ABSTRACT

PURPOSE This study was a systematic review with a quantitative synthesis of the literature examining the overall effect size of practice facilitation and possible moderating factors. The primary outcome was the change in evidence-based practice behavior calculated as a standardized mean difference.

METHODS In this systematic review, we searched 4 electronic databases and the reference lists of published literature reviews to find practice facilitation studies that identified evidence-based guideline implementation within primary care practices as the outcome. We included randomized and nonrandomized controlled trials and prospective cohort studies published from 1966 to December 2010 in English language only peer-reviewed journals. Reviews of each study were conducted and assessed for quality; data were abstracted, and standardized mean difference estimates and 95% confidence intervals (CIs) were calculated using a random-effects model. Publication bias, influence, subgroup, and meta-regression analyses were also conducted.

RESULTS Twenty-three studies contributed to the analysis for a total of 1,398 participating practices: 697 practice facilitation intervention and 701 control group practices. The degree of variability between studies was consistent with what would be expected to occur by chance alone ($I^2 = 20\%$). An overall effect size of 0.56 (95% CI, 0.43-0.68) favored practice facilitation ($z = 8.76; P < .001$), and publication bias was evident. Primary care practices are 2.76 (95% CI, 2.18-3.43) times more likely to adopt evidence-based guidelines through practice facilitation. Meta-regression analysis indicated that tailoring ($P = .05$), the intensity of the intervention ($P = .03$), and the number of intervention practices per facilitator ($P = .004$) modified evidence-based guideline adoption.

CONCLUSION Practice facilitation has a moderately robust effect on evidence-based guideline adoption within primary care. Implementation fidelity factors, such as tailoring, the number of practices per facilitator, and the intensity of the intervention, have important resource implications.

INTRODUCTION

There are many challenges to the adoption of evidence-based guidelines into the clinical practice of primary care physicians, and a consensus has emerged from the literature that having knowledge is rarely sufficient to change practice behavior. Didactic education or passive dissemination strategies are ineffective, whereas interactive education, reminder systems, and multifaceted interventions have a greater effect. Outreach or practice facilitation is a multifaceted approach that involves skilled individuals who enable others, through a range of intervention components and approaches, to address the challenges in implementing evidence-based care guidelines within the primary care set-
Nagykaldi et al. in 2005 conducted a systematic review of practice facilitation and found through a narrative summary of effects that practice facilitation increased preventive service delivery rates, assisted with chronic disease management, and implemented system-level improvements within practice settings.

In this meta-analysis, we examined the overall effect size of practice facilitation using a quantitative synthesis of the literature. We included in the analysis studies that described the intervention as outreach or practice facilitation for the implementation of evidence-based practice guidelines within primary care practice settings. The quantitative analyses were undertaken to describe the range and distribution of effects across studies, to explore probable explanations of the variation, and to demonstrate results quantitatively compared with the descriptive systematic reviews that have been done to date on practice facilitation.

**METHODS**

**Study Design and Primary Outcome**

This study was a systematic review with a quantitative synthesis of the literature examining the overall effect size of practice facilitation and possible moderating factors. The primary outcome was the change in evidence-based practice behavior calculated as a standardized mean difference. We used the guidelines outlined in the PRISMA statement for reporting systematic reviews and meta-analyses and applied the methods of the Cochrane Collaboration.

**Inclusion Criteria and Selection of Studies**

The literature review focused solely on controlled trials or evaluations of facilitation within health care, where an explicit facilitator role was adopted to promote changes in clinical practice. The definition provided by Kitson and colleagues was used to determine study eligibility—a facilitator is an individual carrying out a specific role, either internal or external to the practice, aimed at helping to get evidence-based guidelines into practice. We built on the review of 25 studies conducted by Nagykaldi et al from 1966 to 2004 by adding the following inclusion criteria for study selection: English language only, peer-reviewed journals from December 2004 to December 2010, an intervention study using practice facilitation to improve the adoption of evidence-based practice, and a controlled trial (randomized or not) or a pre- and post-intervention cohort study.

