Collaboratives

Marlies Hulscher, Loes Schouten, and Richard Grol

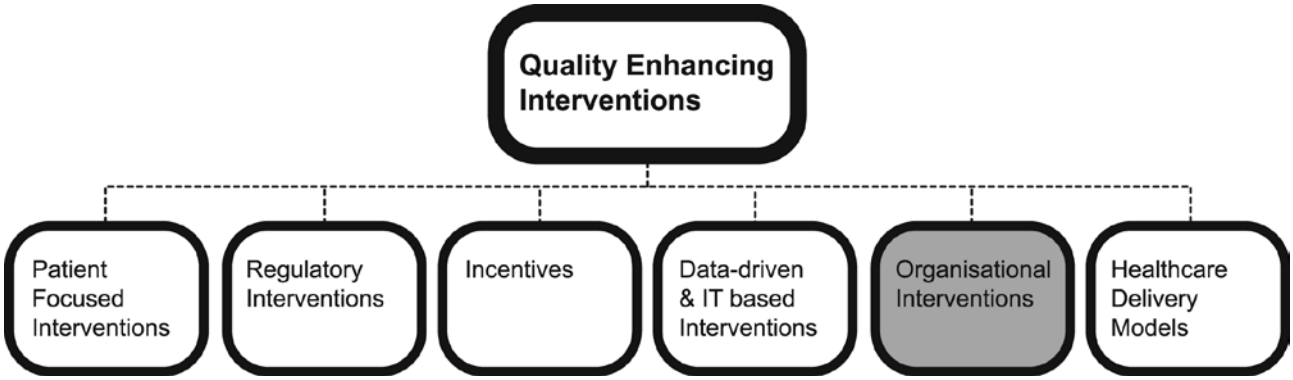
Radboud University Nijmegen

QQUIP Quest for Quality and Improved Performance

QQUIP and the Quality Enhancing Interventions project

QQUIP (Quest for Quality and Improved Performance) is a five-year research initiative of The Health Foundation. QQUIP provides independent reports on a wide range of data about the quality of healthcare in the UK. It draws on the international evidence base to produce information on where healthcare resources are currently being spent, whether they provide value for money and how interventions in the UK and around the world have been used to improve healthcare quality.

The Quality Enhancing Interventions component of the QQUIP initiative provides a series of structured evidence-based reviews of the effectiveness of a wide range of interventions designed to improve the quality of healthcare. The six main categories of Quality Enhancing Interventions for which evidence will be reviewed are shown below.



For more information visit www.health.org.uk/qquip

Acknowledgements

This study was produced as part of the Quest for Quality and Improved Performance (QQUIP), an initiative of The Health Foundation. Thanks are due to anonymous reviewers.

Published by:

The Health Foundation
90 Long Acre
London WC2E 9RA
Telephone: 020 7257 8000
Facsimile: 020 7257 8001

www.health.org.uk

Registered charity number 286967
Registered company number 1714937

First published 2009

ISBN 978-1-906461-11-9

Contents

Foreword – Collaboratives

Twenty years ago, it would not be unusual to find clinical practitioners who had worked only in their own care setting over many years. As variations in care across geographical areas have been documented extensively, so methods have emerged actively to promote exchange between professional teams from multiple sites.

The method that has emerged which has attracted the most public profile is the use of ‘collaboratives’. However, it is difficult to identify other equivalent mechanisms. Most other approaches are based on either more traditional educational approaches directed at individual professionals or the use of extrinsic performance management tools. Yet the evidence points to the need for complex interventions to achieve systematic changes in large scale complex health care settings.

In their recent BMJ article, Schouten and colleagues published a systematic review of the evidence for the impact of quality improvement collaborative. Drawing on five overview papers published between 1997 and 2003, they define the characteristics of collaborative. These articles chart the efforts of those committed to improving quality but without reliable and well developed tools. ‘The collaborative’ is still in development.

Schouten and colleagues have now updated their review of evidence of the effectiveness of collaborative for The Health Foundation and provided an analysis of the factors contributing to success which emerge – a look into the ‘black box’ of the intervention to study the determinants of success or failures.

The Health Foundation has made a considerable investment in quality improvement collaboratives. Our patient safety collaborative began in late 2004 with four hospitals, each of which worked with 15-20 change agents to implement a change package. In late 2006 this was extended to a further 20 hospitals. In 2006 we launched a collaborative to reduce maternal and neonatal mortality in Malawi, working first with 10 hospitals and then in 2008 with 64 health centres across central Malawi. In 2008 we funded a collaborative to improve stroke care in 26 hospitals in the North West of England.

Given that the method of collaborative is still in the early stages of development the effects reported are surprisingly positive, given the different health conditions and health care settings in which the collaboratives were assessed. The report’s findings of the importance of context and the limits of collaboratives in improving individual professional behaviour, for example in prescribing behaviour, are validated by other studies.

We are already making our own significant contribution to the evidence base. We have invested in four evaluations on our own collaboratives – two randomized controlled trials and two controlled observational studies. Our evaluation of these programmes will make a considerable contribution to the size and quality of the knowledge base.

The Health Foundation

Summary

Worldwide, organisations are adopting quality improvement collaborative (QIC) approaches. QICs are used in different clinical areas and organisational contexts, and they have been adopted by numerous and very different healthcare organisations. QIC initiatives represent substantial investments of time, effort and funding in the healthcare delivery system, although estimates of the total investment in and application of QICs are not available. Unequivocal evidence of the effectiveness of the method is lacking, despite the volume of ongoing QIC initiatives, the growing number of published papers, good face validity of the model, and anecdotal opinion claiming support from professionals and institutions.

This report describes the effectiveness of the QIC approach by systematically reviewing empirical studies. It updates our previous paper, published in the BMJ in 2008,[7] on the effectiveness of nine QICs, which concluded that to assess the real value of QICs it is crucial to understand how and why they work, and why they sometimes do not work. This report therefore also focuses on describing the determinants of success and failure of the approach.

Effectiveness of collaboratives

From the theoretical literature on QCIs, we propose the following definition of a 'typical' collaborative:

A QIC is an organised, multifaceted approach to quality improvement that involves five essential features: (1) there is a specified topic; (2) clinical experts and experts in quality improvement provide ideas and support for improvement; (3) multi-professional teams from multiple sites participate; (4) there is a model for improvement (setting targets, collecting data and testing changes); and (5) the collaborative process involves a series of structured activities.

We included 13 papers representing 10 studies in this review. Three types of collaboratives were tested. Eight studies were explicitly based on the Breakthrough Series (BTS) developed by the Institute of Healthcare Improvement (IHI) in Cambridge, Massachusetts. Four of these studies combined BTS with elements of the Chronic Care Model (CCM), which identifies essential elements of a healthcare system that encourage high-quality care for those with chronic disease. The BTS methodology was used in these studies to implement the elements of the CCM. Two studies were based on the Vermont Oxford Network (VON) approach, which is a data-driven, 'non-profit, voluntary collaboration of healthcare professionals dedicated to improving the quality and safety of medical care for newborn infants and their families'.

Our systematic review of these ten controlled studies produced positive but limited results. Three studies reported a positive effect; however, in two of these, it is difficult to attribute the effects solely to the collaborative activities. Five studies (including one randomised controlled trial (RCT)) showed a positive effect on some of the selected effect parameters. Two studies (including one RCT) did not show any significant effect.

Thus, the evidence of the impact of QICs is still limited and the effects cannot be predicted with great certainty. Given the widespread belief in and use of QICs, and this limited available evidence, it is important to acquire a better understanding of the factors associated with a positive impact.

Determinants of success

To understand better what makes a collaborative effective, we first looked at the theoretical literature to identify the basic elements of an effective collaborative. From five overview papers on collaboratives, we extracted expert opinion-based characteristics and success factors of collaboratives, and categorised them using our definition of a collaborative. Determinants (that is, characteristics and success factors) were described and related to: (1) the selection of a collaborative's topic; (2) the selected experts, the information they presented and the methods they used to present this information; (3) the composition of the collaborative group, the requirements set for the participating organisations, and the composition and preparation of the collaborative teams; (4) the improvement steps – setting targets, collecting data and testing changes – taken; and (5) the type of activity, and the improvement and exchange methods chosen.

This activity resulted in an overview of expert opinion-based determinants that can be used as a checklist by organisers of collaboratives to ensure that all elements of a 'good' collaborative have been included. This overview can also be used by evaluators who are interested in explaining success or lack of success following the performance of a collaborative. They can compare different collaboratives or different teams within a collaborative by simultaneously looking at their effectiveness and the presence – or not – of specific determinants.

Next, we looked at the studies included in our review to verify empirically whether any of these theoretical determinants indeed have an impact on the effectiveness of the collaborative. Our comparison showed, first of all, that the QICs' overall effectiveness did not vary between the three types of collaboratives (those based on the BTS, those based on the BTS combined with the CCM, and those based on the VON approach). Data also showed that overall effectiveness differed between the various types of effect parameters chosen: 'process of care' parameters more often showed positive results than 'patient outcomes', probably because of the chosen timeframe.

The huge variation in the topics and settings of the different collaboratives made it impossible to draw any conclusions about whether collaboratives are particularly effective for specific topics or in specific settings. Collaboratives seem to be particularly effective in improving patient education and monitoring of patients. It is, however, unclear whether patient education and monitoring is more susceptible to change or whether change is seen more often because the collaboratives put more focus on this topic.

Finally, comparisons of teams within a specific collaborative showed that 'having a dedicated nurse who provides the desired care' was related to greater effectiveness regarding lifestyle modifications and monitoring outcomes. It was also shown that 'length of time professionals used the CCM model' affected outcomes; outcomes were more favourable in longer use of CCM (three-, two- or one-year use of the CCM model, all including twelve months of the BTS). Several other variables were explored, but their influence remains unclear.

Conclusions and recommendations

Most of the studies included in our systematic review of the effectiveness of QICs showed an effect on some of the selected effect parameters. We conclude that the evidence underlying the strategy is positive but limited and that the effects cannot be predicted with great certainty. Despite the widespread belief in and use of QICs, the limited available evidence should caution organisations and governments against the wide application of collaboratives. The review of controlled studies has shown that collaboratives can work. It is, however, unclear when to apply a collaborative. More insight is needed into the factors that make collaboratives work, so that future collaboratives can be tailored to those factors.

The five overview papers on collaboratives provided a long list of probable determinants of success as assessed by experts. This information can be used by organisers and evaluators of collaboratives to ensure and to understand respectively the success of collaboratives. The ten studies included in our systematic review of the effectiveness of QICs provided little empirical evidence for these determinants of success, as only a very few determinants could be tested. Our comparison of collaboratives, however, showed that collaboratives may be particularly effective in improving patient education and monitoring of patients. In addition, our comparison of teams showed that 'having a dedicated nurse who provides the desired care' was related to the degree of effect. Taking into account the many determinants regarding teams (that is, team building, organisational support of teams, and team composition and preparation) and the theoretical concept that collaboratives are about team processes – and not about individual behaviour and people not performing well – leads to the hypothesis that collaboratives might be particularly effective for those specified changes or improvements where a team approach is vital. Collaboratives might work for those changes that address the organisational context within which the desired care is provided, where a team operates and where the team appoints persons responsible for the different tasks, maybe especially where tasks are delegated to the nurse. This hypothesis also predicts that if collaboratives want to be effective in, for example, prescribing, diagnosing, etc (that is, changes addressing individual behaviour), incentives to provide desired care have to be built into the organisation and the team process to ensure that the individual is not solely responsible for these improvements.

Based on the outcomes of our review of the literature on QICs, we propose the following recommendations for policy makers and for providers, users and evaluators of QICs.

Recommendation 1

Be careful that widespread belief in QICs doesn't result in an over-assessment of their expected impact. Ensure that collaboratives are always evaluated to test whether collaborative goals have been reached.

Recommendation 2

Use the checklist (that is, the overview of probable determinants of success) while developing the collaborative to ensure that all aspects of a 'good' collaborative have been included. In this way, the collaborative has the highest chance of success.

Recommendation 3

Use valid effect parameters and reliable data when evaluating the collaborative. Ensure that these are also used to compare the outcomes of collaborative participants with the outcomes of their colleagues in non-participating sites.

Recommendation 4

Pay attention to the organisational context within which the desired care is provided. A team approach, maybe with the special involvement of a nurse, seems to be an important predictor of success (hypothesis).

Recommendation 5

Aim to improve the organisation of care and to involve the team. If a collaborative is also used to improve individual professional behaviour (for example, prescribing behaviour), incentives to provide desired care have to be built into the organisation and the team process (hypothesis).

Recommendation 6

Use the checklist (that is, the overview of probable determinants of success) to explain the success or the lack of effect of collaboratives. Focus on comparing different teams within a collaborative instead of comparing different collaboratives. Given the huge variation in topics, settings and effect parameters addressed, it is difficult to compare collaboratives.

Recommendation 7

Integrate information on effects and determinants of success and gain insight into why some QIC teams are successful while others fail to change practice. Use this information to revise the collaborative strategy to improve its effectiveness.

Recommendation 8

More research is needed on the costs and the cost-effectiveness of collaboratives. In addition, more research is needed on spin-off effects of collaboratives: for example, does teaching participants about improvement theory and techniques indeed build the internal capacity of organisations to improve beyond the timeframe of the collaborative?

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 (QICs) to achieve rapid improvements. Teams from multiple healthcare units or organisations join forces for several months to work in a structured way to improve the delivery of health services. QICs are being used worldwide, including, for example, in the United States, Canada, Australia, the United Kingdom, Norway, Sweden and the Netherlands. In northern European countries like the UK, Sweden and the Netherlands, governments sponsor quality programmes based on this strategy.

The earliest well-documented activities of QICs are those of the Northern New England Cardiovascular Disease Study Group, established in 1986, and the Vermont Oxford Network (VON), established in 1988. Another well-known approach is the Breakthrough Series (BTS) developed by the Institute of Healthcare Improvement (IHI) in 1995.[1–3]

QICs are used in different clinical areas and organisational contexts, in the public and private sector and at national and local levels.[4] These initiatives represent substantial investments of time, effort and funding in the healthcare delivery system, although estimates of the total investment and application of QICs are not available.[5] The strength of the QIC seems to be its relatively efficient use of experts and peers and the exchange of best practice to facilitate and guide improvement.

Available evidence

A recent non-systematic review concludes that ‘the collaborative methodology, when implemented and conducted according to key conceptual principles, has significant potential to improve patient outcomes and to facilitate sustainability of quality improvement’.[6] The authors make three assertions: (1) the collaborative methodology has significant potential to improve outcomes; (2) ‘institutions working together’ and ‘leadership’ are important characteristics that differentiate QICs from other quality improvement (QI) initiatives; and (3) the collaborative methodology has significant potential to sustain clinical improvement after the completion of the formal collaborative period. 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 does it make clear on what types of QIC the conclusions are based.

Clear evidence of the effectiveness of the method is lacking, despite the volume of QIC initiatives, the growing number of published papers and good face validity of the model.[4] Little is known about the effectiveness of QICs or specific components that enhance their effectiveness, and there is hardly any information about their cost-effectiveness and sustainability.

This report describes the effectiveness of the QIC approach by systematically reviewing empirical studies. It updates our previously published paper on the effectiveness of QICs, and empirical data are summarised.[7] The earlier paper showed the need for a better understanding of the effectiveness of QICs; this report also focuses on describing the determinants of success and failure of the approach.

Definition

There are different types of multi-organisational collaboratives whose purpose it is to improve care.[1–3] The term ‘quality improvement collaborative’ seems to be used for different multifaceted packages that focus on accelerating outcomes.[4] Therefore, for the purpose of this review, we started by distinguishing key components of the QIC. The next step comprised a review of the effectiveness of the QIC as defined.

To distinguish and define key components, studies from the search described below (see ‘Data sources’) were included if they gave an overview of key elements or components of QICs applied in healthcare. Titles of articles and abstracts identified in the search were reviewed for relevance by two researchers (LS and MH). Each potentially eligible study was independently assessed by the two reviewers. We also reviewed the reference lists of the included papers.

Our search identified five studies that met our inclusion criterion in giving an overview of key elements of QIC applied in healthcare.[1,4,8–10] Box 1 shows the results of our analysis of the theoretical literature on QICs. The different authors identify five to eight key components that can be used when describing a QIC. Based on our analysis, we propose the following definition and criteria to select a ‘typical’ collaborative.

