparisons (i.e., all comparison groups and followup timepoints), but there were still too few studies reporting on outcomes besides HbA_{1c} and BMI. The metaregressions used for the subgroup analysis on ethnicity in KQ 6 are limited by comparator (only usual care), and the number of studies did not allow us to capture multiple variables in a single analysis. Moreover, our reliance on study-level data for the subgroup analyses makes these results exploratory. Several outcomes of importance to patients and policymakers, such as quality of life, development of complications, and health care use, were reported by too few studies to confidently support conclusions of effect or to analyze in terms of mediation by implementation factors.

Many trials had methodological limitations introducing some risk of bias. Blinding of participants and personnel is arguably difficult for trials of behavioral programs, especially when the comparator is usual care. According to our decision rules for assessing risk of bias, a low risk of bias for participant and personnel blinding was granted if the comparator was an active control or another program, the authors stated some means to blind the study hypothesis from participants, and personnel followed a structured training and protocol. Participant blinding in this manner was rarely reported. Similarly, blinding of outcome assessors, highly feasible in any situation, was rarely reported or sufficient. These two domains resulted in medium or high risk of bias being assigned for the subjective outcomes of most trials. For both subjective and objective outcomes, medium or high risk of bias was assigned in many cases from lack of intention-to-treat analysis (e.g., reporting only on results for completers) and/or from high participant attrition. Some studies had small sample sizes, and a few failed to achieve baseline comparability in their samples.

Research Gaps

Table D highlights some potential research needs based on our KQs.

Table D. Potential research needs by Key Question

KQ	Potential Research Needs
1 Effectiveness for T1DM	There were limited data to determine the effectiveness of behavioral programs for T1DM at durations of followup beyond 6 months. Future studies should strive to assess outcomes at longer term followup, to better determine the effects of these programs for periods of time that may better influence long-term outcomes of complications and quality of life.
1 Effectiveness for T1DM	There was insufficient evidence to demonstrate whether lifestyle programs (i.e., combining structured physical activity and dietary interventions) are effective for T1DM. Many individuals with T1DM under good glycemic control may have other risk factors (e.g., overweight, hyperlipidemia, hypertension) for which these programs may be warranted. Trials of lifestyle programs enrolling people with both types of diabetes should undertake subgroup analysis.
1 & 3 Effectiveness & moderating factors for T1DM	The effectiveness of adding a clinical, behavioral, psychosocial, or educational support phase to programs for T1DM is unknown. These may be useful for prolonging the effects of behavioral programs and to address some of the psychosocial aspects of the disease (particularly in adolescents) to a greater extent.
3 Moderating factors for T1DM	Only one study in T1DM compared behavioral programs delivered in person with those delivered via some form of technology allowing for interaction between the provider and patient. Transitioning individuals with diabetes between pediatric and adult care facilities and providers can be challenging, hampered by the scheduling structure of traditional clinics at a time in life when contact information and location of home, work, and education are often changing frequently. As a result, further research on providing behavioral programs via technology or creative scheduling is warranted for adolescents and young adults with diabetes.

Table D. Potential research needs by Key Question (continued)

