Methods Research Report

Developing Consistent and Useful Quality Improvement Study Data Extraction for Health Systems

AHRQ
Agency for Healthcare Research and Quality

Methods Research Report

Developing Consistent and Useful Quality Improvement Study Data Extraction for Health Systems

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. 290-2015-00012-I

Prepared by:
University of Connecticut Evidence-based Practice Center
Hartford, CT

Investigators:
Adrian V. Hernandez, M.D., Ph.D.
Yuani M. Roman, M.D., M.P.H.
C. Michael White, Pharm..D., FCP, FCCP

AHRQ Publication No. 19(20)-EHC025-EF
September 2020

Key Messages

Purpose of review

To determine if the consistency of data extraction improves after creating explicit instructions for quality improvement study assessment criteria.

Key messages

- Findings: Creating explicit instructions to accompany quality improvement study assessment criteria greatly enhances the consistency of data extraction.
- Lessons Learned for EPC Program: While inconsistency in data extraction was a major barrier to moving forward with identifying criteria crucial for the understanding of quality improvement studies, an iterative process to assess the consistency and then refine the instructions can be successful.
- Utility for Health Systems: Learning health systems can expect to see extracted data for different criteria that would look consistent regardless of the person doing the extraction and enhancing the consistency from report to report as well.

This report is based on research conducted by the University of Connecticut Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2015-00012I). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

The information in this report is intended to help healthcare decision makers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of healthcare services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

Suggested citation: Hernandez AV, Roman YM, White CM. Developing Consistent and Useful Quality Improvement Study Data Extraction for Health Systems. Methods Research Report. (Prepared by the University of Connecticut Evidence-based Practice Center under Contract No. 290-2015-00012-I.) AHRQ Publication No. 19(20)-EHC025-EF. Rockville, MD: Agency for Healthcare Research and Quality; September 2020. Posted final reports are located on the Effective Health Care Program search page.
DOI: https://doi.org/10.23970/AHRQEPCMETHQUALIMPRCRITERIA.

Preface

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

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although they may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers and the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Methods Research Project. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane Rockville, MD 20857, or by e-mail to epc@ahrq.hhs.gov.

Gopal Khanna, M.B.A.
Director
Agency for Healthcare Research and Quality

Arlene Bierman, M.D., M.S.
Director
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Center Program
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

David W. Niebuhr, M.D., M.P.H., M.Sc.
Medical Officer/Task Order Officer
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Key Informants

The EPC consulted several Key Informants who represent the end-users of the research, health-systems. The EPC sought the Key Informant input on the priority areas for research and synthesis. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any conflicts of interest.

Christina M. Polomoff, PharmD, BCACP, BCGP
Hartford Healthcare Integrated Care Partners
Wethersfield, CT

Kevin Chamberlin, Pharm.D., FASCP
University of Connecticut (UConn) John Dempsey Hospital
At UConn Health
Farmington, CT

Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report does not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential non-financial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential non-financial conflicts of interest identified.

Ethan Balk, M.D., M.P.H.
Center for Evidence Synthesis in Health
Brown University
Providence, RI

Paula J. Santrach, M.D.
Department of Laboratory Medicine and Pathology
Mayo Clinic
Rochester, MN

Developing Consistent and Useful Quality Improvement Study Data Extraction for Health Systems

Structured Abstract

Background. Quality improvement studies can provide important insight to learning health systems. The Agency for Healthcare Research and Quality (AHRQ) could devote resources to collate and assess these quality improvement studies to support learning health systems (LHS) but there is no reliable data on the consistency of data extraction for important criteria.

Methods. We identified quality improvement studies in asthma and evaluated the consistency of data extraction from two experienced independent reviewers at three time points: baseline where only a rudimentary description of the criteria was available, first revision where explicit instructions for each criterion were created, and final revision where the instructions were revised. Six investigators looked at the data extracted by one of the systematic reviewers and then the other for the same criteria and determined the extent of similarity on a scale of 0 to 10 (where 0 represented no similarity and 10 perfect similarity). There were 42 assessments for baseline, 42 assessments for the first revision, and 42 assessments for the final revision. We then asked two LHS participants to assess the relative value of our criteria in a pilot phase.

Results. We went through two refinements of the data extraction instructions for each criterion and were able to improve the consistency of extraction from 1.17 ± 1.85 at baseline to 6.07 ± 2.76 after revision one ($P<0.001$) and to 6.81 ± 1.94 out of 10 for the final revision ($P<0.001$). However, the final revision was not significantly improved over revision one ($p=0.14$). In the pilot phase, our two LHS participants felt that some of our 33 criteria were more valuable than others were.

Discussion/Conclusion. Creating explicit instructions for extracting data for quality improvement study helps enhance the consistency of data extraction. Future studies with a larger cadre of LHS participants should help determine the most important criteria.

Contents

Tables

Appendixes

Evidence Summary

Background

Quality improvement studies can provide important insight to learning health systems (LHS) but are difficult to locate and time intensive to evaluate. Agency for Healthcare Research and Quality (AHRQ) could devote resources to collate and assess these quality improvement studies but there is no reliable data on the consistency of data extraction for important criteria. There is also a lack of LHS feedback on the value of the 33 criteria themselves.

Methods

We used the 33 candidate criteria for characterizing quality improvement studies identified by AHRQ in previous methods work. We utilized studies contained in the AHRQ report entitled "Closing The Quality Gap: A Critical Analysis of Quality Improvement Strategies: Volume 5—Asthma Care" as our data sources.[1] We evaluated the consistency of data extraction from two experienced independent reviewers at three time points: Baseline where only a rudimentary description of the criteria was available, Revision One where the instructions for each criterion were developed, and Final Revision where the instructions were revised. Six investigators looked at the data extracted by one of the systematic reviewers and then the other for the same criteria and determined the extent of similarity on a scale of 0 to 10 (where 0 represented no similarity and 10 perfect similarity of extraction). There were 42 assessments (seven randomly selected criteria with six researchers assessing the consistency of the dual extraction) for baseline, 42 assessments for the first revision, and 42 assessments for the final revision. Since extraction consistency is non-parametric data, it was compared between the groups (baseline vs. first revision, baseline vs. final revision, first revision vs. final revision) using the related-samples Wilcoxon Signed Rank Test (SPSS). A p-value of <0.05 was considered statistically significant.

In a second pilot phase, we completed data extraction for all seventeen studies and then asked two LHS participants to provide feedback on the most and least valuable of the 33 criteria that we proposed.

Results

At baseline, we found very little extraction similarity for the criteria between our two independent systematic reviewers. After we refined the instructions about how to extract data for the criteria the first time, we dramatically improved the consistency of extraction scores between the two reviewers by 5.2-fold versus baseline ($P<0.001$). The final revision of the instructions increased the rating of similarity between the two reviewers by 5.8-fold versus baseline ($P<0.001$) but this is only marginally better than after the first revision ($p=0.14$) (Table A).
In the exploratory phase, the LHS participants both identified the same four criteria as being most valuable (Frequency of Intervention, Number and Description of Components, Types of Intervention Effects, Organizational Setting) with six other criteria selected only by one of them. Two of the criteria were chosen as being least valuable (Organizational Receptivity/ Readiness, Organizational History of Change) by both participants with six others selected only by one of them. No criteria found to be most valuable was subsequently determined least valuable by the other participant.

Table A. Consistency of data extraction at baseline and during two revisions of instructions

Ratings	Mean (Standard Deviation)	Median (25^{th}-75^{th} Percentile)
Baseline Ratings	1.17 (1.85)	0 (0–3)
First Revision Ratings	6.07 (2.76)*	6 (3–9)
Final Revisions Ratings	6.81 (1.94)*,†	7 (6–8)

* $p<0.0001$ versus baseline

† = $p=0.14$ versus first revision. Scores could range from 0 to 10 where 0 = no consistency of extraction between systematic reviewers and 10 = perfect consistency.

