Health Information Exchange Evaluation Toolkit

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract Number: 290-04-0016

Prepared by:

Caitlin M. Cusack, M.D., M.P.H.
Center for IT Leadership

Eric G. Poon, M.D., M.P.H.
Brigham and Women's Hospital

AHRQ Publication No. 08-0026-EF
November 2007

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov
HEALTH IT

Suggested Citation

Cusack CM, Poon EG. Health Information Exchange Evaluation Toolkit. Prepared for the AHRQ National Resource Center for Health Information Technology under contract No. 290-04-0016. AHRQ Publication No. 08-0026-EF. Rockville, MD: Agency for Healthcare Research and Quality. November 2007.

The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Acknowledgments

The authors would like to thank numerous members of the AHRQ National Resource Center's Value and Evaluation Team for their invaluable input and feedback: Davis Bu, M.D., M.A. (Center for IT Leadership), Karen Cheung, M.P.H. (National Opinion Resource Center), Dan Gaylin, M.P.A. (National Opinion Resource Center), Julie McGowan, Ph.D. (Indiana University School of Medicine), Adil Moiduddin, M.P.P. (National Opinion Resource Center), Anita Samarth (eHealth Initiative), Jan Walker, R.N., M.B.A. (Center for IT Leadership), and Atif Zafar, M.D. (Indiana University School of Medicine). Thank you also to Mary Darby, Burness Communications, for editorial review.

Contents

Introduction 1

Section I: Developing an Evaluation Plan 2
- I. Brief Project Description 2
- II. Project Goals 2
- III. Evaluation Goals 3
- IV. Choose Evaluation Metrics 3
- V. Search for Other Easily Measured Metrics 4
- VI. Consider Project Impacts on Potential Metrics 5
- VII. Consider Ongoing Evaluation of Barriers, Facilitators, and Lessons Learned 6
- VIII. Grade Your Chosen Metrics 7
- IX. Determine Which Measurements Are Feasible 7
- X. Determine Your Sample Size 8
- XI. Rank Your Choices on Both Importance and Feasibility 9
- XII. Choose the Metrics You Want to Evaluate 9
- XIII. Draft Your Plan around Each Metric 9
- XIV. Consider Your Evaluation Budget 11
- XV. Write Your Evaluation Plan 12

Section II: Examples of Measures 13
- Table 1: Data Exchange between Outpatient Providers and Laboratories 15
- Table 2: Data Exchange between Outpatient Providers and Pharmacies 18
- Table 3: Data Exchange between Providers 22
- Table 4: Data Exchange between Outpatient Providers and Radiology Centers 24
- Table 5: Data Exchange between Outpatient Providers and Public Health Departments 29

Section III: Examples of Process and Outcomes Measures 31
- Table 1: Clinical Outcomes Measures 31
- Table 2: Clinical Process Measures 33
- Table 3: Provider Adoption and Attitudes Measures 36
- Table 4: Patient Knowledge and Attitudes Measures 38

Section IV: Example 39

Appendixes

Appendix A 41

Introduction

We are pleased to present the AHRQ National Resource Center (NRC) Health Information Exchange Evaluation Toolkit targeted towards health data exchange projects. The intent of the toolkit is to help your team work its way through the process of creating an evaluation plan for this type of Healthcare Information Technology (Health IT) project.

Data exchange projects are relatively new in the world of Health IT and thus, there is a dearth of research data about them. The project your team is carrying out represents an important step in the national effort to use electronic exchange of health care information to improve patient safety, quality, effectiveness and efficiency of care. Since data exchange projects are so new and their impact on safety and quality remains to be fully defined, it is critical for your project to include an evaluation component. Evaluation serves multiple important purposes. First, a continuous evaluation process serves to guide the project itself, as the thoughtful examination of impact will allow your project to fine-tune your approach to data exchange, and may even allow you to elucidate the unintended consequences of electronic data exchange. Second, by carefully documenting the barriers encountered and the lessons learned, others will be able to understand how to best approach their own data exchange projects in the future. In our experience, evaluation efforts have the best chance of fulfilling their promise when they are planned for during the early phases of the project.

This toolkit has been developed to help guide you through the process of devising a realistic and achievable evaluation plan. Section I walks you and your team step by step through the process of determining the goals of your project, what is important to your stakeholders, what needs to be measured to satisfy stakeholders, what is truly feasible to measure, and how to measure these items.

Sections II and III includes lists of measures that may be used to evaluate your project. Each table in these lists includes possible measures, suggested data sources for each measure, cost considerations, potential pitfalls, and general notes. While these tables distill the various experiences of members of the National Resource Center, they should not be considered exhaustive, as there may be many opportunities to explore and learn from various aspects of your data exchange projects. At the same time, you should not pick these measures without carefully considering whether each measure will help you to answer an important question for your stakeholders or whether you have the resources to use the measure. The final section contains an example of a project and measures which could be used in an evaluation of that project.

We invite and encourage feedback on the content, organization and usefulness of this toolkit as it continues to be expanded and developed. If you have any comments or questions about the evaluation toolkit or the AHRQ National Resource Center, please do not hesitate to contact NRC-HealthIT@ahrq.hhs.gov.

SECTION I:
DEVELOPING AN EVALUATION PLAN

I. BRIEF PROJECT DESCRIPTION

This may come straight out of your project plan or proposal.

_____
_____
_____
_____
_____
_____

II. PROJECT GOALS

What is it that you hope to gain from this implementation? What are the goals and expectations of your stakeholders around this project? (Clinicians, laboratories, pharmacies, C-level individuals and so forth). What would need to happen for the project to be deemed a success by your stakeholders? In thinking about your stakeholders, consider the entity which is responsible for the project, the structure of that entity and its governance. Are the goals being proposed in alignment with this entity?