One author (N.B.B.) conducted a systematic literature search on February 1, 2011, using MEDLINE and the Thomson Scientific Web of Science database, which contains the Science Citation Index, the Social Sciences Citation Index, and the Arts and Humanities Citation Index. The following key word search was used:

(Primary care or family medicine or general practice or family physician or practice-based research or audit or prevent or quality improvement or practice enhancement or practice-based education or evidence based or office system) and (facilitator or facilitation) and (controlled trial or clinical trial or evaluation)

The references from the published systematic reviews of practice facilitation, the references from retrieved articles, and other secondary sources that met the inclusion criteria were also consulted to supplement articles found through the initial literature search.

Initial screening of the identified articles was based on their titles and abstracts and conducted by one author (N.B.B.). Two authors (N.B.B., C.L.) and an assistant reviewed in more detail studies that could not be excluded based on the abstract alone to determine whether they met the inclusion criteria.

**Quality Assessment**

Given that no critical appraisal reference standard tool exists, we used a modified version of the Physiotherapy Evidence-Based Database (PEDro) method, which consisted of 12 criteria, each receiving either a yes (reported) or no (not reported) score, for assessing the risk of bias of practice facilitation studies. Compared with the Jadad assessment criteria, PEDro has been shown to provide a more comprehensive picture of methodological quality for studies in which double-blinding is not possible. We added an adequate intervention description and adjusting for interclass correlation (ICC) to the scale, because unit of analysis errors have been identified as a methodological problem in the implementation research literature. The protocol covered the study characteristics considered key by the Cochrane Collaboration of methods, participants, interventions, outcome measures, and results.

Two authors (N.B.B., C.L.) and an assistant independently rated all included studies (n = 45) using the same protocol, and discrepancies were resolved by consensus with the inclusion of a fourth rater (W.H.). Interrater reliability between the 3 raters was assessed to be very good, Fleiss’ $\kappa = 0.78$ (95% CI, 0.73-0.84). With a maximum score of 12, we considered studies from the 44 that had a total quality score of 6 (the average score) or greater (mean = 5.57, 95% CI, 4.79-6.35) to be of high quality.

**Data Analysis and Effect Size Determination**

Selected study measures, such as participation rates and attributes of participating practices, were summarized across all studies descriptively using measures...
of central tendency for continuous data and frequencies for categorical and binomial data. The change in the primary outcome measure from preintervention to postintervention assessment for each study was ascertained by determining the difference between the practice facilitation and comparison group postintervention and the difference from baseline for prospective cohort studies. All statistics were computed using SPSS 18.0 and Comprehensive Meta-Analysis software.31,32

**Effect Sizes**

The standardized mean difference (SMD) for the primary outcome (as identified by the authors of the study) of selected high methodological performance studies was computed using Hedges’ (adjusted) 𝑔.24 Cohen’s categories were used to evaluate the magnitude of the effect size, calculated by the standardized mean difference, with 𝑔 < 0.5 as a small effect size; 𝑔 ≥ 0.5 and ≤ 0.8, medium effect size; and 𝑔 > 0.8, large effect size. When the primary outcome was unspecified or more than 1, the median outcome was selected to calculate the standardized mean difference.33 Methods for determining standard deviations from confidence intervals and 𝑃 values were used when standard deviations were not provided.24 For studies in which the unit of analysis and the unit of randomization did not agree,34–36 verification that the intraclass correlation was taken into consideration was done to avoid including potentially false-positive results.37–39 For studies that provided results for primary outcomes only as odds ratios, the formula proposed by Chinn40 was used to convert the odds ratio to a standardized mean difference and determine the standard error.

The DerSimonian and Laird41 random effects meta-analysis was conducted to determine the overall effect size of practice facilitation interventions and the presence of statistical heterogeneity. Ninety-five percent confidence intervals were calculated for effect sizes based on a generic inverse variance outcome model. The 𝑧 statistic was used to test for significance of the overall effect.

