

Evidence for the impact of quality improvement collaboratives: systematic review

Loes M T Schouten, senior consultant,¹ Marlies E J L Hulscher, senior researcher,² Jannes J E van Everdingen, senior consultant,¹ Robbert Huijsman, professor,³ Richard P T M Grol, professor and director²

¹Dutch Institute for Healthcare Improvement, PO Box 20064, 3502 LB Utrecht, Netherlands

²Centre for Quality of Care Research, University Medical Centre St Radboud, Nijmegen

³Institute of Health Policy and Management, Erasmus MC, University Medical Centre Rotterdam

Correspondence to: L M T Schouten
lschouten@cbo.nl

doi:10.1136/bmj.39570.749884.BE

ABSTRACT

Objective To evaluate the effectiveness of quality improvement collaboratives in improving the quality of care.

Data sources Relevant studies through Medline, Embase, PsycINFO, CINAHL, and Cochrane databases.

Study selection Two reviewers independently extracted data on topics, participants, setting, study design, and outcomes.

Data synthesis Of 1104 articles identified, 72 were included in the study. Twelve reports representing nine studies (including two randomised controlled trials) used a controlled design to measure the effects of the quality improvement collaborative intervention on care processes or outcomes of care. Systematic review of these nine studies showed moderate positive results. Seven studies (including one randomised controlled trial) reported an effect on some of the selected outcome measures. Two studies (including one randomised controlled trial) did not show any significant effect.

Conclusions The evidence underlying quality improvement collaboratives is positive but limited and the effects cannot be predicted with great certainty. Considering that quality improvement collaboratives seem to play a key part in current strategies focused on accelerating improvement, but may have only modest effects on outcomes at best, further knowledge of the basic components effectiveness, cost effectiveness, and success factors is crucial to determine the value of quality improvement collaboratives.

INTRODUCTION

Stimulated by the concerted, ongoing efforts of the Institute of Medicine and other platforms which state that “reform around the margins is inadequate to address system ills” in quality of care, healthcare organisations in many countries are setting up quality improvement collaboratives. Multidisciplinary teams from various healthcare departments or organisations join forces for several months to work in a structured way to improve their provision of care. Quality improvement collaboratives are being used increasingly in the United States, Canada, Australia, and European countries. In northern European countries such as the United Kingdom and the Netherlands

health authorities support nationwide quality programmes based on this strategy.

Different types of multiorganisational collaboratives exist, the purpose of which are to improve care.¹⁻³ The term quality improvement collaborative seems to be used for different multifaceted packages that focus on accelerating better outcomes.⁴ Quality improvement collaboratives are used in different clinical areas and organisational contexts and have been adopted by numerous large and small healthcare systems and individual clinics. These initiatives represent substantial investments of time, effort, and funding in the delivery of health care, although estimates of the total investment and applications of the collaborative are not available.⁵ The strength of the quality improvement collaborative seems to be the relatively efficient use of experts and peers and the exchange of best practices to facilitate and guide improvement.

The earliest well documented activities of quality improvement collaboratives are those of the Northern New England Cardiovascular Disease Study Group, established in 1986, and the Vermont Oxford Network, established in 1988. Another well known approach is the Breakthrough Series developed by the Institute of Healthcare Improvement in 1995. Many of the present approaches of the quality improvement collaborative are based on those of the Breakthrough Series.

A recent non-systematic review⁶ concluded that the collaborative methodology has important potential to improve outcomes for patients and to facilitate sustainability of quality improvement. Unfortunately this introduction to the collaborative methodology neither considers whether the evaluation of effectiveness is based on a controlled or an uncontrolled study design nor makes clear on what types of quality improvement collaborative the conclusions are based.

Clear evidence of the effectiveness of the methods is lacking, despite the ongoing initiatives of the quality improvement collaborative reflecting different multifaceted intervention packages, the growing number of published papers, good face validity of the model, and facilitators claiming that many professionals appreciate taking part in a collaborative for both professional and organisational development.⁴ Little is known about the

effectiveness of quality improvement collaboratives or specific components that enhance the effectiveness of such collaboratives, and there is hardly any information about their cost effectiveness and sustainability. We therefore assessed the effectiveness of the quality improvement collaborative by systematically reviewing empirical studies.

METHODS

We searched Medline, CINAHL, Embase, PsycINFO, and Cochrane databases for literature on quality improvement collaboratives in the period from January 1995 to June 2006 inclusive. Initially we searched Medline using free text terms describing quality improvement collaboratives, and we combined these with the keywords (non-MeSH) “quality and improvement and collaborative” or “(series or project) and breakthrough” (box 1). These steps were repeated for the other databases. We also reviewed the reference lists of the included papers.