> *Example:*
>
> *To improve the quality of care provided to patients by successfully exchanging laboratory data between providers and laboratories.*

_____
_____
_____
_____
_____
_____

III. EVALUATION GOALS

Who is your audience for your evaluation? Do you intend to prepare a report for your stakeholders? If you have received an AHRQ contract, do you intend to prepare a report for AHRQ in order to fulfill the requirements of your contract? Will you use the evaluation to convince late adopters of the value of your implementation? To share lessons learned? To demonstrate the project's return on investment? Or are your goals more external? Would you like to share your experiences with a wider audience and publish your findings? If you plan to publish your findings, that might affect your approach to your evaluation. In addition, look to your funding source, be it from your stakeholders, a grant or a contract. Are there required goals within this funding vehicle that must be met?

Example:

Goal: To prepare a report for our stakeholders, AHRQ, and other groups considering undertaking a data exchange project.

_____
_____
_____
_____
_____
_____

IV. CHOOSE EVALUATION METRICS

Take a good look at your project goals. What needs to be measured in order to demonstrate that the project has met those goals? Brainstorm with your team on everything that could *be measured, without regard to feasibility. These can be around whether or not the ground work for the project has been successfully completed, such as developing a governance structure, coming to a consensus on how to handle privacy and security issues or developing a sustainability model. Perhaps you want to track whether or not the project was able to come up with a minimum data set to share, and the rate at which that data was able to be shared. You can also consider looking at categories organized around type of data exchange such as:*

- *Outpatient providers and laboratories*
- *Outpatient providers and pharmacies*
- *Between provider and other providers*
- *Between outpatient providers and radiology centers*
- *Between outpatient providers and public health departments.*

Your team might find it helpful to break down the measures in a similar fashion. Whatever you choose to evaluate, metrics should map back to your original goals for the project, and as indicated by the examples may be either quantitative or qualitative. In addition, for those projects which are past the implementation phase, you may want to look at evaluating outcomes and process measures, such as:

- *Clinical Outcomes Measure*
- *Clinical Processes Measures*
- *Provider Adoption and Attitudes Measures*
- *Patient Knowledge and Attitudes Measures*
- *Workflow Impact Measures*
- *Financial Impact Measures*

Section II provides a wide range of these potential metrics to give your team ideas about the kinds of metrics they can be looking to evaluate.

Example:

Goal: to successfully exchange laboratory data between providers and laboratories. Possible measures: track progress of completing the architecture necessary to exchange laboratory data, track progress of the actual exchange of data, collect usage statistics.

_____
_____
_____
_____
_____
_____

V. SEARCH FOR OTHER EASILY MEASURED METRICS

Clinicians, laboratory services, pharmacies, hospitals and other such groups collect a tremendous amount of data for multiple purposes: to satisfy various federal and state requirements, to conduct ongoing quality assurance evaluations, to measure patient and staff satisfaction, etc. There are therefore likely teams within your participant groups that are already collecting data that might be useful to you. Reach out to these groups to learn what information they are currently collecting, and determine whether those data can be used as an evaluation metric.

In addition, contact the various groups you are working with to learn the reporting capabilities of their current software programs. There may be opportunities to leverage those reporting capabilities for your evaluation. For example, do your participant labs already track phone

calls from clinicians looking for results? Are the participant pharmacies already evaluating customer satisfaction? Could your evaluation team piggy-back with another group to abstract a bit of additional information? Are there useful measurements that could be taken from existing reports? Likewise, you may find that activities you are planning as part of your evaluation would be helpful to groups within your participants. Cooperation in these activities can increase goodwill on both sides.

Example:

The regions participating pharmacies are contacted and inquires are made regarding reports that are being carried out on a regular basis. It is discovered that the pharmacies actively track calls they make to physicians to clarify information on prescriptions. It is hypothesized that the ability to electronically exchange data regarding patient medications will decrease these calls. Adding this metric to the evaluation plan is easy and helps to measure whether or not the regional project is having a measurable impact.

VI. CONSIDER PROJECT IMPACTS ON POTENTIAL METRICS

Consider the potential metrics on your list and whether and how your project might impact those metrics. Would your implementation truly impact these metrics? You may find that this exercise eliminates some metrics from your list because they will not, in truth, be impacted by your project. In considering the impact of a project, think about where the project will be implemented and what stakeholders it is going to affect directly.

VII. CONSIDER ONGOING EVALUATION OF BARRIERS, FACILITATORS, AND LESSONS LEARNED

Lessons learned are important measures of your project, and typically are captured using qualitative techniques. These lessons may reflect the barriers and facilitators you encountered at various phases of your project. Barriers may include organizational barriers, technology barriers, security/privacy barriers, financial barriers, legal barriers and so forth.

In addition to tracking barriers, also track what steps were taken to overcome those barriers. For example, strong leadership, being impartial across the participants, good training, support in the early stages of implementation, and obtaining buy in from your target community, may serve as important facilitators to your efforts. This type of information is extremely valuable not only to you but also to others undertaking similar projects. Other lessons learned of great interest to others, would be approaches to determining governance, legal, organizational, consumer and technical issues. In formulating a plan for capturing this information, consider scheduling regular meetings with your project team to discuss the issues at hand openly, and to record these discussions.

If there are personnel assigned to support the early implementation stages, they may build a Communications Bridge that will facilitate early feedback on any issues raised so that they might be addressed. Also, the observer may suggest changes to the metrics to better capture the intended data. Moving beyond such discussions, you could conduct focus groups. For example, you could ask physicians who are using data exchange about what has gone well, what has gone poorly, and what the unexpected consequences of the project have been. Consider how you could incorporate these qualitative analysis techniques into your evaluation plan. Clearly state what you want to learn, how you plan to collect the necessary data, and how you would analyze the data.

