Department of Veterans Affairs
Health Services Research & Development Service | Evidence-based Synthesis Program

QUERI

Shared Medical Appointments for Chronic Medical Conditions: A Systematic Review

July 2012

Prepared for:
Department of Veterans Affairs
Veterans Health Administration
Quality Enhancement Research Initiative
Health Services Research & Development Service
Washington, DC 20420

Prepared by:
Evidence-based Synthesis Program (ESP) Center
Durham Veterans Affairs Healthcare System
Durham, NC
John W Williams Jr., M.D., M.H.Sc., Director

Investigators:
Principal Investigator:
David Edelman, M.D.

Co-Investigators:
Jennifer R. McDuffie, Ph.D.
Eugene Oddone, M.D., M.H.Sc.
Jennifer M. Gierisch, Ph.D., M.P.H.
John W. Williams Jr., M.D., M.H.Sc.

Research Associate:
Avishek Nagi, M.S.

Medical Editor:
Liz Wing, M.A.

VA HEALTH CARE Defining EXCELLENCE in the 21st Century

PREFACE

Quality Enhancement Research Initiative's (QUERI's) Evidence-based Synthesis Program (ESP) was established to provide timely and accurate syntheses of targeted healthcare topics of particular importance to Veterans Affairs (VA) managers and policymakers, as they work to improve the health and healthcare of Veterans. The ESP disseminates these reports throughout VA.

QUERI provides funding for four ESP Centers and each Center has an active VA affiliation. The ESP Centers generate evidence syntheses on important clinical practice topics, and these reports help:

- develop clinical policies informed by evidence,
- guide the implementation of effective services to improve patient outcomes and to support VA clinical practice guidelines and performance measures, and
- set the direction for future research to address gaps in clinical knowledge.

In 2009, the ESP Coordinating Center was created to expand the capacity of QUERI Central Office and the four ESP sites by developing and maintaining program processes. In addition, the Center established a Steering Committee comprised of QUERI field-based investigators, VA Patient Care Services, Office of Quality and Performance, and Veterans Integrated Service Networks (VISN) Clinical Management Officers. The Steering Committee provides program oversight, guides strategic planning, coordinates dissemination activities, and develops collaborations with VA leadership to identify new ESP topics of importance to Veterans and the VA healthcare system.

Comments on this evidence report are welcome and can be sent to Nicole Floyd, ESP Coordinating Center Program Manager, at nicole.floyd@va.gov.

Recommended citation: Edelman D, McDuffie JR, Oddone E, Gierisch JM, Nagi A, Williams JW Jr. Shared Medical Appointments for Chronic Medical Conditions: A Systematic Review. VA-ESP Project #09-010; 2012.

This report is based on research conducted by the Evidence-based Synthesis Program (ESP) Center located at the Durham VA Medical Center, Durham, NC, funded by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Quality Enhancement Research Initiative. The findings and conclusions in this document are those of the author(s) who are responsible for its contents; the findings and conclusions do not necessarily represent the views of the Department of Veterans Affairs or the United States government. Therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs. No investigators have any affiliations or financial involvement (e.g., employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that con¬flict with material presented in the report.

TABLE OF CONTENTS

FIGURES

TABLES

EVIDENCE REPORT

INTRODUCTION

Optimal care of chronic illness requires both high-quality care (i.e., excellent clinical outcomes) and ready access to care. However, health care systems struggle with providing access and quality simultaneously—and to achieve both these objectives while simultaneously maintaining staff job satisfaction is a formidable undertaking. Although Veterans Affairs (VA) has made large strides in delivering quality care over the past two decades, quality gaps remain, both gaps in technical quality (e.g., still only 70 to 75% of patients have appropriate blood pressure control[1]) and gaps in timeliness of services. Group medical visits offer the promise of improving the effectiveness, timeliness, and efficiency of health care. Additionally, because clinicians may prefer to work in collaborative, multidisciplinary settings, group medical visits also have the potential to improve staff satisfaction.[2-8]

Group medical visits are defined as multiple patients seen together while in the same clinical setting. A subset of group clinics—referred to as shared medical appointments (SMAs)—is defined by groups of patients meeting over time for comprehensive care, usually involving a practitioner with prescribing privileges, for a defining chronic condition or health care state. SMAs often use educational and/or self-management enhancement strategies, paired with medication management, in an effort to achieve improved disease outcomes.

SMAs have been scientifically studied in an array of primary care settings over the last 10 to 15 years.[3,5,8-23] However, there has been great variability among these studies. In particular, the settings of these studies have been heterogeneous; different chronic health care states have been assessed; and the impact on clinical, cost, and utilization outcomes has been variable. Most important, there has been significant variation in the SMA intervention itself—in particular, which types of clinical, educational, and self-efficacy approaches are included in the specific SMA under evaluation. This uncertainty regarding the optimal design and impact of SMAs led the VA to commission this evidence synthesis report.

BACKGROUND

The SMA approach developed as a care redesign strategy over the last 15 years. SMAs are defined as providing group-based, longitudinal medical care for a number of patients who have a common characteristic such as type 2 diabetes.[3,5,8,13-20,22,23] This commonality may be a disease (e.g., diabetes), a demographic (e.g., patients over 65 years of age), or some other health care–related element (e.g., high utilization of services). The patient group may stay constant, in an attempt to provide group bonding, or patients may be allowed to attend sessions chosen from a schedule at their own convenience to promote attendance.

In general, SMAs will have more than one health care provider involved; often the care team will include a person trained or skilled in delivering patient education or facilitating patient interaction (nurse, psychologist) and a prescribing provider empowered to make and initiate a comprehensive care plan. Like patients, providers can either be constant with the same patients or vary over time.

SMA sessions usually last from 60 to 120 minutes. Sessions usually have part of their time set aside for social integration, part set aside for interactive education, and part committed to changes in the care plan for the common condition. The education piece is designed to improve self-management skills; educators will often be formally trained in skills such as motivational interviewing to help patients enhance their self-management. Because they involve both self-management improvement along with medication intensification, SMAs have the potential to coordinate these strategies to maximize the effects of each.

SMAs have been touted as a way to improve key elements of health care, particularly access, outcomes, and cost. Improved access is thought to occur because in a drop-in group structure, patients get their chronic illness care when they want it and/or because a group visit is usually shorter than the amount of time it takes to see all patients in the group one-on-one, thereby improving the provider's throughput and patients' access to that provider. Improved outcomes are thought to occur because (1) the group provides enhanced self-management education due to more time spent in that education, the use of motivational interviewing by the trained facilitator, or the peer support of members of the group, (2) the group provides a focused environment for care of the common condition or unifying characteristic (e.g., older age) without the distractions of multiple other issues that come up in a brief primary care visit, and/or (3) the group provides access to medication changes performed by a provider with special expertise in the common condition or by a team of providers with synergistic knowledge, thus leading the group to function like a specialty referral. Costs are thought to be lowered because the aforementioned efficiency in throughput leads to lower total costs of care or access to a group keeps patients from using acute-care settings for management of chronic illness, saving the costs associated with unnecessary emergency department visits or hospitalizations.

Early studies of SMAs focused on common demographic characteristics (elderly, high utilization) rather than common illnesses.[9,12] However, most recent studies have focused on a common chronic illness as the unifying theme for the SMA, with diabetes the most commonly studied. This may make clinical sense, given that patients with the same chronic illness require self-efficacy for the same self-management skills (e.g., patients with diabetes need to feel empowered to correctly monitor and record blood sugar readings). The disease focus may also be because of the ease of identifying disease-specific clinical outcomes for research studies.

OBJECTIVE OF THIS REPORT

Our objective in this evidence synthesis was to summarize the results of the diverse studies of SMAs in an effort to understand their impact on staff satisfaction, patient experience, and clinical outcomes along with effects on health care utilization. A second objective was to determine whether the impact of SMA visits varies by clinical condition or specific components of the intervention.

METHODS

TOPIC DEVELOPMENT

This review was commissioned by the VA Evidence-based Synthesis Program. The topic was nominated after a topic refinement process that included a preliminary review of published peer-reviewed literature, consultation with internal partners and investigators, and consultation with key stakeholders. We further developed and refined the key questions based on a preliminary review of published peer-reviewed literature in consultation with VA experts.

The final key questions were:

Key Question 1. For adults with chronic medical conditions, do shared medical appointments (SMAs) compared with usual care improve the following:

- Patient and staff experience?
- Treatment adherence?
- Quality measures such as (a) process of care measures utilized by VA, National Quality Forum, or National Committee for Quality Assurance and (b) biophysical markers (laboratory or physiological markers of health status such as HbA1c and blood pressure)?
- Symptom severity and functional status?
- Utilization of medical resources or health care costs?

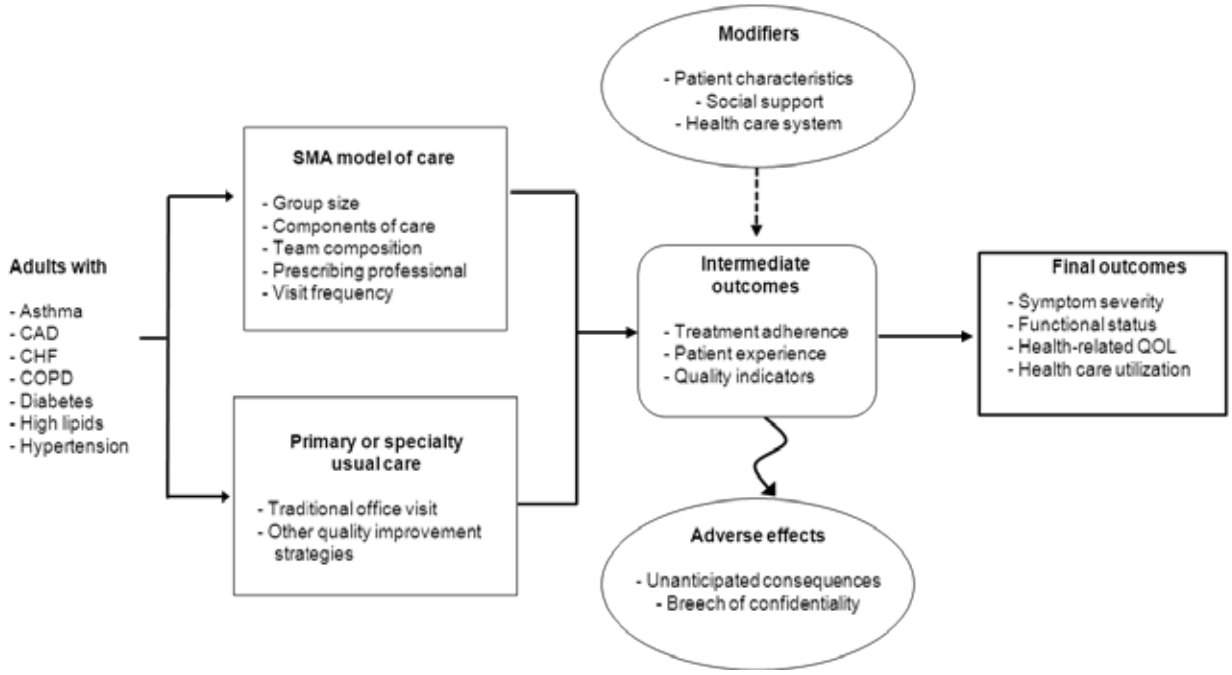
Key Question 2. For adults with chronic medical conditions, do the effects of SMAs vary by patient characteristics such as specific chronic medical conditions and severity of disease?