**Publication Bias and Heterogeneity**

The Cochran’s 𝑄 statistic and the Higgins’ 𝑃 statistic24 were used to determine statistical heterogeneity between studies. A low 𝑃 value (≤ 0.10) for the 𝑄 statistic was considered evidence of heterogeneity of treatment effects. Forest plots were generated to display graphically both the study-specific effect sizes (along with associated 95% confidence intervals) and the pooled effect estimate. A funnel plot was generated to show any evidence of publication bias32,41 along with 2-tailed Begg-Mazumdar rank correlation test44 and Egger regression asymmetry test.45 We also assessed the presence of significant heterogeneity between studies that undertook blinding, allocation concealment, and intention-to-treat and those that did not.46

**Influence, Subgroup and Meta-Regression Analysis**

To investigate the influence of each individual study on the overall effect size estimate, we conducted an influence analysis by computing the estimate and omitting 1 study in each turn. Finally, we conducted a subgroup analysis using a random-effects model, generated scatter plots, and tested the significance of regression equations, using the maximum likelihood method for mixed effects and the calculation of the Q statistic, to determine whether there were any potential effect size modifiers from year of the study, the number of practices per facilitator, duration of the intervention, tailoring of the intervention, and intensity of the intervention.

**RESULTS**

Figure 1 is a flow diagram of the selection of relevant studies. The initial literature search resulted in 207 (1980 to 2010) articles (February 1, 2011), of which 46 articles were determined to be relevant and were added to the 25 outcome studies identified by Nagykaldi et al,22 for a total of 71 articles; 54 were retrieved for closer inspection. From these articles, 44 articles were judged to meet the inclusion criteria and included in the review.17,20,34–36,47,48 The reasons for not including the 10 articles were because the intervention under study did not include an individual with the explicit role of facilitator (n = 4),86–89 the study had already been captured in the facilitation literature with no new information, had a shorter follow-up period, or measured a different outcome on the same cohort (n = 4),90–93 and the article was an editorial or written in the narrative (n = 2).94,95 In the instance where research teams produced several outcome-based studies from the same intervention, all published studies were included when the populations and outcomes being measured were different between studies.35,36,76 Finally, 21 studies were excluded because of limited validity and a quality rating of less than 6, leaving 23 studies with greater validity for the final analysis (Table 1). Supplemental Table 1, available at http://www.annfammed.org/content/10/1/63/suppl/DC1, provides the reasons for exclusion. Seventy-six percent of the 21 studies excluded were nonrandomized trials, case studies, or before-after designs with no control group. Further, 91% did not report conducting an intent-to-treat analysis, 95% did not report blinding outcome assessments, and 100% did not report allocation concea-
Because of unmatched groups at baseline, 7 of the 9 controlled clinical trials and randomized controlled trials were excluded.

**Characteristics of Selected Studies**

The 23 controlled clinical trials and randomized controlled trials included a total 1,398 participating practices: 697 randomized or allocated to the practice facilitation intervention, and 701 to a control group. The mean number of primary care practices participating per study was 59.5 (95% CI, 42.1-77.0).

Table 1 displays the research design characteristics of the 23 trials included in the analysis, along with the effect size for each study rank-ordered by methodological quality. The selected trials were reported from 1992 through 2010, spanning 18 years. Of the 20 studies that were randomized controlled trials, 3 were cluster randomized-controlled trials in which clusters of patients were randomized rather than the practices. Eleven studies reported having adhered to the intention-to-treat principle, 12 reported allocation concealment, and 14 reported blinding of assessment. Eighty-three percent of studies had a form of preventive service as the primary outcome measure (Table 1), and of those studies, 13 used the mean performance and 6 used a percentage performance.

Supplemental Table 2, available at http://www.annfammed.org/content/10/1/63/suppl/DC1, provides an overview of intervention characteristics for the 23 high-quality studies, including targeted behavior, facilitator qualifications, intervention components, and the tools used. Forty-four percent of studies described the qualifications of the facilitator as a registered nurse or masters’ educated person with training. The tools used varied; however, audit with feedback was a component of each intervention study, 91% used interactive consensus building and goal setting, and 39% used a reminder system. Seventy-four percent of the studies reported that the practice facilitator tailored the intervention to the needs of the practice.

**Intervention Effects**

Figure 2 is a forest plot that shows most of the studies have effect size point estimates which favor the intervention condition; the test for an overall effect across the 23 included studies is significant (z = 8.76; P < .001), with an overall moderate effect size point estimate of 0.56 (95% CI, 0.43-0.68) based on a random-effects model. Converting the SMD of 0.56 to an odds ratio (OR) results in an OR = 2.76 (95% CI, 2.18-3.43). Although some statistical heterogeneity is expected given practice facilitation studies with differing intervention components, outcomes, and measures, the final random-effects model was homogenous, with the test for heterogeneity being nonsignificant, χ²(1, n = 22) = 27.55, P = .19. To further understand the percentage of variability in effects caused by the heterogeneity, we computed an I² statistic, which showed that 20% of the variation among the studies could not be explained by chance.