KQ	Potential Research Needs
3 Moderating factors for T1DM	Several studies for T2DM included a small subsample of people with T1DM. Trials of lifestyle programs that incorporate exercise need to perform subgroup analysis by type of diabetes, particularly when evaluating the outcome of glycemic control. Adjustment of insulin for exercise in individuals with T1DM can be challenging and could result in differential effects of lifestyle programs on glycemic control, depending on the medical management of the participants.
3 & 5 Moderating factors for T1DM & T2DM	There was large diversity in the reporting and use of behavior change techniques employed within the programs. An evaluation of the effects of different strategies may shed additional light on the factors (within components) determining effectiveness for behavioral programs.
5 Moderating factors for T2DM	The identification of the combination of providers (e.g., physician, nurse, dietitian, pharmacist, social worker, psychologist, and trained lay individual) that is best for implementation of behavioral programs for T2DM deserves further evaluation.
5 Moderating factors for T2DM	Clinical psychologists are often employed to deliver program components that incorporate advanced behavioral approaches, such as motivational interviewing; this approach may not be feasible for all settings or within all program budgets. More research is required to determine the effectiveness of similar programs when delivered by other personnel trained to use these behavioral techniques.
5 Moderating factors for T2DM	Few trials directly compared interactive programs delivered in person with those delivered via technology. Because a technology-based approach may lessen resource burden, help to reach patients living in rural areas, and/or be desirable for younger adults more familiar with technology, its effectiveness needs further evaluation.
6 Effectiveness for different subgroups in T2DM	Trials including populations of diverse ethnic backgrounds should perform subgroup analysis based on age, race/ethnicity, and baseline glycemic control to further explore outcomes for these groups from programs that are not designed specifically for them, as might be common in most community health settings.
All	Few trials evaluated outcomes important to patients and decisionmakers (e.g., quality of life, micro- and macrovascular complications, health care use) in a manner that allowed pooling of results across studies. Use of widely accepted generic quality-of-life measures would be beneficial.
All	Study attrition rates affected the overall risk of bias substantially. More research on methods for maintaining study participation is required.
All	The risk of bias from participant and personnel blinding was high in most trials. Although many trials compared behavioral programs with active controls (limiting risk of bias because of blinding) comparisons with usual care would benefit from some mechanism to blind participants from the study hypothesis. Blinding of outcome assessors should always be attempted for subjective outcomes.
All	There is a need for consensus on what constitutes clinically important differences in outcomes for behavioral programs, such that they can be interpreted in meaningful ways for clinicians and patients.

KQ = Key Question; T1DM = type 1 diabetes mellitus; T2DM = type 2 diabetes mellitus

Conclusions

Behavioral programs for T1DM offer some benefit for glycemic control when followup extends beyond end of intervention up to 6 months. There was no significant difference at end of intervention or followup longer than 6 months, although our confidence in these findings is low and we cannot rule out benefit. There was no difference in generic HRQL at end of intervention, or in diabetes distress or self-management behaviors at up to 6-month followup, although the SOE was low for these findings with the exception of generic HRQL at end of intervention (moderate SOE). Behavioral programs appear to be acceptable to patients with T1DM, given a 21-percent lower rate of attrition among those in behavioral programs than among those receiving usual care. Data were insufficient to draw any conclusions for other outcomes, including diabetes-specific HRQL, change in body composition or lifestyle behaviors, micro- and macrovascular complications, and mortality. Encouraging patients with T1DM to participate in behavioral programs to improve outcomes apart from HbA_{1c} is not supported by the current evidence.

For T2DM, our analyses showed limited benefit in glycemic control from DSME programs offering ≤10 hours of contact with delivery personnel and suggested that in-person delivery of behavioral programs is more beneficial than incorporation of technology. We found that programs focused on lifestyle or on DSME can have similar benefit in terms of glycemic control, and that lifestyle programs appear to be better for reducing BMI. Whether the behavioral program is delivered by a health care professional or a trained lay person, or via individual or group format, appears to be less important based on the available evidence. Behavioral programs seem to benefit individuals having suboptimal or poor glycemic control more than those with optimal control. Tailoring programs to ethnic minorities—such as offering culturally appropriate materials and incorporating group interaction with peers—appears to be beneficial. While efforts should be made to provide culturally sensitive programs, community health settings that serve populations that are diverse in language and ethnicity may not have the opportunity to provide this flexible programming to meet each group's needs.

Efforts at integrating behavioral programs into care settings that incorporate the latest management guidelines should be prioritized. Program evaluation is an important component to build into the implementation of any behavioral program for diabetes, to ensure that it is the correct fit to be effective for the population that it is meant to serve. At this time, there remains a need for clinicians to evaluate each patient's success after participating in these programs, in case additional means are necessary to control their disease more adequately to prevent devastating complications.

References

1. Renders CM, Valk GD, Griffin SJ, et al. Interventions to improve the management of diabetes mellitus in primary care, outpatient and community settings. Cochrane Database Syst Rev. 2001;(1):CD001481. PMID: 11279717.

2. Landon BE, Hicks LS, O'Malley AJ, et al. Improving the management of chronic disease at community health centers. N Engl J Med. 2007 Mar 1;356(9):921-34. PMID: 17329699.