Key Question 3. Is the intensity of the intervention or the components used by SMAs associated with intervention effects?

ANALYTIC FRAMEWORK

We followed a standard protocol for all steps of this review; certain methods map to the PRISMA checklist.[24] Our approach was guided by the analytic framework shown in Figure 1.

Figure 1. Analytic framework for evaluating shared medical appointments



Abbreviations: CAD=coronary artery disease; CHF=congestive heart failure; COPD=chronic obstructive pulmonary disease; KQ=key question; QOL=quality of life; SMA=shared medical appointment

SEARCH STRATEGY

We searched MEDLINE® (via PubMed®), Embase®, CINAHL® , PsycINFO and Web of Science for peer-reviewed publications comparing shared medical appointments or group visits with usual care from January 1996 through September 2011. Our search strategy used the National Library of Medicine's medical subject headings (MeSH) keyword nomenclature and text words for types of visits or clinic appointments, and validated search terms for both randomized controlled trials[25] and relevant observational studies adapted from the Cochrane Effective Practice & Organization of Care Group search version 1.9. Our final search terms included terms for group visits together with terms for trials or relevant observational designs. We limited the search to articles published in the English language involving human subjects 18 years of age and older. The full search strategy is provided in Appendix A. An updated search for publications in PubMed was conducted in April 2012. We developed our search strategy in consultation with an experienced search librarian.

We supplemented the electronic searches with a manual search of citations from a set of key primary articles, three review articles and one systematic review.[3,4,7,10,11,15,18,19,22,26-28] The reference list for identified pivotal articles was manually hand-searched and cross-referenced against our library in order to retrieve additional manuscripts. All citations were imported into two electronic databases: EndNote® Version X5 (Thomson Reuters; Philadelphia, PA) for referencing and DistillerSR (Evidence Partners; Manotick, ON, Canada) for data abstraction. As a mechanism to assess the risk of publication bias, we searched www.clinicaltrials.gov for completed but unpublished studies in March 2012.

STUDY SELECTION

Using prespecified inclusion and exclusion criteria, two reviewers assessed titles and abstracts for relevance to the Key Questions (KQs). Full-text articles identified by either reviewer as potentially relevant were retrieved for further review. Each article retrieved was examined by two reviewers against the eligibility criteria (Table 1). Disagreements on inclusion, exclusion, or major reason for exclusion were resolved by discussion or by a third reviewer.

The criteria to screen articles for inclusion or exclusion at both the title-and-abstract and full-text screening stages are detailed in Table 1. We modified these criteria to include observational studies designs recommended by the Cochrane Effective Practice and Organization of Care Review Group (i.e., controlled before and after, nonrandomized cluster controlled trials, interrupted time-series). Studies excluded at the full-text review stage are listed with the reasons for exclusion in Appendix B.

Table 1. Summary of inclusion and exclusion criteria

Study characteristic	Inclusion criteria	Exclusion criteria
Population	Adults (≥18 years) of age with asthma, coronary artery disease, congestive heart failure, chronic obstructive pulmonary disease, diabetes mellitus, hyperlipidemia, hypertension, or combinations of these chronic medical conditions	Populations selected for individuals with substance abuse disorders
Intervention	"Exposure" must meet all the following criteria: • A series of medical visits (at least two) where at least one health care professional (including a prescribing clinician[a]) cares for a groups of patients • The medical provider addresses each patient's unique medical needs individually, with the potential to make changes in medications, but in the context of the group setting	Study excluded if exposure meets any of the following criteria: • Not a group visit • No prescribing clinician present at the group visit meeting • No plan for adjusting medications when indicated
Comparator	Usual care or other quality improvement strategy	None; study must have a control condition
Outcome	• Patient and/or staff experience • Treatment adherence (attendance, medications, self-management) • Biophysical markers (e.g., HbA1c, blood pressure) • Symptom severity • Functional status • Utilization of medical resources	None
Timing	Outcomes reported at least 3 months from randomization and initiation of intervention	Outcomes reported less than 3 months from randomization and initiation of intervention
Setting	• Outpatient settings; specifically, primary care or specialty clinic/practice • Conducted in North America, Western Europe, Australia/New Zealand	• Conducted in an inpatient or nonmedical community setting (i.e., senior centers, etc.) • Conducted in countries other than those specifically listed as included

Study characteristic	Inclusion criteria	Exclusion criteria
Study design[b]	• Patient or cluster RCTs • Nonrandomized cluster controlled trials • Controlled before-and-after studies • Interrupted time series designs	Cross-sectional studies and other observational study designs not specifically listed as "included" study designs
Publications	• English-language only • Published from 1996 to present • Peer-reviewed article	• Non-English language publication • Published before 1996[c]

[a]A prescribing clinician may be a medical doctor, doctor of osteopathy, advanced practice registered nurse (APRN), physician assistant, or doctor of pharmacy.
[b]Study designs recommended by the Cochrane Effective Practice and Organization of Care Group.
[c]Shared medical appointments were introduced by Beck et al. with their seminal article in 1997.
Abbreviations: HbA1c=glycosylated hemoglobin; KQ=key question; RCT=randomized controlled trial

DATA ABSTRACTION

Before general use, the abstraction form templates designed specifically for this report were pilot-tested on a sample of included articles and revised to ensure that all relevant data elements (Appendix C) were captured and that there was consistency and reproducibility between abstractors. We designed the data abstraction forms for this project to collect the data required to evaluate the specified eligibility criteria for inclusion in this review, as well as population characteristics and other data needed for determining outcomes (e.g., biophysical markers, resource utilization) and risk of bias. We paid particular attention to describing the details of the intervention, including the clinical team (clinical disciplines represented, team size and continuity), characteristics of the patient group (group size, group continuity, inclusion of family members or peer supports), and group visit processes (individual breakouts, medication changes, visit duration, telephone contacts). In addition, we examined the included articles for subgroup analyses of relevance to our key questions.

One investigator abstracted the data, and the second reviewed the completed abstraction form alongside the original article to check for accuracy and completeness. Disagreements were resolved by consensus or by obtaining a third investigator's opinion if consensus could not be reached by the first two. We supplemented abstraction of published data by contacting authors for missing information. We contacted 11 of 19 authors; of these, 7 replied with the requested information.

QUALITY ASSESSMENT

We also abstracted data necessary for assessing quality and applicability, as described in the Agency for Healthcare Research and Quality's (AHRQ's) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.[29] For RCTs, these key quality criteria consisted of (1) adequacy of randomization and allocation concealment, (2) comparability of groups at baseline, (3) blinding, (4) completeness of followup and differential loss to followup, (5) whether incomplete data were addressed appropriately, (6) validity of outcome measures, and (7) conflicts of interest (Appendix D). Using these quality criteria, we assigned a summary quality score (good, fair, poor) to individual RCTs studies as defined in the *Methods Guide*.

Threats to internal validity of systematic review conclusions based on observational studies were identified through assessment of the body of observational literature as a whole, with an examination of characteristics of individual studies.[29,30] Study-specific issues that were considered include (1) potential selection bias (i.e., degree of similarity between intervention and control patients), (2) performance bias (i.e., differences in care provided to intervention and control patients not related to the study intervention), (3) attribution and detection bias (i.e., whether outcomes were differentially detected between intervention and control groups), and (4) magnitude of reported intervention effects (see the section on "Selecting Observational Studies for Comparing Medical Interventions" in the *Methods Guide*).[29] For each study, one investigator assigned a summary quality rating for "hard outcomes" (e.g., laboratory measures) and a separate rating for "soft" outcomes (e.g., staff experience), which were then reviewed by a second investigator; disagreements were resolved by consensus or by a third investigator if agreement could not be reached.

DATA SYNTHESIS

We critically analyzed studies to compare their characteristics, methods, and findings. We then determined the feasibility of completing a quantitative synthesis (i.e., meta-analysis) by exploring the volume of relevant literature, the completeness of the results reporting and the conceptual homogeneity of the studies. Because the elderly and individuals with diabetes mellitus are high utilizers of the health care system and are distinct groups of clinical patients with distinct primary endpoints, we examined the groups of studies as they pertained to these target conditions separately.

When a meta-analysis was appropriate, we used random-effects models to synthesize the available evidence quantitatively. For other outcomes we analyzed the results qualitatively. The outcomes amenable to meta-analysis were continuous; we therefore summarized these outcomes by a weighted difference of the means when the same scale (e.g., blood pressure) was used and a standardized mean difference when the scales (e.g., health-related quality of life) differed across studies. We present summary estimates (standardized so that a negative value favors SMA) and 95 percent confidence intervals (95% CIs). Heterogeneity was examined among the studies using graphical displays and test statistics (Cochran's Q and I^2); the I^2 describes the percentage of total variation across studies due to heterogeneity rather than to chance.[31] Heterogeneity was categorized as low, moderate, or high based on I^2 values of 25 percent, 50 percent, and 75 percent respectively. We explored heterogeneity in study effects by using subgroup analyses for categorical variables (e.g., study quality) and meta-regression analyses for continuous or discrete variables (e.g., baseline HbA1c, intervention robustness). We constructed a "robustness score" that could range from 0 to 9, based on 7 intervention elements that were chosen a priori: theoretical framework guiding the intervention, individual breakouts, continuity between patients and clinical team, scheduled visits above the median, and medication changes. The latter two characteristics were scored 0 (absent) or 2 (present); all other items were scored as 0 or 1. We conducted a sensitivity analyses by using only studies whose populations had type 2 diabetes. Our subgroup and meta-regression analyses should be considered hypothesis-generating because they consist of indirect comparisons (across studies that may differ in ways other than the target condition) and thus are subject to confounding.