We then conducted an influence analysis to test the sensitivity of the overall 0.56 effect size of any 1 of the 23 studies. The observed impact of any single study on the overall point estimate was negligible; the effect varied from as high as 0.58 (95% CI, 0.46-0.71) with the Cockburn et al. study removed to as low as 0.53 (95% CI 0.41-0.65) with the study by Solberg et al. removed.
Table 1. Research Design Characteristics of Studies with High Methodological Performance Scores (N = 23)

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Score</th>
<th>Trial Characteristics</th>
<th>Outcome Measure</th>
<th>Months Follow-up (% Retention)</th>
<th>Effect Size SMD (SE) 95% CI</th>
</tr>
</thead>
</table>
| Kottke et al, 1992 | 6 | Design: CCT  
Allocation concealed: N  
Blinded: Y  
Intent to treat: Y | Mean percentage of patients advised to quit | 19 (83) | 1.01 (0.52) 0.00 to 2.02 |
| McBride et al, 2000 | 6 | Design: RCT  
Allocation concealed: N  
Blinded: N  
Intent to treat: N | Percentage of records with CVD screening | 18 (100) | 0.82 (0.46) –0.08 to 1.72 |
| Stange et al, 2003 | 6 | Design: RCT  
Allocation concealed: N  
Blinded: N  
Intent to treat: N | Mean rate of preventive service | 24 (NR) | 0.59 (0.23) 0.13- to 1.05 |
| Lobo et al, 2004 | 6 | Design: RCT  
Allocation concealed: Y  
Blinded: N  
Intent to treat: N | Mean health-related quality of life | 21 (57) | 0.44 (0.18) 0.09 to 0.79 |
| Roetzhiem et al, 2005 | 6 | Design: C-RCT  
Allocation concealed: N  
Blinded: Y  
Intent to treat: N | Mean number of CA-screening tests | 24 (100) | 0.84 (0.29) 0.27 to 1.41 |
| Hogg et al, 2008 | 6 | Design: CCT  
Allocation concealed: N  
Blinded: N  
Intent to treat: N | Mean preventive performance index | 6 (87) | 0.73 (0.29) 0.16 to 1.30 |
| Aspy et al, 2008 | 6 | Design: CCT  
Allocation concealed: N  
Blinded: Y  
Intent to treat: Y | Percent given physical inactivity brief intervention | 18 (89) | 1.12 (0.36) 0.42 to 1.82 |
| Jaén et al, 2010 | 6 | Design: RCT  
Allocation concealed: N  
Blinded: N  
Intent to treat: N | Mean prevention service score | 26 (86) | 0.04 (0.37) –0.69 to 0.77 |
| Cockburn et al, 1992 | 7 | Design: RCT  
Allocation concealed: N  
Blinded: Y  
Intent to treat: N | Mean number of cessation cards used | 3 (79) | 0.24 (0.15) –0.06 to 0.54 |
| Modell et al, 1998 | 7 | Design: RCT  
Allocation concealed: N  
Blinded: N  
Intent to treat: N | Median number of hemoglobin tests | 12 (100) | 0.32 (0.40) –0.45 to 1.09 |
| Engels et al, 2006 | 7 | Design: RCT  
Allocation concealed: Y  
Blinded: N  
Intent to treat: N | Mean number of projects initiated | 12 (92) | 1.04 (0.32) 0.41 to 1.67 |
| Aspy et al, 2008 | 7 | Design: RCT  
Allocation concealed: N  
Blinded: Y  
Intent to treat: Y | Mean percent with MMG | 9 (100) | 1.31 (0.57) 0.20 to 2.42 |
| Deitrich et al, 1992 | 18 | Design: RCT  
Allocation concealed: N  
Blinded: Y  
Intent to treat: N | Mean rate of prevention service | 12 (96) | 0.59 (0.29) 0.02 to 1.16 |

Table 1 continues
Figure 3 is a publication bias funnel plot of practice facilitation effect size as represented by the standardized mean difference (x-axis) and the standard error (y-axis) for each of the 23 practice facilitation studies. The funnel plot provides evidence of publication bias, in that there were fewer small studies with small effects included in the meta-analysis, as displayed by the imputed results. Publication bias was confirmed by the Begg and Mazumdar44 test \( (P = .003) \) and the Egger et al45 test \( (P = .003) \).