Example of a 'lesson learned':

You observe early on in the project that the electronic exchange of test orders between ambulatory practices and commercial labs was consistently missing important milestones. You therefore decide to evaluate the barriers involved and try to understand (or even suggest!) ways to overcome these barriers. You set out by conducting semi-structured interviews with the stakeholders involved in the delay. You may discover that several laboratories were concerned about the loss of control and the disruption of existing workflow patterns if they started accepting orders generated by different EMR vendors. You report this finding to the main project team and decide to ask the state medical society to convene a meeting for the major EMR vendors and commercial labs so that the two parties can better understand each other's requirements. This approach was a success and the project began meeting its milestones. A lesson learned was thus to convene the appropriate stakeholders early in the design process so that each stakeholder does not feel threatened by the others.

_____
_____
_____
_____
_____

VIII. GRADE YOUR CHOSEN METRICS IN ORDER OF IMPORTANCE TO YOUR STAKEHOLDERS

Now that your team has a list of metrics to measure, grade each metric in order of importance to your stakeholders, i.e., Clinicians, laboratories, pharmacies, C-level individuals and so forth. You could use a scale such as: 1 = Very Important, 2 = Moderately Important, 3 = Not Important. This will help you begin to filter out those metrics that are interesting to you but will not provide you with information of interest to your stakeholders. Another approach to determining importance of metrics may be to consider your contract requirements. For instance if you are required to be exchanging a given percentage of data by a particular date, this may rise to the top as a 'very important' metric to be measuring.

1. Very Important:_____
   _____
2. Moderately Important:_____
   _____
3. Not Important:_____
   _____

Determining which measurements to use for your evaluation may be difficult for your team. Data exchange projects typically have a variety of stakeholders, across many types of facilities, all with seemingly different goals and priorities. It is best to recognize this up front, and maintain your impartiality as best as you can. If necessary you can bring the players to the table and together determine what is most important to the project as a whole.

IX. DETERMINE WHICH MEASUREMENTS ARE FEASIBLE

Now examine your list to determine which metrics are feasible for you to measure. Be realistic about the resources available to you. Teams frequently are forced to abandon evaluation projects that are labor-intensive and expensive. Instead, focus on what is achievable and on what needs to be measured to determine whether your implementation has met its goals. For example, you might want to know whether your implementation reduces adverse drug events (ADEs). That's a terrific evaluation project, but if you have neither the money nor the individuals needed for chart abstraction, the project will likely fail. Keep your eye on what can

be achieved. Again, you can use a ranking scale: 1 = Feasible, 2 = Feasible with Moderate Effort, 3 = Not Feasible.

1. Feasible:_____
_____

2. Moderate Effort :_____
_____

3. Not Feasible:_____
_____

X. DETERMINE YOUR NEEDED SAMPLE SIZE

The feasibility of an evaluation plan often hinges on the minimal sample size you need for your quantitative measures. In a typical evaluation project, you may be interested in evaluating whether your project has impacted a quantitative metric of interest. In general, if the metric tries to capture rare events, you will need to make many observations in order to observe a sufficient number of events to draw meaningful conclusions. Also, if the impact of the project is small, then you will need to make more observations in order to say with confidence that any measured impact is truly due to the project itself and not random noise. Needless to say, observations cost money, and you may find that some metrics are out of reach given the resources you have at your disposal. Appendix A offers a hypothetical example.

Estimate the number of observations you will need for each metric. You may find this exercise eliminates further metrics from being feasible.

_____
_____
_____
_____
_____
_____

XI. RANK YOUR CHOICES ON BOTH IMPORTANCE AND FEASIBILITY

Place your remaining metrics into the appropriate box in the grid below.

		Feasibility Scale		
		1-*Feasible*	**2-*Moderate Effort***	**3-*Not Feasible***
Importance Scale	**1-*Very important***	**(1)**	**(2)**	
	2-*Moderately important*	**(3)**	**(4)**	
	3-*Not important*	**(5)**		

Those metrics that fall within the green zone (Most important, Most Feasible) are ones you should definitely undertake; the yellow zones are ones you can undertake in the order listed; those in the red zone should be avoided.

XII. CHOOSE THE METRICS YOU WANT TO EVALUATE

You now have a list of metrics ranked by importance and feasibility. Narrow that list down to four or five primary metrics. If you want to measure other metrics and you believe that you will have the required resources available to you, list those as secondary metrics.

_____

_____

_____

_____

_____

XIII. DRAFT YOUR PLAN AROUND EACH METRIC

Map out how you will measure each metric. What is the timeframe for your study? What is your comparison group? If you are doing a quantitative study, what statistical analysis will you use? Having a statistician review you plan at this point may save you time later in your evaluation. If you plan to deploy a survey as part of your evaluation, you may want to conduct a small pilot to save yourself from getting into trouble later as well. In crafting your specific plan around each metric, we suggest that you use the following template to help you flush out the details.

Measure	1st measure	2nd measure	3rd measure	4th measure, etc.
Briefly describe the intervention.				
Describe the expected impact of the intervention and how you think your project will exert this impact.				
What questions do you want to ask to evaluate this impact? These will likely reflect the expected impact (either positive or negative) of your intervention.				
What will you measure in order to answer your questions?				
How will you make your measurements?				
How will you design your study? For a quantitative study, you might consider what comparison group you will use. For a qualitative study, you might consider whether you will make observations or interview users.				
For quantitative measurements only: What types of statistical analysis will you perform on your measurements?				
Estimate the number of observations you need to make in order to demonstrate that the metric has changed statistically.				
How would the answers to your questions change future decision-making and/or implementation?				