All basic analyses were conducted using Review Manager (RevMan) 5.1.4. (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011). Meta-regression analyses were conducted using Comprehensive Meta-Analysis software (Version 2; Biostat, Englewood, NJ).

RATING THE BODY OF EVIDENCE

In addition to rating the quality of individual studies, we evaluated the overall quality of the evidence for each KQ as described in the *Methods Guide*.[29] In brief, this approach requires assessment of four domains: risk of bias, consistency, directness, and precision. Additional domains considered were strength of association (magnitude of effect) and publication bias. For risk of bias, we considered basic (e.g., RCT) and detailed study design (e.g., adequate randomization). We used results from meta-analyses when evaluating consistency (forest plots, tests for heterogeneity), precision (confidence intervals), strength of association (weighted mean difference), and publication bias (clinicaltrials.gov survey). Optimal information size and consideration of whether the confidence interval crossed the clinical-decision threshold using a treatment model were also used when evaluating precision.[32] These domains were considered qualitatively, and a summary rating of high, moderate, low, or insufficient strength of evidence was assigned after discussion by two investigators. This four-level rating scale consists of the following definitions:

- High—Further research is very unlikely to change our confidence on the estimate of effect.
- Moderate—Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
- Low—Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- Insufficient—Evidence on an outcome is absent or too weak, sparse, or inconsistent to estimate an effect.

When a rating of high, moderate, or low was not possible or was imprudent to make, a grade of insufficient was assigned.[33] We also considered the risk of publication bias. Publication bias was addressed through a careful search of www.clinicaltrials.gov for identification of any study completed but unpublished or ongoing. For the single outcome with at least 10 studies, we used graphical (e.g., funnel plots) and test statistics (e.g., Beggs test) to detect publication bias; these methods do not perform well with fewer than 10 studies and thus were not performed for the other outcomes.[34,35]

PEER REVIEW

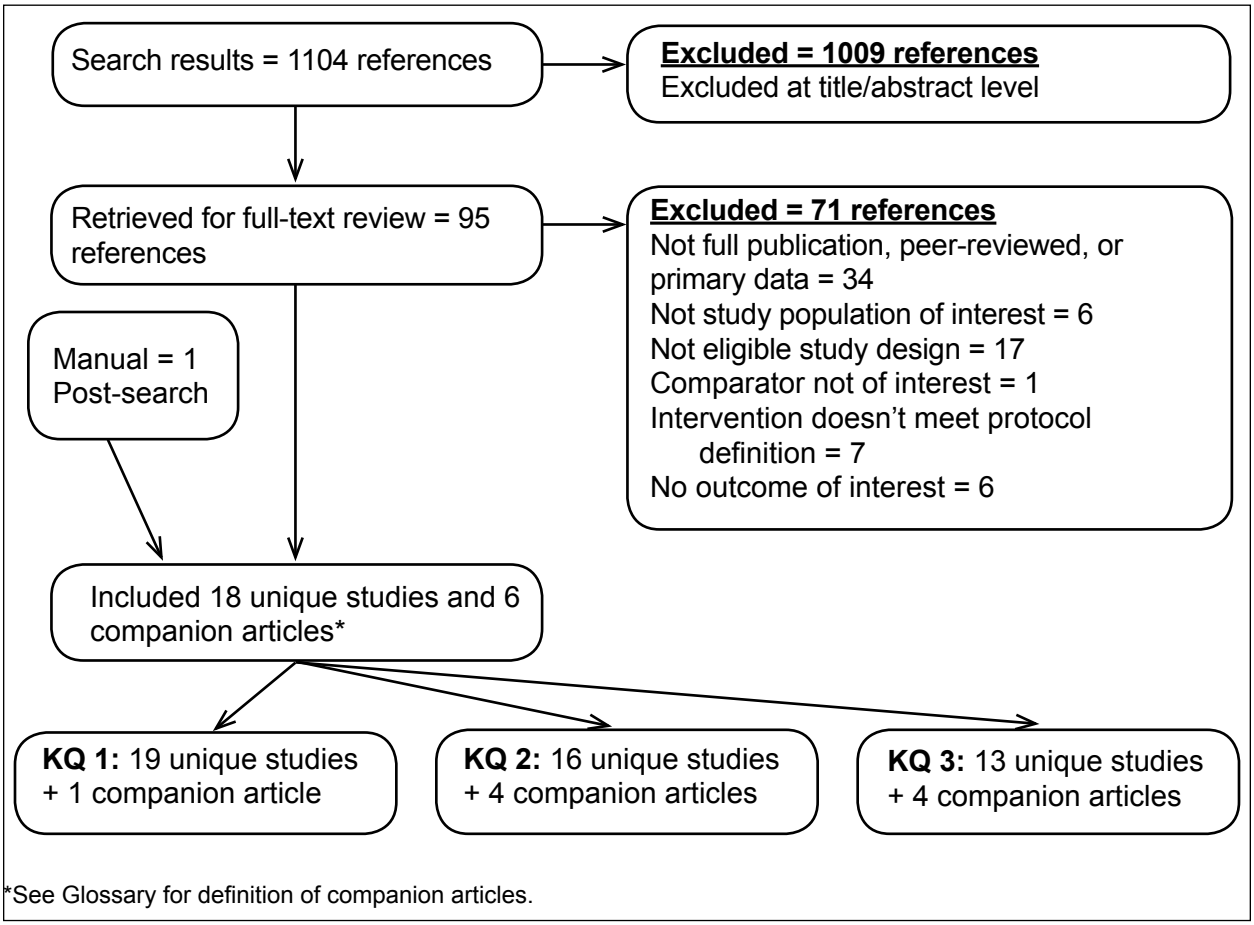
A draft version of the report was reviewed by technical experts and clinical leadership. A transcript of their comments can be found in Appendix E, which elucidates how each comment was considered in the final report.

RESULTS

LITERATURE FLOW

The flow of articles through the literature search and screening process is illustrated in Figure 2. We identified 1101 unique citations from a combined search of MEDLINE (via PubMed, n=323), CINAHL (n=290), Embase (n=145), PsycINFO (n=157) and the Web of Science (n=186). Manual searching of included study bibliographies and review articles identified 2 additional citations for a total of 1104 citations. After applying inclusion/exclusion criteria at the title-and-abstract level, 95 full-text articles were retrieved and screened. Of these, 70 were excluded at the full-text screening stage, leaving 25 articles (representing 19 unique studies) for data abstraction. All studies compared shared medical appointments with usual care or enhanced usual care; there were no direct comparisons between types of quality-improvement strategies. Our search of www.clinicaltrials.gov did not suggest publication bias. There were no completed studies that were unpublished. We found four ongoing studies (Appendix F), one of which had a methods paper. Interestingly, in light of the narrowness of the medical conditions in which SMA has been tested, one study is on patients with heart failure.

Figure 2. Literature flow diagram for randomized controlled trials and observational studies on SMA



Abbreviations: KQ=key question; SMA=shared medical appointment

STUDY CHARACTERISTICS

Of the 19 studies, 16 (13 trials) evaluated SMA interventions in patients with diabetes mellitus and 3 (2 trials) evaluated SMAs in older adults with high utilization of medical resources. Most studies were conducted in primary care settings that are part of integrated health systems in the United States (Table 2). Of the 19 studies, 15 reported outcomes at 1 year or later. Detailed study characteristics are given in Appendix G.

Table 2. Overview of studies evaluating SMA

Study Characteristic	Adults With Diabetes	Older Adults
N studies (participants)	16 (3221)[a]	3 (1851)
Mean age of sample: median (range)	60.8 (27 to 69.8)	74.1 (73.5 to 78.2)
Setting: N studies (participants) Primary care Medical Subspecialty	 13 (2232) 3 (989)	 3 (1851) 0
Health care system: N studies (participants) Government (VA, FQC) Private integrated system (HMO) University-affiliated clinic	 7 (771) 2 (892) 7 (1558)	 0 3 (1851) 0
Country: N studies (participants) United States Europe	 13 (2232) 3 (989)	 3 (1851) 0
Study design: N studies (participants) Randomized controlled trial Observational	 13 (2921) 3 (300)	 2 (615) 1 (1236)
Sites: N studies (participants) Single Multisite	 14 (2106) 2 (1115)	 1 (321) 2 (1530)
Study duration[b]: N studies (participants) 6 to 12 months >12 months	 4 (410) 12 (2811)	 0 3 (1851)

[a]Participant number is based on the number randomized.
[b]Study duration is measured from time of randomization to most distal followup.
Abbreviations: FQC=Federally qualified center; VA=Veterans Administration; HMO=health maintenance organization

Characteristics of Shared Medical Appointments

In the studies we assessed, SMAs were led by teams of 1 to 3 clinicians that included a physician (n=15), clinical pharmacists (n=9; the prescribing clinician in 3 studies) and a registered nurse (Table 3). The clinical team was multidisciplinary in most studies; pharmacists and licensed mental health professionals participated in almost half the studies. Sessions were designed for closed panels of patients in all but three studies; these later studies used drop-in models. Group size was 6 to 10 for most studies, with group size ranging between 10 and 20 in 4 studies and group size as large as 25 members in one study. The planned visit frequency ranged from approximately every 3 weeks to every 3 months. SMA visits were a median of 2 hours (range 1 to 3.5 hours).

At least 16 of 19 studies offered individual breakouts with a physician or clinical pharmacist as part of the SMA design specified that medication changes could be made at group visits. Three studies did not report this information. About half the studies invited participation by family members or friends. Three studies described the educational approach as "patient-centered adult learning,"[20-22] and two studies used the stages-of-change model to design the intervention;[8,26] no other study described a theoretical model. In about half the studies, patients participated in

selecting or prioritizing educational topics, and printed materials were tailored to the individual patient. Few studies used telephone contact as a part of the SMA intervention. Details of the SMA interventions are given in Appendix H.