Discussion

We are unaware of any other studies explicitly looking at the consistency of data extraction of quality improvement studies and how it can be improved. The standard description of the criteria is insufficient to allow for standardized data extraction. Our systematic reviewers are trained members of EPCs with ample experience but in the absence of detailed instructions, the consistency of data extraction from quality improvement studies in our study was very poor. Fortunately, we found that heterogeneous data extraction is surmountable with explicit instructions. We went through two iterative refinements of the data extraction instructions for each criterion and were able to improve the consistency of extraction from 1.17 to 6.81 out of 10. Since the second refinement only increased the consistency of extraction slightly, further refinements are unlikely to provide appreciable enhancements. We believe that the uniqueness of quality improvement studies requires this more standardized approach to data extraction versus more traditional observational studies and randomized trials.

While we have 33 criteria to extract, not all of the criteria are equally valuable to LHS. Future studies should further explore the most valuable criteria to keep and those that can be eliminated because the value of the criteria may change for different diseases and study sets or among different health systems.

Conclusion

In the absence of explicit and detailed instructions, there is very high heterogeneity in the data that independent systematic reviewers extract. This improves considerably with the refinement of the instructions for the criteria using an explicit process. Now that consistency of extraction has been enhanced, a future study should determine more extensively criteria yielding less value to learning health systems.

References

1. Bravata DM, Sundaram V, Lewis R, et al. Asthma Care. Vol 5 of: Shojania KG, McDonald KM, Wachter RM, et al, editors. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies. Technical Review 9 (Prepared by the Stanford University-UCSF Evidence-based Practice Center under Contract No. 290-02-0017). AHRQ Publication No. 04(07)-0051-5. Rockville, MD: Agency for Healthcare Research and Quality. January 2007.

Background

Quality patient care is an important part of health-system accreditation and is increasingly tied to revenue through Pay for Performance (P4P).[1,2,3] In 2017, only 25 percent of health systems added net revenue from P4P while 69 percent of health systems received a financial penalty.[2] In its report, Best Care at Lower Cost: The Path to Continuously Learning Health Care in America, the Health and Medicine Division of the National Academy of Sciences proposed the concept of the learning health system to improve healthcare quality and outcomes.[4] Regular health systems become learning health systems when they internally identify areas of quality weakness and strive to utilize the best knowledge to improve them.[5,6]

The Agency for Healthcare Research and Quality (AHRQ) created a working group in 2017 that interviewed nine leaders in quality and safety improvement in health systems to explore how they used evidence.[7] Health systems looked to other institutions within their buying group, institutions in a formal consortium, or institutions reporting their experiences in the biomedical literature, for quality improvement approaches. However, finding the full spectrum of quality improvement projects completed in an area from the medical literature is difficult. There are many ways to improve the quality of care including purchasing new equipment or software, dedicating staff to champion changes, educational programs, email reminders, electronic health record reminders, internal protocols and guidelines, prior authorization, and limiting utilization to a specialist or specialty service. Faced with multiple approaches to improve, health systems need to identify approaches that overlap with their strengths and are realistic with their fiscal and staffing reality. The best approach in one health system may not translate well to another and in some cases, may not even be feasible.

In 2018, a multi-EPC and AHRQ working group sought to develop a process to assess and present useful information from quality improvement projects to learning health systems.[8] There was progress in defining quality improvement studies and generating a list of 33 candidate criteria describing features of the interventions and the environment in which they were conducted. Unfortunately, there was little similarity to what was being extracted by experienced systematic reviewers, which was an impediment to progress and workgroup activities were postponed. Furthermore, these criteria were not assessed by learning health system participants to identify the ones of greatest value.

Goal/Objective

The University of Connecticut (UConn) EPC sought to move this important evidence review area forward by utilizing the criteria that the 2018 working group identified to better characterize quality improvement studies in order to:

(1) Primary Aim: Develop data extraction instructions for each criterion that can enhance the consistency of data extraction among different systematic reviewers.

(2) Exploratory Secondary Aim: Assess the difficulty health systems have with identifying and using quality improvement studies and the usefulness of the criteria that we developed in aggregate and for individual criteria by health systems.

Methods

Assessing the Consistency of Data Extraction

We utilized studies contained in the AHRQ report entitled "Closing The Quality Gap: A Critical Analysis of Quality Improvement Strategies: Volume 5—Asthma Care" as our data sources.[9] In the first phase of the study, we used the criteria that the 2018 AHRQ working group identified as potentially valuable for health systems to understand how to use and implement QI studies.[8] These criteria were selected from previous work in the quality improvement and implementation literature.[10-13]

The specific criteria include: 1. who is delivering the intervention (e.g., provider types)?; 2. who is receiving the intervention (e.g., patient types)?; 3. provider demographics; 4. recipient demographics; 5. active vs. passive components; 6. discretionary vs. mandatory components; 7. duration of the intervention; 8. frequency of intervention; 9. intensity of the intervention; 10. duration of the effect of the intervention; 11. team composition (people delivering the intervention); 12. external policies and incentives required; 13. required skills/training; 14. number and description of components; 15. a-priori components vs. added later/final; 16. theoretical foundation; 17. which interventions are independent; 18. cost of implementation; 19. leadership commitment and involvement; 20. Clinical champion involvement; 21. physical environmental changes required, 22. incentives; 23. implementation strategies; 24. types of intervention effects; 25. organizational setting; 26. financial setting; 27. organizational receptivity/readiness; 28. population needs/burden of illness; 29. geographic location; 30. external factors; 31. organizational history of change; 32. fidelity; and 33. intervention adaptation.

At baseline, two systematic reviewers extracted data for all 33 criteria from two different studies. There was only a rudimentary description of each criterion at this stage. When examining the consistency of the results, three investigators mutually developed detailed instructions on how to extract study data for each criterion. Two systematic reviewers then independently extracted data for all 33 criteria from two other source studies utilizing these instructions (Revision One). Reviewing inconsistencies among extractors led to research team to refine the instructions for each criterion and reduce ambiguity. Using the refined instructions (Final Revision), data was independently extracted by two systematic reviewers for each of the 33 criterion from two other studies. Our final set of instructions for data extraction for each criterion appears in Appendix A.

To assess the impact of the instructions on the consistency of data extraction, an investigator randomly selected dually extracted data for a specific criterion (termed a data field) from the studies. In all, seven dually extracted data fields were selected at baseline, seven after the first revision of the instructions were completed (first revision), and seven after the final instructions were completed (final revision). Six investigators looked at the data extracted by one of the systematic reviewers and then the other for the same field and determined the extent of similarity on a scale of 0 to 10 (where 0 represented no similarity and 10 perfect similarity of extraction). This meant there were 126 assessments of the similarity of extracted data (42 assessments for baseline, 42 assessments for first revision, and 42 assessments for the final revision). This non-parametric data on extraction consistency was compared between the groups (baseline vs. first revision, baseline vs. final revision, first revision vs. final revision) using the related-samples Wilcoxon Signed Rank Test (SPSS). We provide mean and median extraction consistency values

for Baseline, Revision 1, and Final Revision and a p-value of <0.05 considered statistically significant.

Assessing the Relative Value of the Criteria

In the exploratory phase of the study, we used several studies to create a table with extracted data for each of the criteria (Appendix B). We shared this table with quality improvement clinicians from Hartford Healthcare and the University of Connecticut Health Center. The participants were first asked to rate how difficult it is to find and use quality improvement studies in health system quality improvement endeavors. They were then asked how valuable it would be if AHRQ found and collated the background information about the interventions and the health systems from published quality improvement studies. For both questions, people were asked to use a 0 to 10 rating scale where 0 denotes no difficulty or no value and 10 denotes great difficulty or great value, respectively. Finally, we asked them to select seven of the proposed criteria they found most valuable and five they found least valuable. They provided written comments about why they selected these criteria.