Study selection

We included studies that were written in English, contained data on the effectiveness of care processes or outcomes, were in a healthcare setting, and met the criteria for a quality improvement collaborative derived from a rigorous analysis of the theoretical literature on quality improvement collaboratives (box 2).¹⁴⁷⁻⁹ To include studies we used the following definition and criteria. A quality improvement collaborative is an organised, multifaceted approach to quality improvement that involves five essential features: there is a specified topic—a subject exists with large variations in care or gaps between best and current practice; clinical experts and experts in quality improvement provide ideas and support for improvement—they identify, consolidate, clarify, and share scientific knowledge and best practice as well as knowledge in quality improvement; a critical mass of multiprofessional teams from multiple sites is willing to improve and share care; a model for improvement focuses on setting clear and measurable targets, collecting data, and testing changes on a small scale to advance reinvention and learning by doing; and the collaborative process involves a series of structured

Box 1 Search terms

Set 1—(non-MeSH)

(quality and improvement and collaborative) or ((series or project) and breakthrough)

Set 2—(MeSH headings)

Organizational-Innovation, Models-Organizational, Cooperative-Behavior

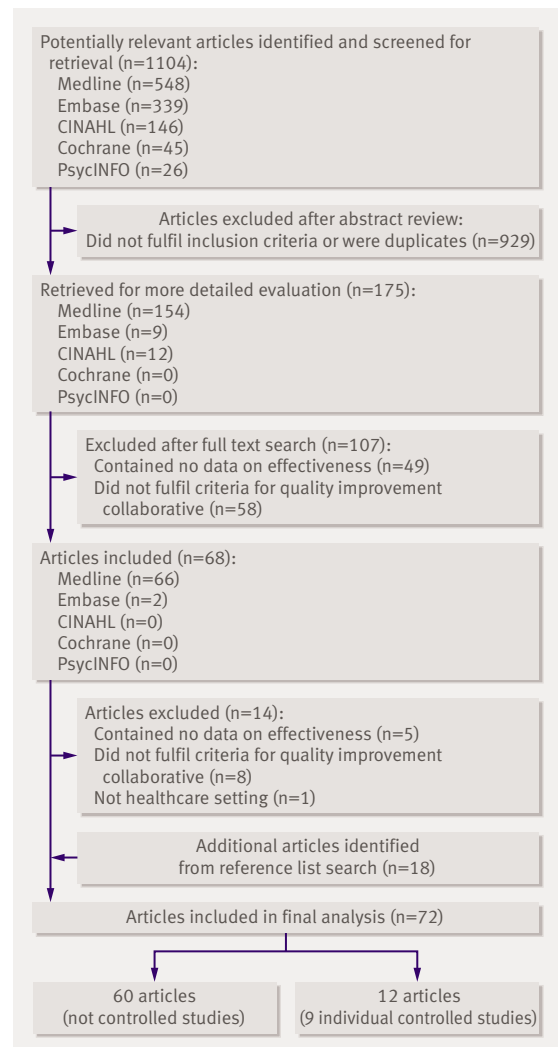
And Program-Evaluation, Total-Quality-Management, Quality-Assurance-Health-Care

And Outcome-and-Process-Assessment-Health-Care

And Health-Services-Research, Regional-Medical-Programs, Organi?ation* near (collabor* or participa*)

Set 3—(MeSH headings)

The steps in set 2 combined with Statistics, Statistics-and-numerical-data



Selection process for studies included in analysis

activities (meetings, an active email list, visits to facilitators) in a given time frame to advance improvement, exchange ideas, and share experiences of the participating teams.

Data extraction

Each potentially eligible study was independently assessed by two reviewers (LMTS and MEJLH) for inclusion and quality. We assessed the methodological quality of the studies by evaluating the design, method of randomisation, characteristics of control sites, protection against bias, reliable outcome measures, and how sites and patients lost to follow-up had been handled in the analysis (see checklist of the Cochrane Effective Practice and Organisation of Care Review Group at www.epoc.cochrane.org). We used a standardised extraction checklist to obtain data on topics, study design, setting, numbers of participants, characteristics of the collaborative strategy, and relevant results. Two of the authors (LMTS and MEJLH) independently completed this checklist for each study. Disagreements on data extraction and classification of study results were resolved by consensus. We

categorised the studies in groups on the basis of their study designs and characteristics of the collaborative strategy used. We could not use formal meta-analytical techniques for pooling results because the studies used many different effect measures.

RESULTS

Figure 1 shows the flow of papers through the review. Overall, 1104 abstracts of studies published from January 1995 to June 2006 were identified. During

screening 929 papers were excluded on the basis of the abstract because they did not fulfil the inclusion criteria or because of duplication. A total of 175 articles were requested for detailed review. An additional 107 articles were subsequently excluded; 49 studies contained no original data on effectiveness or outcomes and 58 did not meet our criteria for a quality improvement collaborative. The criteria most often not fulfilled were the lack of structured activities aimed at exchange of ideas and information between participating teams (collaborative process) and the fact that participants did not set aims, collect data, and test changes in quick succession (model for improvement). Reviewing the reference lists of the remaining 68 studies led to the identification of a further 18 potentially eligible studies. Four of them fulfilled our inclusion criteria and did not duplicate any of the studies already included. A total of 72 papers were eligible for inclusion in the study.