Table 3. Characteristics of shared medical appointment interventions

Characteristic	Diabetes (16 Studies)	Older Adults (3 Studies)
Intervention team disciplines: N studies (participants)		
Medical doctor	12 (2731)	3 (1851)
Nurse practitioner	3 (298)	1 (1236)
Pharmacist	8 (1609)	1 (294)
Registered nurse	10 (2791)	2 (615)
Dietician	4 (1208)	0
Physical therapist/exercise specialist	3 (269)	1 (294)
Psychologist/Behavioral specialist	3 (326)	1 (321)
Health educator	3 (1116)	0
Social worker	2 (164)	0
Other[a]	6 (1238)	2 (1530)
Intervention team size		
Median number of members (range)	2.5 (2 to 7)	2.5 (2 to 3)
Average visit duration		
Median minutes (range)	120 (60 to 210)	120 (90 to 120)
Number of planned visits		
Median (range)	7.5 (4 to 36)	12 (12 to 24)
Medication changes made during sessions		
Yes	13 (2217)	3 (1851)
Not reported/unclear	3 (995)	0
Individual breakouts: N studies (participants)		
Yes	12 (2850)	3 (1851)
No	1 (88)	0
Not reported/unclear	3 (273)	0
Behavioral components		
Licensed mental health professional led group education session		
Yes	6 (1356)	1 (321)
No	9 (1149)	2 (1530)
Not reported/unclear	1 (707)	0
Family/friend participation		
Yes	7 (1346)	2 (615)
No	4 (527)	0
Not reported/unclear	5 (1169)	1 (1236)
Patient-clinician group continuity: N studies (participants)		
Group member continuity[b]		
Closed	13 (2942)	2 (1557)
Open/drop-in	3 (270)	1 (294)
Team continuity		
Consistent care team	12(2120)	2 (1530)
Care team changes/rotates	3 (989)	1 (321)
Not reported/unclear	1 (103)	0
Some intervention components delivered by telephone		
Yes	2 (424)	1 (312)
No	10 (2385)	1 (294)
Not reported/unclear	4 (403)	1 (1236)

[a]Disciplines that were present on only one team: occupational therapist, medical assistant, research assistant or undefined; there were no physician assistants used, although this would be a valid clinical discipline for teams.
[b]Group membership was classified as closed when the same group of patients were scheduled for each SMA visit.

Comparison Condition

In all studies, the comparison condition was some form of usual care. This care was inconsistently described. Three studies by one group[20-22] and one other study[18] used a structured or enhanced form of usual care. In one study,[22] this care consisted of individual visits with a forced interval of 3 months; in another study,[18] this was VA usual care supplemented with a single diabetes education session; and in the other two studies,[20,21] usual primary care was enhanced by one-on-one education sessions with the group facilitator. Three studies conducted in the VA[8,26,36] described usual care at some length, including average visit frequencies of 4 months, online clinical tools, electronic medical records with clinical reminders related to diabetes care, and a full range of referral services including diabetes education. Three other VA studies[5,14,15] very briefly described usual care. The other nine studies did not describe usual care at all.

KEY QUESTION 1: For adults with chronic medical conditions, do shared medical appointments (SMAs) compared with usual care improve the following:

- **Patient and staff experience?**
- **Treatment adherence?**
- **Quality measures such as (a) process of care measures utilized by VA, National Quality Forum, or National Committee for Quality Assurance and (b) biophysical markers (laboratory or physiological markers of health status such as HbA1c and blood pressure)?**
- **Symptom severity and functional status?**
- **Utilization of medical resources or health care costs?**

Effects of Shared Medical Appointments on Clinical, Process, and Economic Outcomes

The outcomes reported varied widely across studies and between studies for adults with diabetes and older adults. We describe the results separately for these two populations.

Effect of SMAs on Outcomes for Adults With Diabetes

Patient selection for SMA studies among patients with diabetes

Patient characteristics are reported in Table 4. Briefly, 10 of 15 studies required patients to be "out of control" with regard to their A1c; however, this inclusion floor varied from a low of 6.5% to a high of 9.0%. Four studies required elevated blood pressure, and two required elevated lipids. Other criteria were used by no more than two studies (e.g., efforts to assure that diabetes was type 2, insulin-requiring, high utilization in past year).

We identified 13 randomized trials that evaluated the effects of SMAs on outcomes for patients with diabetes.[3,8,14,15,17-22,26,36,37] Of these, ten enrolled only patients with type 2 diabetes,[3,8,14,15,17,18,20,22,26,36] two enrolled mixed samples,[19,37] and one enrolled only patients

with type 1 diabetes.[21] Three observational studies evaluated SMAs.[5,13,16] All but one of these 16 studies compared SMAs with usual care. One study[18] compared SMAs with a traditional, two-session, diabetes education intervention. Study quality was rated as good for 6 trials, fair for 6 trials and 2 observational studies, and poor for the two remaining studies. For trials, methodological problems included (1) failure to describe allocation concealment (n=9), (2) outcomes assessed without blinding to intervention (n=6), and (3) an inadequate approach to addressing incomplete data (n=6). Except for the study in patients with type 1 diabetes, patients were older adults with representative gender and racial mixes (Table 4).

Table 4. Study details for SMAs enrolling adults with diabetes

Characteristic	Randomized Trials	Observational studies
N studies (participants)	13 (2921)[a]	3 (300)
Median age of sample (range)[b]	60.8 (29 to 69.8)	59.4 (56.8 to 61.0)
Sex: N (%)		
Male	1585 (54.3%)	93 (31.0%)
Female	1137 (38.9%)	128 (42.7%)
Not reported (3 studies)	190 (6.8%)	79 (26.3%)
Race: N (%)[c]		
African American	425 (16.4%)	–
White	952 (36.7%)	–
Other	127 (4.9%)	–
Not reported	1088 (42.0%)	300 (100%)
Study quality: N (%)		
Good	6 (46%)	0
Fair	6 (46%)	2 (67%)
Poor	1 (8%)	1 (33%)

[a]Participant number is based on the number included in description of population characteristics, which is a smaller sample than those randomized.
[b]Mean age was not reported in one study.
[c]Of studies reporting race, 329 participants were not accounted for; therefore, percentage is of n=2592.

Treatment Experience and Adherence Outcomes

Only two trials[21,37] described the effects on patient experience, and none reported effects on staff experience. Neither of those trials showed greater satisfaction among those in SMAs compared with usual care. One study reported no effects on medication adherence,[3] another reported no effects on blood glucose self-monitoring,[20] and two studies reported mixed effects on self-management behaviors.[19,36] In both studies, patients in the SMA group increased the frequency of home glucose monitoring more than in the usual care group. Foot self-exams increased significantly in one study,[36] and exercise time increased by a statistically nonsignificant degree compared with usual care.

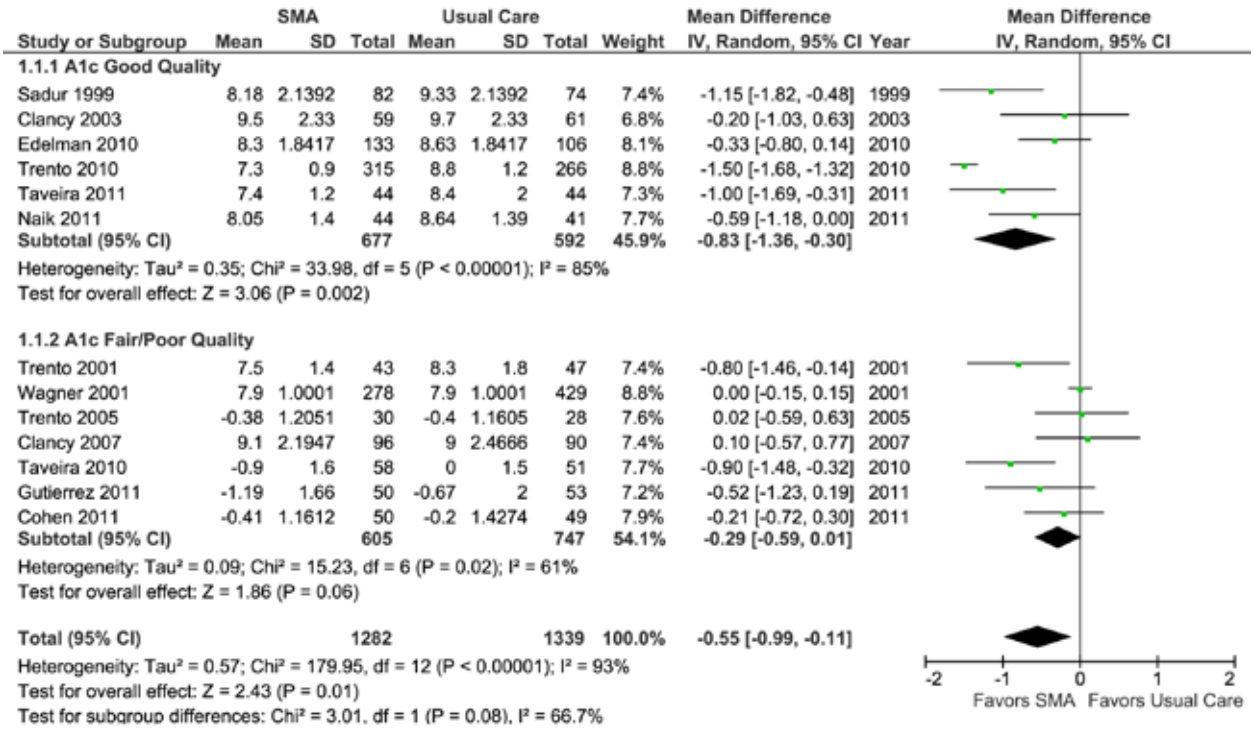
Effects on medication treatment were reported in 8 of 13 studies, but outcomes were reported inconsistently. One of four studies[26] reported more medication starts or dose titrations for oral hypoglycemic medications, and one of two studies[8] reported more insulin starts and increased insulin doses for the SMA group. One of three studies[26] found more antihypertensive medication starts or dose titrations overall in the SMA intervention group, and two studies[8,15] found greater use of dose titrations for selected antihypertensive medications. Only one of five studies[8] found a statistically significant increase in lipid-lowering medications and this was only for niacin. Most of the positive intervention effects were in studies led by clinical pharmacists. Patient or staff experience was not reported in any of the observational studies.