We presented our proposed criteria and the data on how it impacted the similarity of extraction at the June 5, 2019, joint AHRQ EPC and Leaning Health System Panel meeting in Rockville, MD and sought feedback.

Health System and Representative Description

Hartford Healthcare is a fully integrated health system that includes seven hospitals (including an 800-bed hospital with a level I trauma center), Connecticut's most extensive behavioral health network, a large multispecialty physician group, a regional home care system, an array of senior care services, a large physical therapy and rehabilitation network, and an accountable care organization.

Christina M. Polomoff, Pharm.D., BCACP, BCGP is the Population Health Clinical Pharmacist at Integrated Care Partners, Hartford Healthcare's physician-led clinically integrated network. She works with providers to improve pharmacy performance metrics related to cost and quality, and provides direct care to high-modifiable risk patients.

UConn Health is a rural academic health-system including the 224-bed John Dempsey Hospital located in the town of Farmington, CT. The UConn Health-System made important strides in enhancing safety over the past 5 years and is above the median ranking for patient safety by Consumer Reports.

Kevin Chamberlin, Pharm.D., FASCP is a quality improvement inpatient clinical pharmacist and director of the postgraduate pharmacy residency program.

Results

Impact of Iterative Instruction Revision on Consistency

For our primary aim, Table 2 delineates the degree to which raters found similarities in data extracted from studies for each of the randomly selected criteria between our two independent extractors. At baseline, we found very little extraction similarity for the criteria in the source studies by our two independent systematic reviewers. After we refined the instructions about how to extract data for the criteria the first time, we dramatically improved the consistency of extraction scores between the two reviewers by 5.2-fold (P<0.001). The second and final revision of the instructions again increased the rating of similarity between the two reviewers versus baseline by 5.8-fold (P<0.001) but only nominally different from that achieved after the first revision (p=0.14) (Table 1). The mean and median agreement scores of 6.81 and 7 represent a moderate level of agreement with the final set of instructions for extracting data for these criteria.

Table 1. Data extraction consistency ratings

Ratings	Mean (Standard Deviation)	Median (25^{th}-75^{th} Percentile)
Baseline Ratings	1.17 (1.85)	0 (0–3)
First Revision Ratings	6.07 (2.76)*	6 (3–9)
Final Revisions Ratings	6.81 (1.94)*,†	7 (6–8)

* p<0.0001 versus baseline

† = p=0.14 versus first revision. Scores could range from 0 to 10 where 0 = no consistency of extraction between systematic reviewers and 10 = perfect consistency.

Assessing the Value of the Criteria

In the exploratory phase, the two health system representatives rated how difficult it is to find and use quality improvement studies in health system quality improvement endeavors in general. On a scale of 0 to 10, the first key informant rated it a 6 (moderately difficult) while the second reviewer rated it a 4 (moderately difficult), justified by the paucity of published quality improvement literature that they can locate and the time involved in evaluating them. They were then asked how valuable it would be if AHRQ found and collated the demographic information about the health systems and the interventions used in published quality improvement studies. On a scale or 0 to 10, the first key informant rated it a 9 (highly valuable) and the second rated it a 6 (moderately valuable).

The criteria selected by at least one of the health system participants as being of most value and least value are included in Tables 3 and 4, respectively, along with the rationale the participants gave when selecting them. Both participants chose the same four most valuable criteria (i.e. Frequency of Intervention, Number and Description of Components, Types of Intervention Effects, Organizational Setting) while six others criteria were selected by only one of them. Both key informants selected two criteria as being least valuable (i.e. Organizational Receptivity/ Readiness, Organizational History of Change) with six others solely selected by one of them. No criteria selected as being most valuable by one participant was selected least valuable by the other. All other criteria not selected most or least valuable by default were determined to be of intermediate value to health systems.

We presented our findings to the AHRQ joint EPC and Learning Health System panel meeting. The general sentiment of the group was that the criteria to characterize quality improvement studies are comprehensive. One member of the Learning Health System Panel cautioned that while some of the criteria may be less valuable in a topic like asthma, it could be more valuable elsewhere. Three people stressed that further work would be needed to assess the value of the criteria for learning health systems. Another meeting participant referred us to a tool called the Minimum Quality Criteria Set (QI-MQCS) and a method to determine the most valuable criteria that we describe in the discussion section.

Table 2. Most valuable criteria selected by health system participants

Criteria	Key Informant Direct Quotes (Two Quotes Means Both Participants Selected the Criterion)
8. Frequency of Intervention	*"Sustainability of the effect is a significant consideration to the intervention. Resources involved cost money, as does time."* *"The frequency of the intervention is helpful to assess feasibility of carrying out a similar intervention in our setting. If the frequency is unrealistic based on our team composition, and number of team members and recipients, or would require a significant amount of time beyond what we could allocate, and then the study would not translate to our health system setting."*
14. Number and Description of Components	*"Number of components to the intervention will influence acceptability and repeatability of the intervention. It is perceived that too many components will limit external validity/uptake of the intervention due to thoughts of difficulty of implementation. Similarly, lack of detail in the description to each of the components of the intervention could affect uptake of the intervention."* *"Knowing the structure of care/intervention/timeline in the study provides a bird's eye view of the components for implementation. This could be used strategically in the brainstorming session of the health system."*
24. Types of Intervention Effects	*"Type of benefits (or negatives) – morbidity/mortality outcomes, changes in medication use, changes in disease severity or classification, functionality changes, impact on quality of life, what is the NNT vs. NNH, etc. – are ultimately what the clinician cares most about from the intervention being made."* *"The primary outcomes and clinical/process measures allows the organization to hone in on the relevance of the study to their target measures. For example, if the study's primary outcome is parents' view of physician behavior, this would not be as relevant for our target measure of emergency department visits or Asthma Medication Ratio $\geq$0.5."*
25. Organizational Setting	*"This is important for consideration of feasibility of applicability between different organizational settings and support structures, thus influencing external validity of the study being considered."* *"The type of setting (hospital vs. group of clinics vs. clinically integrated network vs. ACO) and number of practices is important to determine scalability, coordination, integration, and communication. For example, if the study takes place in 1 primary care office versus a clinically integrated network comprised of 8 hospitals."*
1. Who is Delivering the Intervention	*"The executer of the intervention is necessary to determine if the study would translate to a health system setting. For example, if the main executer is a health coach, this would likely not be feasible as we do not have health coaches in current state."*
2. Who is Receiving the Intervention.	*"The recipient of the intervention is one of the most important details of a study because that defines the population needing the quality improvement. For example if our health system would like to implement a quality improvement project for patients with asthma, and use a HEDIS measure (such as Asthma Medication Ratio $\geq$ 0.5 for members 5-64 years of age who were diagnosed with persistent asthma), we need to know if the recipients in the study align with our target population. If the study is for children 1-4 years of age, we would likely not extrapolate the interventions/findings of the study to our population."*
4. Recipient Demographics	*"Recipient demographics allow us to further narrow if the study is applicable to our population. Just knowing the population is "patients with asthma" would not be sufficient. For our CMS Part D star rating measures, we would reference studies whose patient composition is older vs. younger to reflect our targeted Medicare population."*
7. Duration of Intervention	*"Sustainability of the effect is a significant consideration to the intervention. Resources involved cost money, as does time."*