3. Couch R, Jetha M, Dryden DM, et al. Diabetes education for children with type 1 diabetes mellitus and their families. Evid Rep Technol Assess (Full Rep). 2008 Apr;(166):1-144. PMID: 18620470.

4. Gage H, Hampson SE, Skinner TC, et al. Educational and psychosocial programmes for adolescents with diabetes: approaches, outcomes and cost-effectiveness. Patient Educ Couns. 2004 Jun;53(3):333-46. PMID: 15186872.

5. Hood KK, Rohan JM, Peterson CM, et al. Interventions with adherence-promoting components in pediatric type 1 diabetes: meta-analysis of their impact on glycemic control. Diabetes Care. 2010 Jul;33(7):1658-64. PMID: 20587726.

6. Murphy HR, Rayman G, Skinner TC. Psycho-educational interventions for children and young people with type 1 diabetes. Diabet Med. 2006 Sep;23(9):935-43. PMID: 16922699.

7. Urban AD, Berry D, Grey M. Optimizing outcomes in adolescents with type 1 diabetes and their families. J Clin Outcomes Manag. 2004;11(5):299-306.

8. Centers for Disease Control and Prevention. National Diabetes Statistics Report, 2014. www.cdc.gov/diabetes/data/statistics/2014statisticsreport.html. Accessed November 27, 2014.

9. Pettitt DJ, Talton J, Dabelea D, et al. Prevalence of diabetes in U.S. youth in 2009: the Search for Diabetes in Youth Study. Diabetes Care. 2014 Feb;37(2):402-8. PMID: 24041677.

10. Search for Diabetes in Youth Study Group. The burden of diabetes mellitus among US youth: prevalence estimates from the Search for Diabetes in Youth Study. Pediatrics. 2006 Oct;118(4):1510-8. PMID: 17015542.

11. Dall TM, Zhang Y, Chen YJ, et al. The economic burden of diabetes. Health Aff (Millwood). 2010 Feb;29(2):297-303. PMID: 20075080.

12. Centers for Disease Control and Prevention. Diabetes Report Card 2012. www.cdc.gov/diabetes/pubs/pdf/diabetesreportcard.pdf. Accessed November 15, 2013.

13. Bystritsky A, Danial J, Kronemyer D. Interaction between diabetes and anxiety and depression: implications for treatment. Endocrinol Metab Clin N Am. 2014;43(1):269-83. PMID: 24582102.

14. American Diabetes Association. Standards of medical care in diabetes-2014. Diabetes Care. 2014 Jan;37 Suppl 1:S14-80. PMID: 24357209.

15. Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med. 1993 Sep 30;329(14):977-86. PMID: 8366922.

16. Nathan DM. The Diabetes Control and Complications Trial/epidemiology of diabetes interventions and complications study at 30 years: overview. Diabetes Care. 2014 Jan;37(1):9-16. PMID: 24356592.

17. Orchard TJ, Nathan DM, Zinman B, et al. Association between 7 years of intensive treatment of type 1 diabetes and long-term mortality. JAMA. 2015 Jan 6;313(1):45-53. PMID: 25562265.

18. Fullerton B, Jeitler K, Seitz M, et al. Intensive glucose control versus conventional glucose control for type 1 diabetes mellitus. Cochrane Database Syst Rev. 2014;2:CD009122. PMID: 24526393.

19. UK 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 Sep 12;352(9131):837-53. PMID: 9742976.

20. UK Prospective Diabetes Study (UKPDS) Group. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). Lancet. 1998 Sep 12;352(9131):854-65. PMID: 9742977.

21. Hemmingsen B, Lund SS, Gluud C, et al. Targeting intensive glycaemic control versus targeting conventional glycaemic control for type 2 diabetes mellitus. Cochrane Database Syst Rev. 2013;11:CD008143. PMID: 24214280.

22. Colhoun HM, Betteridge DJ, Durrington PN, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. Lancet. 2004 Aug 21-27;364(9435):685-96. PMID: 15325833.

23. Heart Protection Study Collaborative Group. MRC/ BHF heart protection study of cholesterol-lowering with simvastatin in 5963 people with diabetes: a randomised placebo-controlled trial. Lancet. 2003 Jun 14;361(9374):2005-16. PMID: 12814710.