A QIC is an organised, multifaceted approach to quality improvement that involves five essential features:

1. There is a **specified topic**, that is, a subject with large variations in care or gaps between best and current practice.
2. 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 improvement knowledge.
3. A critical mass of multi-professional teams from **multiple sites**, who are willing to improve care and share, participate.
4. There is a **model for improvement** that focuses on setting clear and measurable targets, collecting data and testing changes on a small scale to advance reinvention and learning by doing.
5. The **collaborative process** involves a series of structured activities (meetings, ‘listserv’ activity, visiting facilitators) in a given time frame to advance improvement, the exchange of ideas and sharing of experiences between the participating teams.

Indeed, as Øvretveit describes:[4]

> A quality improvement team which is not part of a collaborative – a traditional QI team – uses similar methods to plan and test changes but chooses its own problem and spends time working on diagnosing the problem and analysing causes before planning and testing changes. It may work on subjects for which there is no previous research on the potential performance that could be achieved or evidence of effective treatments or organizational forms. A team in a collaborative has already decided on the problem to work on (i.e. criterion 1) and is given the evidence and change ideas, thus cutting out much of the investigation work of a traditional quality project (i.e. criterion 2 and 4). It also gets expert support (i.e. criterion 2) and peer stimulus (i.e. criterion 3 and 5) which might not otherwise be available.

Box 1: Key components of the quality improvement collaborative as described by the various authors

Plsek[1]	Kilo[8]	Kilo[9]
Key concepts behind multi-organisational collaborative improvement efforts (7)	***Key elements of BTS (5)***	***Key components BTS (7)***
1. ***Multiple organisations*** (group forms around a focused clinical topic) 2. ***Quantified variability in process or outcome*** (group defines a consistent data set and associated collection methods and are confronted by variation in practice) 3. ***Open sharing*** (all parties agree to full disclosure, within the group, of both outcome data and details of practice) 4. ***Internal process characterisation*** (group commits to understanding its own practices and processes. To do this, it needs quality/ change management skills) 5. ***Formal benchmarking visits and identification of 'best practice'*** (to discover practices that lead to superior performance, on-site visits by teams, literature review, expert panels, focused conferences are performed) 6. ***Replication efforts*** (adaptation of improvement ideas to fit the context and resources of a specific organisation) 7. ***Measured improvement*** (group tracks continuously throughout its life span the set of indicators it intends to impact)	1. ***Topic selection*** (criteria: elimination of gap can lead to demonstrably higher quality of care, high-level performance examples, significant financial savings through performance improvement, leader to chair the collaborative) 2. ***Theory for improvement*** (expressed in terms of aims (what can be achieved), measures (end points processes, and outcomes to be monitored) and key changes (what changes have the highest likelihood of improving performance)) 3. ***Multiple sites*** (critical mass (25–40) of participating sites is needed to cultivate a useful exchange of ideas, experience and learning) 4. ***The model for improvement*** (focus on clinical topic, continuous improvement at the background, theory for improvement, Shewhart or Plan Do Study Act cycle) 5. ***Focus*** (determine valuable content, learning session handouts of 10–20 pages with critical information, discussion of key changes with examples, one-page report format) 6. ***Maintenance of tension for change*** (through senior leadership support, shared deadlines, monthly assessment of progress, peer pressure) 7. ***Spread*** ('how to' knowledge is placed within the public domain; National Congress, public report, journal articles, conferences and meetings)	1. ***Identification of topics*** (criteria: gap, examples of improved practices and outcomes, benefit of closing the gap, national tension for improvement) 2. ***Identification and consolidation*** of relevant scientific (what to do) and improvement of ***knowledge*** (how to do it) 3. Recruitment and preparation of ***collaborative participants*** (multi-institutional or multi-site work groups) 4. ***The collaborative process*** (series of structured activities designed to advance the improvement work as rapidly as possible: learning sessions, conference calls, email 'listserv' discussion, monthly exchange of printed reports, visits to collaborative institutions) 5. ***Deployment of knowledge*** (concluding National Congress, monograph, peer-reviewed articles) 6. Limitations of time and focus create ***tension for change*** 7. Each organisation ***collects and maintains its own data*** (establishment of performance measurement systems to collect data and test change)

Øvretveit et al[4]	**Wilson et al[10]**
Features of collaboratives (8)	***Features of collaboratives as critical determinants of effectiveness (7)***
1. Participation of ***a number of multi-professional teams*** with commitment to improve and share 2. A ***focused clinical or administrative subject*** 3. Evidence of large ***variations*** in care or ***gaps*** between best and current practice 4. ***Participants learn from experts*** about evidence for improvement, change concepts, practical changes that worked at other sites and about quality improvement methods 5. Participants use a ***change-testing method*** to plan, implement and evaluate small changes in quick succession, eg the rapid cycle improvement method 6. Teams set ***measurable targets*** and ***collect data*** to track performance 7. ***Participants meet*** at least twice to learn the methods, report changes and results, share experiences and consider spread 8. ***Between meetings participants continue to exchange*** ideas, organisers provide extra support through visits, email, conference calls	1. ***Sponsorship*** (to improve an effort's credibility or perceived importance) 2. ***Topic*** (choice of an ideal topic) 3. ***Ideas for improvement*** (expert groups (quality improvement, system thinking, practical knowledge) to suggest improvements) 4. ***Participants*** (careful selection of volunteers, multidisciplinary teams including clinicians, inclusion of all personnel (clinical and non-clinical) directly affected by the changes, 2–8 people) 5. ***Senior organisational leadership support*** (eg chief executive officer of a hospital or health system, head of a department) 6. ***Preliminary work*** (having participants collect data or analyse the system before the collaborative started was important to understand the organisation/the problem and to provide baseline data for later comparison) 7. ***Strategies for learning about and making improvements*** (quality improvement theory and techniques, an improvement model, sharing improvement strategies, competition, 2–7 carefully planned meetings, balanced activities to meet varying needs and preferences, social exchange within and between teams, optimal central support, formal report)

Effectiveness of collaboratives: methods

Data sources

We searched the Medline, Cinahl, Embase, Cochrane and Psychinfo databases for literature about QICs in the period from January 1995 to June 2006 inclusive. We started with a Medline search for free text terms describing QICs, and we combined the keywords (non-MeSH): 'quality and improvement and collaborative' or '(series or project) and breakthrough' (Appendix 1). The same steps were repeated for the other databases. We also reviewed the reference lists of the included papers.[7]

In March 2008 (week 2) an update of the literature search was performed. We limited our search to the Medline database, as an evaluation of the efficacy of our search method showed that a combination of Medline with a review of reference lists was as effective as searching all data sources. The systematic search of Cinahl, Embase, Cochrane and Psychinfo (1995–2006) resulted in the inclusion of one extra (uncontrolled) study that was ultimately also found in the reference lists of the included studies in Medline.

Study selection

We included studies that (1) were written in English, (2) contained data about the effectiveness of care processes or outcomes, (3) had a healthcare setting, and (4) met the criteria for a QIC as described above.

Data extraction

Each potentially eligible study was independently assessed by the two reviewers 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 (for definitions, see the data collection 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 about topics, study designs, settings, numbers of participants, characteristics of the collaborative strategy, and relevant results. Two authors (LS and MH) independently completed this checklist for each study. Disagreements about 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-analytic techniques for pooling results because the studies used many different effect parameters.

Included studies

We identified 1,104 abstracts of studies published between January 1995 and June 2006 inclusive. While screening the abstracts, we excluded 929 papers because they did not fulfil our inclusion criteria or because of duplication. A total of 175 articles were requested for detailed review. We excluded an additional 107 articles; 49 papers were studies containing no original data on effectiveness or outcomes and 58 studies did not meet our QIC criteria. The QIC criteria most often not fulfilled were criterion 5, the 'collaborative process', where there was a lack of structured activities aimed at exchange of ideas and information, and criterion 4, the 'model for improvement', where participants did not set aims, collect data

and test changes. Reviewing the reference lists of the remaining 68 studies led to 18 more studies. Four of them fulfilled our inclusion criteria and did not duplicate any of the studies already included. A total of 72 papers entered our study.

In our 2008 update, we identified 66 additional abstracts, including the non-systematic review[6] described above. The reference list of the review provided one additional paper (uncontrolled study) that met our inclusion criteria. Eleven articles were requested for detailed review, of which one was not available and three did not 'contain data about the effectiveness of care processes or outcomes'. Reviewing the reference lists of the remaining seven articles provided no additional papers. Eight additional papers entered the study.

Uncontrolled studies

Of the 80 papers included in our study, 67 reports[11–77] (84 per cent) used an uncontrolled study design (see Appendix 2). These reports were concentrated in management and practitioner-oriented journals. The first report was published in 1996; almost half of the reports (32/67) were published between 2004 and 2007.

Fifty-seven (85 per cent) of these reports described a collaborative explicitly based on the BTS. Several studies contained elements that went beyond a simple QIC. The QICs in these reports were embedded in ongoing quality initiatives or used comprehensive systems for measuring performance.

The reports addressed a wide range of topics including, for example, diabetes care, patient access, emergency care and patient safety. Forty-five (67 per cent) of the reports described activities performed in the hospital setting.

Fifty-nine (88 per cent) of the 67 uncontrolled reports highlighted specific improvements made in patient care and organisational performance that resulted from participating in a QIC. Several reports demonstrated dramatic improvements of 30 to 80 per cent.

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 parameters rather than reviews of medical records, included only anecdotal information, or selected samples from self-selected sites. Thus, almost all the uncontrolled reports suffered from design limitations, were methodologically weak, and were probably biased in favour of positive findings in successful teams. Conclusions on effectiveness could not be drawn from these reports owing to a lack of adequate reporting procedures on data collection, analysis and objective evaluations.

Effectiveness of collaboratives: results

Description of studies

Twelve reports (representing nine individual studies[78–89]) used a 'comparison group' study design. The thirteenth paper[90] described one study that used an interrupted time series design (Table 1). Most of these ten individual studies were published in 2004 and 2005. The characteristics of the collaborative strategy used in the studies varied. Eight studies[78–85,89,90] were explicitly based on the BTS. Two studies[86–88] were based on the VON. Box 2 gives an overview of the different collaborative strategies.

Box 2: Overview of collaborative strategies used in ten controlled studies

Controlled studies	**Collaborative based on Breakthrough Series (BTS)**	**Collaborative based on Vermont Oxford Network (VON)**
Reference	www.ihi.org	www.vtoxford.org
Number of studies	8	2
Description of collaborative strategy	A BTS is a short-term (6–15 months) learning system that brings together a large number of teams from hospitals or clinics to seek improvement in a focused topic area. In 1995, the Institute of Healthcare Improvement (IHI) in Cambridge, Massachusetts developed this series of collaborative projects. The driving vision behind the BTS 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 lies fallow and unused in daily work. The BTS 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 topic areas where they want to make improvements.	The VON is a data-driven, 'non-profit, voluntary collaboration of healthcare 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 VON includes more than 400 neonatal intensive care units (NICUs), predominantly in the United States. VON facilitates a coordinated programme of research, education, and quality improvement. To support this programme, the Network maintains a database for 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 VON receive centre-specific, routinely prepared confidential reports.

Four of the eight studies explicitly based on the BTS combined it with elements of the CCM (www.improvingchroniccare.org).[79, 81–85] The CCM identifies six essential elements of a healthcare system that encourage high-quality care for those with chronic disease. These elements are the community, the health system, self-management support, the delivery system design, decision support, and clinical information systems. The elements aim to foster productive interactions between prepared, proactive

practice teams and well-informed, motivated patients. Provider roles, standards of care and treatment aims are explicit and evidence-based. Care management is linked to a patient registry, which provides reminders, data collection, scheduling of care and performance data to caregivers. Patients are supported through self-management education, participatory goal setting and written care plans. The BTS methodology was used in these studies to implement the elements of the CCM.

Evidence from randomised controlled trials (RCT) (n = 2)

There is inconclusive evidence from RCTs for the impact of QICs. One RCT showed no effect on any of the nine effect parameters.[81] The second RCT showed effect on two of three specific process of care parameters, but not on two patient outcomes, nor on twenty of the twenty-three secondary process of care parameters (Table 2).[88]

Homer et al (2005)[81] evaluated the impact of a 12-month quality improvement programme (BTS and CCM) on care (3 process of care parameters) and outcomes for children with asthma (6 patient outcomes). Using telephone interviews, information was gathered from 43 primary care practices, with 13,878 paediatric patients with asthma, randomised to intervention (22) and control groups (21). Change from baseline (before/after intervention = 294/236 children, before/after control = 337/254) in the proportion of children with persistent asthma who received appropriate medication therapy for asthma (two process of care parameters) and in the proportion of children whose parents received a written management plan for their child's asthma were determined. No overall effect of the intervention on these primary effect parameters regarding processes of care was found. No significant differences between intervention and control group children were found on the secondary effect parameters regarding patient outcomes (% any asthma attack in past 12 months; mean limitation from strenuous exercise; % any asthma hospitalisation in the past 12 months; % any asthma ED visit in the past 12 months; patient experience of care; functional outcomes).

Horbar et al (2004)[88] examined the effects of a multifaceted collaborative quality improvement advice (VON) among 114 NICUs, randomised to intervention (57 NICUs) and control group (57 NICUs). The exact duration of the collaborative activities was unclear. The VON maintains a database for very low birth weight infants at the 300 member hospitals. This database was used to determine the effects of the intervention (intervention before/after = 3,332/3,313 infants of 23–29 weeks' gestation, control before/after = 2,850/2,726 infants) on primary (3 process of care parameters and 2 patient outcomes) and secondary outcomes (10 process of care parameters and 13 patient outcomes). Compared with those in control hospitals, infants in intervention hospitals were more likely to receive surfactant in the delivery room (adjusted odds ratio (OR) 5.38 (95 per cent confidence interval (CI) 2.84 to 10.20)), were less likely to receive their first dose more than 2 hours after birth (adjusted OR 0.35 (95% CI 0.24 to 0.53)) and received the first dose of surfactant sooner after birth (median of 21 minutes v 78 minutes, $p < 0.001$, adjusted hazard ratio 1.57 (95% CI 1.42 to 2.07)). There were no significant differences in infant outcome parameters (death before discharge from hospital and pneumothorax). Secondary effect parameters included common morbidities and co-interventions for preterm infants. Of these 23 parameters, only 3 showed significant differences between intervention and control group. Intervention arm infants were significantly more likely to be intubated in the delivery room (adjusted OR 1.65 (95% CI 1.19 to 2.29)) and to receive surfactant at any time (adjusted OR 1.55 (95% CI 1.08 to 2.23)). Severe intraventricular haemorrhage (grades 3 and 4) was significantly lower in the intervention arm (adjusted OR 0.70 (95 per cent CI 0.56 to 0.87)). No effect was found in 20 of the 23 remaining secondary effect parameters.

Evidence from controlled before/after (CBA) and interrupted time series (ITS) studies (n = 8)

The outcomes of 7 of the 8 non-randomised trials indicated some positive impact of QICs. These 7 studies[78,79,82–87,89,90] showed a positive effect on some of the selected effect parameters. One study[80] did not show any significant effect (Table 2).

Pierce-Bulger et al (2001)[90] performed an ongoing evaluation of one of the 9 teams that participated in a community-wide learning collaborative (ITS). They examined whether post-neonatal infant mortality for Anchorage natives could be reduced (patient outcome); their death rate was 3 times higher than that for other Alaskan infant population groups. In 1993 the team participated in the IHI's collaborative. The team embedded the collaborative into longitudinal quality activities, including a 'one-stop shopping' clinic to provide convenient prenatal, neonatal and postpartum care, and home-visiting services until the infant reached 12 months of age. These latter activities started in June 1994. During the 7 years of the programme (1993–1999), there was an increase from a pre-programme average of 55 days (1989–1994) to an average of 114 days between neonatal deaths (1995–2000). It is difficult to disentangle the effects of these longitudinal quality improvement activities from those of the collaborative.