Of the 72 papers included in our study, 60 (83%) reports^{w1-w60} used an uncontrolled study design. Fifty (83%) of these 60 reports were based on the Breakthrough Series. Several studies contained elements that went beyond a simple quality improvement collaborative. The quality improvement collaboratives in these reports were embedded in ongoing quality initiatives or used comprehensive systems for measuring performance. Of the 60 reports of uncontrolled studies 37 (almost 62%) were based on self report measures of the participating teams or sites and 14 (23%) of the studies were single case reports describing changes in an individual facility or team (a detailed table listing the uncontrolled studies is available from the authors). The study designs of the uncontrolled reports relied almost entirely on post measurement, used before and after studies without being able to account for secular trends, made use of self report measures rather than reviews of medical records, included only anecdotal information, or selected samples from self selected sites. Conclusions on effectiveness could not be drawn from these reports owing to the lack of adequate reporting procedures on data collection, analysis, and objective evaluations.

Eleven reports (eight studies)^{w61-w71} used a “comparison group” study design. One study^{w72} used an interrupted time series design (table 1). Most studies were published in 2004 and 2005. The characteristics of the collaborative strategy used in these nine studies varied. Seven studies^{w61-w68 w72} were explicitly based on the Breakthrough Series with elements of the chronic care model (www.improvingchroniccare.org).^{w62 w64-w68} Two studies^{w69-w71} were based on the Vermont Oxford Network. Box 3 gives an overview of the different collaborative strategies. Table 1 shows the methodological quality of the nine studies. Most studies had important flaws. Limitations included possible differences in baseline measurement, limited data on characteristics of control sites, no specification of blinded assessment, and possible contamination.

Box 2 Key components of quality improvement collaboratives

Key concepts behind efforts to improve multiorganisational collaboratives (Plsek, 1997¹)

- Multiple organisations
- Variability in performance
- Benchmarking efforts and identification of “best practices”
- Purposeful replication efforts of ideas to fit the context
- Characterisation of internal processes
- Measured improvement
- Open sharing

Elements of Breakthrough Series (Kilo, 1998⁷)

- Multiple sites
- Topic selection
- Theory
- Model for improvement
- Spread, focus, and tension for change

Key components of Breakthrough Series (Kilo, 1999⁸)

- Collaborative participants
- Identification of topics
- Identification and consolidation of relevant knowledge
- Limitations of time and focus, creating tension for change
- Collection and maintenance of data by each organisation
- The collaborative process
- Deployment of knowledge

Features of collaboratives (Øvretveit et al, 2002⁴)

- Participation of several multiprofessional teams
- A focused clinical or administrative subject
- Evidence of large variations in care (gaps)
- Participants learn from experts about evidence, concepts of change, practical changes
- Participants use a method for testing change
- Teams set measurable targets and collect data
- Participants meet at least twice
- Participants continue to exchange information between meetings

Key components (Wilson et al, 2003⁹)

- Sponsorship
- Topic
- Ideas for improvement
- Participants
- Support by senior leadership
- Preliminary work
- Learning about and making improvements

Randomised controlled trials

Two of the nine studies were randomised controlled trials, one of which was a trial of a quality improvement collaborative for children with asthma and based on a combination of the Breakthrough Series and the chronic care model.^{w64} The other was a large trial of a quality improvement collaborative for neonatal intensive care and based on the Vermont Oxford Network method.^{w71} The trial based on the Breakthrough Series and chronic care model^{w64} did not show effects on any of the key processes or intermediate outcomes of care for children with asthma. The trial based on the Vermont Oxford Network method^{w71} showed significant improvement in two specific processes of care but no significant improvement in patient outcomes (mortality and pneumothorax), nor in 20 of the 23 secondary process measures. In this cluster randomised controlled trial, infants in the intervention hospitals ($n=57$) were more likely to receive a surfactant in the delivery room (54.7% *v* 18.2%) and were less likely to receive the first dose more than two hours after birth (9.4% *v* 24.9%) than infants in the control hospitals (table 2).

Controlled before and after studies and interrupted time series studies

Five of the six controlled before and after studies based on the Breakthrough Series^{w61 w62 w65-w68 w72} showed significant improvements in the outcomes of care. Pierce-Bulger et al^{w72} showed a significant improvement in the number of days between neonatal deaths— increase from an average of 55 days (1989-94) before the programme to an average of 114 days (1995-2000)—during a seven year quality programme including a Breakthrough Series project in 1993. In Baier et al's study,^{w61} the prevalence of residents with pain in 21 nursing homes diminished significantly (7.2% *v* 11.2%

of patients) after participation in a Breakthrough Series. Landon et al's study^{w63} of 9986 patients with HIV infection did not show any significant effect on virological outcomes or process measures, such as screening, prophylaxis, and access to care.