24. UK Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. BMJ. 1998;317(7160):703-13. PMID: 9732337.

25. Lv J, Neal B, Ehteshami P, et al. Effects of intensive blood pressure lowering on cardiovascular and renal outcomes: a systematic review and meta-analysis. PLoS Med. 2012;9(8):e1001293. PMID: 22927798.

26. Norris SL, Zhang X, Avenell A, et al. Long-term effectiveness of lifestyle and behavioral weight loss interventions in adults with type 2 diabetes: a meta-analysis. Am J Med. 2004 Nov 15;117(10):762-74. PMID: 15541326.

27. Schellenberg ES, Dryden DM, Vandermeer B, et al. Lifestyle interventions for patients with and at risk for type 2 diabetes: a systematic review and meta-analysis. Ann Intern Med. 2013 Oct 15;159(8):543-51. PMID: 24126648.

28. Thomas DE, Elliott EJ, Naughton GA. Exercise for type 2 diabetes mellitus. Cochrane Database Syst Rev. 2006(3):CD002968. PMID: 16855995.

29. Watts NB, Spanheimer RG, Digirolamo M, et al. Prediction of glucose response to weight loss in patients with non-insulin-dependent diabetes mellitus. Arch Intern Med. 1990 Apr;150(4):803-6. PMID: 2327840.

30. Funnell M. Beyond the data: moving towards a new dawn in diabetes. Diabet Med. 2013 Jul;30(7):765-6. PMID: 23710971.

31. Knight KM, Dornan T, Bundy C. The diabetes educator: trying hard, but must concentrate more on behaviour. Diabet Med. 2006;23(5):485-501. PMID: 16681557.

32. Haas L, Maryniuk M, Beck J, et al. National standards for diabetes self-management education and support. Diabetes Care. 2013;36(Suppl 1):S100-S8. PMID: 23264420.

33. International Diabetes Federation. Position Statement: Self-Management Education. 2011. www.idf.org/ education/self-management-education. Accessed November 15, 2013.

34. Jones H, Berard LD, MacNeill G, et al. Clinical practice guidelines: self-management education. Can J Diabetes. 2013;37(Suppl 1):S26-S30. PMID: 24070958.

35. National Institute for Health and Clinical Excellence. Type 2 diabetes: the management of type 2 diabetes. NICE Clinical Guideline 87. London: National Collaborating Centre for Chronic Conditions, Centre for Clinical Practice; 2009.

36. Hoerger TJ, Segel JE, Gregg EW, et al. Is glycemic control improving in U.S. adults? Diabetes Care. 2008 Jan;31(1):81-6. PMID: 17934153.

37. Chodosh J, Morton SC, Mojica W, et al. Meta-analysis: chronic disease self-management programs for older adults. Ann Intern Med. 2005 Sep 20;143(6):427-38. PMID: 16172441.

38. Deakin T, McShane CE, Cade JE, et al. Group based training for self-management strategies in people with type 2 diabetes mellitus. Cochrane Database Syst Rev. 2005(2):CD003417. PMID: 15846663.

39. Fan L, Sidani S. Effectiveness of diabetes self-management education intervention elements: a meta-analysis. Can J Diabetes. 2009;33(1):18-26.

40. Gary TL, Genkinger JM, Guallar E, et al. Meta-analysis of randomized educational and behavioral interventions in type 2 diabetes. Diabetes Educ. 2003 May-Jun;29(3):488-501. PMID: 12854339.

41. Medical Advisory Secretariat. Behavioural interventions for type 2 diabetes: an evidence-based analysis. Ont Health Technol Assess Ser. 2009;9(21):1-45. PMID: 23074526.

42. Minet L, Moller S, Vach W, et al. Mediating the effect of self-care management intervention in type 2 diabetes: a meta-analysis of 47 randomised controlled trials. Patient Educ Couns. 2010 Jul;80(1):29-41. PMID: 19906503.

43. Norris SL, Lau J, Smith SJ, et al. Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control. Diabetes Care. 2002 Jul;25(7):1159-71. PMID: 12087014.

44. Warsi A, Wang PS, Lavalley MP, et al. Self-management education programs in chronic disease: a systematic review and methodological critique of the literature. Arch Intern Med. 2004 Aug 9-23;164(15):1641-9. PMID: 15302634.