Baier et al (2004)[78] examined the outcomes of 21 nursing home facilities that participated in a 15-month collaborative intervention. Of these facilities, 17 completed the project. Using chart review and the minimum data set (www.cms.hhs.gov/medicaid/mds20), they compared the pre-test/post-test pain outcomes (patient outcome) of the intervention group (n = 15 nursing homes, 276 patients) to those of the remaining Rhode Island facilities (control, n = 72) (CBA). Facilities with denominators less than 30 were excluded from the analyses of this quasi-experiment. The decrease in prevalence of 'residents with pain' in the intervention group was significantly greater than the decrease in the remaining facilities in Rhode Island (–5% v –1.5%, p = 0.003). The authors also evaluated (uncontrolled) 7 parameters in pain management processes of care. Non-pharmacological processes of care (including pain assessment and treatment) showed significant improvement between the baseline and follow-up periods for all 3 process of care parameters. One of the 4 effect parameters regarding pharmacological processes of care (including pain medication prescription or change in medication) showed a trend toward improved adherence ('prescriptions of ATC or as needed WHO step II or step III pain medications in residents with daily pain and moderate or severe pain', p = 0.06).

Landon et al (2004)[80] evaluated, in a controlled pre- and post-intervention study (CBA), the effectiveness of a 16-month quality improvement collaborative in improving the quality of care for HIV-infected patients. They measured changes in 8 process of care parameters abstracted from medical records of pre- and post-intervention samples of patients at each study clinic. They used 44 intervention clinics (before/after = 3,190/3,216 HIV-infected patients) and 25 control clinics (before/after = 1,761/1,919 patients) matched by location (urban or rural), region, size and clinic type. Differences in changes in the process of care parameters were not statistically significant. The proportion of patients with appropriate antiretroviral therapy (2 parameters: receipt of highly active anti-retroviral therapy and controlled viral load on last visit), appropriate screening tests and prophylaxis (5 parameters: tuberculosis screening, influenza shot, hepatitis C status, pap smear, prophylaxis against Pneumocystis carinii pneumonia) or with appropriate access to care (1 parameter: visits in 3 or 4 quarters) did not differ between intervention and control sites.

Benedetti et al (2004)[79] examined in 1 organisation, Rockwood Clinic, the effects of participation (12 months) in the Washington state Diabetes Care Collaborative. This Collaborative promoted spread of the CCM. Rockwood is a multispecialty group practice with 9 primary care sites (47 primary care providers). The paper compared the results of 11 participating primary care providers (managing 698 patients with diabetes) with those of the 19 primary care physicians (managing 1,300 patients with

diabetes) who did not participate (CBA). The authors did not describe how data were gathered, nor did they clearly describe how many patients were measured. The results showed that 7 of the 12 parameters were significantly better for participating providers (see below: only 6 parameters are specified in the paper). Of the 7 process of care parameters measured, 3 showed significant differences. Patients of participating providers received significantly more often an annual eye (exact % intervention patients v % control patients = unclear (histogram); 15% v 43%) and foot exam (exact % intervention patients v % control patients = unclear (histogram); 71% v 32%). Participants had a significantly higher percentage of patients older than 40 years who were taking acetylsalicylic acid (exact % intervention patients v % control patients = unclear (histogram); 63% v 29%). Three out of the 5 patient outcomes showed significant improvements. Compared to non-participating providers, participating providers experienced significantly better rates of patients with glycohaemoglobin (A1C) < 9.5 (exact % intervention patients v % control patients = unclear (histogram); 92% v 83%), low-density lipoprotein (LDL) < 130 (75% v 45%, $p < 0.05$) and blood pressures < 130/85 mm Hg (49% v 35%, $p < 0.05$). The authors showed that length of time in the CCM affected outcomes; outcomes were more favourable for the highest participation level (level I = three years' use of the model, 2 providers; level II = two years' use of the model, 3 providers; level III = 1 year's use of the model, 6 providers).

Mangione-Smith et al (2005)[82] and **Schonlau et al (2005)**[83] examined whether a 12-month collaborative to improve asthma care positively influenced processes and outcomes of that care. They reported the results of their QIC in 2 papers, splitting the outcomes for children[82] and adults.[83] A total of 26 primary practices participated in the collaborative. Of these 26 practices, 9 participated in the evaluation of pediatric asthma care and 6 participated in the adult care group. They choose internal comparison sites to ensure that control patients had similar personnel and organisational factors (CBA). Four of the 9 organisations focusing on children and 3 of the 6 organisations focusing on adults provided such comparison sites. In both evaluations a medical records review was supplemented with information from telephone interviews.

Mangione-Smith[82] et al conclude that the intervention improved some important aspects of processes of pediatric care that have previously been linked to better outcomes. For the 8 selected process of care parameters, the medical records of 348 children in the intervention group were, at pre- and post-measurement, compared with the information of 153 children in the control group. The overall process of asthma care improved significantly in the intervention group but remained unchanged in the control group (change in process score +13% v 0%; $p < 0.0001$). Six of the 8 process of care parameters showed significantly greater improvement in the intervention than in the control group: children should have a peak expiratory flow rate measurement at least annually (post-measurement intervention v control 49% v 4%; $p < 0.0001$); they should have written action plans in their medical records (42% v 3%; $p < 0.0001$); they should have ≥ 2 follow-up visits annually (86% v 78%; $p < 0.004$); they should be educated in self-management (41% v 17%; $p < 0.0001$); they should be instructed in use of metered dose inhalers (30% v 9%; $p < 0.002$); and they should have a record of collaborative goal setting in their records (10% v 0%; $p < 0.003$). No improvement was shown regarding the prescription of ß-2 antagonists, nor in the performance of follow-up visits after the medication was changed. Two of the 3 process of care parameters regarding self-management were supported by the outcomes of the interviews (intervention group 385 parents; control group 126 parents); peak flow monitoring (post-measurement intervention v control 70% v 43%; $p < 0.0001$) and written action plans (41% v 22%; $p = 0.001$) improved, while goal setting did not. In these interviews, parents of the intervention group reported significantly better patient outcomes: general health-related quality of life (80.2 v 77.0; $p = 0.05$) and asthma-specific quality of life related to treatment problems (88.6 v 85.3; $p < 0.05$). No impact was found on asthma-specific quality of life related to symptoms, on family functioning, parent satisfaction or adolescent satisfaction, on acute care service use, on missed school days or parent lost work days, on asthma knowledge or on use of long-term controller medications. All these parameters except for the latter were patient outcomes.

Similar outcomes were reported by Schonlau et al[83] regarding adult care. For 8 selected process of care parameters, the medical records of 109 adults in the intervention group were, at pre and post-measurement, compared with the information of 76 adults in the control group. The overall process of asthma care improved significantly in the intervention group but remained unchanged in the control group (change in process score +10% v 1%; $p < 0.003$). Four of the 8 process of care parameters showed significantly greater improvement in the intervention than in the control group: peak expiratory flow rate measurement at least annually (post-measurement intervention v control 28% v 14%; $p < 0.03$); written action plans (27% v 0%; $p < 0.0001$); instructions on use of metered dose inhalers (22% v 7%; $p < 0.04$); and collaborative goal setting (7% v 0%; $p < 0.03$). No improvement was reached in ≥ 2 follow-up visits annually; self-management education; and the prescription of ß-2 antagonists or ß-blockers. One of the 4 process of care effect parameters regarding self-management was supported by the outcomes of the interviews (intervention group, 123 adults; control group, 62 adults): more adults in the intervention group attended educational sessions (20% v 5%; $p = 0.03$); peak flow monitoring, goal setting and written action plans did not improve. In these interviews, adults in the intervention group reported significantly better patient outcomes on satisfaction with communication (62% v 39%; $p = 0.02$). None of the remaining patient outcomes on quality of life and use of acute care or bed days resulting from asthma-related illness differed between groups.

Asch et al (2005)[84] and **Baker et al (2005)**[85] described a collaborative in which 14 sites participated during a 12-month period to achieve rapid changes in chronic heart failure (CHF) care based on the CCM. Asch[84] conducted a quasi-experiment in 4 volunteer organisations participating in the collaborative and 4 comparable control organisations (CBA). A computerised chart abstraction tool was used to review a total of 489 medical records obtained from these sites (from 261 intervention patients and 228 control patients). A list of 25 effect parameters (including 21 process of care parameters (4 diagnostic, 4 medication, 5 follow-up, 8 counselling) and 4 patient outcomes) was compiled and each parameter was operationalised into its component data elements for chart abstraction. Participating sites showed greater improvement than control sites for 9 of the 21 process of care parameters. Two of the medication parameters (use of lipid lowering and angiotensin converting enzyme inhibition therapy) and 7 of the 8 counselling parameters showed greater improvement, but none of the diagnostic parameters or follow-up parameters showed greater improvement. The overall improvement was greatest for parameters of education and counselling (24% v –1%, $p < 0.0001$). Participant sites had higher rates of counselling regarding medications, diet, exercise, weight control, disease management, water weight management and goal setting compared with control sites. The participating sites showed no improvement over the control sites for any patient outcomes.

Baker et al[85] conducted a cross-sectional telephone survey of patients in 6 of the participating sites (n = 387) and 6 comparable control organisations (n = 414). If the above mentioned counselling activities were truly performed more often by participant sites, the authors expected that participant patients had greater knowledge (15 patient outcomes), received more educational components (14 process of care parameters), had better communication with their providers (4 process of care parameters), had better self-management activities (4 process of care parameters), had higher self-efficacy (3 patient outcomes), satisfaction (4 patient outcomes) and quality of life (physical and mental scores, patient outcomes), and had less heart failure symptoms (7 patient outcomes) and healthcare use (3 patient outcomes). Patients in the intervention group had higher knowledge scores on 8 of the 15 items, received more patient education (self-reported) (5 of 14 items) , were more positive about communication (all 4 items) and reported more self-management behaviours (2 of 4 items). Participants had similar numbers of outpatient visits but fewer emergency department visits (adjusted difference –0.25, $p = 0.02$) and hospitalisations (adjusted difference –0.13, $p = 0.007$) than controls. No differences in relation to self-efficacy, satisfaction, quality of life or heart failure symptoms were found.

Horbar et al (2001)[86] and **Rogowski et al (2001)**[87] evaluated, in a CBA, whether their 3-year collaborative improved the quality[86] and cost[87] of neonatal intensive care among 10 NICUs. Using the VON database, information for infants with birth weight 501g to 1,500g was collected. Six participating NICUs worked on infection and 4 focused on chronic lung disease.

Horbar et al[86] compared the infection group NICUs (patient outcomes: infection rate with coagulase-negative staphylococcal pathogens and infection rate with other pathogens) and the lung disease group NICUs (process of care parameter: rate of supplemental oxygen administration at 36 weeks and patient outcome: rate of death before 36 weeks) to 66 other NICUs that served as a contemporaneous comparison group (note: one of the comparison NICUs did not participate in the 1997 database). Logistic regression modelling showed a significant greater change for the participating NICUs relative to the comparison NICUs for the infection rate with coagulase-negative staphylococcal pathogens ($p = 0.001$), but not for the rate of infection with other bacterial pathogens ($p = 0.68$). The rates of infection after the third day of life with coagulase-negative staphylococcal pathogens for preterm infants decreased from 22.0 per cent in 1994 (n = 745 infants) to 12.3 per cent in 1997 (n =789 infants) at the 6 infection group NICUs. At the comparison NICUs, the infection rate with coagulase-negative staphylococcal pathogens changed from 15.4 per cent in 1994 (n = 5,108 infants at the 66 infection group NICUs) to 16.5 per cent in 1997 (n = 5,572 infants at the 65 infection group NICUs).

Similarly, logistic regression modelling showed that the change at the participating NICUs was significantly different from that at the comparison NICUs for the rate of supplemental oxygen administration at 36 weeks ($p = 0.049$) but not for the rate of death by 36 weeks ($p = 0.14$). The rates of supplemental oxygen administration at 36 weeks decreased from 43.5 per cent in 1994 (n = 188 preterm infants) to 34.0 per cent in 1997 (n =185 infants) at the 4 lung disease group NICUs. At the comparison NICUs, the rates of supplemental oxygen administration rose slightly, from 36.3 per cent in 1994 (n = 2,411 preterm infants) to 38.7 per cent in 1997 (n = 2,719 infants).

Rogowski et al[87] compared the median treatment cost per infant with birth weight 501 to 1,500 g at the collaborative NICUs (patient outcome) to the cost per infant at 9 comparison NICUs. The VON database was supplemented by the collection of data on treatment costs to evaluate the results of the collaborative on the costs of patient care. When patient mix and other factors known to affect treatment costs are controlled for in a multivariate regression, treatment costs in 1997 at NICUs in the infection group declined significantly ($p = 0.0037$) relative to the comparison group. For NICUs in the lung disease group, however, there was no significant difference in treatment costs.

Howard et al (2007)[89] describe an organ donation breakthrough collaborative to encourage adoption of 'best practice' for identifying potential donors and obtaining consent for deceased organ donation. They evaluated the impact of the first phase (12 months) on organ donation rates (patient outcome). Their CBA included 95 hospitals that participated and a control group of 125 hospitals. Administrative data from the Organ Procurement and Transplantation Network to compute the conversion rate in each hospital and time period were used. The pre-period was the year before the start of the collaborative (September 2002 to August 2003); the post-period was the final 6 months of the first phase (March 2004 to August 2004). The collaborative also included feedback on donation rates. The conversion rate is the proportion of eligible donors who became actual donors. Pre-period conversion rates in both groups were similar: 52 (participants) and 51 per cent (control hospitals). In the post-period, the conversion rate increased to 60 per cent among collaborative participants and remained unchanged among control hospitals. The relative change was 8 per cent (95% CI: 2–13%, $p < 0.001$). The impact of the collaborative varied at the hospital level. Of the collaborative hospitals, 67 per cent experienced an increase in conversion rates and 33 per cent a decrease. Of the control hospitals, 44 per cent experienced an increase and 56 per cent a decrease. The authors conclude that the collaborative led to an increase in donation rates at participating hospitals.

The first phase was followed by a second phase. The authors note that the conversion rate among collaborative hospitals increased during the first phase and continued increasing during the second phase. The conversion rate among control hospitals was fairly constant during the first phase but increased steeply during the second phase. Nationwide, there were large increases in the total number of organ donors during the first phase. The conversion rate was rising in collaborative hospitals before the start of the collaborative. It is unclear whether the conversion rate would have continued to rise in the absence of the collaborative.

Effectiveness of collaboratives: summary and conclusions

Our systematic review of the ten studies produced some positive results. Three studies reported a positive effect.[78,89,90] In two of them there are some concerns about whether the effect can be attributed solely to the collaborative activities.[89,90] Five studies (including one RCT) showed a positive effect on some of the selected effect parameters.[79,82–88] Two studies (including one RCT) did not show any significant positive effect.[80,81] This is the first systematic review of the effectiveness of QICs. It shows that the evidence underlying the strategy is still limited and the effects cannot be predicted with great certainty.

QIC participants are not only provided with information on evidence and best practice regarding a medical topic, they are also provided with 'improvement knowledge' that focuses on quality improvement and change management knowledge and skills. Teaching participants about improvement theory and techniques is thought to be extremely important, building the internal capacity of organisations to improve beyond the timeframe of the collaborative. These benefits to an organisation were, however, not measured in the included studies.

None of the studies looked at the cost-effectiveness of the collaborative approach. Only one of them provided information on costs: the VON collaborative (neonatal intensive care infections and chronic lung disease) compared the median treatment cost per infant at the collaborative NICUs to the cost at the comparison sites.[87]

Fifty-nine (88 per cent) of the 67 uncontrolled reports[11–77] highlighted specific improvements made in patient care and organisational performance that resulted from participating in a QIC. Several reports demonstrated dramatic improvements of 30 to 80 per cent. However, almost all uncontrolled reports suffered from design limitations, were methodologically weak, and were probably biased in favour of positive findings in successful teams.

There are limitations that should be considered when interpreting the results of this review. First, as in any systematic review, it may be possible that we have missed relevant studies. However, we searched multiple databases and we checked our search with free text words using a strategy that included MeSH terms based on keywords in the relevant studies. These searches did not add new studies. On the other hand, confining the search to the English language may mean that some studies have been missed, particularly given the extensive uptake of this approach in some Scandinavian countries. Similarly, we did not review the grey literature, which means that some evaluation reports – for example, those commissioned by government agencies investing in this approach – have been missed. This literature is, however, unlikely to reveal any more controlled studies.