Biophysical Outcomes

*Hemoglobin A1c***.** Figure 3 shows the forest plots for the random-effects meta-analyses of the effect of SMAs on glucose. All studies reported effects on average glucose (A1c) at the end of the intervention, assessed at 6 months to 4 years. SMAs were associated with lower A1c than usual care (mean difference=-0.55; 95% CI, -0.99 to -0.11). However, effects varied significantly across studies (Q=179.9, df=12, $p < 0.001$; I^2 =93%)—variability that was not explained by study quality. Because of the variability in effects between studies, we conducted analyses to evaluate this variability. First, we conducted a sensitivity analysis, excluding the study in patients with type 1 diabetes,[21] but variability remained high (I^2=94%). Next, we used meta-regression analyses to evaluate the association between baseline A1c and intervention robustness with treatment effects. Neither baseline A1c nor intervention robustness (B=0.02 decrease in A1c per 1 point increase in robustness; CI, -0.23 to 0.26) was associated with treatment effects (p=0.90). Thus, SMAs were associated with a mean decrease in A1c, but effects varied markedly and were not explained by factors we hypothesized a priori to be associated with variation in treatment effect.

Effects of SMAs on glucose from the observational studies were generally consistent with the trial data. Two of the three observational studies[5,13] found statistically significant reductions in A1c from baseline to followup among patients participating in SMAs. Only one study[5] compared this change with a control group, finding a statistically significant benefit from SMA participation (p=0.002).

Figure 3. Effects of shared medical appointments on hemoglobin A1c



Cholesterol. Figures 4 and 5 show the forest plots for the random-effects analyses of the effect of SMAs on total cholesterol (5 studies) and LDL cholesterol (5 studies). For both outcomes, SMAs were associated with a statistically nonsignificant decrease in cholesterol. For each outcome, treatment effects varied significantly across studies. Because of the small number of studies, we did not complete meta-regression analyses to examine variability in treatment effects. One additional study[17] reported a statistically nonsignificant increase in the proportion of patients achieving an LDL of less than 100—findings that are consistent with the analysis of mean change in LDL. Only two of the observational studies reported effects on cholesterol. Both found reductions in LDL cholesterol, but only one[5] compared the SMA with the control group, and the differences were not statistically significant.

Figure 4. Effects of shared medical appointments on total cholesterol

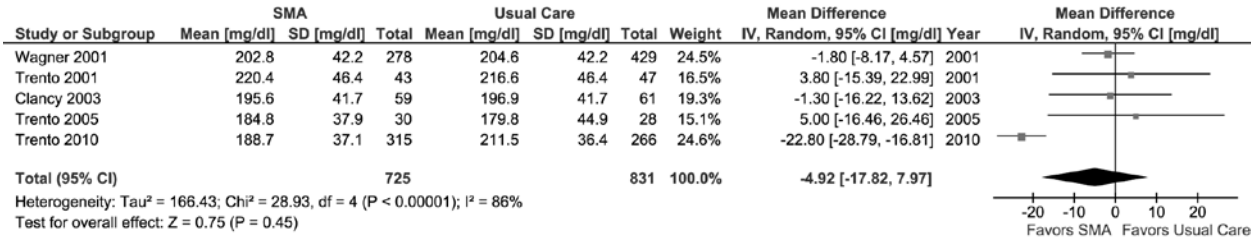
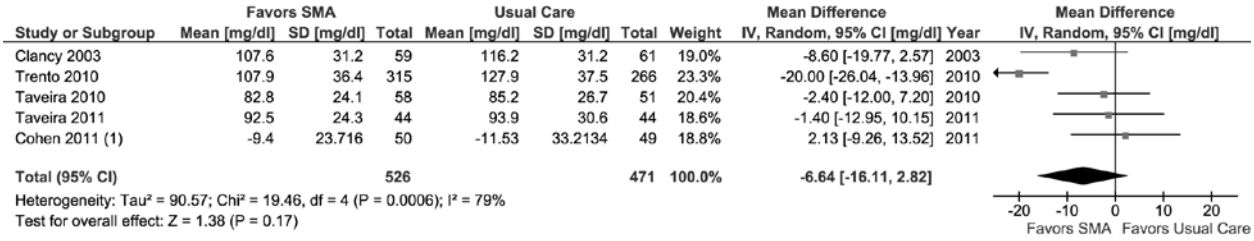
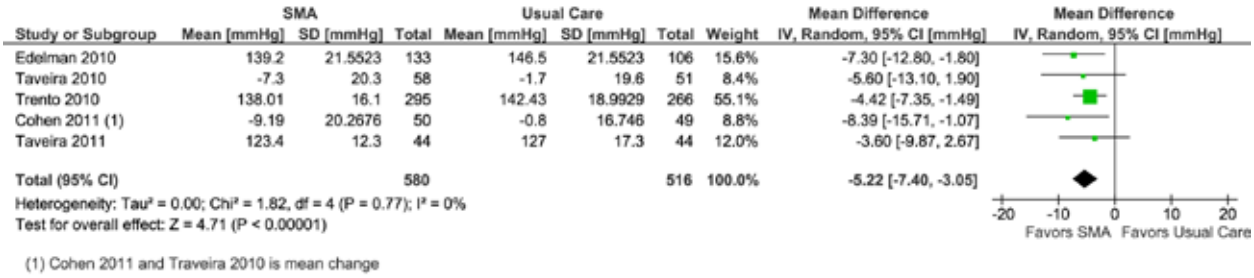


Figure 5. Effects of shared medical appointments on LDL cholesterol



Blood pressure. Figure 6 shows the forest plots for the random-effects analyses of the effect of SMAs on systolic blood pressure. Five studies reported effects on systolic blood pressure;[3,8,22,26,36] four of these were conducted in VA. SMAs were associated with improved blood pressure control (MD, -5.22; 95% CI, -7.40 to -3.05). Results were consistent across studies (Q=1.82, df=4, p=0.77, I^2=0%). Of the three observational studies, only one[5] found a statistically significant pre–post change in systolic blood pressure for the SMA participants. In this study, the blood pressure effects were also greater for the SMA group (-14.93 mmHg) than for the control group (-2.54 mmHg, p=0.04).

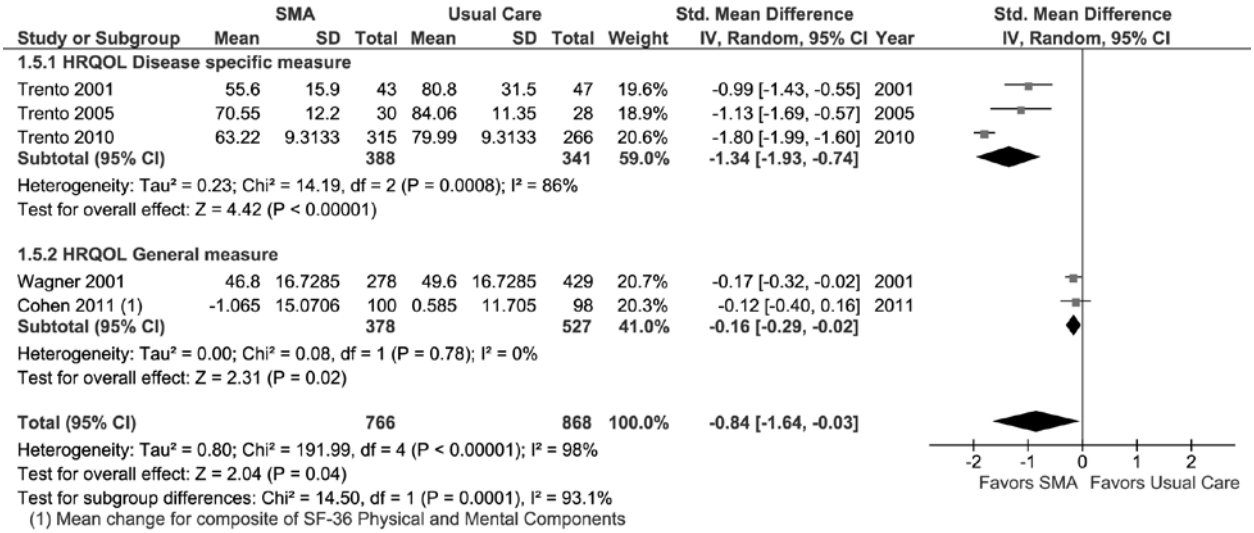
Figure 6. Effects of shared medical appointments on systolic blood pressure



Health-Related Quality-of-Life Outcomes

Figure 7 shows the random-effects meta-analysis of the effect of SMAs on health-related quality of life (HRQOL). Six studies[17,20-22,36,37] reported measuring HRQOL, but only five of these reported outcomes.[20-22,36,37] The studies by Trento et al. measured HRQOL with the Diabetes Quality-of-Life Measure, Cohen et al. reported the mental and physical components of the SF-36, and Wagner et al. reported the general health subscale of the SF-36. Because these measures differ, we analyzed the data using standardized mean difference. SMAs were associated with a large improvement in HRQOL (SMD -0.84; 95% CI, -1.64 to -0.03), but effects varied substantially across studies (Q=191.99, df=4, p<0.001; I^2=98%). There were too few studies to evaluate the variability in treatment effects quantitatively. However, the studies with the smallest effects[36,37] used general rather than disease-specific measures.

Figure 7. Effects of shared medical appointments on health-related quality of life



Economic Outcomes

Rates for hospital admissions and emergency department visits. The effect of SMAs on hospital admissions was reported in five studies.[3,14,19,26,37] Four studies reported admission rates involving 603 patients followed for 6 to 18 months. In three of these, admission rates were lower with SMAs, but the result was statistically significant in only one study.[19] The fifth study[37] followed 707 patients for 2 years and reported a statistically nonsignificant lower proportion of patients with a hospital admission who were randomized to SMAs (16.9% versus 21.0%, p=0.10).

Effects on emergency department visits were reported in the same five studies. Two studies reported significantly lower visit rates[3] or the proportion with an emergency department visit.[37] Rates were not significantly different in the other three studies. Observational studies did not report comparative effects on admission rates or emergency department visits.