All three controlled before and after studies combining the Breakthrough Series with the chronic care model^{w62 w65-w68} showed significant improvement of some of the selected process and outcome measures of care. Benedetti et al^{w62} compared participating providers with non-participating providers and reported significant improvements in, for example, rates of annual diabetes examinations for eyes and feet, and better outcomes for haemoglobin A_{1c} and blood pressure. Mangione-Smith et al^{w65} and Schonlau et al^{w66} showed significant improvements for specific items of patient self management (for example, peak flow monitoring in 70% versus 43% of patients) and education (such as instructions for the use of metered dose inhalers in 30% versus 9% of patients) in asthma care. Mangione-Smith et al^{w65} also reported higher scores on quality of life measures for the intervention group in a survey carried out after measurement. The levels of asthma severity between the intervention group and the control group, however, differed. In the same study Schonlau et al^{w66} showed significant improvement in satisfaction with communication (overall score 62% *v* 39%). Asch et al^{w67} reported significant improvement in specific items on counselling and education (related to diet, drugs, exercise, weight loss, disease management, water weight, and goal setting) in the care of patients with chronic heart failure. The process measures with greater improvement were those with initially low performance rates, and the rates remained below 50% for most educational processes. Significant improvement in two of four measures for the appropriate use of drugs—where baseline rates were good—was less dramatic (appropriate angiotensin converting enzyme inhibitors, 93% *v* 87% and lipid lowering therapy, 66% *v* 64%). The changes in the rates of counselling were confirmed in a cross sectional survey among patients.^{w68}

The controlled before and after study based on the Vermont Oxford Network^{w69 w70} reported a significant decrease in the rate of infection at six intervention neonatal intensive care units (12.3% *v* 16.5%) and a decrease in the rate of supplemental oxygen at four neonatal intensive care units (34% *v* 38.7%). Although the effect among the units on both arms was heterogeneous, the changes show that the intervention units improved at a significantly faster rate in a four year period than did the 66 comparison units. In the same study, Rogowski et al^{w70} presented data showing that costs may be reduced as a result of participation in the quality improvement collaborative, although the cost savings across hospitals were heterogeneous. Table 3 gives a summary of effectiveness found in the nine controlled studies.

Box 3 Overview of collaborative strategies used in nine controlled studies

Collaborative based on Breakthrough Series (www.ih.org); seven studies

A Breakthrough Series is a short term (6 to 15 months) learning system that brings together a large number of teams from hospitals or clinics to seek improvement in a focused topic area. The driving vision behind the Breakthrough Series is that sound science exists on the basis of which the costs and outcomes of current healthcare practices can be greatly improved, but much of this science is unused in daily work. The Breakthrough Series is designed to help organisations make “breakthrough” improvements to close this gap by creating a structure in which interested organisations can easily learn from each other and from recognised experts in topics where they want to make improvements.

Collaborative based on Vermont Oxford Network (www.vtoxford.org); two studies

The Vermont Oxford Network is a data driven, “non-profit, voluntary collaboration of health care professionals dedicated to improving the quality and safety of medical care for newborn infants and their families.” It was established in 1988, and today the network includes more than 400 neonatal intensive care units, predominantly in the United States. The network facilitates a coordinated programme of research, education, and quality improvement. To support this programme, the network maintains a database of infants with very low birth weight at the member hospitals. The database includes information concerning medical practices and patient outcomes such as morbidity, mortality, and length of stay. Members of the network receive centre specific, routinely prepared, confidential reports.

Table 1 | Overview of types of quality improvement collaborative and methods in included randomised controlled studies