Second, our search was limited to QICs involving the five essential features as described in our inclusion criteria. This might introduce bias if the effectiveness described in these studies differs systematically from those involving other features. Third, the key components of some QICs could have been misclassified, although our abstraction process showed good inter-rater reliability.

Determinants of success: methods

The evidence of the impact of QICs is still limited. The apparent inconsistency between the widespread belief in and use of QICs and the available evidence heightens the importance of a deeper understanding of the relative strength of the QIC approach. QICs are, by their very nature, complex and applied in many different ways. Considering that QICs seem to play a key role in current strategies focused on accelerating improvement, but may have only modest effects on outcomes at best, it is crucial to gain further knowledge of QICs' effectiveness, inter- and intra- collaborative variability and success factors to determine their value.

To understand how and why QICs work, it is necessary to understand in more detail why some QICs or collaborative teams are successful while others fail to change practice. What mechanisms are responsible for the results and their variation? Does effectiveness depend on the topic chosen? Are there specific components, supportive contextual factors or site characteristics that enhance the effectiveness of QICs? Is it possible that a QIC works for some teams but not for others due to inherent organisational heterogeneity in history and culture?

To reach a better understanding of the QIC's effectiveness, two approaches were taken. We started with looking at the theoretical literature to identify the elements of an effective collaborative. What are the success and fail factors as mentioned in the theoretical literature on collaboratives? In this manner we may be able to identify the determinants that the different experts agree most commonly influence the effectiveness of a collaborative. For this approach, a systematic search for determinants of success was performed using the overview papers included in our systematic review of the impact of collaboratives. Two authors (MH and LS) independently extracted collaboratives' characteristics and success factors from this theoretical literature, and categorised them using our definition of a collaborative.

Next, we looked at the ten studies included in our systematic review of the effectiveness of QICs to verify empirically whether any of these theoretical determinants indeed have an impact on the effectiveness of a collaborative. We compared all collaboratives to detect similarities between the collaboratives and their effectiveness. In this manner we can answer questions such as 'are specific types of collaboratives more effective then others?', 'are specific effect parameters or specific topics more susceptible to change?' and 'are specific organisations or settings more susceptible to change?' Similarly, we compared teams within a specific collaborative to detect similarities in the various teams participating within the collaborative and their effectiveness. In this manner we can answer the question, 'what characteristics of a team make it more effective than others?' This question can only be answered if the authors objectively measured – for example, in subgroup analyses – what aspects of the team resulted in more (or less) change and, therefore, what were the success factors that contributed to success?

Determinants of success: results of the theoretical literature

In the five overview papers,[1,4,8–10] characteristics and success factors of collaboratives were formulated by the authors using expert opinion, either from experience or from theory (adult learning theory,[8] theories on organisational change,[10] diffusion of innovation theory[10] adult learning theory[10] and theory on effective teams[10]). Presence or not of any of these determinants of success could influence the outcomes of a collaborative.

The topic of the collaborative

The five overview papers describe the following expert opinion-based characteristics and success factors of collaboratives regarding 'the topic of the collaborative'. These determinants recommend ***how to select*** a successful topic to collaborative organisers. Differences in effectiveness between different collaboratives may be explained by these determinants.

Topic selection criteria

- 'Ripe' topics for improvement are selected using criterion: national tension for improvement exists in the topic area.[9]
- A focused clinical or administrative subject.[4]
- Professionals feel the proposed improvement is important, are motivated to achieve it.[4]
- Criterion for topic selection: leader with national reputation and expertise exists to chair the collaborative.[8]
- The subject is likely to be strategically important to organisations.[4]
- Criterion for topic selection: significant financial savings.[8]
- There is evidence of large variations in care (gaps).[4,9]
- Criterion for topic selection: sound scientific evidence suggests improved approaches compared with prevailing practices (ie a 'gap' exists between knowledge and prevailing practice).[8,4,9]
- Criterion for topic selection: high-level performance examples already exist (model for others).[8,4,9]
- 'Ripe' topics for improvement are selected using criterion: both organisations and patients would benefit from 'closing the gap'.[9]
- Participants can exchange ideas and suggestions which can be applied in different settings and stimulate ideas and motivation to change.[4]
- Sponsorship is an important determinant of the success of a collaborative as it can improve credibility or perceived importance. Evaluate the extent to which participants respect and agree with the priorities of the sponsor.[10]
- The choice of the topic is an important determinant of success of a collaborative, but opinion varied on the characteristics of an ideal topic (broad topic, simple, familiar topic, national priority?) The perceived salience and complexity of the topic should be important characteristics to be considered when comparing collaboratives.[10]

The ideas and support for improvement by experts

The overview papers provided the following expert opinion-based characteristics and success factors regarding 'the ideas and support for improvement by experts'. The determinants recommend ***what experts*** to select, and ***what knowledge/information*** to present to the participants using ***what method***, to collaborative organisers. Differences in effectiveness between different collaboratives may be explained by these determinants.

Expert characteristics

- To find ideas for improvement: having an expert panel to legitimise the knowledge was thought to be most likely to lead to adoption.[10]
- Planning group of approximately five individuals who have first-hand experience in leading substantive improvement in the topic area (the greater their personal experience in successful frontline improvement, the more effective they will be).[8] Input from established experts.[9] Credible experts need to be carefully chosen and briefed about their roles.[4]
- Collaboratives involve a sponsor, programme organisers, facilitators, subject experts, change experts and quality improvement experts.[4] Researchers familiar with the topic, along with practitioners who have achieved documented cutting-edge performance, are invited to join an expert planning group whose role is to clarify the nature of the gap and to consolidate the scientific and improve knowledge.[9]
- Composition: most collaboratives used outside experts (national or local) rather than participants themselves to suggest improvements; experts were occasionally too different from the participants for their advice to appear relevant; most interviewees included experts in quality improvement and systems thinking to identify knowledge to provide general concepts that allowed local changes for implementation and an expert with practical knowledge.[10]
- Experts, in particular the chair, should be opinion leaders, ie part of the system in which change is to occur, not too innovative, with a high level of technical competence.[10] A leader with national reputation and expertise to chair the collaborative.[8,9]
- Build the team's confidence through peers giving examples of changes which they can then translate.[4]

Expert information

Information on evidence and best practice

- Participants learn from experts about evidence, change concepts, practical changes.[4]
- The theory is expressed first in terms of collaborative aims: what outcomes can be achieved based on available knowledge.[8] The collaborative helps practitioners define optimal practice. Planning group identifies high-leverage changes (the ones with the greatest likelihood of improving performance).[8] Ideas that are most likely to be adopted have a relative advantage, are compatible with local practices and it is helpful if local reinvention is permitted, are of practical relevance to daily work.[10]
- The collaborative helps to find and describe best practices and diffuse them throughout collaborative organisations.[8] The challenge lies in understanding and disseminating the knowledge of 'how to' – application or improvement knowledge that helps assure that patients receive the best possible care. Find, describe and diffuse best practices throughout collaborative organisations.[9]
- The collaborative improves outcomes by understanding systems of care.[8,9]

- The collaborative entails the formation of new knowledge: works to codify the highest leverage ideas and changes in each topic area from disparate sources, where new knowledge also pertains to methods of real-time measurement.[8]
- Planning group develops a theory for improvement expressed in terms of aims, measures and key changes (knowledge is codified from the literature as well as data-supported experience).[8]

Information on data and data collection

- Organisers sometimes fail to recognise that teams often do not see how important data collection is, or fail to see the difficulties which teams experience in planning and collecting data.[4]
- Planning groups explore and suggest measurement systems specific to the topic.[9]
- Teams must be equipped to deal with data and change.[4] Each organisation collects and maintains its own data. Our emphasis is on building the internal capabilities of participants to collect, analyse and display data beyond the lifetime of the collaborative.[9]

Information on 'change' theory

- It is the synergy between scientific knowledge and improvement knowledge that enables improvement to occur at the patient level.[9]
- Collaboratives need to give sufficient time to developing the understanding of change theories and issues of individuals, as well as specific change skills. These include skills for breaking down problems, for undertaking project management, and for analysing and managing the politics of change. Both the knowledge and the skills need to be developed.[4] Collaboratives need to develop a team's belief in their ability to make change as well as their skills in planning and implementation.[4] Teaching participants about quality improvement theory and techniques was thought to be extremely important. Each used the model of Langley.[1] Develop healthcare expertise in the science of improvement within each topic area.[9] The collaborative develops national and local expertise in the science of change.[8] The collaborative builds the internal capacity of organisations to improve beyond the time frame of the collaborative.[8]
- Make time to learn about and plan how to sustain improvements.[4]
- Teach teams how to institutionalise the changes to survive individuals leaving, how to recognise when further changes are needed and how to make those changes.[4]

Information on team building

- Organisers can give guidance about team building (management, decision making, effective teamwork) as teams have to be formed and built/prepared.[4]

Expert methods

- Most collaboratives provided part-time central support for every three or four teams (ranged from one part-time person for the entire collaborative to a full-time project officer per team).[10] Central support varied between a very tight level of centralised control and a looser style.[10]
- A collaborative will not be successful if the central programme organisers and facilitators do not have a sufficient amount of time or the skills to plan and organise the work and to give support to the team.[4]
- The collaborative treats improvement as part of the work process; primary focus is on the clinical matter, not quality improvement methodologies.[8] Clinical issues and knowledge remain in the foreground; in the background, continuous improvement is used as the method through which clinical knowledge is employed.[8]

- Available knowledge guides the end points, processes and outcomes that should be monitored while pursuing the aim.[8]
- Collaborative organisers need to clarify at the outset the many different reasons that participants have for taking part and to agree what the primary objectives are going to be.[4]
- Collaborative organisers need to 'fine tune' the collaborative to meeting the needs of their participant customers and to state openly those which the collaborative cannot meet.[4]
- Organisers should give prospective participants a method to self-assess how much they might benefit from participating as well as guidance about the different benefits and about what would be required to benefit.[4]
- Build motivation (getting credible experts, having good evidence, showing how patients may be suffering unnecessarily is not enough) by stimulating a strong sense of purpose and mission.[4]
- The need for local adaptation means that the collaborative group must state its recommendations as concepts, rather than specific mechanisms. The collaborative group's recommendations should therefore seek only to capture the core concept and the critically necessary details.[1]
- The planning group elaborates in detail on concepts in lectures, conference calls, written materials and coaching sessions.[9]
- To continue to use the methods, teams will need to learn how to use them flexibly and be convinced of the value of doing so (deep learning).[4]
- Each such planning group establishes numerical goals that could be achieved if the best available science were used.[9]

The critical mass of multi-professional teams from multiple sites

From the overview papers we collected the following expert opinion-based characteristics and success factors regarding 'the critical mass of participating teams from multiple sites'. These determinants recommend how the collaborative organisers can compose a successful collaborative group and what requirements the participating organisations have to meet to enhance the chance of a successful collaborative. In addition, determinants of success regarding teams (composition and preparation) are described. Differences in any of these determinants may explain differences in effectiveness between and within collaboratives.

The collaborative group

- National call for participants from organisations serious about making demonstrable, major improvements at an unprecedented rate.[8,9] Participation of a number of multi-professional teams with a commitment to improve services and to share how they made improvements.[4]
- 20 to 40 organisations working together.[8,9] Critical mass of participating sites is needed to cultivate a useful exchange of ideas, experience and learning (IHI's experience suggests 25–40).[8] Multiple organisations must work together to define a consistent data set and associated collection methods.[1]
- There is no evidence that collaboratives which have selected teams have been more successful than those accepting all-comers.[4]
- Organisations without a background in basic quality management often have difficulty in both understanding the reasons for and in executing this analysis. Lack of relevant experience with quality management in some organisations may affect the pace of improvement in the collaborative as a whole.[1]

- Good and poor performers were encouraged to work together.[10] Both individuals and organisations must be ready for change. Selection is a way of assessing such readiness. Early adopters should be included in anticipation that diffusion will occur naturally from that group to another.[10] Diversity in practice is so important in a successful collaborative – we should have some data to verify it. The goal in forming the collaborative should be to have a range of performance on the indicators of interest so that we can explore both the factors that impede performance and those that improve it.[1] The improvement potential of a collaborative group will naturally be limited by the best practices within the group.[1]
- Collaborative groups that are formed on the basis of data from existing, well-defined indicators start with this challenge (note: the difficulties regarding data collection) behind them.[1]
- Existing professional boundaries inhibit improvement. Healthcare professionals, physicians in particular, have not been trained to, and often do not, work well across professional boundaries. Collaboration enables the synergism to solve difficult clinical problems.[9]
- Ideas are most likely be adopted from the same professional group. The BTS encourages peer-to-peer spread from innovators to early adopters.[10]

Organisational support of participating teams

- There is anecdotal evidence that teams who have been sent by their management and who are less motivated appear less likely to make improvements.[4] The majority of collaboratives relied on volunteers (versus mandatory).[10]
- Interested organisations must complete an application.[9] An application requires senior leadership support and demonstrates commitment to change.[8] Senior leaders must sign the collaborative application and are encouraged to participate actively in the work (frequent progress reviews).[8] Organisers could agree a contract with the organisation.[4]
- Interested organisations must be willing to contribute to the budgetary requirements of the collaborative.[8,9] In terms of funding strategies, in some collaboratives, participants covered all costs; in others, no fees were required and extra resources were provided. Both individuals and organisations must be ready for change. Requiring payment is a way of assessing such readiness. Early adopters should be included in the anticipation that diffusion would occur naturally from that group to another.[10]
- Time and resources necessary to achieve the aim.[8] Theory and observation suggest that the ability to make improvements depends mostly on the organisational context for the team, their time and their motivation. Teams and management need to recognise how many resources and how much work and supporting conditions are needed to make improvements and to achieve all the objectives.[4] Critical to success is the ability of an organisation to secure the time and resources necessary to achieve the aim.[9]
- Interested organisations must show evidence of senior executive commitment to the aim,[9] gain agreement and involvement of senior clinical and managerial leaders, and ensure that they understand what is required of them, of the team, and of the organisation, if the team is to make improvements. Quality improvement needs time and attention from local senior clinicians and managers as well as from project team managers and members.[4] Necessary support for the team is secured by discussing needs with the managers.[8] Without visible and real sponsorship and support from senior leaders, it is unlikely that any improvement will be significant or sustained.[4] Support from senior organisational leaders (chief executive officer of a hospital or health system, head of a department) was crucial.[10] Management and a supportive culture amplifies teams' capabilities.[4] Organisational ownership of data.[9]

- Teams which are clear about their objectives for taking part and have discussed this with others in their 'home' organisation appear to be more successful; teams and their management need separately and together to examine and agree what they want to achieve from taking part.[4]
- Alignment of the team's objectives with organisational priorities.[4]

Collaborative team

Composition

- Gather the right individuals on an improvement team, based on their knowledge of and involvement in the aim and the system of interest.[8]
- Teams attending the learning session were usually composed of three or four people (ranged from two to eight people)[10]
- Necessary components for team: system leadership (with authority to allocate time and resources), technical expertise (who knows the subject and understands the processes of care, if necessary added technical support by an expert on improvement methods), day-to-day leadership (critical driving component of the team: person who works in the process on a daily basis and who can work well with the champion).[8] Project teams typically include a team leader, at least one clinician, an opinion leader and a quality specialist. A collaborative will not be successful if the team leaders do not have the sufficient time or the skills to plan and organise the work and to give support to the team.[4] Individuals should be chosen based on their knowledge of the system and interest in the aim. Teams vary in size and composition from organisation to organisation, but each should have representation from three different domains: system leadership (someone with enough clout in the organisation to leverage resources and change), technical expertise (someone who knows the subject intimately, is trusted by other clinicians in the organisation and understands the processes of care), and day-to-day leadership (a person who works in the process on a daily basis and who will be persistent in assuring that the necessary work is completed). Additional technical expertise may be provided by an improvement expert who can help a team understand what needs to be measured, design simple, effective measures, and collect, display and interpret data. All teams should have at least one active physician champion. The greater this individual's influence with other clinicians, the more effective they will be in engaging them in the process of change.[9] Every collaborative had multidisciplinary teams including both clinical and non-clinical personnel that were to be directly affected by the changes; managers were less often included; clinicians were always included; inclusion of quality improvement officers and project managers varied.[10]
- Heterogeneous membership of improvement groups is important, as is involving physicians, while learning in teams that work together is more likely to be effective. Theory suggests that five to seven members are most likely to bring about change and that a senior leader is necessary to provide a mission for the project.[10]
- Team: at least one physician champion with a good working relationship with colleagues and with the day-to-day leader, interested in driving change.[8]
- Many teams had additional members who did not attend the collaborative but participated in the improvement process.[10]
- Continuity of role of the team leader: a change of team leader during a collaborative appears to be a strong predictor of whether the team will drop out.[4]

Preparation

- Teams have to make preparations to benefit more from meetings.[4] Having participants collect data, perform audit work or analyse the system they were in before the collaboration started was important in understanding their organisation, the nature of the problems and in developing baseline data for later comparison. Work before meetings are important to increase tension or motivation for change by re-emphasising the relative advantage.[10] Members of successful collaborative improvement groups commit first to understanding their own practices and processes or they will have little to share with others. Team members report that the breakthroughs in knowledge about the variability in practice within their own organisations were just as important as their learning about practices within other organisations.[1]

The model for improvement

The five overview papers produced the following expert opinion-based characteristics and success factors of collaboratives regarding 'the model for improvement'. The determinants recommend to collaborative organisers what improvement steps need to be taken by participating teams to successfully improve their processes of care or patient outcomes. Steps involve setting ***targets***, collecting ***data*** and ***testing changes***. Differences in effectiveness between and within collaboratives may be explained by these determinants.