Costs. Four studies reported effects on total costs, one in a large HMO,[37] two in a university-affiliated general medical clinic serving low-income patients,[14,15] and another in an Italian diabetes clinic.[20] Findings were mixed. In the largest trial testing a low-intensity intervention,[37] the total health care costs (excluding the clinical study personnel) did not differ significantly. The studies by Clancy et al.[14,15] tested more robust interventions. The earlier study found significantly higher total costs (inpatient, outpatient, and emergency department costs) for SMAs compared with usual care ($2,886 versus $1,490 per patient over six months; p=0.0003). Total costs were heavily influence by higher inpatient costs for the SMA group. In the later study, 1-year charges were significantly lower for the SMA group ($5,869 versus $8,412 per patient, p<0.05). Lower modeled charges were driven primarily by lower outpatient charges, in particular for specialty visits. The study by Trento et al.,[20] conducted in Italy, reports costs that may not be applicable to the U.S. health system. An evaluation that included staff costs, medications, and transportation costs for diabetes care showed a small increase for SMA patients ($597 versus $570 over 4 years, p=NR). Observational studies did not report comparative costs.

Effect of SMAs on Outcomes for Older Adults

Patient selection for SMA studies among older adults

Only three studies evaluated SMA interventions in older adults. Two of the four studies required a minimum age of 60; the other two used 65. All studies required some elevated use of health care in the past year; two operationalized that directly, while the third required a hospitalization in the past year.

We identified two randomized trials[9,11] that evaluated the effects of SMAs in 615 older adults with a recent hospitalization or other criteria for increased utilization. One observational study evaluated a similar population of 2251 older adults.[10] All studies were conducted in primary care, in group-model HMO settings in the United States, and compared SMAs with usual care. The mean age of participants ranged from 73.5 to 78.2 years of age. The most common chronic conditions were arthritis, hypertension, difficulty hearing, heart disease, liver disease, and bladder/kidney disease. All studies reported effects on utilization or costs at 1 year or greater. One trial was rated fair quality[11] and one poor quality;[9] the observational study was rated fair quality.[10] In the trial by Scott et al.,[11] only participants expressing a strong interest in group care (37% of those eligible) were randomized. Methodological problems included failure to

describe allocation concealment, outcomes assessed without blinding to intervention, and poor specification of outcome measures. Additional study details are in Appendix G.

The design of SMA visits was similar to the diabetes studies, except that fewer disciplines participated in the clinical teams. Detailed intervention descriptions are in Appendix H.

Treatment Experience and Adherence Outcomes

All studies reported a measure of patient experience. The two trials reported patient perceptions of quality of care, and both reported higher quality ratings with SMAs compared with usual care. In the study by Beck et al.,[9] more patients rated the overall quality of care as excellent (37% versus 27%, p=0.019), and Scott et al.[11] found that patients assigned to SMAs rated the quality of care 0.3 points higher on a 1-to-4 scale than usual care patients did (p=0.048). In the observational study, only SMA participants rated satisfaction, and 90 percent of participants reported satisfaction with four aspects of group visits, including the visit overall. In aggregate, these results support high levels of satisfaction with group visits among older adults. No study evaluated staff satisfaction using a validated measure, and no study reported comparative data on medication adherence. In the study by Levine et al.,[10] 90 percent of SMA providers agreed or strongly agreed that they felt a lot of satisfaction from group visits, and 50 percent endorsed that group visits enhance their practice. Beck et al.[9] reported that participants attended 55 percent of scheduled SMA visits. Among participants with a high interest in group visits, Scott et al.[11] reported 2 or fewer visits over 24 months by approximately 25 percent of patients.

Biophysical Outcomes

Biophysical outcomes were not reported, likely because patients were selected on the basis of age and health care utilization rather than a particular illness.

Health-Related Quality-of-Life Outcomes

Both trials reported effects on overall health status (via the Likert scale) and functional status using activities of daily living or instrumental activities of daily living; there were no differences in outcomes for any of these measures. Scott et al.[11] reported effects on HRQOL using a 10-point scale with 10 indicating the highest quality of life possible. Participants randomized to SMAs rated HRQOL higher at 24-month followup (mean score, SMA 7.2 [1.8] versus usual care 6.3 [2.0]; p=0.002). The single observational study did not reported effects on HRQOL or functional status.

Economic Outcomes

Rates for hospital admissions and emergency department visits. All studies showed fewer admissions in the SMA group, but the difference was statistically significant in only one study (mean admissions/patient, 0.44 [0.89] versus 0.82 [1.7]; p=0.013).[11] SMA visits were also associated with a statistically significant decrease in emergency department visits in both trials (mean difference in visit rates/year, 0.22 to 0.26); the observational study did not report emergency department visits. Other outpatient utilization was not significantly lower in the SMA groups. Primary care visits were not lower in any of the three studies, and only one of two studies[9] found significantly lower specialty visits.

Costs. The specific approach to cost analyses varied, but all studies included estimated costs of SMA visits. Total costs were lower for the SMA group in each study (range in mean difference in annual costs, -$178 to -$1599) but varied substantially across studies and did not reach statistical significance for any study. The two trials reported lower hospital costs, ranging from -$178/person per year (p=NR) to -$1145/person per year (p=0.07); the observational study did not report hospital costs. Other cost data were not reported consistently across studies.

KEY QUESTION 2: For adults with chronic medical conditions, do the effects of SMAs vary by patient characteristics (e.g., specific chronic medical conditions and severity of disease)?

We planned to address this question using two approaches, beginning with comparing the effects of SMAs across conditions. However, studies did not examine subgroups within their populations, and there was too little variability in diagnosis across studies for analysis—all condition-specific studies enrolled patients with diabetes. The single study enrolling adults with type 1 diabetes found similar treatment effects compared with those enrolling adults with type 2 diabetes. Second, we planned and conducted an evaluation of the association between treatment effects and baseline severity of disease. This analysis was possible only for the studies enrolling patients with diabetes. We used meta-regression analysis to examine the baseline association between A1c and treatment effects on glucose control. Baseline A1c was not associated with treatment effects (B=0.14 increase in A1c per 1 point increase in baseline A1c; 95% CI, -0.47 to 0.75; p=0.66). However, this analysis is limited by the relatively small number of studies, indirect comparisons, and potential for ecological fallacy since only the average baseline A1c for the study sample was available. A more robust approach would be a meta-analysis at the patient level, where baseline A1c is evaluated for each patient; however, these data were not available.

KEY QUESTION 3: Is the intensity of the intervention or the components used by SMAs associated with intervention effects?

Characteristics of the SMA interventions are summarized in Table 3 (KQ 1). Detailed descriptions for each study are given in Appendix H. As described in the Methods section, we developed a measure of intervention robustness based on seven intervention components. Two of the components (involving a behavioral health specialist or a medication change during SMA visits) were weighted double, and thus scores could range from zero to nine. For these analyses, we limited the sample to the trials in patients with diabetes and used A1c as the outcome, yielding a set of studies with similar characteristics except for the independent variable of interest (intervention robustness). We used meta-regression analyses to examine the relationship between robustness and intervention effects on A1c. For the 12 trials, robustness scores ranged from 3 to 8 (median=5). There was no association between intervention robustness score and treatment effects (B=0.02 decrease in A1c per 1-point increase in robustness score; 95% CI, -0.30 to 0.25; p=0.88).

SUMMARY AND DISCUSSION

SMAs have the potential to offer chronic disease care that is more efficient while improving staff satisfaction and patient outcomes. We identified 15 RCTs and 4 observational studies of varying quality comparing SMAs with usual care or enhanced usual care. Studies were conducted exclusively in patients with diabetes or in older adults with higher than average medical utilization. No eligible studies enrolled patients with the other chronic conditions of interest: coronary artery disease, chronic heart failure, asthma, chronic obstructive pulmonary disease, hyperlipidemia, or hypertension. This limited diversity in patient populations compromised our ability to determine if effects varied by condition. However, the included studies reported outcomes ranging from patient experience to biophysical and economic outcomes. These findings and the overall strength of evidence are summarized and discussed by key question.

SUMMARY OF EVIDENCE BY KEY QUESTION

Key Question 1

Few studies (0 to 3) reported effects on staff experience, patient experience, or treatment adherence. The strength of evidence for each of these outcomes was judged to be insufficient to estimate an effect of the SMA intervention in both patients with diabetes and older adults.

The most robust finding of this evidence synthesis is that SMAs for patients with diabetes appear to have a significant impact on biophysical outcomes. Hemoglobin A1c improved by approximately 0.6 percentage points, and systolic blood pressure by about 5 mmHg; both these findings were statistically significant. LDL-C improved by approximately 7 mg/dl, but this was not statistically significant. While each individual finding is only moderately robust given the limitations in study quality and unexplained variability in intervention effects, the constellation of findings taken together indicates that SMAs help intermediate clinical outcomes for type 2 diabetes. Similar outcomes were not reported in older adults.

For patients with diabetes, there was significant improvement on HRQOL, measured in 3 of 4 studies with a relatively sensitive, disease-specific, quality-of-life scale. Positive effects on HRQOL were found in one trial conducted in older adults, but functional status was not affected in these studies. Studies in older adults show a pattern of lower health care utilization, but the number of studies and participants are relatively few and these results should be considered preliminary. In patients with diabetes, lower hospitalization was the most consistent effect, but effects on other economic outcomes were too preliminary to estimate an effect. Our judgments about the strength of evidence (SOE) prioritized data from RCTs.