Study, reference (year)	Intervention		Study design (control condition)	Unit of analysis (project sample size); study sample size	Methods	
	Topic, disease, or condition	Type of collaborative; length (months) of intervention			Method; No of patients	Methodological quality (status)
1, Pierce-Bulger et al ^{w72} (2001)	Infant mortality in community	Breakthrough Series embedded in longitudinal quality activities (clinic and home visiting services); 12	Interrupted time series (no)	Medical centre (NS); 1 (intervention)	Chart review; NS	Protection from secular changes (not done); data analysed appropriately (NS); reason for points before and after design (NS); shape of intervention effect (NS); protection against bias (NS); blinded assessment (NS); completeness of dataset (NS)
2, Baier et al ^{w61} (2004)	Pain	Based on Breakthrough Series; 15	Controlled before and after study (no)	Nursing homes (21); 15 (intervention), 72 (control)	Minimal dataset www.cms.hhs.gov/medicaid/mds20; 276 (intervention)	Baseline measurement (no differences); characteristics of control sites (differences); blinded assessment (NS); contamination (unlikely); follow-up sites or teams, 71% intervention, 92% control
3, Benedetti et al ^{w62} (2004)	Diabetes in primary care	Breakthrough Series, chronic care model; 12-36	Controlled before and after study (no)	Providers (NS); 11 (intervention), 19 (control)	Method unclear	Baseline measurement (not reported); characteristics of control sites (no data, but seemed similar); blinded assessment (NS); contamination (possible); follow-up of sites or teams (NS)
4, Landon et al ^{w63} (2004)	HIV	Breakthrough Series; 16	Controlled before and after study (no)	Clinics (62); 44 (intervention), 25 (control)	Review of medical records: 6406 (intervention), 3580 (control)	Baseline measurement (no differences); characteristics of control sites (similar); blinded assessment (NS); contamination (unlikely); follow-up of sites or teams (NS)
5, Homer et al ^{w64} (2005)	Asthma: children	Breakthrough Series, chronic care model; 12	Randomised controlled trial (no)	Primary care practice (22); 22 (intervention), 21 (control)	Telephone interview: 530 (intervention), 591 (control)	Unit of allocation (by practice); randomisation (no process); follow-up of sites (100%); blinded assessment (NS); baseline measurement (no differences); contamination (possible)
6, Mangione-Smith et al ^{w65} (2005)	Asthma: children	Breakthrough Series, chronic care model; 12	Controlled before and after study (no)	Primary care practice (26); 9 (intervention), 4 (control)	Review of medical records: 348 (intervention), 153 (control); telephone survey (after study): 385 (intervention), 126 (control)	Baseline measurement (possibly different); characteristics of control sites (limited data); blinded assessment (NS); contamination (possible); follow-up of sites or teams (100%);
6, Schonlau et al ^{w66} (2005)	Asthma: adults	Breakthrough Series, chronic care model; 12	Controlled before and after study (no)	Primary care practice (26); 6 (intervention), 3 (control)	Review of medical records, 109 (intervention), 76 (control); telephone survey (after study): 123 (intervention), 62 (control)	Baseline measurement (possibly different); characteristics of control sites (limited data); blinded assessment (done); contamination (unlikely); follow-up of sites or teams (100%)
7, Asch et al ^{w67} (2005)	Chronic heart failure	Breakthrough Series, chronic care model; 12	Controlled before and after study (no)	Clinics (14); 4 (intervention), 4 (control):	Review of medical records: 261 (intervention), 228 (control); survey (after study): 301 (intervention +control):	Baseline measurement (22 of 25 measures similar); characteristics of control sites (limited data); blinded assessment (done); contamination (unlikely); follow-up of sites or teams (100%)
7, Baker et al ^{w68} (2005)	Chronic heart failure	Breakthrough Series, chronic care model; 12	Controlled cross sectional study (no)	Clinics (13); 6 (intervention), 6 (control)	Telephone survey (after study): 367 (intervention), 414 (control)	Baseline measurement (NA); characteristics of control sites (limited data); blinded assessment (NA); contamination (NA); follow-up of sites or teams (NA)
8, Horbar et al ^{w69} (2001)	Preterm infants: infection or chronic lung disease	Vermont Oxford Network; 36	Controlled before and after study (feedback in routine reports)	Neonatal intensive care units (10); 10 (intervention), 66 (control)	Database: 3801 (intervention), 21 509 (control)	Baseline measurement (possibly different); characteristics of control sites (limited data); blinded assessment (NS); contamination (unlikely); follow-up of sites or teams (100%)
8, Rogowski et al ^{w70} (2001)	Preterm infants: infection or chronic lung disease	Vermont Oxford Network; 36	Controlled before and after study (feedback in routine reports)	Neonatal intensive care units (10); 10 (intervention), 9 (control)	Database on infection: 2993 (intervention), 2203 (control); database on chronic lung disorder 663 (intervention), 1007 (control)	Baseline measurement (possibly different); characteristics of control sites (limited data); blinded assessment (NS); contamination (unlikely); follow-up of sites or teams (100%)
9, Horbar et al ^{w71} (2004)	Preterm infants: surfactant treatment	Vermont Oxford Network; NS	Randomised controlled trial (feedback in routine reports)	Neonatal intensive care units (57); 57 (intervention), 57 (control)	Database: 6645 (intervention), 5576 (control)	Unit of allocation (by hospital); randomisation (computer generated); follow-up of sites (100%); blinded assessment (NS); baseline measurement (no differences); contamination (unlikely)

NS=not specified; NA=not applicable.