The model

- The collaborative helps organisations make rapid, measurable, sustainable improvement.[8]
- Model for improvement (Langley) is the engine for change.[8]
- Action-oriented learning: the primary learning occurs after a change has been made in the context of a small-scale test, not from data collection and analysis before a change has been made.[8]
- Participants use a change testing method to plan, implement and evaluate many small changes in quick succession.[4]

Targets

- Clear aims.[9] Teams must have measurable and time-specified targets, including outcome targets and process or care activity targets.[4] Each organisation must establish specific organisational goals that are clearly stated, data-based, numerical.[8] Teams set measurable targets.[4] IHI works with each organisation to establish specific organisational goals that are clearly stated, data-based, numerical, and grounded in the knowledge developed by the chair and planning group.[9]
- Participants have to define their objectives and assess their capacity to benefit.[4]
- Teams must have measurable and achievable targets.[4] Teams need to be encouraged to set challenging targets, but also need to feel that targets are achievable.[4]
- Teams must define their targets early.[4]

Data

- The rigorous but parsimonious use of measurement is a key aspect that accelerates improvement.[8]
- Teams need to take stock of which data they have access to and consider which they could use in their quality improvement work before the first meeting.[4]
- Teams collect data to track performance, to measure progress towards targets.[4]

- Each team has to find a cost-effective way to collect relevant baseline data and then to organise the collection, analysis and reporting of the data to follow progress in reaching the target. Next, they have to decide which other data are needed.[4]
- Strong, early emphasis on establishing performance measurement systems within participating institutions.[9]
- Precise measurement systems.[9]

Testing of 'change'

- During the collaborative, organisations are taught to test changes on a small scale before fully implementing them.[8] The focus on rapid testing of small changes is probably a critical component.10 Rapid tests of change.[9]
- A focus on action and results.[9]

The structured activities for improvement, exchange and sharing

The overview papers described expert opinion-based characteristics and success factors of collaboratives regarding the collaboratives' 'structured activities for improvement, exchange and sharing'. The authors of the overview papers distinguish determinants related to the ***type of activity***, the ***methods*** and the ***data*** used to advance improvement, exchange of ideas and sharing of experiences of the participating teams. Differences between and within collaboratives in relation to these determinants may influence the chance of successful improvement of collaboratives and participating teams.

Type of activities

- Once an application has been accepted, IHI sends preparatory work and arranges an initial conference call.[9]
- Organisations working together for 9 to 12 months.[8] Each BTS collaborative consists of organisations working together for approximately 6 to 12 months on a specific topic.[9]
- Three two-day learning sessions.[8] Meeting three or more times for two to three days.[4] Most collaboratives had three meetings of two days each (range in number: two to seven meetings, range in length: half day to three days).[10] Each BTS collaborative utilises three two-day learning sessions.[9]
- Participants meet at least twice for one to three days to learn the methods, report their changes and results, share experiences, and consider how to spread their innovations to other services.[4] Participating organisations meet together at three two-day 'learning sessions' (LS) during this period, where they learn from the planning group, clarify the science, plan their own tests of change and learn from each other's efforts.[9] First LS: focus on developing an understanding of the critical changes necessary to achieve significant improvement in the topic area. Second LS: focus on assessing progress, sharing key findings within the topic area and planning how to spread progress beyond the initial test site(s). Third LS: lessons learned and ways to further disseminate knowledge are compiled.[8] Repeated exchanges among teams.[9]
- Some leaders added early sessions, eg to teach participants setting aims and quality improvement techniques.[10]
- During action periods: organisations and experts share knowledge and experience using regular visits, written reports.[8] Between meetings, participants continue to exchange ideas and collaborative organisers provide extra support, sometimes through visiting facilitators, email, and conference calls.[4] Between learning sessions, the collaborative experience is sustained.

Participants remain linked with each other through weekly conference calls (often involving experts on specific topics), email, 'listserv' discussion groups, and monthly exchange of printed reports. They visit other collaborative institutions, examining first-hand differences in systems of care – an extremely powerful and under-utilised learning and motivating experience.[9]

- Evaluations suggest that all teams should report their performance to the collaborative leaders on a regular (perhaps quarterly) basis. However, researchers took different views about the value of each team presenting their performance to the full meeting.[4]
- One-page reports are disseminated for group learning.[8]
- Self-assessments and planning group assessments are performed on a monthly basis to monitor participants and collaborative progress (IHI 5-point scale).[8] Some researchers also felt that self-assessment rating was less useful than actual performance level reporting.[4]
- At the end of each collaborative, the fundamentally new 'how to' knowledge (about achievable aims, practical real-time measurement and high-leverage changes) is placed within the public domain (national congress, public report): collaboratives build.[8] Each collaborative concludes with a public 'National Congress' where participants become faculty and national experts on improvement of the topic. The most useful scientific and improvement knowledge, measurement methodologies, lessons and results are also consolidated into a monograph, Guide to Improvement, and articles are prepared for publication in peer-reviewed journals. Disseminate and deploy knowledge gained during the collaboratives as broadly as possible to others in the healthcare community.[9]

Improvement methods

- Sharp focus on achieving the aim (learning session handouts with only critical information/one-page reports with only vital information).[8]
- Tension for change is a critical element of successful change.[8] Central support and reporting acted as feedback mechanisms for maintaining tension for change.[10]
- Constructive peer pressure to succeed.[8] Competition was important.[10]
- Shared deadlines help to maintain forward momentum even in the press of everyday events.[8]
- BTS collaboratives are limited in time (6 to 12 months) and focused on using joint learning to foster rapid, intra-organisational improvement. This creates tension for change.[9]

Exchange and sharing methods

- Collaborative fosters meaningful and effective collaboration between individuals and institutions.[8] Collaborative helped professionals to make links with colleagues in other organisations.[4]
- Plenary sessions should be kept to a minimum: most useful part of the meeting was the time available to spend with their own team and with colleagues from other organisations.[4]
- Give enough time for informal socialising so that participants can contact each other after meetings to follow up ideas. Collaborative organisers need to build in time and to engineer opportunities for informal contact as much 'know how' is transferred in this way.[4]
- Plan for a network to continue after the collaborative so that project team managers and members are still sharing ideas and experiences with others who participated in the collaborative.[4]
- Spread depends on effective contact and exchange between teams inside and outside the meetings.[4]
- Collaborative organisers should give guidance to teams about how formally to present their changes at the meetings, give structured opportunities for exchange, as well as make informal exchange easier.[4]

- Matching the educational needs is more successful than using a set curriculum.[10]
- Having participants share improvement strategies facilitated improvement. Ideas are most likely be adopted from the same professional group. The BTS encourages peer-to-peer spread from innovators to early adopters.[10]
- Meetings needed careful planning and clear agendas and objectives.[10]
- Collaboratives need to understand the different learning needs and styles to adapt the balance of the activities within the meetings to meet varying needs and preferences.[10]
- Social exchange (meals, parties, time out, mixing) within and between teams was very important. Some participants arranged extra meetings to continue social interaction. Social support and development of social networks are important for diffusion of ideas.[10]
- Three factors have to be in place before successful change can take place: curiosity, forgiveness (allowing for mistakes and encouraging teams to learn from them) and trust.[10]
- When clinicians from different organisations meet to discuss the details of their practices, they are confronted by the variation in practice and quickly realise that 'the right way of doing things' is not so clear as it once seemed.[1]
- One ground rule in a successful collaborative improvement effort is that all parties agree to full disclosure, within the group, of both outcome data and details of practice. Without such open sharing (because of fear of being put at a competitive disadvantage or fear of legal action), there is little chance of learning.[1] Collaboration accelerates the pace of improvement by exposing individuals and organisations to each other in an intellectually safe environment where sharing is the norm. This creates tremendous motivation for improvement.[9]
- An enabling social climate.[9]
- Give enough time to facilitating learning by practice and to allowing teams to discuss how to apply ideas in the team's home setting to help them to develop the ability to judge when quality methods are necessary and to adapt them flexibly to the setting and problem. Lectures alone do not develop this competence and other learning methods need to be used – for example, simulations and supervised practice applying the methods.[4] Large variations in the balance between plenary and workshop sessions.[10]

Data for improvement, exchange and sharing methods

- Research suggests that there is value in agreeing a common set of measures which all teams will track. This helps monitoring and evaluation, as well as enabling teams to learn from each other. It reduces the complexity which can come in a collaborative that tries to facilitate improvement on many different measures. To allow for differences among teams, teams can define a second set of measures which reflect their other objectives.[4]
- Reporting progress on targets keeps teams focused on the collaborative objective and on the need for measurement, and helps them to learn the importance of objective assessment.[4]
- Reporting progress on targets helps organisers to track progress and to decide which teams may need extra support.[4]
- Organisers often did not give enough time or examples and learning experiences to help develop a team's capabilities in the area of data collection.[4]
- All collaboratives used formal reporting; half of the interviewees used validated self-administered rating scales to assess progress towards improvement.[10]

- The collaborative group's main objective is to discover practices in the complex system of 'know how' that lead to superior performance. This requires some form of comparative benchmarking that goes beyond the simple exchange of indicator data; we must delve deeply into the details of practice (eg by on-site visits, literature reviews and expert panels, focused conferences by invitation).[1]
- It is wise for a collaborative group to track continuously throughout its lifespan the set of indicators it intends to impact. Collaborative groups formed around existing indicators in well-established benchmarking databases have an obvious advantage here.[1]

Determinants of success: results of the 'systematic review' studies

To verify empirically whether any of the theoretical determinants mentioned above have an impact on the effectiveness of the collaborative, we looked at the ten studies included in our systematic review of the effectiveness of QICs. We compared all collaboratives to detect similarities between the collaboratives and their effectiveness. Similarly, we compared teams within a specific collaborative to detect similarities between the various teams participating within the collaborative and their effectiveness. The first two sets of theoretical determinants ('the topic of the collaborative' and 'the ideas and support for improvement by experts') can be used to compare the various collaboratives. The last three sets ('the critical mass of multi-professional teams from multiple sites', 'the model for improvement' and 'the structured activities for improvement, exchange and sharing') can be used to compare both the various collaboratives and the various teams within a specific collaborative.

Comparison of the various collaboratives

The papers on the ten controlled studies enabled us to make four comparisons. Collaboratives and their effectiveness could be compared in relation to type of collaborative, outcome, topic and type of setting.

Type of collaborative

The studies included described three different kind of collaboratives: collaboratives based on the IHI Breakthrough Series (BTS), collaboratives based on a combination of the Breakthrough Series and the Chronic Care Model (CCM) and collaboratives based on the Vermont Oxford Network (VON). The QICs' overall effectiveness did not seem to vary between these different types of collaboratives; all three groups of collaboratives showed similar results (Table 3).

BTS

Three[78,89,90] of four studies based on the BTS[78,80,89,90] showed significant improvements in the effect parameters chosen.

Pierce-Bulger et al[90] showed a significant improvement in the number of days between neonatal deaths during a seven-year quality programme (1993–1999) including a BTS project in 1993.

In **Baier et al's study**,[78] the prevalence of residents with pain in 15 nursing homes diminished significantly (7.2 versus 11.2% of patients) after participation in a BTS.

Landon et al's study[80] of 9,986 patients with HIV infection did not show any significant effect on eight process of care parameters, such as screening, prophylaxis and access to care.

Howard et al[89] presented a significant change of 8 per cent (95% CI: 2–13%, $p < 0.001$) in the proportion of eligible donors who became actual donors during the first phase of the collaborative. The conversion rate among control hospitals, however, increased steeply during the second phase. Nationwide, there were large increases in the total number of organ donors during the first phase.

BTS and CCM

Three[79,82–85] of four studies combining the BTS with the CCM[79,81–85] showed significant improvement in some of the selected process of care and patient outcome parameters.

The Homer study[81] did not show effects on any of the three key processes of care parameters and five patient outcomes for children with asthma.

Benedetti et al[79] compared participating providers to non-participating providers and reported significant improvement in seven out of twelve effect parameters (note: only six parameters are specified in the paper). For example, patients from participants had higher rates of annually diabetes eye and foot exams, and better outcomes for hemoglobin A1c and blood pressure.

Mangione-Smith et al (twenty-three effect parameters regarding children with asthma)[82] and **Schonlau et al** (twenty effect parameters regarding adults with asthma)[83] showed significant improvements for several medical records process of care parameters (six out of eight process parameters and four out of eight respectively). Regarding patient self-management, the authors reported significant differences in two out of three and one out of four process parameters respectively. Mangione-Smith et al[82] also reported higher scores on two out of three quality-of-life patient outcome parameters for the intervention group in a post-measurement survey. However, the asthma severity levels between the intervention group and the control group differed. Schonlau et al[83] showed significant improvement in satisfaction with communication (overall score: 62 per cent versus 39 per cent, one patient outcome).

Asch et al[84] reported significant improvement in performance regarding nine out of twenty-one process of care parameters for the care of patients with chronic heart failure (seven out of eight counselling parameters and two out of four medication parameters). The counselling parameters showed improvement mainly varying between 20 and 40 per cent; they had, however, initially low performance rates, and the rates remained – for most processes – below 50 per cent after intervention. Significant improvement in the two parameters for the appropriate use of medication – where baseline rates were quite good – was less dramatic: appropriate angiotensin-converting enzyme inhibitors (93 per cent in the intervention group versus 87 per cent in the control group post-intervention) and appropriate lipid-lowering therapy (66 per cent in the intervention group versus 64 per cent in the control group post-intervention). The changes in the rates of counselling were confirmed in a cross-sectional survey among patients by Baker et al.[85]

VON

Both studies based on the VON[85–87] showed significant improvement in some of the selected process of care and patient outcome parameters.

Horbar et al (2004)[88] demonstrated significant improvement in two specific processes of care, but no significant improvement in patient outcomes (mortality and pneumothorax), nor in twenty of the twenty-three secondary effect parameters.

Horbar et al (2001)[86] reported a significant decrease in the rate of infection at six intervention NICUs (12.3 per cent versus 16.5 per cent) and a decrease in the rate of supplemental oxygen at four NICUs (34 per cent versus 38.7 per cent). Although the effect among NICUs on both arms was heterogeneous, the changes showed that the intervention NICUs improved at a significantly faster rate in a four-year period than the sixty-six comparison NICUs. In the same study, Rogowski et al (2001)[87] presented cost data showing that costs may be reduced as a result of participation in the QIC, although the cost savings across hospitals were heterogeneous.

Type of effect parameter

The included studies selected both 'process of care parameters' and 'patient outcomes' as effect parameters. Process of care parameters more often showed positive results than patient outcomes. This might be because there was too small a sample size (power problem) or because of an inadequate duration of the measurement period. Given the chosen timeframe of QICs, process of care parameters are probably more susceptible to change then patient outcomes.