What is the planned timeframe for your project?				
Who will take the lead for the project? For data collection? Data analysis? Presentation of the findings? Final write-up?				
Estimate the cost for evaluating the metrics. Take into consideration planning, meetings, travel, analysis, consult time with a statistician and time to prepare your final report on your findings.				

XIV. CONSIDER YOUR EVALUATION BUDGET

Having mapped out the metrics you intend to measure, take another look at the costs involved in evaluating these metrics. Are there metrics which will put your budget at risk? Are there ways to reduce the costs of these measurements? If it is clear that you can not meet your budget with your planned metrics, have your team work through the importance and feasibility matrix a second time. Are some metrics too expensive and therefore drop in your team's estimation as to whether or not they are feasible? Are some metrics expensive, but so important as to cause you to drop several of the less important metrics in order to afford the more expensive metrics? The team must come up with an evaluation plan which is financially feasible that lies within your planned budget. Your plan should have some discussion around budget justification indicating that you have taken costs into consideration.

XV. WRITE YOUR EVALUATION PLAN

You now have everything you need to write your evaluation plan: project description, goals, metrics, and methodology for your evaluation. We suggest you follow the following structure:

I. *Short Description of the Project*

II. *Goals of the Project*

III. *Questions to be Answered by the Evaluation Effort*

IV. *First Measure to be Evaluated — Quantitative*

   a. *Overview – General Considerations*

   b. *Timeframe*

   c. *Study Design/Comparison Group*

   d. *Data Collection Plan*

   e. *Analysis Plan*

   f. *Power/Sample Size Calculations*

V. *Second Measure to be Evaluated – Qualitative*

   a. *Overview – General Considerations*

   b. *Timeframe*

   c. *Study Design*

   d. *Data Collection Plan*

   e. *Analysis Plan*

VI. *Subsequent Measures to be Evaluated in Same Format*

VII. *Budget Justification*

VIII. *Conclusion*

SECTION II: EXAMPLES OF MEASURES

Section II and Section III includes lists of measures that may be used to evaluate your project. Each table in these lists includes possible measures, suggested data sources for each measure, cost considerations, potential pitfalls, and general notes. While these tables distill the various experiences of members of the National Resource Center, they should not be considered exhaustive, as there may be many opportunities to explore and learn from many aspects of your data exchange projects. At the same time, you should not pick these measures without carefully considering whether each measure will help you to answer an important question for your stakeholders or whether you have the resources to use the measure.

If you are working under an AHRQ contract for a SRD project AHRQ has asked that the following be considered as measurements.

Measure Domains Stipulated in Contract

Quality & Safety	Organizational Efficiency & Effectiveness	Financial
• Advances in care processes • Improved patient outcomes • Better monitoring of diseases and other health risks • Reduced medication errors	• Work and quality improvement processes • Communication among individuals, groups, and organizations • Satisfaction of needs and expectations of patients, providers, and other stakeholders • Organizational risk mitigation Reduced ordering of redundant laboratory and radiology examinations	• Cost reductions • Revenue enhancement • Productivity gains • Cost savings resulting from redundant test ordering • Greater use of lower cost medications

In order to meet these items AHRQ suggests that each SRD measure at the minimum the following:

	Measure	**Contract**
1	Measure the volume of discrete clinical data elements moved	Organizational Efficiency & Effectiveness
2	Measure usage: the number of data elements that were available versus how many data elements were viewed by clinicians	Organizational Efficiency & Effectiveness
3	Measure usage: the number of patients for which data was available versus the number of patients for which data was viewed by clinicians	Organizational Efficiency & Effectiveness
4	Measure timeliness: the time from which data was generated to when that data was able to be viewed	Quality & Safety
5	Measure costs: choose a measure to evaluate costs	Financial
6	Measure satisfaction: conduct a satisfaction survey	Quality & Safety; Organizational Efficiency & Effectiveness

Table 1: Data Exchange between Outpatient Providers and Laboratories

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Was electronic ordering of laboratory tests between outpatient providers and laboratories achieved?	Implementation team	Low as data should be readily available		
Are providers using	Usage statistics.	Low-IT team should be able to readily collect this data	There are several different ways you might want to measure this. First would be the number of discrete providers using the system as the numerator and the number of total providers as the denominator. A second approach might be how frequently individual providers are accessing the system with hit rates as the numerator and an individual provider as the denominator. A third approach might be to look at hit rates divided by total number of providers to get an overall average rate. Providers might be defined as nurses and/or physicians. Tracking this information over time would give an interesting view of your project. Can also track the number of paper transactions still be used: i.e.: clinical staff putting labs into records.	Finding baseline provider rates might be difficult. I.E.-what is your pool of physicians who could be using the system? You could consider getting this information from local medical societies, FOLIOS, and Boards of Medicine.
What percentage of laboratory orders is sent electronically?	Usage statistics	Medium if it requires counting paper orders	Denominator = all orders Numerator = electronic orders Can do this both on the laboratory	

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
			side and on the provider side	
Reduction in calls to providers to clarify an order	Call logs	Low		
Costs to send orders to lab	Pre and post implementation check of logs; time motion/workflow analysis in a sample of various settings	Medium depending on whether or not these statistics have been tracked	Estimate first what these costs are (labor costs to prepare forms, costs to send forms) and multiply by the number of orders sent out. Compare paper and electronic methods using time motion studies: how much time spent looking for results, writing orders, transcribing, etc.	
Impact on duplicate laboratory tests	Pre and post implementation chart reviews	High due to chart review.	If you are rolling out your project in stages you could consider using those who haven't come on line yet as your control group. In this manner you could collect your data without needing to do a chart review retrospectively. May be able to use billing data to help focus the search for redundant tests.	Need to define 'duplicate' as this would be different for a CA-125 versus a Hct, and also different if the initial test were normal versus abnormal.
Was electronic exchange of laboratory results between outpatient providers and laboratories achieved?	Implementation team	Low as data should be readily available		
Impact on the number of results calls to the lab	Laboratory call logs	Low	A reduction in the number of calls necessary to the laboratory for results implies that providers are able to find their results in a more timely fashion. This combined	These measurements need to be adjusted for the volume of labs done by each of the participating labs so that one can compare the data in a meaningful manner. Also