Criteria	Key Informant Direct Quotes (Two Quotes Means Both Participants Selected the Criterion)
9. Intensity of the Intervention	*"Combined with the duration of effect, the overall perceived and actual intensiveness of the intervention will likely affect patient / provider willingness to embrace the intervention, accept it, and carry through with it."*
10. Duration of the Effect of the Intervention	*"Perhaps the ultimate testament to the overall input/output aspect of the intervention. How much time, resources, etc. went in to gain X time of effect from the intervention? The longer the duration, the perceived greater the outcome and benefit and thus buy-in from those involved."*

Table 3. Least valuable criteria selected by health system participants

Criteria	Key Informant Direct Quotes (Two Quotes Means Both Participants Selected the Criterion)
27. Organizational Receptivity/ Readiness	*"Often, QI projects are done not because an organization wants to make the change, but more likely because some regulatory body mandates the change. Most institutions embrace QI projects as a method of piloting or demonstrating effectiveness of change in this case. Thus, receptivity does not necessarily seem like a highly valued criterion."* *"Reading about the organizational receptivity/readiness in a study would not be helpful. Internal factors in our health system would facilitate or prohibit receptivity/readiness to undertake a quality improvement project."*
31. Organizational History of Change	*"As for criteria 27, this criterion doesn't seem to relate to any sort of importance within the acceptability of the intervention."* *"Knowing the organizational history of change would be less helpful. Having an awareness of the current status of the organizational setting to determine translatability would be more meaningful."*
17. Which Interventions are Independent	*"The type and frequency of interventions are more valuable than whether or not the intervention is "independent". The components of the intervention including who delivers the intervention is more important."*
19. Leadership Commitment and Involvement	*"Leadership commitment/involvement is a moot point because it can be assumed that our health system will need leadership approval to move forward with a quality improvement project. If the study did not have leadership commitment/involvement, the health system would still need leadership buy-in regardless."*
22. Clinical Champion Involvement	*"Incentives used in the study would likely not be realistic or applicable to our health system setting. The health system would likely be incentivized by the projected improvement in cost/quality outcomes, possibility of team expansion, and potential to improve patient/provider satisfaction."*
23. Implementation Strategies	*"Only 1 of 7 studies had these as a component for consideration. It would seem as though these would be already incorporated or described in either criterion 14 or 9 and thus not necessary to identify as a separate criterion."*
32. Fidelity	*"While important for understanding the extent to which delivery of an intervention adheres to the protocol, sound research implies this was done. QI projects are not typically powered or designed to target statistical significance so much as to identify clinical and/or financial impact."*
33. Intervention Adaptation	*"When intervention adaptation is described within the discussion of the study, external validity could most definitely be enhanced. That said, the more adaptation is carried out, the further away from the original intervention and thus repeatability and intent can become clouded."*

Discussion

This is the first study, that we are aware of, that explicitly looked at the consistency of data extraction from quality improvement studies. Our systematic reviewers are trained members of EPCs with ample experience but in the absence of detailed instructions, the consistency of data extraction from quality improvement studies in our study was very poor. Fortunately, we found that heterogeneous data extraction is surmountable with explicit instructions developed in an iterative fashion. We went through two refinements of the data extraction instructions for each criterion and were able to improve the consistency of extraction from baseline to the final revision from 1.17 to 6.81 out of 10. Since the second refinement only increased the consistency of extraction slightly over the first, further refinements are unlikely to provide appreciable enhancements. We believe that the uniqueness of quality improvement studies requires this standardized approach to data extraction versus more traditional observational studies and randomized trials. Since we only looked at one disease state and a relatively few number of studies, there are limitations to our approach and perhaps some issues of applicability as well. As such, further research looking at other disease states would be beneficial.

In the exploratory phase of our study, we had two learning health system participants extensively review our criteria and a representative table with data extracted from quality improvement studies in asthma. They were asked to identify those criteria that they felt were of greatest and least value. While this is interesting pilot data, definitive determinations of the final criteria to retain cannot be made. The least valuable criteria (particularly 19, 22, 27, 31 and 33) represent "diagnostic criteria" for suboptimal outcomes and/or improvement failure. They relate to real or potential barriers to change. While our representatives did not find them as valuable other larger health systems or those from other geographic areas might. Intervention adaptation is increasingly important as health systems get larger and more diverse. With multiple sites and care delivery models in a large geographic footprint, adaptation is the rule and not the exception. That is even more so in the early days of a merged organization arising from multiple previously independent entities. As such, only more extensive research into the value of these criteria will identify those of greatest important or identify scenarios where an extended or contracted criteria set could be used.

Our criteria overlaps with that of the QI-MQCS. The QI-MQCS criteria is a pared down version of criteria created by the Standards for Quality Improvement Reporting Excellence (SQUIRE) group.[10,14] The QI-MQCS tool, developed with input from nine expert panelists, selected 14 criteria that the panelists gave a mean rating of 2.0 or greater in terms of importance (scale from 1 to 3 where 3 denoted it should be included, 2 denoted it may be included, and 1 denoted it should not be included) and two additional criteria not vetted through the expert panel. Of the 16 QI-MQCS criteria, our criteria set includes 12 of them (organizational motivation, organizational readiness, intervention, intervention rationale, organizational characteristics, implementation, timing, adherence/fidelity, penetration or reach, sustainability, comparator, and data source). Our criteria did not include the study design, health outcomes, ability for the intervention to be replicated, or inclusion of study limitations criteria. However, our quality improvement applicability table would accompany the standard information presented in EPC evidence reviews where the study design, health outcomes, and qualitative or quantitative synthesis of the results appear. In total, our criteria encompasses and expands on their criteria. This is not surprising since we both relied on the SQUIRE 2.0 criteria while we also used other sources to identify criteria that we felt was valuable.[10-13]

We agree that following the QI-MQCS process and asking several members of learning health systems to rate the importance of our selected criteria on the same three-point scale is an important next step. If we can pare down the number of criteria that need to be extracted without losing information important to a learning health system, it would make it more efficient for data extractors and ultimately, the readers of the report. However, we do not believe that just having the learning health systems rate the value of a criterion purely in abstract without looking at actual data extracted from studies for each criterion is adequate. We do agree though that, as one member of the Learning Health System Panel stated, the value of the criteria might change given the broad nature of the quality improvement literature so extracting data from more than one disease would be valuable.

Conclusions

Our study suggests that learning health systems need support in identifying quality improvement studies and the key features of the interventions and the institutions that carried them out. In the absence of explicit and detailed instructions, there is very high heterogeneity in data extraction among independent reviewers that improves considerably with the refinement of the criteria using an explicit process. Now that consistency of extraction has been enhanced for each of our candidate criteria, a future study should determine the relative value of each criterion to learning health systems.