Table 5. Summary of the intervention effects and SOE for KQ 1

Population	Number of Studies[a] (Subjects)	Domains Pertaining to SOE				SOE
		Risk of Bias: Study Design/ Quality	Consistency	Directness	Precision	Effect Estimate (95% CI)
Staff experience						**Insufficient**
Diabetes	0	NA	NA	NA	NA	Not estimable
Older adults	1 (1236)	Obs/Fair	NA	Direct	Imprecise	Not estimable
Patient experience						**Insufficient**
Diabetes	2 (769)	RCT/Fair	Consistent	Direct	Imprecise	No effect
Older adults	2 (444)	RCT/Fair	Inconsistent	Direct	Imprecise	Small to large positive effect
Treatment adherence						**Insufficient**
Diabetes	3 (536)	RCT/Fair	Some inconsistency	Direct	Imprecise	Not estimable
Older adults	0	NA	NA	NA	NA	Not estimable
Biophysical						
Diabetes: A1c	13 (2921)	RCT/Good	Inconsistent	Direct	Some imprecision	MD = -0.55 (-0.99 to -0.11) Moderate SOE
Diabetes: Total Cholesterol	5 (1556)	RCT/Fair	Inconsistent	Direct	Imprecise	MD = -4.9 (-17.8 to 7.9) Low SOE
LDL Cholesterol	5 (997)	RCT/Fair	Inconsistent	Direct	Imprecise	MD -6.6 (-16.1 to 2.8) Low SOE
Diabetes: Blood pressure	5 (1125)	RCT/Good	Consistent	Direct	Some imprecision	MD = -5.2 (-7.4 to -3.1) Moderate SOE
Older adults	0	NA	NA	NA	NA	Not estimable
Health-related quality of life or functional status						
Diabetes	5 (1561)	RCT/Fair	Inconsistent	Direct	Imprecise	SMD = -0.84 (-1.6 to -0.03) Low SOE
Older adults	2 (615)	RCT/Fair	Inconsistent	Direct	Imprecise	Not estimable
Economic						
Diabetes	5 (1339)	RCT/Good	Inconsistent	Direct	Imprecise	*ED visits* lower rates in 2 of 5 studies Insufficient SOE
	5 (1339)	RCT/Good	Consistent	Direct	Some imprecision	*Hospitalizations* lower in 4 of 5 studies Low SOE
	4 (1125)	RCT/Fair	Inconsistent	Direct	Imprecise	*Total costs* range from lower to higher Insufficient SOE

Population	Number of Studies[a] (Subjects)	Domains Pertaining to SOE				SOE
		Risk of Bias: Study Design/ Quality	Consistency	Directness	Precision	Effect Estimate (95% CI)
Older adults	2 (615)	RCT/Fair	Consistent	Direct	Imprecise	*ED visits* lower rates in 2 of 2 studies Low SOE
	2 (615)	RCT/Fair	Some inconsistency	Direct	Imprecise	*Hospitalizations* lower in 1 of 2 studies Insufficient SOE
	2 (615)	RCT/Fair	Inconsistent	Direct	Imprecise	*Total costs* lower but not statistically significant Insufficient SOE

[a]Studies (subjects) given are for randomized trials; observational studies were also considered in SOE ratings but are not listed separately in the table.

Abbreviations: CI=confidence interval; ED=emergency department; MD=mean difference; NA=not applicable; RCT=randomized controlled trial; RD=risk difference; RR=risk ratio; SMD=standardized mean difference; SOE=strength of evidence

Key Questions 2 and 3

No studies explored KQ 2 (identifying the subgroups of patients that would benefit most from an SMA intervention) or KQ 3 (identifying the specific components of an SMA intervention that were most potent). We devised a robustness score to attempt to address KQ 3, but it was not able to discriminate degrees of effectiveness among intervention components. More than 70 percent of all studies were similar on six of the seven variables used in the robustness score: (1) whether the team was continuous, (2) whether the group was closed, (3) whether individual breakout sessions were conducted, (4) whether medication changes were made, (5) how long each session was, and (6) whether there was contact outside the session. It is possible that there are other more important variables that are not being measured with current approaches. The strength of evidence for both questions was judged to be insufficient.

CLINICAL AND POLICY IMPLICATIONS

A key finding is that SMAs have been evaluated primarily in patients with diabetes, and to a lesser extent and with a narrow range of outcomes for older adults with high utilization. Even where the data are more robust in those with diabetes, it is challenging to place into context the improvements seen in biophysical parameters with SMAs. However, we can discuss the clinical importance of these findings in at least two ways. First, we can compare the results to clinical trial data relative to starting any agent for these conditions. The improvement seen in one year on systolic blood pressure across all arms of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT), trial, after adding the chosen first medication, was approximately 6.6 mmHg; patients in SMAs achieved approximately 75 percent of that level of improvement.[38] Similarly, adding a first-line oral hypoglycemic agent at a maximally tolerated dose usually lowers A1c by 1 to 1.5 percentage points;[38] patients in SMAs achieve 33 to 50 percent of that goal. The change in LDL-C of 7 mg/dl is much smaller compared with drug effect, approximately 15 percent of what would be expected with clinical trial doses of an HMG-CoA reductase inhibitor ("statin"). However, each drug comparison

is made relative to placebo controls. For SMA interventions, the comparator is usual care, which typically includes medication treatment, and thus one would expect the effects to be smaller.

Another way to evaluate the improvements observed with SMA is against the known standard deviations for the outcomes in the population of patients with disease, and then calculate effect sizes. While many different values for standard deviations for the relevant parameters are reported in the literature, effect sizes of SMA interventions for systolic BP, A1c, and LDL-C are approximately 0.5, 0.33, and 0.25, respectively. These are considered moderate to small effect sizes, but all would be considered important.[39]

The improvements in A1c and blood pressure, and the more modest improvement in LDL-C are possibly synergistic, or at least additive, in prevention of the macrovascular and microvascular complications of diabetes.[40] Thus, as a whole, SMAs may impact the risk of complications among patients with diabetes. Even if half the effect were lost in translation due to lower treatment fidelity when implemented outside of clinical trials, there would still likely be an important improvement in complication risk for patients enrolled in a diabetes SMA intervention. However, it is important to remember that the degree of synergy in the context of improvements in multiple outcomes is guesswork at best; SMAs—and indeed multicomponent health services in general—have not been studied with enough patients to determine their actual effects on major cardiovascular or microvascular complications.

Finally, many authors propose that SMAs are more satisfying than standard outpatient visits for both patients and providers, but few have measured patient and staff satisfaction. Because SMAs are a major shift in clinic organization, more data are needed on these variables as well as cost-to-benefit ratios before a general policy recommendation can be made.

Generalizability of Findings

The results of the diabetes studies have limitations to their external validity. Using the PICOTS framework (population, intervention, comparator, outcome, timing, setting), the applicability of the findings appears strong with respect to (1) population because a reasonable balance of race and sex was achieved among patients, (2) outcomes because there is general consensus that A1c, blood pressure, and LDL-C are the important outcomes in diabetes, and (3) timeframes because there is general consensus that improvement of 6 months or longer is clinically relevant. However, none of the studies examined maintenance of effect after the intervention ended. Although similar in many aspects, there were enough differences in intervention process that a conclusion as to what makes an SMA intervention particularly successful could not be drawn. In addition, what constituted usual care was inconsistently defined. Therefore, intervention heterogeneity and the types of usual care comparators, may also be important limitations to the generalizability of our results.

The heterogeneity of the studies is concerning. Complex health services interventions are often a black box; that is, they contain many components that are hard to capture and tease out even in a well-conducted analysis. If there was a particular aspect of these interventions that was critical to predicting improved clinical outcomes, we were unable to capture that with the available data. This raises the question of a possible uncaptured element of SMAs that is important for potency, effectiveness, or generalizability. Without further, more mechanistic studies that attempt

to elucidate the key components of an SMA intervention, implementation of a diabetes SMA or design of an SMA for another condition will be at least partially based on reasoned judgment rather than strict evidence-based decision making.

An additional concern is that none of these studies was conducted in "real world" settings. All of the diabetes studies were conducted in academic, government, or vertically integrated systems. There are two potential reasons for this. First, all complex chronic care redesign interventions are easier to implement in systems that are either highly controlled or in which there is interest in research. Second, SMAs are difficult to implement in fee-for-service, independent clinics because they are unlikely to derive any financial benefits from improved quality of care but would have to absorb the cost in time and money of implementing the SMA. It is possible that this barrier could be relieved by Accountable Care Organizations, but this theory is still untested. Lastly, academic, government, or vertically integrated systems may also have very high quality usual care. While factors related to setting may not negatively impact the generalizability of these results for implementation of diabetes or other SMAs within vertically integrated systems such as the VA, they do suggest caution when considering the use of SMAs outside such systems.

Should SMAs Be Implemented?

The clearest finding of this evidence synthesis is that the existing knowledge base does not provide enough evidence to make a strictly evidence-driven decision about implementation of SMAs in any context except diabetes. Regarding diabetes SMA implementation, this evidence synthesis raises several key issues summarized in Table 6.

Table 6. Implementation issues

Issue for Implementation of Diabetes SMAs	Potential Solution
Enrollment There was no clear indication of which patients will receive the most benefit from this alternative structure.	Allow selection criteria for SMAs to fit specific local needs.
Elements of intervention There was no clear indication of which elements were most effective.	Use the most prevalent common elements: Prescribing clinician A consistent clinical lead for the SMA group At least 3 team members Closed group participants Individual time with clinician (brief) Medications evaluated Group duration of 90 to 120 minutes Variable elements that could be tailored to clinic or patient population: Group size Participation of family and friends Contact with participants outside of group
Potential mechanisms of intervention Very few studies reported any intermediate or mechanistic outcomes such as self-management, medication change, or access to care.	Measure these at implementation; use Plan-Do-Study-Act approach to allow these factors to change intervention over time.

Issue for Implementation of Diabetes SMAs	Potential Solution
Infrastructure changes There was no clear indication whether the change in clinic structure was more effective or efficient.	In already vertically integrated settings, such as VA, these changes are not as difficult. The broad-based improvement seen was clinically meaningful balanced against satisfaction and cost, especially for older adults.

SMAs also have costs. These costs are not just the labor cost of redirecting providers away from their existing clinical responsibilities to conduct an SMA; there is also the time and labor cost to establish a new structure for care. This lack of information about both direct costs and changes in utilization in all but older adults who are high utilizers of the health care system is a key gap in the existing literature. For those patients, hospital admissions, ER visits and total costs were consistently lower with SMAs. Also, implementation of SMAs will not succeed if either patients or providers are unsatisfied with the new structure, and effects on patient and staff experience remain largely unknown, again with the exception of older adults who expressed increased satisfaction with SMAs.