Table 2 | Overview of effectiveness of quality improvement collaboratives in included controlled studies

Study, reference, study outcome measures	No of measures (No of significant measures)†	Significant outcomes* (post measurement intervention group v control group)
1, Pierce-Bulger et al^{w72}		
Infant mortality	1 (1)	Annual days between deaths: increase from pre-programme average of 55 days (1989-94) to an average of 114 days (1995-2000)‡
2, Baier et al^{w61}		
Residents with pain	1 (1)	Prevalence: 7.2% v 11.2% patients
3, Benedetti et al^{w62}		
End organ surveillance or therapy	5 (3)	Eye examination annually (NS); foot examination annually (NS); patients aged >40 years taking acetylsalicylic acid (NS)
Glycaemic control	2 (1)	Haemoglobin A _{1c} <9.5 (NS)
Dyslipidaemia therapy	2 (1)	Low density lipoprotein test <130 (75% v 45% patients)
Hypertension control	2 (1)	Blood pressure <130/85 (49% v 35% patients)
4, Landon et al^{w63}		
Antiretroviral therapy	2 (0)	
Screening and prophylaxis	5 (0)	
Access to care	1 (0)	
5, Homer et al^{w64}		
Primary study outcomes:		
Written asthma management plan	1 (0)	
Daily use of inhaled steroids,	1 (0)	
Daily use of controller drugs [Author: Please clarify dash]	1 (0)	
Secondary study outcomes:		
Asthma attack	1 (0)	
Parent report of limited activities	1 (0)	
Parents' experience of care	1 (0)	
Parent reported functional status	1 (0)	
Admission to hospital or emergency department for asthma	1 (0)	
6, Mangione-Smith et al^{w65}		
Medical records process indicators	9 (7)	Peak expiratory flow rate measured annually (49% v 4% patients); written action plan (42% v 3% patients); ≥2 follow-up visits annually (86% v 78% patients); educated in self management (41% v 17% patients); instructed in use of metered dose inhalers (30% v 9% patients); collaborative goal setting between patient and provider (10% v 0% patients); overall asthma process of care summary score (56% v 40% patients)
Quality of life	3 (2)	General quality of life (80.2 v 77.0); asthma specific quality of life: treatment problems (88.6 v 85.3)
Patient self management	3 (2)	Peak flow monitoring (70% v 43% patients); written action plan (41% v 22% patients)
Satisfaction with care, impact on family functioning, acute care service use, missed school days, parent lost work days, asthma knowledge, use of long term drugs for control	2 (0), 1 (0), 1 (0), 1 (0), 1 (0), 1 (0), 1 (0)	
6, Schonlau et al^{w66}		
Medical records process indicators	9 (5)	Peak expiratory flow rate measured annually (28% v 14% patients); written action plan (27% v 0% patients); instructed in use of metered dose inhalers (22% v 7% patients); collaborative goal setting between patient and provider (7% v 0% patients); overall asthma process of care summary score (46% v 38% patients)
Self management	4 (1)	Attended educational session (20% v 5% patients)
Satisfaction with communication	1 (1)	With clinician or lay educator communication (overall score 62% v 39% patients)
Quality of life, use of acute care, bed days resulting from asthma related illness, asthma knowledge, drugs for asthma control	2 (0), 1 (0), 1 (0), 1 (0), 1 (0), 1 (0)	
7, Asch et al^{w67}		
Counselling indicators	8 (7)	Drugs (44% v 17% patients); diet (46% v 11% patients); exercise (42% v 12% patients); weight loss (42% v 7% patients); disease management (61% v 23% patients); water weight management (42% v 4% patients); goal setting (5% v 4% patients)
Drug indicators	4 (2)	Angiotensin converting enzyme inhibitor for left ventricular ejection fraction ≤40% (93% v 87% patients); lipid lowering therapy for coronary artery disease (66% v 64% patients)
Outcome indicators, diagnostic indicators, follow-up indicators	4 (0), 4 (0), 5 (0)	
7, Baker et al^{w68}		
Knowledge	15 (8)	Not to drink more fluids than normal (69% v 53% patients); check weight (84% v 44% patients); heart not pumping blood as well as it should (64% v 53% patients); high salt food (canned vegetables) (89% v 85% patients); high salt food (cheese) (76% v 68% patients); shortness of breath (61% v 55% patients); swelling of legs or ankles (80% v 70% patients); weight gain (76% v 61% patients)