References

1. The Joint Commission. Quality Accreditation Standards Information. Available at: https://www.jointcommission.org/standards_information/jcfaq.aspx. Accessed, 8/6/18.
2. Advisory Board, Inc. See your hospital's 2017 pay-for-performance penalty or bonus. January 15, 2018. Available at: https://www.advisory.com/daily-briefing/2017/01/23/hospital-p-4-p. Accessed 8/6/18.
3. Health Services Advisory Group. National Impact Assessment of the Centers for Medicare & Medicaid Services (CMS) Quality Measures Report. Feb 28, 2018. Available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures/Downloads/2018-Impact-Assessment-Report.pdf. Accessed 8/6/18.
4. Institute of Medicine. 2013. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. Washington, DC: The National Academies Press. Available at: https://doi.org/10.17226/13444. Accessed. 8/6/18
5. The Learning Healthcare Project. Available at: http://www.learninghealthcareproject.org/section/background/learning-healthcare-system. Accessed 8/6/18.
6. McGinnis JM. Evidence-based medicine - engineering the Learning Healthcare System. Stud Health Technol Inform. 2010;153:145-57.
7. White CM, Sanders-Schmidler GD, Butler M, et al. Understanding health systems' use of and need for evidence to inform decisionmaking. Agency for Healthcare Research and Quality (US); 2017 Oct. Report 17(18)-EHC035-EF. Rockville, MD PMID 29611913.
8. AHRQ Quality Improvement Working Group. Synthesizing Evidence for Quality Improvement. Agency for Healthcare Research and Quality. August 8, 2018. Available at: https://effectivehealthcare.ahrq.gov/topics/health-systems/quality-improvement. Accessed 6/10/2019.
9. Bravata DM, Sundaram V, Lewis R, et al. Asthma Care. Vol 5 of: Shojania KG, McDonald KM, Wachter RM, et al, editors. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies. Technical Review 9 (Prepared by the Stanford University-UCSF Evidence-based Practice Center under Contract No. 290-02-0017). AHRQ Publication No. 04(07)-0051-5. Rockville, MD: Agency for Healthcare Research and Quality. January 2007.
10. Ogrinc G, Davies L, Goodman D, et al. Standards for Quality Improvement Reporting Excellence 2.0: revised publication guidelines from a detailed consensus process. Journal of Surgical Research. 2016;200(2):676–682. doi: 10.1016/j.jss.2015.09.015.
11. Pinnock H, Barwick M, Carpenter CR, et al. Standards for Reporting Implementation Studies (StaRI): explanation and elaboration document. BMJ Open. 2017;7(4). http://dx.doi.org/10.1136/bmjopen-2016-013318.
12. Hoffmann T, Glasziou P, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348: g1687. doi: 10.1136/bmj.g1687.
13. Guise J-M, Butler ME, Chang C, et al. AHRQ series on complex intervention systematic reviews—paper 6: PRISMA-CI extension statement and checklist. Journal of Clinical Epidemiology. 2017;90:43–50. doi: 10.1016/j.jclinepi.2017.06.016.
14. Hempel S, Shekelle PG, Liu JL, et al. Development of the Quality Improvement Minimum Quality Criteria Set (QI-MQCS): a tool for critical appraisal of quality improvement intervention publications. BMJ Qual Saf 2015;24:796-804. doi: 10.1136/bmjqs-2014-003151.

Abbreviations and Acronyms

AHRQ	Agency for Healthcare Research and Quality
EPC	Evidence-based Practice Center
LHS	Learning Health System
QI-MQCS	Minimum Quality Criteria Set
TOO	Task Order Officer
SQUIRE	Standards for Quality Improvement Reporting Excellence
UConn	University of Connecticut

Appendix A. Final Extraction Instructions for Each Criterion

Criteria	Description of Instructions for "user's guide"
Population	**Population Instructions**
1. Who is Delivering the Intervention (e.g., provider types)?	For each individual intervention described in the manuscript under review, you will describe the specific type of person delivering the intervention (e.g. environmental services technician, diabetes educator), including specific skills, education or training required, as applicable. If unclear, authors may be contacted for extra information. Please add negative info according to the instructions (e.g., no info on skills, education, or training).
2. Who is Receiving the Intervention (e.g., patient types)	For each individual intervention described in the manuscript under review, you will describe the specific type of person receiving the intervention (e.g. advanced practice provider in an inpatient oncology setting, patient with NYHA Class IV heart failure in a general medicine clinic), including specific skills, education or training required, as applicable. Please clarify all individuals receiving intervention (e.g. physician, parents, and children).
3. Providers Demographics	For each individual intervention described in the manuscript under review, you will describe the demographics of those delivering the intervention, including age, gender, ethnicity, language and education characteristics as available. State first absence of info, then actual existing info.
4. Recipient Demographics	For each individual intervention described in the manuscript under review, you will describe the demographics of those receiving the intervention, including age, gender, ethnicity, language and education characteristics as available. State first absence of info, then actual existing info. Please be as specific as possible on characteristics.
Intervention	**Intervention Instructions**

5. Active vs. Passive Components	For each individual intervention described in the manuscript under review, you will determine whether it is active or passive. Active interventions would be selected if it requires ongoing expenditure of resources or time. Examples of active components would include: the health-system needs nurses to apply a special bandage to each patient, housekeeping needs to intensify their cleaning procedures, a physician's assistant needs to risk stratify each patient, and the continual use of consultants to implement the intervention. Passive components would be those that do not require ongoing expenditure of resources or time. Some examples of passive components would include: adding a warning or suggestion to the medical record, changing the labeling or signage for clarity, and adding or removing a drug from the formulary.
6. Discretionary vs. Mandatory Components	For each individual intervention described in the manuscript under review, you will determine whether it is discretionary or mandatory. Discretionary components are those that clinicians or support personnel can choose to undertake or not. Hospital guidelines are the best example of discretionary components although interventions described as voluntary or encouraged would also apply. Mandatory components would include things like hospital protocols and interventions that are stated to be compulsory. For example, a standard order set for all patients undergoing bypass surgery would be mandatory but a physician order entry alert with a suggested alternative therapy would be discretionary. We anticipate that when the intervention was randomized, the intervention is mandatory by definition.
7. Duration of the Intervention	For each individual intervention described in the manuscript under review, you will determine the duration of the intervention. There are two components for the duration of the intervention. The duration of each intervention and then the length of time this new intervention has been in place. For example, a health-system begins a new cleaning procedure for patient's rooms to reduce hospital-acquired infections 2 years ago. If it takes an additional 15 minutes per cleaning, that would be the duration of intervention. However, the length of time the intervention has been in place was 2 years. Duration of intervention should specify total hours/days of intervention. It comes from the multiplication of frequency and intensity. Length of intervention is the calendar time the intervention has been in place.

8. Frequency of Intervention	For each individual intervention described in the manuscript under review, you will determine how frequently it is administered or conducted. The creation of signage, an educational video or a formulary decision would be denoted as occurring once. A screening during admission would also occur once but an assessment occurring at the changing of every 8-hour nursing shift would occur three times daily. An intervention in a health-system whereby handwashing education occurs every 3 months would be classified as such, not every time anyone washes their hands before and after touching a patient. It is linked to the intervention, what the health-system is doing differently to cause a change in outcome.
9. Intensity of the Intervention	What is the intensity of the intervention (e.g., dose for drug, duration of exposure in each session for behavioral intervention)? Specify for each component.
10. Duration of the Effect of the Intervention	For most interventions, you will not be able to determine the duration of the effect of the intervention. For each individual intervention described in the manuscript under review, you will determine if you can judge the duration of the effect of the intervention or not. In some manuscripts authors will make a one-time intervention (educational program) and then determine how long it resulted in changed practice. The time from the roll out to when it ceased to change practice patterns would be what is entered here.
11. Team composition (People Delivering the Intervention)	For each individual intervention described in the manuscript under review, you will identify the team members who are actually delivering it. For example, the head of infectious disease and the CMO advocate to purchase new equipment to sterilize products and it is approved by the board of directors. The nurse requests the use of the equipment for an individual room and the custodian uses the device to sterilize the room and equipment. The nurse certifies the room is clean before allowing a new patient in the room. The “team” would only include the nurse and the custodian.
12. External Policies & Incentives Required	For each individual intervention described in the manuscript under review, you will determine whether a policy change or incentive was used. If policies were adopted (policy, guideline, protocol, order set, etc) to encourage compliance with an intervention, it would be recorded here. In addition, discretionary interventions that are incentivized by recognition programs, monetary incentives, or other valued items would be recorded here.
13. Required Skills/Training	This row is linked to the team composition row. For each member identified in the team delivering the intervention, what level of skill or training was required? For the above example, the skills would include someone with a nursing degree plus a 3-hour course on detecting infection risk in a room and validating its cleanliness. For the custodian, it is standard custodial training plus training on using the new piece of equipment. Here should go details on training/skills per team member.