There was no association between the methodological characteristics of studies as determined by the methodologic performance score and effect size \( (\beta = –0.04; \ldots) \).
Further, Jüni et al. have shown that 3 key domains have been associated with biased effect size estimates in meta-analysis. Effect sizes for the 23 included studies did not differ significantly in terms of allocation concealment ($P = .77$), blinding of outcome assessments ($P = .80$), and the handling of attrition through intent to treat ($P = .85$).

### Practice Facilitation Effect Size Moderators

There was no significant difference between studies published in or after 2001 when compared with studies published before 2001 ($P = .69$), and the relationship between duration of the intervention and effect size was not significant ($P = .94$). Those practice facilitation studies that reported an intervention tailored to the context and needs of the practice had a significantly larger overall effect size of 0.62 (95% CI, 0.48-0.75; $P = .05$) compared with studies that did not report tailoring (SMD = 0.37; 95% CI, 0.16-0.58).

The scatter plot in Figure 4 depicts the relationship between the ratio of practices per facilitator (Supplemental Table 2) and effect size for each study. It shows the fitted regression line and a significant negative association between the number of practices per facilitator and effect size ($\beta = -0.02; P = .004$). Each selected study is shown on the graph as a bubble, and the size of the bubble represents the amount of weight associated with the results of that study. Data were not available for 2 of the studies.

Intensity of practice facilitation was calculated by multi-
plying the average number of contacts with a practice by the average meeting time in hours (Supplemental Table 1). Figure 5 depicts a significant trend between the intensity of the intervention and the effect size ($\beta = 0.008; P = .03$).

**DISCUSSION**

The translation of evidence-based guidelines into practice is complex, and research continues to find major gaps between research evidence and practice. Alternative intervention models are being advanced to address the numerous challenges that face traditional primary care practices in providing high-quality care. This systematic review and meta-analysis has shown the potential for practice facilitation to address the challenges of translating evidence into practice. Primary care practices are 2.76 (95% CI, 2.18-3.43) times more likely to adopt evidence-based guidelines through practice facilitation.
These findings should prove important to health policy makers and those involved in practice-based research networks when designing quality improvement programs. We know that practice facilitation improves adoption of guidelines in multiple clinical practice areas that focus on prevention. Prevention activities in health care organizations are well suited to a practice facilitation approach, because much of the uptake can be improved by a focus on the organization of care, such as using simple reminder systems, recalls, and team care that need not involve the physician. We do not know whether facilitation can be translated to other areas that will require more direct physician uptake, such as clinical management requiring medication optimization or chronic illness care.

All the studies included audit with feedback, practice consensus building, and goal setting as key components, as well as basing the change approach on the system level and the organization change on common quality improvement tools, such as plan-do-study-act. Many also incorporated collaborative meetings, whether face to face or virtually. Such collaborative meetings can add costs to the programs, and it is not known whether these resource-intensive meetings increase effectiveness. There was variation in the process of implementation among the studies related to the facilitator qualifications, training, number of practices, intensity, and duration of intervention.

We found that as the number of practices per facilitator increases, the overall effect of facilitation diminishes but did not plateau. The intensity of the intervention is associated with larger effects as well. In addition, whether the intervention was tailored for the practice also affected effectiveness, and we found that a larger effect is associated with tailored interventions. Previous research has shown that tailored interventions are key to improving performance, and this study has confirmed this finding.

Implementing practice facilitation into routine quality improvement programs for organizations can be challenging. These findings support the need to tailor to context, to incorporate audit and feedback with goal setting, and to consider intensity of the intervention. The key components appear to be processes of care and organization of care with less focus on content knowledge. These findings, in addition to the qualifications of the facilitator, hold important resource implications that will complicate adequately funding facilitator interventions.