45. Ellis SE, Speroff T, Dittus RS, et al. Diabetes patient education: a meta-analysis and meta-regression. Patient Educ Couns. 2004 Jan;52(1):97-105. PMID: 14729296.

46. Glazier RH, Bajcar J, Kennie NR, et al. A systematic review of interventions to improve diabetes care in socially disadvantaged populations. Diabetes Care. 2006 Jul;29(7):1675-88. PMID: 16801602.

47. Michie S, Ashford S, Sniehotta FF, et al. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the calo-re taxonomy. Psychol Health. 2011 Nov;26(11):1479-98. PMID: 21678185.

48. Malik K, Human Development Report 2013 Team. Human Development Report 2013. The Rise of the South: Human Progress in a Diverse World. United Nations Development Programme. www.undp.org/content/undp/en/home/. Accessed September 3, 2015.

49. Ohkubo Y, Kishikawa H, Araki E, et al. Intensive insulin therapy prevents the progression of diabetic microvascular complications in Japanese patients with non-insulin-dependent diabetes mellitus: a randomized prospective 6-year study. Diabetes Res Clin Pract. 1995 May;28(2):103-17. PMID: 7587918.

50. Higgins JPT, Green S, eds. Section 8. Assessing risk of bias in included studies. In: Cochrane Handbook for Systematic Reviews of Interventions. The Cochrane Collaboration; 2011. www.handbook.cochrane.org. Accessed March 4, 2014.

51. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. www.ohri.ca/programs/clinical_epidemiology/nosgen.pdf. Accessed September 11, 2015.

52. Ip S, Hadar N, Keefe S, et al. A Web-based archive of systematic review data. Syst Rev. 2012;1:15. PMID: 22588052.

53. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research. Guidance for Industry: Diabetes Mellitus: Developing Drugs and Therapeutic Biologics for Treatment and Prevention. 2008. www.fda.gov/downloads/ Drugs/Guidances/ucm071624.pdf. Accessed May 27, 2014.

54. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003 May;41(5):582-92. PMID: 12719681.

55. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008 Feb;61(2):102-9. PMID: 18177782.

56. Inthout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard Dersimonian-Laird method. BMC Med Res Methodol. 2014;14:25. PMID: 24548571.

57. Sidik K, Jonkman JN. A simple confidence interval for meta-analysis. Stat Med. 2002 Nov 15;21(21):3153-9. PMID: 12375296.

58. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at www.effectivehealthcare.ahrq.gov.

59. Court JM, Cameron FJ, Berg-Kelly K, et al. Diabetes in adolescence. Pediatr Diabetes. 2009 Sep;10 Suppl 12:185-94. PMID: 19754629.

60. Ingerski LM, Anderson BJ, Dolan LM, et al. Blood glucose monitoring and glycemic control in adolescence: contribution of diabetes-specific responsibility and family conflict. J Adolesc Health. 2010 Aug;47(2):191-7. PMID: 20638012.

61. Schilling LS, Knafl KA, Grey M. Changing patterns of self-management in youth with type I diabetes. J Pediatr Nurs. 2006 Dec;21(6):412-24. PMID: 17101399.

62. Peters A, Laffel L. Diabetes care for emerging adults: recommendations for transition from pediatric to adult diabetes care systems: a position statement of the American Diabetes Association, with representation by the American College of Osteopathic Family Physicians, the American Academy of Pediatrics, the American Association of Clinical Endocrinologists, the American Osteopathic Association, the Centers for Disease Control and Prevention, Children with Diabetes, the Endocrine Society, the International Society for Pediatric and Adolescent Diabetes, Juvenile Diabetes Research Foundation International, the National Diabetes Education Program, and the Pediatric Endocrine Society (formerly Lawson Wilkins Pediatric Endocrine Society). Diabetes Care. 2011 Nov;34(11):2477-85. PMID: 22025785.

63. Moher D, Pham B, Klassen TP, et al. What contributions do languages other than English make on the results of meta-analyses? J Clin Epidemiol. 2000 Sep;53(9):964-72. PMID: 11004423.

Full Report