14. Number & Description of Components	If the manuscript has 3 new interventions included in an arrhythmia mitigation strategy (baseline labs and mandatory correction of low K+, Mg++, use of balloon pumps before catecholamines for cardiogenic hypotension, aggressive us of beta-blockers for all peri-myocardial infarction patients) the total number of interventions would be recorded as 3 and then a description of each intervention would be made separately. Training, education, material, instruments that happen before randomization are part of the intervention and should be included.
15. A-priori Components versus Added Later/Final	Describe the number and nature of intervention components as planned *(a-priori)*. If planned and delivered components are identical, please note. If an intervention program has been going on for 2 years and three components were used initially and then another was added a year later, the first three would be a-priori and the newest would be added later. If an a-priori intervention was used and then abandoned during the followup, that should be identified in this section as well.
16. Theoretical Foundation	What is the theoretical foundation for the intervention? Does the report provide an explanation of the theoretical basis for the intervention? (yes/no) If yes, what is the main theoretical behavior change model or what are the main elements of the supporting theoretical framework? For each individual intervention, describe the rationale the investigators gave for believing it would work. For example, ultraviolet light kills bacteria so total room irradiation would reduce hospital acquired infections. This is usually limited to one to two sentences per intervention. Please first include absence of info, then give theoretical foundation per component of the intervention.
17. Which Interventions are Independent	Are any elements of the intervention administered independently of each other? (yes/no). If yes, which elements of the intervention are administered independently and thus could have independent effects on the outcomes of interest? Was the evaluation designed to determine whether elements of the intervention have independent effects?

18. Cost of Implementation	What is the total direct cost to the health system of implementing the intervention per interval of time (including cost of labor and other resources spent by the health system entity in implementing the intervention, but excluding the associated effect on costs incurred by patients participating in the intervention)? [Note: cost savings and cost-effectiveness should be considered with outcomes, rather than here in the description of the intervention.] For each intervention, if available, or for the total program, if that is all that is described, identify the monetary or personnel costs required to implement and sustain the intervention. This could include the time per patient to intervene, the cost of new equipment, the differential cost of one cleaning solution or one drug versus another, or the time it takes to develop and deliver an educational program. This item should include both monetary and personal costs separately. This item does not refer to study funding.
19. Leadership Commitment and Involvement	While most interventions are done by the clinicians themselves, new equipment requires the C-suite to appropriate the money. Protocols and guidelines need to be approved by the Pharmacy and Therapeutics and the Medical Executive Committee. Those leadership members need to be identified in this section.
20. Clinical Champion Involvement	Usually there are one or two people who are the drivers or overseers of the intervention. They are the ones doing the education, collecting the data, and being the point person for questions and concerns. Those people should be identified here by their job title and function in the health-system.
21. Physical Environmental Changes Required	For many interventions, no redesign of a building, lighting, surgical suite, or environmental controls (air conditioning or heating), are needed. If there are construction changes needed to have the intervention be feasible, that should be entered here.
22. Incentives	Did the intervention include any financial incentives for providers or patients to participate in the intervention? If yes, what were the incentives?
23. Implementation Strategies	In this section, we want a description of how the intervention moved from idea to being implemented in the health-system. Was there a working group who formulated possible solutions that were presented to the P&T committee or to the Quality Improvement Department. When the final list of interventions was agreed upon, what happened to alert the clinicians and to create the environment from which the intervention could begin. If there was a CQI process once it was launched that refined the interventions, that should be described here as well.

24. Types of Intervention Effects	Describe each effect of the intervention (e.g., hospitalization, visits to emergency department, number of days without symptoms, functional health status, cost-savings, cost-effectiveness)
Setting/Context	**Setting/Context Instructions**
25. Organizational Setting	Describe the structure of the organization implementing the intervention. Note if implementation was due to organizational change. If not specified in the manuscript, you might have to assess the website of the health system.
26. Financial Setting	Describe the payment/insurance structure, such as fee for service, capitation, Medicare/Medicaid, uninsured. (This may appear as a population descriptor as well.) Note if implementation was due to change in financial structures. In several settings, the financial structure is mixed. A Federally qualified health center (FQHC) may have Medicaid and uninsured, a hospital might have all of these.
27. Organizational Receptivity / Readiness	Describe the overall organizational state of receptivity (willingness) and the readiness (ability) to implement the intervention.
28. Population Needs/ Burden of Illness	Was the intervention modified to adapt to population differences, such as condition severity, multimorbidity, disabilities, or cultural factors? If yes, describe.
29. Geographic Location	Single State (if applicable), region if it is a regional health system (Northeast, Southwest), multiple regions for larger health systems (Southeast and Midwest), or the country(ies) and region(s) if international. It should also be described as urban, rural, suburban, or mixed based on the location(s) of the health system where the intervention is provided. If not specified in the manuscript, this information can be culled from a search of the internet for the health system in question.
30. External Factors	Factors affecting the system under which the QI intervention was implemented but that are external to the intervention itself. Ideally note how the external factors may impact the ability to transfer a QI strategy from one setting to another (i.e., may impact the effectiveness and generalizability of the intervention). For example, regulatory requirements or incentive systems.
31. Organizational History of Change	QI interventions or organizational changes that occurred prior to implementing the current QI protocol that may impact how the QI intervention should be evaluated and the ability to transfer a QI strategy from one setting to another (i.e., may impact the effectiveness and generalizability of the intervention).

32. Fidelity	Information on how closely implementation of the QI intervention adhered to the QI protocol; i.e., the treatment integrity. Information on adherence or compliance of providers, participants, organizations, and other players.
33. Intervention Adaptation	Components of the QI intervention that required personalization, individualization, tailoring, or adaptation. Particularly those that may impact the ability to transfer a QI strategy from one setting to another (i.e., may affect the effectiveness and generalizability of the intervention).

Appendix B. Applicability Table for Asthma Quality Improvement Projects

Item	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7
1. Who is Delivering the Intervention (e.g., provider types)?	Pediatricians from Ann Arbor and New York who responded to request of investigators, who had general pediatrics as the primary specialty; be licensed no earlier than 1960; provide direct patient care; if board specialized, be certified only in pediatrics; and be willing to take part in the interactive seminar if randomized to the treatment group.	Faculty. No info on specific skills, education or training.	Columbia physician in a tutorial, nurse educator. No info on specific skills, education or training.	Nurse with extensive asthma experience (Peer Leader education intervention); Trained asthma nurse (Planned care intervention).	Chief investigator, GP	Investigators and pediatricians of the NRCPM	Not described
2. Who is Receiving the Intervention (e.g., patient types)	Children 1-12 years of age; diagnosis of asthma made by a physician; no other chronic disorders with pulmonary com-plications; and at least one emergency medical visit for asthma in the previous year.	3-member multidisciplinary team (physician, a nurse, and a front office staff person); No info on specific skills or education. Asthma patients.	All staff from Bureau of Children Health , including clerks, public health assistants, lab technicians, physicians, nurses. Children with asthma.	Physician becoming an asthma champion; asthma patients (Peer leader education intervention); PLE recipients + family (patients (Planned care intervention).	GP, Children with moderate to severe asthma	Children and parents	Children with asthma, parents, doctors, pharmacists, community nurses, school teachers.