There are several important limitations to this study. First, in an effort to focus and limit the scope of work involved, only published journal literature was included. We did not search for unpublished studies or original data. Second, although the variability between studies was consistent with what would be expected to occur by chance alone, the differing outcome measures, settings, and the diversity of guidelines being implemented and the potential modifying effect of such factors warrant caution. Third, not all of the study characteristics were analyzed in terms of the relationship to effects, and further research and meta-regression analysis are recommended. Finally, there is evidence of publication bias for practice facilitation research. Researchers should publish good-quality studies with null effects to better understand the limits of practice facilitation, as it is unlikely to be able to change successfully every type of targeted evidence-based behavior in all contexts.

In conclusion, despite the professional, organizational, and broader environmental challenges of getting evidence into practice, this study has found that practice facilitation can work. An understanding of the conceptual model for practice facilitation exists, and more randomized controlled trials to test the model are not required. Instead, large-scale collaborative, practice-based evaluation research is needed to understand the impact of facilitation on the adoption of guidelines, the relationship between context and the components of facilitation, sustainability, and the costs to the health system. This study has provided information on the empirical effects of practice facilitation that can be used to adjust expectation for what is realistic based on the current evidence and to move forward.

To read or post commentaries in response to this article, see it online at http://www.annfammed.org/content/10/1/63.

Key words: Practice facilitation; outreach facilitation; primary care; implementation research; evidence-based guidelines; meta-analysis; systematic review; knowledge translation; behavior change


Acknowledgments: Roy Cameron, PhD, Walter Rosser, MD, CCFP, FCFP, MRCGP (UK); Steve Brown, PhD, Paul W. McDonald, PhD, FRIPH, and, Ian McKillop, PhD, contributed to the review of the article. Dianne Zakaria for contributed to the review and methodological assessment of the practice facilitation literature.

References


31. SPSS Inc. SPSS (Version 18.0.0) [Computer software]. Chicago, IL: SPSS Inc; 2009.


62. Carney PA, Dietrich AJ, Keller AA, Landgraf J, O'Connor GT. Tools, 
60. Rebelsky MS, Sox CH, Dietrich AJ, Schwab BR, Labaree CE, Brown- 
58. Margolis PA, Lannon CM, Stuart JM, Fried BJ, Keyes-Elstein L, 
57. Lobo CM, Frijling BD, Hulscher ME, et al. Improving quality of 
55. Modell M, Wonke B, Anionwu E, et al. Setting up improvement 
54. Nagykyaldi Z, Mold JW. Diabetes Patient Tracker, a personal digital 
53. Cox S, Wilcock P, Young J. Improving the repeat prescribing pro 
51. Crotty M, Litt JC, Ramsay TC, Weller DP. Will facilitators be accept 
49. Lawrence M, Packwood T. Adapting total quality management 
48. Stange KC, Goodwin MA, Zyzanski SJ, Dietrich AJ. Sustainability of 
47. Bryson FP, Neville RG, Crombie IK, Clark RA, McKenzie P. Controlled 
46. Jones JN, Marsden P. Improved diabetes care in a UK health dis 
44. Bordley WC, Margolis PA, Stuart J, Lannon C, Keyes L. Improv 
43. Solberg LI, Kottke TE, Brekke ML. Will primary care clinics organize 
42. Aspy CB, Mold JW, Thompson DM, et al. Integrating screening and 
41. Mold JW, Aspy CA, Nagykaldi Z; Oklahoma Physicians Resource/ 
39. Solberg LI, Kottke TE, Brekke ML. Improving a quality improvement 
38. Stange KC, Goodwin MA, Zyzanski SJ, Dietrich AJ. Sustainability of 
35. Engels Y, van den Hombergh P, Mokkink H, van den Hoogen H, 
34. Frijling BD, Lobo CM, Hulscher MEJL, et al. Intensive support to 
33. McKenzie A, Grylls J. Diabetic retinal photographic screening: a 
32. Carney PA, Dietrich AJ, Keller AA, Landgraf J, O'Connor GT, Tools, 
30. Carney PA, Dietrich AJ, Keller AA, Landgraf J, O'Connor GT, Tools, 
27. Carney PA, Dietrich AJ, Keller AA, Landgraf J, O'Connor GT. Too 
15. Carney PA, Dietrich AJ, Keller AA, Landgraf J, O'Connor GT. Too 
2. Carney PA, Dietrich AJ, Keller AA, Landgraf J, O'Connor GT. Too 