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
			with the loss of errors which occur in transmitting results orally should lead to a reduction in errors.	need to record whether or not there were significant changes in market share, or significant problems in running the labs themselves (e.g. if a machine broke down resulting in the a particular test not being able to be run for a period of time)
Decrease in time to report critical results	Call logs pre and post-implementation	Low as long as these statistics have been kept	A great measure to consider given the interest that JCAHO has in this topic	
Costs avoided to receive results	Logs; time motion/workflow analysis	Medium depending on whether or not these statistics have been tracked	Could estimate costs associated with receiving results (labor to open mail, sort, distribute to clinicians, and post on patient chart) and multiply by number of laboratory results received.	If the users are still printing out electronic results to put them in paper charts this cost must be considered as well
Laboratory costs avoided to send results	Logs	Low	Look at costs traditionally used to prepare mailings and send out results	
Impact on the satisfaction of clinicians	Survey: their perception of usability, how easy it was for them to learn to use the system, do they feel more/less efficient as a result of the data exchange	Medium	You might consider sampling both your users as well as those who could be involved in the project but who have chosen not to participate. Going to statewide/region wide MD databases from local medical societies, FOLIOS, board of registrations, and so forth might be one way to determine your target survey group. Consider questions such as asking them how often they were able to find the result they were looking for in a timely manner. Could compare responses before and after (early/late) in implementation	

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Satisfaction of laboratory personnel	Survey	Medium	Your survey could sample the laboratory technicians, or the administrative personnel including those who are responsible for taking phone calls. The survey would need to be designed to be distributed to all involved laboratories	Be careful to survey personnel affected by data exchange. It maybe invisible to some staff.
How much data was able to be exchanged?	Implementation team	Low as data should be readily available	Look at the number of discrete HL-7/OBX elements that were exchanged	

Table 2: Data Exchange between Outpatient Providers and Pharmacies

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Was electronic exchange of information about medication orders and prescriptions between outpatient providers and pharmacies achieved?	Implementation team	Low as data should be readily available		

Are providers using?	Usage statistics.	Low-IT team should be able to readily collect this data	Could collect this information electronically. Alternatively could look at the number of electronic prescriptions received as the numerator and the total number of prescriptions received as the denominator. A second approach would be to look at the number of physicians submitting prescriptions electronically as the numerator divided by the total number of users of the system. The third would be using the number of physicians submitting prescriptions electronically as the numerator and the total number of physicians in the catchment area.	
How much data was able to be exchanged?	Implementation team	Low as data should be readily available	Look at the number of scripts done, the number of eRX messages sent	
Impact on calls to pharmacies	Logs	Low		
Impact on calls to providers to clarify a prescription	Logs	Low		
Impact on calls to patients to clarify their information	Logs	Low		
Impact on costs due to improved formulary compliance	IT team or Chart reviews	Low to High	If the new system has decision support the system may have the data to show how often a switch is made from a non-formulary name choice to a formulary alternative	Could be difficult to find the pre-implementation compliance rate

Impact on costs by switching to generics	IT team or Chart reviews	Low to High	If the new system has decision support the system may have the data to show how often a switch is made from a brand name choice to a generic alternative. Evaluating formulary and brand to generic patterns may be more feasible if you focus on a single drug class or narrow down to a subset of patients.	
Impact on adverse drug events	Chart reviews	High-chart reviews are labor and resource intensive		This can be very difficult to measure and might be a measure best avoided. The teams must come together to decide what constitutes an ADE and how it is going to be measured. ADEs are relatively rare and it takes many chart reviews to be confident about the results.
Clinician Satisfaction	Survey	Medium	You might consider sampling both your users as well as those who could be involved in the project but who have chosen not to participate. Going to statewide/region wide MD databases from local medical societies, FOLIOS, board of registrations, and so forth might be one way to determine your target survey group.	

Pharmacist satisfaction	Survey	Medium	Your survey could sample the pharmacists, the technicians, or the administrative personnel including those who are responsible for taking phone calls. The survey would need to be designed to be distributed to all involved pharmacies	
Patient satisfaction	Survey	Medium	Could include surveys with prescriptions.	

Table 3: Data Exchange between Providers

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Was electronic exchange of information between providers achieved?	Implementation team	Low as data should be readily available		
Are providers using?	Usage statistics; surveys	Medium	Need to consider how you define providers exchanging information with other providers. Would you define it as email communication? Or does it need to be something more? The ability to send referrals electronically? The ability to electronically send a chart of a patient for a referral?	
How much data was able to be exchanged?	Implementation team	Low as data should be readily available	Look at the number of discrete HL-7/OBX elements that were exchanged	
How much of the total health data was exchanged electronically versus other methods such as by fax, mail and courier?	Implementation team, logs	Medium to high		The measurement of the amount of data being exchanged by non-electronic means might be difficult to determine.
Impact on costs of chart pull	Logs; time/motion analysis; chart reviews	Medium	Estimate the labor cost of a pull and multiply by number of referrals in a given time period. Could also review a sample of charts to determine the % of consultant notes that are captured electronically for a sample of patients.	