Item	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7
3. Providers Demographics	Male (60%). The age distribution of the clinician sample was as follows: 30± 39 (22%); 40±49 (37%); 50±59 (27%); and >=60 (14%). Study physicians were divided between solo (57%) and group practice (37%), with an additional 6% in multi-specialty practices. In their practices, 54% of the physicians spoke English only, two physicians (3%) Spanish only, and the rest spoke both English and another language (43%). A high proportion of solo practitioners and bilingual physicians were located in the New York City area.	No info on age, gender, ethnicity, language, or training of providers	No info on age, gender, ethnicity, language, or training of providers	No info on age, gender, ethnicity, language, or training of providers	No info on age, gender, ethnicity, language or training for both.	No info on age, gender, ethnicity, language, or training of providers	No info on age, gender, ethnicity, language, or training of providers

Item	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7
4. Recipient demographics	70% of the patients were males, 7% were aged <2 years, 59% aged 2±7 years, and 34% were 8±12 years of age at baseline. Sixty per cent of the parents were 30±39 years old, 75% were married, and ~90% had a high school education or above. Approximately 20% of the families had <=$20,000 annual income, 16% were at or below the poverty level (<=$15,000 a year), and 17% were on some form of government assistance for healthcare during the period of baseline data collection. 30% of families were nonwhite (Latino/Hispanic: 15%; African American: 15%).	No info on age, gender, ethnicity, language, or training of multidisciplinary team. Children: median age 8.5y, 65% male, 16% Medicaid insured, 50% white.	No info on age, gender, language, or training of recipients. Children were African American (45%) or Latino (34%); 36% were under Medicaid.	Only for asthma patients in both interventions. PLE: mean age 9.3y, 57% male, 58% White, no info on language. PCI: mean age 9.4y, 62% male, 69% White, no info on language.	GPs: 50-58% between 40 and 49y, 67% male; Children: 50% between 9 and 12y; 55% male. No info on ethnicity, language.	Only for children: mean age 9.4y, 64% male. No info on ethnicity, language. No info on parent's demographics.	Only for children: mean age 9.2y, 68% male; 54% low income (<AUS$12000/year). No info on ethnicity, language. No info on parents, doctors, pharmacists, community nurses, school teacher's demographics.
5. Active vs. Passive Components	Active	Active	Active	Active	Active	Active	Active
6. Discretionary vs. Mandatory Components	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	Attendance to sessions and workshops was discretionary, not a RCT. All teachers and community nurses attended educational sessions. 74% of families attended sessions; 21% of pharmacists, and 20% doctors attended workshops.

Item	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7
7. Duration of the Intervention	Duration of intervention: two face-to- face group meetings lasting ~2.5 h each. Length of time of intervention: 2 years.	Duration of intervention: Three 1-day learning sessions. Length of intervention: 12 months.	Duration of intervention: First component (Five 3-h sessions over 5 months, plus 2 additional 3-h sessions at end of 1st year); second component (3-h in first year); third component (no time specified for monthly visits up to 2y follow-up). Total for first and second component was 24h; unknown for third component. Length of intervention: 2 years.	Duration of intervention: PLE 4-6h of workshops for physicians. PCI: 65min per PAC visit (total 260min to 325min), telephone sessions of 20min (unspecified number). Length of intervention: 2 years for both interventions	Duration of intervention: by Chief investigator: unknown, by GP: unknown; length of intervention: 12 months	Duration of intervention: 4 sessions of 1 to 1.25h, one every week. Length of intervention: 1 year	Children and family: 2 sessions of two hours each, one week apart. One session for teachers and community nurses (no time described). No details of other sessions/workshops. Length of intervention: 6 months.
8. Frequency of Intervention	Two face-to-face group meetings held over a 2-3 week period after randomization. No frequency of pediatrician evaluations described.	First session in 2nd month (February), second and third sessions in the last 10 months of a 12-month period.	First component (Five 3-h sessions over 5 months, plus 2 additional 3-h sessions at end of 1st year); second component (3-h in first year); third component (no time specified for monthly visits up to 2y follow-up).	PLE: 2 workshops for training of physician; PCI: 4 to 5 PAC visits during 2 years.	One session + supplementary session for chief investigator; 3-4 visits for GP	Once a week for 4 weeks	Children and family: two sessions one week apart. Teachers and community nurses: one session. No details of other sessions/workshops.
9.Intensity of the intervention	Intervention only done twice at the beginning of study.	Each session was done during one day, although specific hours were not provided	3-h sessions for first and second component; unspecified for third component	PLE: 2-3h per workshops for physicians; PCI: full day training sessions, weekly or biweekly 1h conference call meeting with one of PIs. 4-5 planned asthma visits and telephone follow up during the 2 years. PAC visits averaged 65 minutes (including physician visit) and telephone sessions averaged 20 min.	Unspecified for both chief investigator and GP.	1-1.25h per session.	Children and family: 2hours per session. No details for other groups.
10. Duration of the effect of the intervention	2 years	12 months	2 years	2 years	1 year	1 year	6 months

Item	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7
11. Team composition (People Delivering the Intervention)	Educators, pediatricians	Faculty, without details	Columbia physician in a tutorial, nurse educators for second and third components, respectively. No info about who is delivering the first component.	PLE: Nurse with extensive asthma experience + physician; PCI: Asthma nurse + physician.	Chief investigator, GP	Investigators and pediatricians, without further details.	Not described
12. External Policies & Incentives Required	No external policies or incentives required.	No external policies or incentives required.	No external policies or incentives required.	No external policies or incentives required.	Guidelines endorsing proactive asthma care have been promoted in Australia in the form of the National Asthma Council's six step plan. The six step plan has been adapted for the fee for service environment of Australian general practice with the introduction of the 3+ visit plan (3+ plan). Components of the six step plan are scheduled over three or more general practice visits, with two or more of these visits occurring outside the treatment of an acute exacerbation.	No external policies or incentives required	No external policies or incentives required
13. Required Skills/Training	General pediatrics as primary specialty; be licensed no earlier than 1960; provide direct patient care; if board specialized, be certified only in pediatrics; and be willing to take part in the interactive seminar if randomized to the treatment group.	Not described	Not described	Not described	Not described	Not described	Not described

14. Number & Description of Components	There were two seminar components: optimal clinical practice based on National Asthma Education and Prevention Program (NAEPP) expert panel guidelines, and patient teaching and communication. The training: 1) used interactive methods; and 2) focused on help- ing physicians to create conversation between themselves and the patients to promote partnership by building the following: 1) deriving information for making therapeutic decisions; 2) creating a congenial and supportive atmosphere so patients would be candid; 3) reinforcing positive efforts of families to self-manage; 4) providing a supportive climate for mutual problem-solving; 5) strengthening patients skills in using medicines; 6) providing the patient with a view of the long-term therapeutic plan; and 7) building patients confidence at controlling symptoms.	One intervention: In January 2001, intervention practices were asked to collect baseline data to identify "performance gaps" (the difference between current and desirable performance) in their practice. At the first learning session in February, teams were taught a comprehensive method to proactively care for patients with asthma using the Chronic Care Model and concepts of quality improvement including the Model for Improvement (a specific approach to quality improvement that emphasizes the use of small, incremental tests of change). They were provided materials and information based on the guidelines from the National Asthma Education and Prevention Program, Bethesda, MD, and tools to support implementation of these practices (such as encounter forms and an electronic patient registry). During the next 10 months, coaching and support was provided through 2 additional learning sessions, biweekly conference calls, an active e-mail list, and periodic performance feedback	First component was education in 7 3-h sessions (5 in first 5 months, 2 at end of first year); Second component was tutorial sessions with during one 3-h session during the first year; Third component was problem solving and continuing education by nurse educator monthly up to 2 years.	PLE: Education component only by nurse to physician. It included training in pharmacotherapy, physician behavior change strategy, help from ongoing learning network for peer leaders, reception of materials, summaries, articles with recommendations, and laminated pocket cards, kit of patient education materials. PCI: PLE plus planned asthma care (PAC) visits, training in EpR2 and self-management support techniques, weekly or biweekly calls with PI; each PAC included standardized evaluation of patients, support and participation in care planning, and self-management support to families. Also, in between PAC visits, telephone follow-up.	Chief investigator: academic session to all GPs. GP: Administration of a system of structured asthma care in three or more GP visits, with participating families reminded to attend the general practitioner (3+ care). We classified a 3+ plan as fully complete if over the 12 month trial period the child attended their general practitioner for at least two asthma related visits, one of which was proactive, and if each of the two visits showed at least one different 3+ plan content item. We classified 3+ plans as partially complete if during the same period the child attended their general practitioner for at least two asthma related visits, one of which was proactive. The 3+ plan does not demand an active recall system.	The education program was based on the teaching manuals 'Open Airways' and 'Air Power', developed for educating children and parents. Lectures and handouts were translated from English into Russian, rewritten and adapted to Russian culture and knowledge. Each program consisted of four sessions. The contents of seven sessions of the 'Open Airways' program were condensed to four so that the families could attend all the sessions. During the education session's instructions on the treatment of asthma based on the US 'Guidelines for the Diagnosis and Management of Asthma' were given also.	Information about asthma mechanisms, treatment, management plan, and inhalation technique to children and parent. Information about management plan to doctors. Information about appropriate asthma management to pharmacists. Education about asthma to community nurses and school teachers. Those not attending received education materials by mail.