All studies

Very large numbers of effect parameters were used in the ten studies, both describing processes of care and patient outcomes of care. The authors of the studies tested 88 patient outcomes of which 21 (24 per cent) were significantly higher in the collaborative groups following the intervention. In addition, they tested 101 processes of care parameters, of which 43 (43 per cent) showed significant improvement in the collaborative teams after intervention.

Type of topic

Collaboratives seem to be particularly effective in improving patient education/counselling and monitoring of patients; all three studies that addressed and measured these topics showed significant results in these domains.[79,82–85] This conclusion is based on studies in which the BTS was combined with the CCM.

Potential explanations lie, in the opinion of Asch et al,[84] in the flexibility in focus between the participating groups, and the emphasis on educational patient activation interventions in the BTS. Similarly, when explaining the effects on education and self-management, Mangione-Smith et al[82] and Schonlau et al[83] emphasised that the faculty stressed self-management strategies during the collaborative learning sessions because of their previously established relationship to improved health outcomes in asthmatics. It therefore remains unclear whether collaboratives are especially effective in improving patient education/counselling or whether change is more often shown in relation to this topic because the collaborative puts more focus on it.

All studies

Regarding the topics addressed (and measured) in the studies, it must again be concluded that the studies addressed a wide range of different topics. They addressed infant mortality in the community,[90] pain in nursing homes,[78] diabetes in primary care,[79] HIV care in community health centres, hospital-based clinics or HIV specialty clinics,[80] children[81,82] and adults[83] with asthma in primary care practices, chronic heart failure in different settings/clinics,[84,85] care for preterm infants born or admitted to NICUs86–88 and organ donation in hospitals.[89]

This diversity of topics addressed made it difficult to draw conclusions in relation to the question of whether collaboratives are especially effective for specific topics. However, when looking at the effect parameters that showed significant improvements, it stood out that in two studies[82–85] improvements were merely seen in patient education/counselling and monitoring of patients. We looked at the rest of the studies that addressed similar topics with this hypothesis in mind, and it was confirmed in the Benedetti[79] study of diabetes. In this study, annual eye and foot exams ('monitoring of patients' topics) improved, while annual urine protein tests approached significance. Annual glycohemoglobin (A1C) and LDL tests did not improve compared to the control group; baseline percentages were, however, already high (> 70%) for these outcome measures. All three studies combined the collaborative approach with the CCM.

Type of setting

Given the huge diversity in the type of setting in which the collaboratives were performed, it was impossible to draw conclusions regarding the question whether collaboratives are especially effective in specific settings.

All studies

As described above, in the selected studies, change was seen to take place in various settings. The studies addressed infant mortality in the community,[90] pain in nursing homes,[78] diabetes in primary care,[79] HIV care in community health centres, hospital-based clinics or HIV specialty clinics,[80] children[81,82] and adults[83] with asthma in primary care practices, chronic heart failure in different settings/ clinics,[84,85] care for preterm infants born or admitted to NICUs[86–88] and organ donation in hospitals.[89] This diversity of settings made it difficult to draw conclusions on the possible susceptibility for change of a specific type of setting.

Comparison of the various teams within a specific collaborative

The papers on the ten controlled studies provided little information on possible determinants and their impact on the effectiveness of the various collaborative teams. Although authors noted that 'participation in the collaborative programme may have varied among the participating sites' or 'the focus may have varied between the participating groups' or 'it remains unclear which intervention component proved successful',[78,80,81,84–86,88,89] only a few of them linked variation in participation, focus or use of programme components to variation in team effectiveness. Variation in team effectiveness, for example, by providing information on confidence intervals, standard deviations of organisations or teams etc, was not presented by the authors.

When looking for combinations of team characteristics and team effectiveness, the following factors were positively related to the degree of effect: having a dedicated nurse who provides the desired care[85] and length of time using the CCM model.[79]

Authors also tested whether the degree of participation,[80,81] baseline performance,[80] already or newly receiving funding for providing the intended care[80] and mandated or voluntary participation[80] were related to effectiveness. No or unclear differences between intervention and control group data were found regarding these variables.

Homer et al (2005)[81] evaluated the engagement level of the participating practices, showing that attendance at the 3 learning sessions declined progressively from the first to the third (for example, 34 participants at the first session in 1 of the 2 states, 24 at the third). On average, only 42 per cent of the practices submitted performance data and 39 per cent reported their progress in any given month of the intervention, with fewer practices reporting in the later months of the intervention. To evaluate the possibility that the lack of an intervention effect could be due to limited engagement by some of the practices, they reran their analyses including data only from practices that attended all three learning sessions (a minimal indicator of engagement). It was shown that the percentage of children requiring an emergency department visit was reduced more in the children in the intervention group than in the control group. The proportion of intervention group requiring an emergency department visit dropped from 36 per cent to 22 per cent for the entire intervention group, but dropped from 51 per cent to 22 per cent when retaining only the children in practices attending all the learning sessions. This restriction removed all but 9 of 22 intervention practices.

Homer et al[80] described the degree of participation as being related to effectiveness, that is, those nine intervention practices that attended all three learning sessions showed greater improvements (Δ before–after = 29%) as compared to all intervention practices (Δ before–after = 19%) or the control practices (Δ before–after = 14%). However, the authors overlooked the possibility that by only including engaged participants they may have introduced selection bias. The data seem to indicate such a bias: in the control group, the percentage of children with asthma emergency department visits dropped from 36 per cent to 22 per cent. In the collaborative group these percentages dropped from 36 per cent to 17 per cent when all participants were included. When only engaged participants were included, the percentage dropped from 51 per cent to 22 per cent. This could indicate that especially low scoring practices were engaged to change.

Landon et al (2004)[80] evaluated participation in the collaborative and the types of interventions attempted. All clinics attended at least three of the four learning sessions, and 75 per cent submitted monthly reports seven or more times (mean 10.6). All clinics also participated in conference calls and the collaborative 'listserv', although they had no way of quantifying the level of participation in these activities, and described the number of initiatives tried in specific clinical areas. In addition, they explored whether effectiveness was related to specific organisational characteristics: they showed that the effectiveness of the intervention did not differ significantly for clinics that were newly awarded funding under Title III of the CARE Act versus ongoing Title III clinics, nor was the intervention more effective for clinics that were poor performers at baseline. Finally, in their discussion, they described that a subgroup analysis did not show any differences between mandated and voluntary clinics.

Although Landon et al measured per site whether they had at least one intervention that was being refined or implemented per specific clinical area, they unfortunately did not relate this 'focus or not' to the outcomes that were measured in the same clinical area.

Benedetti et al (2004)[79] showed that length of time of using the CCM model affected outcomes; outcomes were more favourable for the highest participation level (level I = three years' use of the model, two providers; level II = two years' use of the model, three providers; level III = one year's use of the model, six providers).

Baker et al (2005)[85] provided additional objective information to explain the greater effectiveness regarding the lifestyle modifications and monitoring outcomes. As the CCM model promotes having a dedicated heart failure (HF) nurse, they analysed whether differences between the participant and control groups were explained by participant patients being more likely to have a HF nurse. Having a HF nurse and more frequent contact with the HF nurse was associated with higher levels of education, knowledge and self-management behaviours. They showed that about 10 per cent of the higher education on lifestyle modification and monitoring was explained by higher rates of contact with the HF nurse.

Howard et al (2007)[89] described how the impact of the collaborative varied at the hospital level. Of the collaborative hospitals, 67 per cent experienced an increase in conversion rates (in contrast with 44 per cent of the control hospitals) and 33 per cent a decrease (in contrast with 56 per cent of controls) From 'informal conversations' they concluded that 'better focusing on the mission of organ procurement organisations' and 'providing a structural mission' were success factors.

Determinants of success: summary and conclusions

To understand better the effectiveness of collaboratives, we started by looking at the theoretical literature to identify the elements of an effective collaborative. From five overview papers on collaboratives, we extracted expert opinion-based characteristics and success factors of collaboratives and categorised them using our definition of a collaborative. Determinants (that is, characteristics and success factors) were described in relation to: (1) the selection of a collaborative's topic; (2) the selected experts, the information they present and the methods they use to present this information; (3) the composition of the collaborative group, the requirements set for the participating organisations, and the composition and preparation of the collaborative teams; (4) the improvement steps – setting targets, collecting data and testing changes – taken; and (5) the type of activity, and the improvement and exchange methods chosen.

This activity resulted in an overview of expert opinion-based determinants that can be used as a checklist by organisers of a collaborative to ensure that all aspects of a 'good' collaborative have been included. This overview of determinants can also be used by evaluators of collaboratives who are interested in explaining success or lack of success after the performance of the collaborative. They can compare different collaboratives or different teams within a collaborative by simultaneously looking at their effectiveness and the presence – or not – of specific determinants. The first two sets of theoretical determinants ('topic' and 'experts') can be used to compare various collaboratives. The last three sets ('collaborative group/teams', 'improvement steps' and 'activities and methods') can be used to compare both various collaboratives and various teams within a specific collaborative.

Next, we looked at the studies included in our review of the effectiveness of QICs, to verify empirically whether any of these determinants have indeed had an impact on the effectiveness of the collaborative. Collaboratives and their effectiveness could be compared in relation to type of collaborative, outcome, topic and type of setting. This comparison of collaboratives showed first of all that the overall effectiveness of QICs did not vary between the three types of collaboratives (those based on the BTS, those based on the BTS combined with the CCM and those based on the VON approach).

The ten studies used a huge number of effect parameters, all describing 'processes of care' and 'patient outcomes of care'. Data showed that overall effectiveness differed between the various types of effect parameters chosen: 'process of care' parameters more often showed positive results than 'patient outcomes', probably because of the chosen timeframe.

The huge variation in the topics and settings addressed in the different collaboratives made it impossible to draw any conclusions regarding the question of whether collaboratives are especially effective for specific topics or in specific settings. Collaboratives seem to be particularly effective in improving patient education and the monitoring of patients. It is, however, unclear whether patient education and monitoring is a topic that is more susceptible to change or whether change is more often shown in relation to this topic because the collaboratives put more focus on it.

Our comparisons of teams within a specific collaborative showed that 'having a dedicated nurse who provides the desired care' was related to greater effectiveness regarding lifestyle modifications and monitoring parameters. It was also shown that 'length of time professionals used the CCM model' affected outcomes; outcomes were more favourable, the longer the CCM was used (three-, two- or one-year use of the CCM model, all including 12 months of BTS). Several other variables (degree of participation, baseline performance, already or newly receiving funding for providing intended care and mandated or voluntary participation) were explored, but their influence remained unclear.

In sum, the exploration of the theoretical literature resulted in a long list of probable determinants of success as mentioned by experts. The studies currently included in our systematic review provided little empirical evidence for these determinants of success, as only very few determinants could be tested. Our comparison of collaboratives, however, showed that collaboratives may be particularly effective in improving patient education and monitoring of patients. In addition, the comparison of teams showed that 'having a dedicated nurse who provides the desired care' was related to the degree of effect. Taking into account the many determinants regarding teams (that is, team building, organisational support of teams and team composition and preparation) and the theoretical concept that collaboratives are about team processes – and not about individual behaviour and people not performing well – leads to the hypothesis that collaboratives might be particularly effective for those specified changes or improvements where a team approach is vital. Collaboratives might work for those changes that address the organisational context within which the desired care is provided, where a team operates and where the team appoints persons responsible for the different tasks, maybe especially where tasks are delegated to the nurse. This hypothesis also predicts that if collaboratives want to be effective in, for example, prescribing and diagnosing (that is, changes addressing individual behaviour), incentives have to be build into the organisation, and into the team process, to ensure that the individual is not solely responsible for these improvements.

As already concluded, the controlled studies included in our systematic review of the effectiveness of QICs provided little information on the theoretical determinants and their impact on the effectiveness of the various collaboratives and the various collaborative teams. More research is therefore needed to study these probable determinants of success or failure and to improve our understanding of how and why QICs work. Collaboratives are performed on many different topics, in many different settings, using many different effect parameters, making it difficult to compare them. Evaluations might therefore specifically focus on comparing teams within a collaborative, as at least the collaborative's topic, the selected experts, the knowledge/information they present, the manner in which they present this and the composition of the collaborative group do not vary between the different participating teams within the same collaborative. The studies needed to answer this 'determinants of success or failure' question balance between rigorously controlled designs and uncontrolled process-oriented reports. Each collaborative therefore has to include an effect evaluation using valid effect parameters and reliable data that can also be used to compare outcomes of the collaborative participants with those of their colleagues in non-participating sites. In addition, a process evaluation has to be carried out to gather information from each participating team regarding presence or not of the collaborative's characteristics and success factors. By integrating these process and effect data, insight is gained into why some QIC teams are successful while others fail to change practice. Outcomes from this research can be used to revise the collaborative strategy to improve its effectiveness or cost-effectiveness.

Table 1: Overview of methods and types of collaborative in controlled studies

Study		Method					Intervention	
Study, reference	**Year**	**Topic, disease or condition**	**Study design (control condition)**	**Unit of analysis Sample size project Sample size study**	**Method and number of patients**	**Length of measurement in months**	**Type of collaborative**	**Length of intervention period (months)**
Pierce-Bulger et al[89]	2001	Infant mortality in community	ITS (No)	Medical centre (?) 1 (I)	Chart review (?)	48 (B) 48 (A)	BTS embedded in longitudinal quality activities 1993 BTS 1995–2000 Clinic and home-visiting services	?
Baier et al[77]	2004	Pain	CBA (No)	Nursing homes (21) 15 (I) 72 (C)	Minimal dataset www.cms.hhs.gov/medicaid/mds20 (I) 276	12 (B) 15 (A)	Based on BTS	15
Benedetti et al[78]	2004	Diabetes primary care	CBA (No)	Providers (?) 11 (I) 19 (C)	Method unclear (I) ? (C) ?	Unclear	BTS CCM	12
Landon et al[79]	2004	HIV	CBA (No)	Clinics (62) 44 (I) 25 (C)	Medical records review (I) 3,190 (B) 3,216 (A) (C) 1,761 (B) 1,819 (A)	12 (B) 18 (A)	BTS extended by 4 months (1 extra learning session)	12 + 4

Table 1: Overview of methods and types of collaborative in controlled studies (continued)

Homer et al[80]	2005	Asthma (children)	RCT (No)	Primary care practice (22) 22 (I) 21 (C)	Telephone interview (I) 294 (B) 236 (A) (C) 337 (B) 254 (A)	12 (B) 12 (A)	BTS CCM	12
Mangione et al[81] Schonlau et al[82]	2005	Asthma (children)	CBA (No)	Primary care practice (26) 9 (I) 4 (C)	Medical records review (I) 348 (B+A) (C) 153 (B+A) Telephone survey (A) (I) 385 (C)126	12 (B) 12 (A)	BTS CCM	12
		Asthma (adults)	CBA (No)	Primary care practice (26) 6 (I) 3 (C)	Medical records review (I) 109 (B+A) (C) 76 (B+A) Telephone survey (A) (I) 123 (C) 62	12 (B) 12 (A)	BTS CCM	12
Asch et al[83] Baker et al[84]	2005	Chronic heart failure	CBA (No)	Clinics (14) 4 (I) 4 (C)	Medical records review (I) 261 (B+A) (C) 228 (B+A) Survey (A) 301 (I+C)	11 (B) 12 (A)	BTS CCM	12
		Heart failure	Controlled cross-sectional study (No)	Clinics (13) 6 (I) 6 (C)	Telephone survey (A) 367 (I) 414 (C)	?	BTS CCM	12

Table 1: Overview of methods and types of collaborative in controlled studies (continued)

Horbar et al[85] Rogowski et al[86]	2001	Preterm infants	CBA (feedback in routine reports)	NICUs (10) 10 (I) (4 infection / 6 chronic lung disease) 66 (C)	Database 3,801 (I) 21,509 (C)	12 (B) 36 (A) (-)	VON	36
		Preterm infants	CBA (feedback in routine reports)	NICUs (10) 10 (I) (4 infection / 6 chronic lung disease) 9 (C)	Database infection 2,993 (I) 2,203 (C) Database chronic lung 663 (I) 1,007 (C)	12 (B) 36 (A) (-)	VON	36
Horbar et al[87]	2004	Preterm infants	RCT (feedback in routine reports)	NICUs (57) 57(I) 57 (C)	Database (I) 3,332 (B) 3,313 (A) (C) 2,850 (B) 2,726 (A)	12 (B) 12 (A)	VON	?
Howard et al[88]	2007	Organ donation	CBA	Hospitals (95) 95 (I) 125 (C)	Administrative data from Organ Procurement and Transplantation Network 'eligible donors per month': (I) 195 (B) 196 (A) (C) 151 (B) 156 (A)	12 (B) 6 (A, last 6 months of the collaborative)	BTS	First phase: 12