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Impact on costs of duplicating paper charts	Logs; time/motion analysis	Medium	Estimate cost of duplicating-finding the chart, copying the chart, preparing for mailing and mailing times the number of charts duplicated.	
Impact on inter-provider calls requesting results	Logs	Low if this information is tracked		Suspect this type of information has not been tracked which would make this difficult to measure
Impact on costs for referral letters (time to write, sending)	Logs	Medium	Estimate labor cost to review chart, dictate referral letter, transcribe letter, mail letter and multiply by number of referrals.	
Satisfaction of providers	Survey	Medium	You might consider sampling both your users as well as those who could be involved in the project but who have chosen not to participate. Going to statewide/region wide MD databases from local medical societies, FOLIOS, board of registrations, and so forth might be one way to determine your target survey group.	

Table 4: Data Exchange between Outpatient Providers and Radiology Centers

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Was electronic ordering of radiology tests between outpatient providers and radiology centers achieved?	Implementation team	Low as data should be readily available		
Was electronic exchange of radiology results between outpatient providers and radiology centers achieved?	Implementation team	Low as data should be readily available		
How much data was able to be exchanged?	Implementation team	Low as data should be readily available	Look at the number of discrete HL-7/OBX elements that were exchanged. Look at the number of DICOM images that were exchanged	
Are providers using?	Usage statistics	Low-IT team should be able to readily collect this data	There are several different ways you might want to measure this. First would be the number of discrete providers using the system as the numerator and the number of total providers as the denominator. A second approach might be how frequently individual providers are accessing the system with hit rates as the numerator and an individual provider as the denominator. A third approach might be to look at hit rates divided by total number of providers to get an overall average rate. Providers might be defined as	Finding baseline provider rates might be difficult. I.E.-what is your pool of physicians who could be using the system You could consider getting this information from local medical societies, FOLIOS, and Boards of Medicine.

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
			nurses and/or physicians. Tracking this information over time would give an interesting view of your project.	
Impact on duplicate radiology tests	Pre and post implementation chart reviews	High due to chart review	If you are rolling out your project in stages you could consider using those who haven't come on line yet as your control group. In this manner you could collect your data without needing to do a chart review retrospectively.	
Impact on costs to send orders (provider)	Pre and post implementation check of logs; time motion/workflow analysis	Medium depending on whether or not these statistics have been tracked	Estimate the labor costs needed to prepare forms, and send them out; multiply by the number of orders sent out	
Impact on costs to receive orders (radiology)	Pre and post implementation check of logs; time motion/workflow analysis	Medium depending on whether or not these statistics have been tracked	Estimate the costs to open forms, and process those forms; multiply by the number of orders sent out	
Impact on results requests from providers	Phone logs	low	A reduction in the number of calls necessary to the radiology center for results implies that providers are able to find their results in a more timely fashion. This combined with the loss of errors which occur in transmitting results orally should lead to a reduction in errors.	These measurements need to be adjusted for the volume of exams done by each center so that one can compare the data in a meaningful manner.
Impact on calls to providers to clarify an order	Phone logs	low		
Impact on time to report critical results	Call logs pre and post-implementation	Low as long as these statistics have been kept	Again, a great measure to consider given the interest that JCAHO has in this topic	

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Satisfaction of radiology personnel	Survey	Medium	Your survey could sample the radiologists, the radiology technicians and/or the administrative personnel including those who are responsible for taking phone calls. The survey would need to be designed to be distributed to all involved radiology centers	
Satisfaction of clinicians	Survey	Medium	You might consider sampling both your users as well as those who could be involved in the project but who have chosen not to participate. Going to statewide/region wide MD databases from local medical societies, FOLIOS, board of registrations, and so forth might be one way to determine your target survey group.	
PACs				
Impact on film costs	Finance tracking (balance sheet, receipts etc), pre and post-implementation	Low		
Impact on chemical costs	Finance tracking (balance sheet, receipts etc), pre and post-implementation	Low		
Impact on file room costs	Labor costs, overtime costs, pre and post-implementation	Low		
Impact on duplication of films for referrals	Logs	Low		
Impact on costs to receive films for	Pre and post implementation check of	Medium depending on whether or not these statistics have	Determine labor costs to open films, distribute to provider, collect films from provider, package for	

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
review (provider)	logs	been tracked	radiology, and return to radiology-then multiply by number of films received. A group may or may not track films received-again our group had a process for this so it wouldn't be hard to know how many films we received a year	
Impact on costs to send films (radiology)	Pre and post implementation check of logs; time motion/workflow analysis in a sample of various	Medium depending on whether or not these statistics have been tracked	Determine labor costs to receive request, copy film, package film, and mail film -then multiply by number of requests received.	
Impact on costs to re-file films received after having sent films out	Pre and post implementation check of logs; time motion/workflow analysis in a sample of various	Medium depending on whether or not these statistics have been tracked	Determine labor costs to receive returned film and re-file and multiply by number received	
Scheduling/workflow				
Impact on images performed due to more efficient scheduling	Pre and post review of schedules	Medium due to labor intensity to review schedules	On line ordering/scheduling leads to increased efficiencies and a resultant increase in the number of tests that can be done. Tests can be more easily grouped by type, and fewer errors are made in resource scheduling	
Impact on time to schedule appointments	Time/motion studies	Medium	This can be done on the provider side doing the scheduling or the receiving side scheduling	
Impact on lost films	Logs	Low	The post-PACs loss rate should be close to zero	
Impact on cancelled exams due to better prep (online instructions available	Pre and post review of schedules	Medium due to labor intensity to review schedules		Groups may or may not have this information in their schedules depending on whether or not they are tracking cancellation reasons

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
to scheduler) and avoidance of contraindications (iodine allergy known at time of scheduling)				

Table 5: Data Exchange between Outpatient Providers and Public Health Departments