Study, reference, study outcome measures	No of measures (No of significant measures)†	Significant outcomes* (post measurement intervention group v control group)
Educational components received	14 (5)	Not to add salt (91% v 83% patients); avoid drinking large amounts (59% v 38% patients); weigh yourself (87% v 34% patients); regular exercise (90% v 83% patients); use pillbox (58% v 49% patients)
Communication	4 (4)	Given choices and options about treatment (3.9 v 3.7); have given me confidence (4.1 v 3.9); are interested (4.2 v 4.1); review self management (4.0 v 3.9)
Self management behaviours	4 (2)	Have functioning scale at home (93% v 81% patients); how frequently weighed oneself (4.2 v 3.2)
Healthcare use	3 (2)	Emergency department visits (adjusted difference -0.25); number of admissions to hospital (0.85 v 0.95)
Quality of life, self efficacy, satisfaction, heart failure symptoms	2 (0), 3 (0), 4 (0), 7 (0)	
8, Horbar et al^{w69}		
Rates of infection	2 (1)	Coagulase negative staphylococcal infection rate (12.3% v 16.5% infants)
Rates of oxygen supplementation	1 (1)	Oxygen supplementation rate (34% v 38.7% infants)
Death	1 (0)	
8, Rogowski et al^{w70}		
Median treatment cost per infant	2 (1)	Median treatment cost per infant (\$57 606 before measurement v \$45 874 after measurement)
9, Horbar et al^{w71}		
Primary study outcomes:		
Surfactant treatment in delivery room	1 (1)	Surfactant treatment in delivery room (54.7% v 18.2%)
First dose surfactant given after 2 hours	1 (1)	First dose surfactant given after two hours (9.4% v 24.9%)
Mortality (death before discharge), pneumothorax	1 (0), 1 (0)	
Secondary study outcomes:		
Delivery room resuscitation	7 (1)	Endotracheal intubation (78% v 69.8%)
Respiratory	6 (1)	Surfactant given at any time (85.1% v 79.5%)
Neurological	3 (1)	Severe intraventricular haemorrhage (10.1% v 14.2%)
Infection, other	4 (0), 3 (0)	

NS=not specified.

£1 (€1.27; \$1.97).

*P<0.05.

†Number of measures with significant improvement at intervention sites.

‡Since clinic and home visiting services began.

DISCUSSION

Worldwide, organisations are adopting the approach of the quality improvement collaborative in different settings. The evidence underlying the strategy is positive but limited and the effects cannot be predicted with certainty.

Twelve reports representing nine studies (including two recent randomised controlled trials) used a controlled design to measure the effects of the quality improvement collaboratives intervention on processes of care or outcomes of care. The studies were based on different collaborative strategies. Seven studies evaluated the Breakthrough Series, four of these were studies on the Breakthrough Series combined with chronic care model and two were based on the Vermont Oxford Network method. A systematic review of the studies produced moderate positive results. Seven studies (including one randomised controlled trial) showed at least a positive effect of a specific selection of processes of care studied. Two studies (including one randomised controlled trial) did not show any significant effect.

As a result of flaws in the methodological quality of the studies and the heterogeneity of the intervention itself, there is no certainty that the quality improvement collaborative was responsible for an effect. Six studies reported possible differences in baseline measurement. One of the controlled studies was a Breakthrough

Series embedded in a seven year quality improvement programme. Four of the studies contained elements of the chronic care model in the intervention. Two of the controlled studies were based on the Vermont Oxford Network. This type of quality improvement collaborative differs from the Breakthrough Series in that it is long term: efforts are led and supported by ongoing data collection of individual member organisations and an ongoing infrastructure of communication and meetings exists that goes beyond a particular limited time frame of a quality improvement collaborative initiative. We were unable to disentangle the different components of an intervention or to assess interactions between longitudinal activities for quality improvement or elements of the chronic care model and collaborative components.

Fifty three (88%) of the 60 uncontrolled reports highlighted specific improvements in patient care and organisational performance that resulted from participating in a quality improvement collaborative. Several reports showed dramatic improvements of 30% to 80%. Almost all of the uncontrolled reports, however, had design limitations, were methodologically weak, and were probably biased in favour of positive findings in successful teams.

The evidence of the impact of quality improvement collaboratives is positive but limited. The apparent inconsistency between the widespread belief in and use

Table 3 | Summary of effectiveness in nine controlled studies

Controlled studies (n=9)	Studies based on Breakthrough Series		Studies based on Vermont Oxford Network
	Breakthrough Series	Breakthrough Series and chronic care model	
Two randomised controlled trials		Homer et al ^{w64} (no effect)	Horbar et al ^{w71} (mixed effect)
Six controlled before and after studies	Landon et al ^{w63} (no effect), Baier et al ^{w61} (positive effect)	Benedetti et al ^{w62} (mixed effect), Mangione-Smith ^{w65} /Schonlau et al ^{w66} (mixed effect), Asch ^{w67} /Baker et al ^{w68} (mixed effect)	Horbar ^{w69} /Rogowski et al ^{w70} (mixed effect)
One interrupted time series	Pierce-Bulger et al ^{w72} (positive effect)		