Abbreviations: (A), after; (B), before; BTS, Breakthrough Series; (C), control group; CBA, controlled before–after study; CCM, Chronic Care Model; HIV, human immunodeficiency virus; (I), intervention group; ITS, interrupted time series; NICUs, neonatal intensive care units; QIC, quality improvement collaborative; RCT, randomised controlled trial; VON, Vermont Oxford Network

Table 2: Overview of effectiveness and methodological quality of controlled studies

Study, reference	Process of care parameters	Patient outcomes	Methodological quality	
Pierce-Bulger et al[89]		**Infant mortality (annual days between deaths increased from a pre-programme average of 55 days (1989–1994) to an average of 114 days (1995–2000) since (!) clinic and home-visiting services began)#**	Protection from secular changes	Not done
			Data analysed appropriately	Not specified
			Reason for points pre and post	Not specified
			Shape of intervention effect	Not specified
			Protection detection bias	Not specified
			Blinded assessment	Not specified
			Completeness of dataset	Not specified
Baier et al[77]		**Prevalence (residents with pain, 7.2 versus 11.2 patients)§**	Baseline measurement	No differences
			Characteristics of control sites	Differences
			Blinded assessment	Not specified
			Contamination	Unlikely
			Follow-up sites/teams	I 71% / C 92%
Benedetti et al[78]	HbA1C test annually	Hb_{A1C} < 8	Baseline measurement	Not reported
	LDL test annually	**Hb_{A1C} < 9.5 *** (?)	Characteristics of control sites	Seem similar, no data
	Urine protein test annually	**LDL test < 130 (75% versus 45% patients; p < .05)**	Blinded assessment	Not specified
	Eye exam annually (?)	**Blood pressure < 130/85 (49% versus 35% patients; p < .05)**	Contamination	Possible
	Foot exam annually (?)		Follow-up sites/teams	Not specified
	Patients > 40 years taking acetylsalicylic acid (?)	Blood pressure < 140/90		
	Self-management goal			
Landon et al[79]	Antiretroviral therapy (2 items)		Baseline measurement	No differences
	Screening and prophylaxis (5 items)		Characteristics of control sites	Similar
	Access to care (1 item)		Blinded assessment	Not specified
			Contamination	Unlikely
			Follow-up sites/teams	Not specified

Table 2: Overview of effectiveness and methodological quality of controlled studies (continued)

Homer et al[80]	Written asthma management plan Daily use of inhaled steroids Daily use of controller medications	Asthma attack Parent report of limited activities Parent experience of care Parent-reported functional status Asthma hospitalisation / ED use	Unit of allocation, randomisation Follow up of sites Blinded assessment Baseline measurement Contamination	By practice, no process 100% Not specified No differences Possible
Mangione et al[81] Schonlau et al[82]	**Survey** Patient self-management **Peak flow monitoring (70% versus 43% patients; p < .0001)** **Written action plan (41% versus 22% patients; p = .001** Goal setting with provider Asthma knowledge Use of long-term controller medications **Medical records** **Peak expiratory flow rate measured annually (49% versus 4% patients; p < .0001)** **Written action plan (42% versus 3% patients; p < .0001)** **≥ 2 follow-up visits annually (86% versus 78% patients; p < .004)** **Educated in self-management (41% versus 17% patients; p < .0001)** **Instructed in use of metered dose inhalers (30% versus 9% patients; p < .002)** **Collaborative goal setting between patient and provider (10% versus 0% patients; p < .003)** Beta-2 antagonist prescribed for symptomatic relief If medication changed: follow-up visit within 6 weeks	**Survey** Quality of life **General quality of life (80.2 versus 77.0; p = .05)** **Asthma-specific quality of life: treatment problems (88.6 versus 85.3; p < .05)** Asthma-specific quality of life: symptoms Impact on family functioning Parent satisfaction with provider satisfaction Adolescent satisfaction with care Acute care service use Missed schools days Parent lost work days	Baseline measurement Characteristics of control sites Blinded assessment Contamination Follow-up sites/teams	May substantially differ Limited data Not specified Possible 100%

Table 2: Overview of effectiveness and methodological quality of controlled studies (continued)

	Survey	**Survey**	Baseline measurement		May be different
	Self–management	Quality of life		Characteristics of control sites	Limited data
	Attended educational session (20% versus 5%; p = .03)	General quality of life		Blinded assessment	Done
	Peak flow monitoring	Asthma-specific quality of life		Contamination	Unlikely
	Goal setting	**Satisfaction with clinician/lay educator communication (overall score 62% versus 39%; p = .02)**		Follow-up sites/teams	100%
	Written action plan	Use of acute care			
	Asthma knowledge	Bed days resulting from asthma-related illness			
	Asthma-control medication				
	Medical records				
	Peak expiratory flow rate measured annually (28% versus 14%; p < .03)				
	Written action plan (27% versus 0%; p < .0001)				
	Instructed in use of metered dose inhalers (22% versus 7%; p < .04)				
	Collaborative goal setting between patient and provider (7% versus 0%; p < .03)				
	β-2 antagonist prescribed for symptomatic relief				
	No – blocker prescribed				
	≥ 2 follow-up visits annually				
	Educated in self-management				

Table 2: Overview of effectiveness and methodological quality of controlled studies (continued)

Asch et al[83] Baker et al[84]	Diagnoses LVEF ever measured Cr measured if on digoxin BP measured LDL measured if CAD Medication **ACEI for LVEF ≤40% (93% versus 87%; p < .0001)** **Lipid-lowering therapy for CAD (66% versus 64%; p < .0002)** Beta blockade for LVEF ≤40% Anticoagulation for atrial fibrillation Follow-up Electrolyte monitoring during ACE Rx Electrolyte monitoring during diuretic Rx Electrolyte monitoring on ACE initiation Electrolyte monitoring on diuretic initiation Visit within 4 weeks after discharge Counselling on* **Medication (44% versus 17%; p < .0001)** **Diet (46% versus 11%; p < .0001)** **Exercise (42% versus 12%; p < .0001)** **Weight loss (42% versus 7%; p < .0001)** **Disease management (61% versus 23%; p< .0001)** **Water weight management (42% versus 4%; p < .0001)** **Goal setting (5% versus 4%; p < .0001)** Smoking	BP < 130/80 mm Hg post-MI or LVEF < 4 BP < 140/90 mm Hg post-MI or LVEF > 4 INR 20.0 – 30.0 in atrial fibrillation LDL < 100 if CAD	Baseline measurement Characteristics of control sites Blinded assessment Contamination Follow-up sites/teams	22 of 25 measures similar Limited data Done Unlikely 100%

Table 2: Overview of effectiveness and methodological quality of controlled studies (continued)

Horbar et al[85] Rogowski et al[86]	Chronic lung disease group **Supplemental oxygen rate (34% versus 38.7%; p = .049** (although mixed effects: n = 2 reduction, n = 2 increase)	Nosocomial infection group **Coagulase-negative staphylococcal infection rate (12.3 % versus 16.5%; p = .001)** although mixed effects: n = 4 reduction, n = 2 increase Other bacterial pathogens infection rate Chronic lung disease group Death	Baseline measurement Characteristics of control sites Blinded assessment Contamination Follow-up sites/teams	May be different Limited data Not specified Unclear 100%
		Nosocomial infection group **Median treatment cost per infant ($57,606 versus $45,874; p = .0037)** Chronic lung disease group Median treatment cost per infant ($85,959–$75,084) Treatment cost in control hospitals rose	Baseline measurement Characteristics of control sites Blinded assessment Contamination Follow-up sites/teams	May be different Limited data Not specified Unlikely 100%

Table 2: Overview of effectiveness and methodological quality of controlled studies (continued)

Horbar et al[87]	Primary study outcomes **Surfactant given in delivery room (54.7 versus 18.2 %)** **First dose surfactant given after 2 hours (9.4 versus 24.9%)** Secondary study outcomes Delivery room resuscitation **Endotracheal intubation (78% versus 69.8%)** Oxygen Bag or mask Adrenaline Cardiac compression Cardiopulmonary resuscitation Apgar score ≤ 3 at 1 minute Respiratory **Surfactant given at any time (85.1 versus 79.5%)** Chronic lung disease at 36 weeks Respiratory distress syndrome Conventional ventilation High-frequency ventilation Corticosteroids given for chronic lung disease Infection Late bacterial Coagulase negative staphylococci Fungal Nosocomial Neurological **Severe intraventricular hemorrhage (10.1 versus 14.2%)** Cystic periventricular leucomalacia Intraventricular hemorrhage Other Patent ductus arteriosus Retinopathy of prematurity Severe retinopathy of prematurity	Mortality (death before discharge) Pneumothorax	Unit of allocation, randomisation Blinded assessment Baseline measurement Contamination Follow-up of sites	By hospital, computer Not specified No differences Unlikely 100%

Table 2: Overview of effectiveness and methodological quality of controlled studies (continued)

Howard et al[88]		**Conversion rate (proportion of eligible donors who became actual donors) 60% vs 51%, p < 0.001**	Baseline measurement	No differences
			Characteristics of control sites	Differences
			Blinded assessment	Not specified
			Contamination	Unclear
			Follow-up sites/teams	100%

*The changes in counselling in Asch et al's study[83] were confirmed in a cross-sectional survey in Baker et al's study[84]

#Bolded text = statistically significant effect

§Versus = post-measurement intervention group versus control group

Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; BP, blood pressure; C, control group; CAD, coronary artery disease; Cr, creatinine; I, intervention group; INR, international normalised ratio; k, kappa; MI, myocardial infarction; HbA1C, hemoglobinA1C; LDL, low-density lipoprotein; LVEF, left-ventricular ejection fraction

Table 3: Summary of effectiveness in controlled studies

Controlled studies (n = 10)	Studies based on BTS	Studies combining BTS and CCM	Studies based on VON
2 RCTs		Homer et al[80] (-)	Horbar et al[87] (±)
7 CBAs	Landon et al[79] (-) Baier et al[77] (+) Howard et al[88] (+?)	Benedetti et al[78] (±) Mangione-Smith[81] / Schonlau et al[82] (±) Asch[83] / Baker et al[84] (±)	Horbar[85] / Rogowski et al[86] (±)
1 ITS	Pierce-Bulger et al[89] (+?)		

- no effect

± mixed effect

+ positive effect

+? some concerns whether the effect can be attributed solely to the collaborative activities

Abbreviations: BTS, Breakthrough Series; CBA, controlled before/after; CCM, Chronic Care Model; ITS, interrupted time series; RCT, randomised controlled trial; VON, Vermont Oxford Network

References

1. Plsek PE (1997). 'Collaborating across organizational boundaries to improve the quality of care'. *American Journal of Infection Control*, vol 25(2), pp 85–95.

2. Solberg LI (2005). 'If you've seen one quality improvement collaborative'. *Annals of Family Medicine*, vol 3(3), pp 198–199.

3. Ayers LR, Beyea SC, Godfrey MM et al (2005). 'Quality improvement learning collaboratives'. *Quality Management in Health Care*, vol 14(4), pp 234–247.

4. Øvretveit J, Bate P, Cleary P et al (2002). 'Quality collaboratives: lessons from research'. *Quality & Safety in Health Care*, vol 11(4), pp 345–351.

5. Mittman BS (2004). 'Improving patient care. Creating the evidence base for quality improvement collaboratives'. *Annals of Internal Medicine*, vol 140(11), pp 897–901.

6. Newton PJ, Davidson PM, Halcomb EJ et al (2006). 'An introduction to the collaborative methodology and its potential use for the management of heart failure'. *The Journal of Cardiovascular Nursing*, vol 21(3), pp 161–168.

7. Schouten LMT, Hulscher MEJL, van Everdingen JJE et al (2008). 'Evidence for the impact of quality improvement collaboratives: systematic review'. *British Medical Journal*, vol 336, pp 1491–1494.

8. Kilo CM (1998). 'A framework for collaborative improvement: lessons from the Institute for Healthcare Improvement's Breakthrough Series'. *Quality Management in Health Care*, vol 6(4):1–13.

9. Kilo CM (1999). 'Improving care through collaboration. *Pediatrics*, vol 103 (1 Suppl E), pp 384–393.

10. Wilson T, Berwick DM and Cleary PD (2003). 'What do collaborative improvement projects do? Experience from seven countries'. *Joint Commission Journal on Quality and Safety*, vol 29(2), pp 85–93.

11. No authors listed. '"Breakthrough" project helps trim OR delays'. *OR Manager*, vol 13(5), pp 21–22.

12. Armstrong B, Levesque O, Perlin JB et al (2005). 'Reinventing Veterans Health Administration: focus on primary care'. *Journal of Healthcare Management / American College of Healthcare Executives*, vol 50(6), pp 399–408.

13. Baier RR, Gifford DR, Lyder CH et al (2003). 'Quality improvement for pressure ulcer care in the nursing home setting: the Northeast Pressure Ulcer Project'. *Journal of the American Medical Directors Association*, vol 4(6), pp 291–301.

14. Baird RW (2001). 'Health care research and improvement. Quality improvement efforts in the intensive care unit: development of a new heparin protocol'. Includes commentary by Brock WA and Clemmer TP. Proceedings (*Baylor University Medical Center*), vol 14(3), pp 294–298.

15. Bartlett J, Cameron P, Cisera M (2002). 'The Victorian emergency department collaboration'. *International Journal for Quality in Health Care*, vol 14(6), pp 463–470.

16. Berry BB, Geary DL and Jaff MR (1998). 'A model for collaboration in quality improvement projects: implementing a weight-based heparin dosing nomogram across an integrated health care delivery system'. *The Joint Commission Journal on Quality Improvement*, vol 24(9), pp 459–469.

17. Besserman E, Brennan M, Brown PA et al)1998). 'Multidisciplinary achievement: the collaborative approach to rapid cycle ICU and hospital change'. *Quality Management in Health Care*, vol 6(4), pp 43–51.

18. Bonomi AE, Wagner EH, Glasgow RE et al (2002). 'Assessment of chronic illness care (ACIC): a practical tool to measure quality improvement'. *Health Services Research*, vol 37(3), pp 791–820.

19. Boushon B, Provost L, Gagnon J et al (2006). 'Using a virtual Breakthrough Series collaborative to improve access in primary care'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 32(10), pp 573–584.

20. Brattebo G, Hofoss D, Flaatten H et al (2002). 'Quality improvement report: Effect of a scoring system and protocol for sedation on duration of patients' need for ventilator support in a surgical intensive care unit'. *British Medical Journal*, vol 324(7350), pp 1386–1389.

21. Bundy DG, Randolph GD, Murray et al (2005). 'Open access in primary care: results of a North Carolina pilot project'. *Pediatrics*, vol 116(1), pp 82–87.

22. Burch K, Rhine W, Baker R et al (2003). 'Implementing potentially better practices to reduce lung injury in neonates'. *Pediatrics*, vol 111(4 Pt 2), e 432–e 436.

23. Chin MH, Cook S, Drum ML et al 2004. 'Improving diabetes care in midwest community health centers with the health disparities collaborative'. *Diabetes Care*, vol 27(1), pp 2–8.

24. Cleeland CS, Reyes-Gibby CC, Schall M et al (2003). 'Rapid improvement in pain management: the Veterans Health Administration and the Institute for Healthcare Improvement collaborative'. *Clinical Journal of Pain*, vol 19(5), pp 298–305.

25. Cole SA, Farber NC, Weiner JS et al (2006). 'Double-disease management or one care manager for two chronic conditions: pilot feasibility study of nurse telephonic disease management for depression and congestive heart failure'. *Disease Management*, vol 9, pp 266–276.

26. Daniel DM, Norman J, Davis C et al (2004). 'A state-level application of the chronic illness Breakthrough Series: results from two collaboratives on diabetes in Washington State'. *Joint Commission Journal on Quality and Safety*, vol 30(2), pp 69–79.