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Was electronic exchange of public health information between providers and public health departments achieved?	Implementation team	Low as data should be readily available		
How much data was able to be exchanged?	Implementation team	Low as data should be readily available	Look at the number of discrete HL-7/OBX elements that were exchanged	
Impact on costs to prepare reports manually	Reports prepared; time motion analysis	Medium	Labor costs to find information, prepare report multiplied by the number of reports prepared	
Impact on costs to send paper reports	Reports prepared; time/motion analysis	Medium	Cost to send reports multiplied by the number of reports prepared	
Impact on costs to receive reports (public health)	Logs; time/motion analysis	Medium	Estimate the costs in receiving a report, opening reports multiplied by volume received	
Impact on costs to process paper reports	Logs; time/motion analysis	Medium	Estimate costs in processing a report multiplied by the volume received	
Impact on reportable diseases reported	Logs	Low		
Impact on time to report events	Report review	High	Pre and post implementation sample-track time interval from date of event to time logged into public health database	
Impact on time to detection of an adverse event	Report review pre and post implementation	High	Pre and post implementation review of reports of adverse events or outbreaks to determine if there has been an improvement in the	

Measure	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
			early detection of these events	
Satisfaction of clinicians	Survey	Medium	You might consider sampling both your users as well as those who could be involved in the project but who have chosen not to participate. Going to statewide/region wide MD databases from local medical societies, FOLIOS, board of registrations, and so forth might be one way to determine your target survey group.	
Public health personnel satisfaction	Survey	Medium	Your survey could sample the clinicians, public health practitioners, or the administrative personnel including those who are responsible for collating paper reports. The survey would need to be designed to be distributed to all involved public health departments	

SECTION III: EXAMPLES OF PROCESS AND OUTCOMES MEASURES

For those of you further along with your data exchange process, you may want to look at some measures around care processes and patient outcomes affected by your data exchange. We have included this set of metrics to give you ideas around what can be measured in the areas of: clinical outcomes measures, clinical process measures, provider adoption and attitudes measures, patient knowledge and attitude measures, workflow impact measures, and financial impact measures We understand that many of these measures are expensive to measure, and you should tailor your evaluation plans according to the needs of your stakeholders and the resources at your disposal.

Table 1: Clinical Outcomes Measures

Measure	Quality Domain(s)	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Preventable adverse drug events (ADEs)	• Patient Safety	• Chart review • Prescription review • Direct observations • May also consider patient phone interviews	Very high: events are rare and likely need clinicians to perform reviews.	Errors can be divided by stage of medication use: • Ordering • Transcribing • Dispensing • Administering • Monitoring Can be assessed in both inpatient and outpatient settings.	• Preventable ADEs are relatively rare. • Will need to collect large amount of data to show statistical differences.
Inpatient mortality	• Patient Safety • Effectiveness	• Medical records • Billing data	Medium: (especially if risk adjustment tools are not readily available)		• Need to risk-adjust. • May be very difficult to find statistically significant differences in mortality rates, since death rates tend to be relatively low.
Hospital complication rates	• Patient Safety	• Some can be obtained from ICD-9 codes, although chart review (at least for a sample of charts) is preferable.	Low: if data are already being collected. Medium: if chart review is needed.	Common targets: • Nosocomial infections • PE/DVT • Falls • Pressure ulcers • Catheter-related infections	• Watch out for documentation effect (e.g., falls may become more reliably documented because the measure makes it easier to document falls).

Measure	Quality Domain(s)	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
		• Some measures may already be collected for external reporting purposes.		• Post-op infections • Operative organ/vessel/nerve injury • Post-op MI • Post-op respiratory distress • Post-op shock • Pneumothorax	
Length of stay	• Patient Safety • Efficiency	• Medical records • Billing data	Low: if data are already being collected.		• Need to adjust for disease severity and diagnosis. • Watch out for secular trend, (e.g., financial pressures to discharge patients early, other concurrent QI programs, etc.)
Readmission rates after discharge	• Patient Safety • Effectiveness • Efficiency • Patient-Centeredness	• Medical records • Billing data	Low	7 days, 30 days	• Need to adjust for changes in patient/diagnosis mix over time.
Inpatient admission rates/ED visits for populations with chronic diseases	• Patient Safety • Effectiveness • Efficiency • Patient-Centeredness	• Medical records • Billing data • Patient registries	Low: if patient registries exist.	Common targets: • CHF • Asthma • DM • ESRD • CAD	• Watch out for secular trend (e.g., change in admission criteria).

Table 2: Clinical Process Measures

Measure	Quality	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
• Potential adverse drug events ("near misses") • Medication errors	• Patient Safety	• Chart review • Prescription review • Direct observations • May also consider patient phone interviews	High: since events will likely need chart review by clinicians. However, cost is lower than for ADEs, since these events are more common.	Errors can be divided by stage of medication use: • Ordering • Transcribing • Dispensing • Administering • Monitoring Can be assessed in both inpatient and outpatient settings.	Chart reviews do not capture all errors (especially dispensing and administration errors). Also, chart reviews probably need to be backed up with patient interviews in the outpatient setting, as documentation of adverse events in the ambulatory setting typically is not very reliable.
Number of pharmacist interventions per medication order	• Patient Safety • Efficiency	• Pharmacy intervention logs	Low: if data are already being collected.		Might change threshold for pharmacy intervention
Number of orders ordered verbally	• Patient Safety	• Medical records • Pharmacy records	Low: if medical records department or pharmacy already collect data.		Might be impacted by local policies
Time to complete co-signature of verbal orders	• Patient Safety • Efficiency	• Medical records	Low: if medical records department already collects data		Check reliability of time measurements on paper records.
Chronic disease management targets	• Effectiveness • Patient-Centeredness	• Electronic data repository (if available), chart reviews.	Low: if data are captured reliably in data repository. Medium to High: if chart reviews are needed.	• DM: A1c within goals, LDL within goals, annual foot exam, annual nephropathy screening, annual ophthalmologic exam • HTN: Percent of patients controlled, medication use within guidelines • Depression: appropriate monitoring after starting	Check for documentation effect of measure (e.g., smoking cessation might be better documented than before even though it is not more commonly performed). Also, check for inaccuracies in problem and/or medication lists.