Item	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7
		based on expert review of monthly project team reports.					
15. A-priori Components versus Added Later/Final	Planned and delivered components were the same.	Planned and delivered components were the same.	Planned and delivered components were the same.	Planned and delivered components were the same.	Planned and delivered components were the same.	Planned and delivered components were the same.	Planned and delivered components were the same.
16. Theoretical Foundation	The interactive seminar was based on the theory of self- regulation [6, 7], guiding physicians to examine their own behavior and to identify ways that they could develop a better partnership with their patients. Physicians were re- trained to observe, evaluate, and react to their own efforts to treat and educate their patients.	Continuous quality improvement, an improvement approach adapted from industry to the healthcare setting.	Theory-based approaches to planned organizational change; learning-centered teaching to 1. help staff link the goals of continuing care for asthma to preventive care; 2. help staff solve organizational problems, and 3. build teamwork and sense of owning the program.	PLE: Not described; PCI: Chronic model care, a multiple intervention strategy.	Proactive care (regular review), in conjunction with written asthma action plans and training in self-management, improves outcomes for adults with asthma. No similar evidence base exists for children with asthma.	Not provided	Not provided
17. Which Interventions are Independent	None	None	None	None	None	None	None
18. Cost of Implementation	No info on costs of providing intervention. The direct cost of providing the interactive seminar to physicians is approximately $150 per physician or approximately $795,000 if given to all the pediatric primary care providers in the state.	No information on costs of implementation	No information on costs of implementation	Annual cost of PLE $591; Annual cost of PCI: $1591, per patient.	No information on costs of implementation	No information on costs of implementation	No information on costs of implementation
19. Leadership Commitment and Involvement	None	None	None	None	None	None	None
20. Clinical Champion Involvement	None	None	None	PLE: Physician was the clinical champion; PCI: PLE was also a component of this intervention	None	None	None

Item	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7
21. Physical Environmental Changes Required	None	None	None	None	None	None	None
22. Incentives	None	None	None	None	None	None	None
23. Implementation Strategies	No description	No description	No description	No description	No description	Translation of materials into Russian	None
24. Types of Intervention Effects	Physician behavior, parents' view of physician behavior, patients' healthcare use.	Primary outcomes: any written asthma management plan reported as received by a parent in the past 12 months, daily use of inhaled steroids in the past 4 weeks, and daily use of controller medications in the past 4 weeks. Secondary measures: any asthma hospitalization and any emergency department visit for asthma in the past 12 months, any asthma attack in the past 12 months, and parent report of how limited their child had been from asthma in the previous 2 weeks from very strenuous activities (such as running fast or playing hard).	Patient visits, rate of patient identification per 1000 clinic patients, rate of returning patients per 1000 clinic patients, proportion of patients treated for asthma in 1 year who returned for treatment the next year, use of new medications and delivery devices, degree of health education by physician and nurses.	Asthma symptom days, period of frequent symptoms, persistent asthma, asthma-specific functional status (CHSA tool), use of controller medication.	Process measures: asthma related visits to GP, proactive GP consultations, reactive GP consultations, have written action plan, complete 3+ plan. Clinical measures: emergency department visits, did not missing school days due to wheezing or asthma, >=4 wheezing episodes, severe wheezing, normal activities restricted by respiratory symptoms, sports restricted by respiratory symptoms, use of preventers, relievers, nebulizer.	Over a 10-month period after the educational classes or 1 year after the baseline period in terms of the parents' and children's asthma self-management steps and combined asthma inventory scores, utilization of medical care for asthma (out-patient and hospital visits for asthma care), functional limitations (school or pre-school days off), family's compliance with medical recommendations (medications used for asthma care), children's peak expiratory flow rates, including symptom score and medications used at the time of peak flow measurements.	FEV1, bronchial responsiveness, air flow meter variability, symptom frequency, knowledge of asthma, days absent from school, unscheduled visits to doctors or emergency room.

Item	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7
25. Organizational Setting	74 outpatient pediatric practices in Ann Arbor and New York (no details provided)	43 practices from 2 states: Massachusetts practices were a mix of hospital clinics, independent community health centers, and private practices. In Michigan, all sites were a part of the Henry Ford Health System; 1 was hospital-based.	22 pediatric Health's Bureau of Children Health clinics in NY.	42 primary care pediatric practices affiliated with 4 managed care organizations (MCOs) in 3 states: Washington, Illinois, Massachusetts	24 GP practices in northern region of the Australian capital territory.	4 public health clinics in Moscow	Unknown number of Schools from 2 areas of Sydney, Australia
26. Financial Setting	Not specified in papers	Massachusetts: Medicaid (10%). No info about Michigan.	Not specified in papers	All from private insurers (WA: Group Health Cooperative, IL: Prudential Health Plan, MA: Harvard Pilgrim Health Care and Blue Cross Blue Shield).	Not specified in paper	Not specified in paper	Not specified in paper
27. Organizational Receptivity / Readiness	No description	No description	No description	No description	No description	No description	No description
28. Population Needs/ Burden of Illness	None	None	None	None	None	None	None
29. Geographic Location	Pediatric practices in two cities of two US states	Two states: Massachusetts and Michigan, urban.	One state: New York	3 US states: Washington, Illinois, Massachusetts.	Northern region of the Australian capital territory	Only clinic in Kalininsky district and from three clinics of Cheremushkinsky district, Russia	Sydney, Australia
30. External Factors	No regulatory requirements, no incentive systems	No regulatory requirements, no incentive systems	No regulatory requirements, no incentive systems	No regulatory requirements, no incentive systems	No regulatory requirements, no incentive systems	No regulatory requirements, no incentive systems	No regulatory requirements, no incentive systems
31. Organizational History of Change	No description	No description	No description	No description	No description	No description	No description
32. Fidelity	No description	No description	No description	No description	No description	No description	No description
33. Intervention Adaptation	No adaptation of interventions	No adaptation of interventions	No adaptation of interventions	No adaptation of interventions	No adaptation of interventions	No adaptation of interventions	No adaptation of interventions

Abbreviations: Aus = Australia; CHSA = Children's Health Service tool for Asthma; FEV1 = forced expiratory volume in one second; GP = general practitioner; H = hour; MCO = Managed Care Organization; PCI = planned care intervention; PLE = peer leader education; RCT = randomized controlled trial; US = United States; WK = week; Y = year.