of quality improvement collaboratives and the available evidence heightens the importance of a deeper understanding of the relative strength of this intervention. Quality improvement collaboratives are, by their nature, complex and applied in many different ways. Considering that quality improvement collaboratives seem to play a key part in current strategies focused on accelerating improvement, represent substantial investments of time and funding, but may have only modest effects on outcomes at best, then further knowledge of the effectiveness of the basic components, cost effectiveness, variability within collaboratives, and success factors is crucial for determining their value. What mechanisms are responsible for the results and their variations: for example, does effectiveness depend on the topic chosen and are there specific components, supportive contextual factors, or site characteristics that enhance the effectiveness of quality improvement collaboratives? It is possible that a quality improvement collaborative works for some organisations but not for others because of inherent differences in the history and culture of organisations. The data collected in the included studies did not provide the information needed to understand and explain the findings. To understand how and why quality improvement collaboratives work it is necessary to look into the “black box” of the intervention and to study the determinants of success or failure. A detailed formative evaluation of the projects might provide additional insight into these problems. The studies needed balance between uncontrolled process oriented reports and rigorously controlled designs,¹⁰⁻¹² and a sound integration of process and effect data is needed to understand in more detail why some quality improvement collaboratives and some organisations

participating in such a collaborative are successful while others fail to change practice. We look forward to studies adding to this body of knowledge.¹³⁻¹⁵

Limitations should be considered in interpreting the results of this review. Firstly, as in any systematic review we may have missed relevant studies. We searched multiple databases, however, and checked our search with free text words with a strategy that included MeSH terms (box 1) based on key words in the relevant studies. These searches did not add new studies. Secondly, our search was limited to quality improvement collaboratives involving the five essential features described in our inclusion criteria and to English language journals. This might have introduced bias if the effectiveness described in these studies differed systematically from those involving other features and appearing in other languages. Thirdly, the key components of some quality improvement collaboratives could have been misclassified, although our abstraction process showed good inter-rater reliability.

Despite these limitations, this review shows that the evidence underlying quality improvement collaboratives is positive but still limited and that the effects cannot be predicted with great certainty.

Contributors: LMTS and MEJLH analysed the data. JJEV, RH, and RPTMG interpreted the data. All authors conceived and designed the study, drafted and revised the manuscript, and approved the final version. LMTS is guarantor.

Funding: None.

Competing interests: LMTS and JJEV work at the Dutch Institute for Quality Improvement. The views and opinions expressed are those of the authors and do not necessarily reflect those of the institute.

Ethical approval: Not required.

Provenance and peer review: Not commissioned; externally peer reviewed.

WHAT IS ALREADY KNOWN ON THIS TOPIC

The multi-institutional, quality improvement collaborative is widely accepted as a strategy in health care

Its widespread acceptance and use are not, however, based on a systematic assessment of effectiveness

WHAT THIS STUDY ADDS

Quality improvement collaboratives are complex and are applied in many different ways

The evidence underlying the strategy is positive but limited and the effects cannot be predicted with certainty

- 1 Plsek PE. Collaborating across organizational boundaries to improve the quality of care. *Am J Infect Control* 1997;25:85-95.
- 2 Solberg LI. If you've seen one quality improvement collaborative . . . *Ann Fam Med* 2005;3:198-9.
- 3 Ayers LR, Beyea SC, Godfrey MM, Harper DC, Nelson EC, Batalden PB. Quality improvement learning collaboratives. *Qual Manag Health Care* 2005;14:234-47.
- 4 Øvretveit J, Bate P, Cleary P, Cretin S, Gustafson D, McInnes K et al. Quality collaboratives: lessons from research. *Qual Saf Health Care* 2002;11:345-51.
- 5 Mittman BS. Improving patient care. Creating the evidence base for quality improvement collaboratives. *Ann Intern Med* 2004;140:897-901.
- 6 Newton PJ, Davidson PM, Halcomb EJ, Denniss AR, Westgarth F. An introduction to the collaborative methodology and its potential use for the management of heart failure. *J Cardiovasc Nurs* 2006;21:161-8.

- 7 Kilo CM. A framework for collaborative improvement: lessons from the Institute for Healthcare Improvement's Breakthrough Series. *Qual Manag Health Care* 1998;6:1-13.
- 8 Kilo CM. Improving care through collaboration. *Pediatrics* 1999;103(1 suppl E):384-93.
- 9 Wilson T, Berwick DM, Cleary PD. What do collaborative improvement projects do? Experience from seven countries. *Jt Comm J Qual Saf* 2003;29:85-93.
- 10 Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215-8.
- 11 McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C. Methods in health services research: interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ* 1999;319:312-5.
- 12 Kirkwood B. Making public interventions more evidence based. *BMJ* 2004;328:966-7.
- 13 Marsteller JA, Shortell SM, Lin M, Mendel P, Dell E, Wang S, et al. How do teams in quality improvement collaboratives interact? *Jt Comm J Qual Patient Saf* 2007;33:267-76.
- 14 Shortell SM, Marsteller JA, Lin M, Pearson ML, Wu SY, Mendel P, et al. The role of perceived team effectiveness in improving chronic illness care. *Med Care* 2004;42:1040-8.
- 15 Schouten LMT, Hulscher MEJL, Akkermans R, van Everdingen JJE, Grol RPTM, Huijsman R. Factors that influence the stroke care team's effectiveness in reducing the length of hospital stay. *Stroke* (in press).

Accepted: 21 April 2008