STRENGTHS AND LIMITATIONS

Our study has a number of strengths, including a protocol-driven review, a comprehensive search, careful quality assessment, and rigorous quantitative synthesis methods. Our report, and the literature, also has limitations. An important limitation is the lack of breadth to the types of patients and illnesses that have been studied in the context of an SMA. The evidence synthesis found no explicit data regarding system-level, as opposed to patient-level, benefits of SMAs; the fact that as many studies viewed the SMA as an add-on to, rather than a replacement for, usual primary care suggests that improvements in access may not be as great as desired. In addition, the components of the interventions were often not described adequately for replication, especially the content of the group education time. Finally, outcomes reported varied substantially across studies and our attempts to explain the observed variability in intervention effects were unsuccessful. With unexplained variability, summary measures of treatment effect may not adequately describe the expected effects of the intervention.

RECOMMENDATIONS FOR FUTURE RESEARCH

We used the framework recommended Robinson et al.[41] to identify gaps in evidence and classify why these gaps exist (Table 7). The next generation of research in SMAs for patients with diabetes and other conditions should close the gaps outlined in the previous section.

Table 7. Evidence gaps and future research

Evidence Gap	Reason	Type of Studies to Consider
Patients		
Absence of data for patients with conditions other than diabetes mellitus and high utilization	Insufficient information	Single and multisite RCTs Quasi-experimental studies
Interventions		

Evidence Gap	Reason	Type of Studies to Consider
Uncertain which elements of an SMA intervention are most effective and efficient	Insufficient information	RCTs of head-to-head comparisons of different types of SMAs; Disaggregation trials
Outcomes		
Uncertain effects on patient and staff satisfaction	Insufficient information	Nonrandomized or cluster randomized, multisite implementation studies, qualitative studies
Uncertain effects on physiological variables other than HbA1c	Insufficient information	Large scale RCTs Nonrandomized, cluster controlled trials, controlled before-and-after studies, interrupted time series
Uncertain effects on health system costs with the exception of the elderly high utilizers of the health system	Insufficient information	Costs analyses
Uncertain whether there would be unintended consequences to other aspects of the health care system if SMAs were implemented	Insufficient information	Multisite observational studies

Abbreviation: RCT=randomized controlled trial

Our review shows that SMAs, typically using closed panels with individual breakouts and the opportunity for medication management, help intermediate clinical outcomes for type 2 diabetes. A smaller literature shows positive effects on patient experience in older adults and the possibility of lower health care utilization. SMAs may be most effective for illnesses such as diabetes that have a phase in which the risk of complication is relatively high while the disease is simultaneously asymptomatic, and in which medication titration and self-management are important. Until further studies are done that allow for comparisons across conditions, the targeting of SMA for chronic conditions other than diabetes will remain speculative. Finally, repeating the existing diabetes SMA efficacy trials in fee-for-service settings would be important to understand the extent to which SMAs work when the profit motive is essential to the practice model.

REFERENCES

1. Fletcher RD, Amdur RL, Kolodner R, et al. Blood Pressure Control Among US Veterans: A Large Multi-Year Analysis of Blood Pressure Data from the VA Health Data Repository. *Circulation*. 2012.
2. King H, Aubert RE, Herman WH. Global burden of diabetes, 1995-2025: prevalence, numerical estimates, and projections. *Diabetes Care*. 1998;21(9):1414-31.
3. Edelman D, Fredrickson SK, Melnyk SD, et al. Medical clinics versus usual care for patients with both diabetes and hypertension: a randomized trial. *Ann Intern Med*. 2010;152(11):689-96.
4. Jaber R, Braksmajer A, Trilling JS. Group visits: A qualitative review of current research. *Journal of the American Board of Family Medicine*. 2006;19(3):276-290.
5. Kirsh S, Watts S, Pascuzzi K, et al. Shared medical appointments based on the chronic care model: a quality improvement project to address the challenges of patients with diabetes with high cardiovascular risk. *Qual Saf Health Care*. 2007;16(5):349-53.
6. Noffsinger E. Increasing quality care and access while reducing cost through drop-in group medical appointments (DIGMAs). *Group Pract J*. 1999;48(1):12-18.
7. Riley SB, Marshall ES. Group visits in diabetes care: a systematic review. *Diabetes Educ*. 2010;36(6):936-44.
8. Taveira TH, Friedmann PD, Cohen LB, et al. Pharmacist-led group medical appointment model in type 2 diabetes. *Diabetes Educ*. 2010;36(1):109-17.
9. Beck A, Scott J, Williams P, et al. A randomized trial of group outpatient visits for chronically ill older HMO members: the Cooperative Health Care Clinic. *J Am Geriatr Soc*. 1997;45(5):543-9.
10. Levine MD, Ross TR, Balderson BH, et al. Implementing group medical visits for older adults at group health cooperative. *J Am Geriatr Soc*. 2010;58(1):168-72.
11. Scott JC, Conner DA, Venohr I, et al. Effectiveness of a group outpatient visit model for chronically ill older health maintenance organization members: a 2-year randomized trial of the cooperative health care clinic. *J Am Geriatr Soc*. 2004;52(9):1463-70.
12. Coleman EA, Eilertsen TB, Kramer AM, et al. Reducing emergency visits in older adults with chronic illness. A randomized, controlled trial of group visits. *Eff Clin Pract*. 2001;4(2):49-57.
13. Bray P, Thompson D, Wynn JD, et al. Confronting disparities in diabetes care: the clinical effectiveness of redesigning care management for minority patients in rural primary care practices. *J Rural Health*. 2005;21(4):317-21.
14. Clancy DE, Cope DW, Magruder KM, et al. Evaluating concordance to American Diabetes Association standards of care for type 2 diabetes through group visits in an uninsured or inadequately insured patient population. *Diabetes Care*. 2003;26(7):2032-6.

15. Clancy DE, Huang P, Okonofua E, et al. Group visits: promoting adherence to diabetes guidelines. *J Gen Intern Med*. 2007;22(5):620-4.

16. Culhane-Pera K, Peterson KA, Crain AL, et al. Group visits for Hmong adults with type 2 diabetes mellitus: a pre-post analysis. *J Health Care Poor Underserved*. 2005;16(2):315-27.

17. Gutierrez N, Gimple NE, Dallo FJ, et al. Shared medical appointments in a residency clinic: an exploratory study among Hispanics with diabetes. *Am J Manag Care*. 2011;17(6 Spec No.):e212-4.

18. Naik AD, Palmer N, Petersen NJ, et al. Comparative effectiveness of goal setting in diabetes mellitus group clinics: randomized clinical trial. *Arch Intern Med*. 2011;171(5):453-9.

19. Sadur CN, Moline N, Costa M, et al. Diabetes management in a health maintenance organization. Efficacy of care management using cluster visits. *Diabetes Care*. 1999;22(12):2011-7.

20. Trento M, Passera P, Tomalino M, et al. Group visits improve metabolic control in type 2 diabetes: a 2-year follow-up. *Diabetes Care*. 2001;24(6):995-1000.

21. Trento M, Passera P, Borgo E, et al. A 3-year prospective randomized controlled clinical trial of group care in type 1 diabetes. *Nutr Metab Cardiovasc Dis*. 2005;15(4):293-301.

22. Trento M, Gamba S, Gentile L, et al. Rethink Organization to iMprove Education and Outcomes (ROMEO): a multicenter randomized trial of lifestyle intervention by group care to manage type 2 diabetes. *Diabetes Care*. 2010;33(4):745-7.

23. Krzywkowski-Mohn SM. Diabetic control and patient perception of the scheduled in group medical appointment at the Cincinnati Veterans Administration Medical Center. ProQuest Information & Learning; 2009.

24. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.

25. Wilczynski NL, McKibbon KA, Haynes RB. Response to Glanville et al.: How to identify randomized controlled trials in MEDLINE: ten years on. *J Med Libr Assoc*. 2007;95(2):117-8; author reply 119-20.

26. Taveira TH, Dooley AG, Cohen LB, et al. Pharmacist-led group medical appointments for the management of type 2 diabetes with comorbid depression in older adults. *Ann Pharmacother*. 2011;45(11):1346-55.

27. Loney-Hutchinson LM, Provilus AD, Jean-Louis G, et al. Group visits in the management of diabetes and hypertension: effect on glycemic and blood pressure control. *Current Diabetes Reports*. 2009;9(3):238-242.

28. Kirsh SR, Aron DC. Integrating the chronic-care model and the ACGME competencies: using shared medical appointments to focus on systems-based practice... Accreditation Council for Graduate Medical Education. *Quality & Safety in Health Care*. 2008;17(1):15-19.

29. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality. Available at: http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318. Accessed May 10, 2012.

30. Viswanathan M, Berkman ND. Development of the RTI Item Bank on Risk of Bias and Precision of Observational Studies. Methods Research Report. (Prepared by the RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290-2007-0056-I.) AHRQ Publication No. 11-EHC028-EF. Rockville, MD: Agency for Healthcare Research and Quality. Available at: http://www.effectivehealthcare.ahrq.gov/ehc/products/350/784/RTI-Risk-of-Bias_Final-Report_20110916.pdf. Accessed May 15, 2012.

31. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-58.

32. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol*. 2011;64(12):1283-93.

33. Schunemann HJ, Oxman AD, Brozek J, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*. 2008;336(7653):1106-10.

34. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-34.

35. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ*. 2001;323(7304):101-5.

36. Cohen LB, Taveira TH, Khatana SA, et al. Pharmacist-led shared medical appointments for multiple cardiovascular risk reduction in patients with type 2 diabetes. *Diabetes Educ*. 2011;37(6):801-12.

37. Wagner EH, Grothaus LC, Sandhu N, et al. Chronic care clinics for diabetes in primary care: a system-wide randomized trial. *Diabetes Care*. 2001;24(4):695-700.

38. Fowler MJ. Diabetes Treatment, Part 2:Oral Agents for Glycemic Management. Clinical Diabetes. October 2007 vol. 25 no. 4 131-134.

39. Cohen J. *Statistical Power Analysis for the Behavioral Sciences (second ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

40. Sehestedt T, Hansen TW, Li Y, et al. Are blood pressure and diabetes additive or synergistic risk factors? Outcome in 8494 subjects randomly recruited from 10 populations. *Hypertens Res*. 2011;34(6):714-21.

41. Robinson KA, Saldanha IJ, Mckoy NA. Frameworks for Determining Research Gaps During Systematic Reviews. Methods Future Research Needs Report No. 2. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. HHSA 290-2007-10061-I.) AHRQ Publication No. 11-EHC043-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2011. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm. Accessed May 22, 2012.