27. Daniel DM, Norman J, Davis C et al (2004). 'Case studies from two collaboratives on diabetes in Washington State'. *Joint Commission Journal on Quality and Safety*, vol 30(2), pp 103–108, 57.

28. Dellinger EP, Hausmann SM, Bratzler DW et al (2005). 'Hospitals collaborate to decrease surgical site infections'. *American Journal of Surgery*, vol 190(1), pp 9-15.

29. Doran KA, Henry SA and Anderson BJ (1998). 'Breakthrough change for adult cardiac surgery in a community-based cardiovascular program'. *Quality Management in Health Care*, vol 6(4), pp 29–36.

30. Dushay KM and Walker SD (2007). 'The National Organ Transplantation Breakthrough Collaborative – a Rhode Island hospital perspective'. *Medicine and Health, Rhode Island*, vol 90(3), pp 91–93.

31. Farbstein K and Clough J (2001). 'Improving medication safety across a multihospital system'. *The Joint Commission Journal on Quality Improvement*, vol 27(3), pp 123–137.

32. Flamm BL, Berwick DM and Kabcenell A (1998). 'Reducing cesarean section rates safely: lessons from a "Breakthrough Series" collaborative'. *Birth (Berkely, Calif.)*, vol 25(2), pp 117–124.

33. Gitomer RS (2005). 'Improving access, quality of care, and patient satisfaction in a general internal medicine practice'. *Journal of Clinical Outcomes Management*, vol 12(5), pp 245–249.

34. Glasgow RE, Funnell MM, Bonomi AE et al (2002). 'Self-management aspects of the improving chronic illness care Breakthrough Series: implementation with diabetes and heart failure teams'. *Annals of Behavioral Medicine*, vol 24(2), pp 80–87.

35. Gould DA, Lynn J, Halper D et al (2007). 'The New York City Palliative Care Quality Improvement Collaborative'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 33(6), pp 307–316.

36. Green PL and Plsek PE (2002). 'Coaching and leadership for the diffusion of innovation in health care: a different type of multi-organization improvement collaborative'. *The Joint Commission Journal on Quality Improvement*, vol 28(2), pp 55–71.

37. Gregory KD, Hackmeyer P, Gold L et al (1999). 'Using the continuous quality improvement process to safely lower the cesarean section rate'. *The Joint Commission Journal on Quality Improvement*, vol 25(12), pp 619–629.

38. Griffith C and Turner J (2004). 'United Kingdom National Health Service, Cancer Services Collaborative "improvement partnership", redesign of cancer services: a national approach'. *European Journal of Surgical Oncology*, vol 30, suppl 1, pp 1–86.

39. Hankinson MT, Faraone D and Blumenfrucht M (2006). 'Sustained improvement for specialty clinic access'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 32(3), pp 142–151.

40. Henderson D, Dempsey C, Larson K et al (2003). 'The impact of IMPACT on St. John's Regional Health Center'. *Missouri Medicine*, vol 100(6), pp 590–592.

41. Henderson D, Dempsey C and Appleby D (2004). 'A case study of successful patient flow methods: St. John's Hospital'. *Frontiers of Health Services Management*, vol 20(4), pp 25–30.

42. Joshi M, Kazandjian V, Martin P et al (2005). 'John M. Eisenberg Patient Safety and Quality Awards: a comprehensive grassroots model for statewide safety improvement'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 31(12), pp 671–677.

43. Kaempf JW, Campbell B, Sklar RS et al (2003). 'Implementing potentially better practices to improve neonatal outcomes after reducing postnatal dexamethasone use in infants born between 501 and 1250 grams'. *Pediatrics*, vol 111(4 Pt 2), e534–e541.

44. Katzelnick DJ, Von-Korff M, Chung H et al (2005). 'Applying depression-specific change concepts in a collaborative Breakthrough Series'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 31(7), pp 386–397.

45. Kerr D, Bevan H, Gowland B et al (2002). 'Redesigning cancer care'. *British Medical Journal*, vol 324, pp 164–166.

46. Kilbride HW, Wirtschafter DD, Powers RJ et al (2003). 'Implementation of evidence-based potentially better practices to decrease nosocomial infections'. *Pediatrics*, 111(4 Pt 2), e 519–e 533.

47. Kosseff AL and Niemeier S (2001). 'SSM Health Care clinical collaboratives: improving the value of patient care in a health care system'. *The Joint Commission Journal on Quality Improvement*, vol 27(1), pp 5–19.

48. LaBresh KA, Ellrodt AG, Gliklich R et al (2004). 'Get with the guidelines for cardiovascular secondary prevention: pilot results'. *Archives of Internal Medicine*, vol 164(2), pp 203–209.

49. Lain J, Young C, Dworkin G et al (1998) 'Improving efficiencies and reducing costs in adult cardiac surgery: a team approach'. *Quality Management in Health Care*, vol 6(4), pp 37–41.

50. Landis SE, Schwarz M and Curran DR (2006). 'North Carolina family medicine residency programs' diabetes learning collaborative'. *Family Medicine*, vol 38(3), pp 190–195.

51. Leape LL, Kabcenell AI, Gandhi TK et (2000). 'Reducing adverse drug events: lessons from a Breakthrough Series collaborative'. *The Joint Commission Journal on Quality Improvement*, vol 26(6), pp 321–331.

52. Lynn J (1999). 'Reforming the care system to support those coming to the end of life'. *Cancer Control*, vol 6(2), pp 131–135.

53. Lynn J, Nolan K, Kabcenell A et al (2002). 'Reforming care for persons near the end of life: the promise of quality improvement'. *Annals of Internal Medicine*, vol 137(2), pp 117–122.

54. Malenka DJ and O'Connor GT (1998). 'The Northern New England Cardiovascular Disease Study Group: a regional collaborative effort for continuous quality improvement in cardiovascular disease'. The Jo*int Commission Journal on Quality Improvement*, vol 24(10), pp 594–600.

55. Mercier CE, Barry SE, Paul K et al 2007. 'Improving newborn preventive services at the birth hospitalization: a collaborative, hospital-based quality improvement project'. *Pediatrics*, vol 120, pp 481–488.

56. Mills PD, Weeks WB and Surott-Kimberly BC (2003). 'A multihospital safety improvement effort and the dissemination of new knowledge'. *Joint Commission Journal on Quality and Safety*, vol 29(3), pp 124–133.

57. Mills PD and Weeks WB (2004). 'Characteristics of successful quality improvement teams: lessons from five collaborative projects in the VHA'. *Joint Commission Journal on Quality and Safety*, vol 30(3), pp 152–162.

58. Neily J, Howard K, Quigley P and Mills PD (2005). 'One-year follow-up after a collaborative Breakthrough Series on reducing falls and fall-related injuries'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 31(5), pp 275–285.

59. Nolan K, Schall MW and Erb F, Nolan T (2005). 'Using a framework for spread: the case of patient access in the Veterans Health Administration'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 31(6), pp 339–347.

60. Nugent WC (2005). 'Building and supporting sustainable improvement in cardiac surgery: the Northern New England experience'. *Seminars in Cardiothorac and Vascular Anesthesia*, vol 9(2), pp 115–118.

61. O'Connor GT, Plume SK, Olmstead EM et al (1996). 'A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group'. *Journal of the American Medical Association*, vol 275(11) pp 841–846.

62. Pearson ML, Wu S, Schaefer J et al (2005). 'Assessing the implementation of the chronic care model in quality improvement collaboratives'. *Health Services Research*, vol 40(4), pp 978–996.

63. Resar R, Pronovost P, Haraden C et al (2005). 'Using a bundle approach to improve ventilator care processes and reduce ventilator-associated pneumonia'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 31(5), pp 243–248.

64. Rogers G, Alper E, Brunelle D et al (2006). 'Reconciling medications at admission: safe practice recommendations and implementation strategies'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 32(1), pp 37–50.

65. Saunders RP, Abraham MR, Crosby MJ et al (2003). 'Evaluation and development of potentially better practices for improving family-centered care in neonatal intensive care units'. *Pediatrics*, vol 111(4 Pt 2), e 437– e 449.

66. Schiff GD, Wisniewski M, Bult J et al (2001). 'Improving inpatient antibiotic prescribing: insights from participation in a national collaborative'. *The Joint Commission Journal on Quality Improvement*, vol 27(8), pp 387–402.

67. Silver MP and Antonow JA (2000). 'Reducing medication errors in hospitals: a peer review organization collaboration'. *The Joint Commission Journal on Quality Improvement*, vol 26(6), pp 332–340.

68. Siomos EE, Newsom RS, Camponeschi J et al (2005). 'A statewide collaboration to monitor diabetes quality improvement among Wisconsin health plans'. *American Journal of Managed Care*, vol 11(5), pp 332–336.

69. Stoeckle-Roberts S, Reeves MJ, Jacobs BS et al (2006). 'Closing gaps between evidence-based stroke care guidelines and practices with a collaborative quality improvement project'. *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources*, vol 32(9), pp 517–527.

70. Toncich G, Cameron P, Virtue E et al (2000). 'Institute for Health Care Improvement collaborative trial to improve process times in an Australian emergency department'. *Journal of Quality of Clinical Practice*, vol 20(2-3), pp 79–86.

71. Wagner EH, Glasgow RE, Davis C et al (2001). 'Quality improvement in chronic illness care: a collaborative approach'. *The Joint Commission Journal on Quality Improvement*, vol 27(2), pp 63–80.

72. Wang A, Wolf M, Carlyle R et al (2004). 'The North Carolina experience with the diabetes health disparities collaboratives'. *Joint Commission Journal on Quality and Safety*, vol 30(7), pp 396–404.

73. Weeks WB, Mills PD, Dittus RS et al (2001). 'Using an improvement model to reduce adverse drug events in VA facilities'. *The Joint Commission Journal on Quality Improvement*, vol 27(5), pp 243–254.

74. Whittington J and Cohen H (2004). 'OSF healthcare's journey in patient safety'. *Quality Management in Health Care*, vol 13(1), pp 53–59.

75. Womer RB, Tracy E, Soo-Hoo W et al (2002). 'Multidisciplinary systems approach to chemotherapy safety: rebuilding processes and holding the gains'. *Journal of Clinical Oncology*, vol 20(24), pp 4705–4712.

76. Zhang H, Alexander JA, Luttrell J et al 2005. 'Data feedback and clinical process improvement in acute myocardial infarction'. *American Heart Journal*, vol 149(5), pp 856–861.

77. Young PC, Glade GB, Stoddard GJ et al (2006). 'Evaluation of a learning collaborative to improve the delivery of preventive services by pediatric practices'. *Pediatrics*, vol 117(5), pp 1469–1476.

78. Baier RR, Gifford DR, Patry G et al 2004. 'Ameliorating pain in nursing homes: a collaborative quality-improvement project'. *Journal of the American Geriatrics Society*, vol 52(12), pp 1988–1995.

79. Benedetti R, Flock B, Pedersen S et al (2004). 'Improved clinical outcomes for fee-for-service physician practices participating in a diabetes care collaborative'. *Joint Commission Journal of Quality and Safety*, vol 30(4), pp 187–194.

80. Landon BE, Wilson IB, McInnes K et al (2004). 'Improving patient care. Effects of a quality improvement collaborative on the outcome of care of patients with HIV infection: the EQHIV study'. *Annals of Internal Medicine*, vol 140(11), pp 887–896, I92.

81. Homer CJ, Forbes P, Horvitz L et al (2005). 'Impact of a quality improvement program on care and outcomes for children with asthma'. *Archives of Pediatrics & Adolescent Medicine*, vol 159(5), pp 464–469.

82. Mangione-Smith R, Schonlau M, Chan KS et al (2005). ‘Measuring the effectiveness of a collaborative for quality improvement in pediatric asthma care: does implementing the chronic care model improve processes and outcomes of care?’ *Ambulatory Pediatrics*, vol 5(2), pp 75–82.

83. Schonlau M, Mangione-Smith R, Chan KS et al (2005). ‘Evaluation of a quality improvement collaborative in asthma care: does it improve processes and outcomes of care?’ *Annals of Family Medicine*, vol 3(3), pp 200–208.

84. Asch SM, Baker DW, Keesey JW et al (2005). ‘Does the collaborative model improve care for chronic heart failure?’ *Medical Care*, vol 43(7), pp 667–675.

85. Baker DW, Asch SM, Keesey JW et al (2005). ‘Differences in education, knowledge, self-management activities, and health outcomes for patients with heart failure cared for under the chronic disease model: the improving chronic illness care evaluation’. *Journal of Cardiac Failure*, vol 11(6), pp 405–413.

86. Horbar JD, Rogowski J, Plsek PE et al (2001). ‘Collaborative quality improvement for neonatal intensive care’. *Pediatrics*, vol 107(1), pp 14–22.

87. Rogowski JA, Horbar JD, Plsek PE et al (2001). ‘Economic implications of neonatal intensive care unit collaborative quality improvement’. *Pediatrics*, vol 107(1), pp 23–29.

88. Horbar JD, Carpenter JH, Buzas J et al (2004). ‘Collaborative quality improvement to promote evidence based surfactant for preterm infants: a cluster randomised trial’. *British Medical Journal*, vol 329(7473), p 1004.

89. Howard DH, Siminoff LA, McBride V et al (2007). ‘Does quality improvement work? Evaluation of the organ donation breakthrough collaborative’. *Health Services Research*, vol 42, pp 2160–2173.

90. Pierce-Bulger M and Nighswander T (2001). ‘Nutaqsiivik – an approach to reducing infant mortality using quality improvement principles’. *Quality Management in Health Care*, vol 9(3), pp 40–46.

Appendix 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)	Combined the steps in set 2 with: Statistics Statistics-and-numerical-data

Appendix 2: Uncontrolled studies

Author	Year	Topic
Anonymous	1997	Emergency department / waits and delays
Armstrong	2005	Access
Baier	2003	Pressure sores
Baird	2001	Intensive care
Bartlett	2002	Emergency department
Berry	1998	Heparin dosing
Besserman	1998	Intensive care
Bonomi	2002	ACIC
Boushon	2006	Access (virtual collaborative)
Brattebo	2002	Intensive care
Bundy	2005	Access
Burch	2003	Neonates
Chin	2004	Diabetes CCM
Cleeland	2003	Pain management
Cole	2006	Depression and congestive heart failure
Daniel	2004	Diabetes
Daniel	2004	Diabetes
Dellinger	2005	Surgical site infections
Doran	1998	Cardiac surgery
Dushay	2007	Organ transplantation
Farbstein	2001	Medication safety
Flamm	1998	Cesarean section
Gitomer	2005	Access
Glasgow	2002	Diabetes and heart failure
Gould	2007	Palliative care
Green	2002	Coaching and leadership
Gregory	1999	Cesarean section
Griffith	2004	Cancer care
Hankinson	2006	Clinical access
Henderson	2003	Patient flow
Henderson	2004	Patient flow
Joshi	2005	Patient safety / ICU
Kaempf	2003	Ventilation neonatal care
Katzelnick	2005	Depression

Kerr	2002	Cancer
Kilbride	2003	Nosocomial infections in neonatal care
Kosseff	2001	Divers
Labresh	2004	Cardiovascular secondary prevention
Lain	1998	Cardiac surgery
Landis	2006	Diabetes
Leape	2000	Adverse drug events
Lynn	1999	End-of-life care
Lynn	2002	End-of-life care
Malenka	1998	Cardiovascular disease
Mercier	2007	Newborn preventive services
Mills	2003	Patient safety and dissemination
Mills	2004	Team success
Neily	2005	One-year follow-up after BTS falls and fall-related injuries
Nolan	2005	Patient access and framework for spread
Nugent	2005	Cardiac surgery
O'connor	1996	Coronary artery bypass graft surgery
Pearson	2005	Implementation of CCM in BTS heart failure, diabetes, depression, asthma
Resar	2005	Intensive care
Rogers	2006	Medication safety
Saunders	2003	Family-centred care
Schiff	2001	Antibiotic prescribing
Silver	2000	Medication errors
Siomos	2005	Diabetes
Stoeckle-Roberts	2006	Stroke
Toncich	2000	Emergency department
Wagner	2001	Chronic illness care
Wang	2004	Diabetes
Weeks	2001	Medication safety
Whittington	2004	Patient safety
Womer	2002	Medication safety
Young	2006	Pediatric care
Zhang	2005	Acute myocardial infarction