Measure	Quality Domain(s)	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
				SSRI • ESRD/Chronic kidney diseases: Care consistent with K-DOQI guidelines • CAD: Aspirin use, beta-blocker use, smoking cessation counseling • CHF: ACE inhibitor use, appropriate beta-blocker use • Asthma: smoking cessation counseling • Childhood ADHD • Childhood obesity	
Health maintenance target		• HEDIS measures, electronic data repository (if available), chart reviews.	Low: if data are captured reliably in data repository or by health plans. Medium to High: if chart reviews needed.	• Immunizations (adult and childhood) • Cancer screening (mammogram, Pap smears, etc.) • Counseling (e.g., smoking cessation)	Watch out for documentation effect of measure. Billing data may be more resistant to this effect.
Appropriate Actions/usage: • Percent of alerts or reminders that resulted in desired plan/action • Percent of tests ordered inappropriately (for target tests) • Percent of blood products used appropriately	• Patient Safety • Effectiveness	• Electronic data repository • usage logs	Low: if data captured electronically, although additional resources may be needed to handle the control group. Higher: if control group evaluation requires chart review.	Best to let the alerts trigger equally for both the intervention and control groups, and then prevent the alerts from being displayed to control group users. That would easily track opportunities to carry out the desired action equally between the intervention and control groups.	Need to assess and monitor quality of data used to trigger the alerts and reminders.

Measure	Quality	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Documentation of key clinical data elements	• Patient Safety	• Likely will need chart reviews for paper-records group.	Medium	Examples include: • Allergy on admission • Follow-up plan on discharge • Care plan for next phase of care • Complete pre- and post-admission med list Should also assess clinician perception of data quality.	

Table 3: Provider Adoption and Attitudes Measures

Measure	Quality Domain(s)	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Percent of orders entered by physicians on CPOE	• Patient Safety	• CPOE usage logs • Pharmacy logs	Low		
Frequency of order set use	• Efficiency • Patient Safety • Effectiveness	• CPOE usage logs	Low	Would be helpful to present data in context of how many times order sets could have been used in the same period (e.g. number of patients admitted with CHF).	
Percent of outpatient prescriptions generated electronically	• Patient Safety • Effectiveness	• EMR usage logs	Medium		Getting the denominator may require chart review.
Percent of notes online	• Patient Safety	• EMR usage logs	Medium		Getting the denominator may require chart review.
Percent of practices or patient units that have gone paperless	• Efficiency	• EMR usage logs • Training logs	Low		Likely a gradual progress that takes many months, if not years.
Percent of physicians and nurses who have undergone training for target IT intervention	• N/A	• Training logs	Low	Indirect measure	Some experts believe that classroom training is not the ideal form of training for physicians.
Use of help desk	• N/A	• Help desk logs	Low		May be confounded by quality of up-front training, continued support, usability of application.

Measure	Quality	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Time to resolution of reported problems	• N/A	• Help desk logs	Low		May be confounded by nature of reported problems/
Provider satisfaction towards specific interventions	• N/A	Satisfaction surveys and interviews: • Ease of use • Usefulness • Impact on quality and time savings • Suggestions for improvement	Low for surveys, higher for interviews.		Difficult to achieve good response rates from physicians.
Provider satisfaction towards own job	• N/A	• Direct surveys (human resources may administer already)	Low		Many potential confounders.
Turnover of staff	• N/A	• Human resources log	Low		Many potential confounders.

Note: May be helpful to correlate patient clinical outcomes with adoption of measure, either at the physician or practice unit level. Need to collect baseline data for comparison.

Table 4: Patient Knowledge and Attitudes Measures

Measure	Quality Domain(s)	Data Source(s)	Relative Cost	Notes	Potential Pitfalls
Patient knowledge	• Patient-Centeredness	Patient surveys and interviews	Medium	• Knowledge of own medications (regimen, indications, potential side effects), other prescribed care • Knowledge of own health maintenance schedules • Knowledge of own medical history • Knowledge of own family's medical history	Important to do iterative cognitive testing/piloting of surveys developed internally. Methodologies leading to good survey response rates may be expensive. On-line surveys might lower cost, but may bias results because on-line patients may be different from the general population. May be able to add customized questions to standard surveys such as CAHPS.
Patient attitudes	• Patient-Centeredness	• Patient surveys • Patient interviews • Focus groups and other qualitative methodologies	Medium	• Comfort level • Barriers and facilitators for use	
Patient satisfaction	• Patient-Centeredness	External surveys (CAHPS, commercial)	Low to Medium		
		Internally developed survey	Medium		

SECTION IV: EXAMPLE

Briefly describe the intervention.	**Our project is to allow for the exchange of laboratory data from commercial labs to providers via the web**			
	1	**2**	**3**	**4**
Describe the expected impact of the intervention and briefly describe how you think your project will exert this impact.	Laboratory data will be able to be exchanged.	Laboratory data will be exchanged in a timely fashion	Providers will use the system to review their patients laboratory results	Providers will perceive benefit from the data exchange project
What questions do you want to ask to evaluate this impact? These will likely reflect the expected impact (either positive or negative) of your intervention.	How much data was moved? How many elements were available? How many elements did people look at?	How much time elapsed between the time of lab result generation at the laboratory and the time when the result can be viewed by a provider?	What percentage of the clinicians in the catchment area participate in the project?	How satisfied are the clinicians with the system? How does the system affect their ability to deliver care? Do clinicians spend less time tracking data down on their patients or more time?
What will you measure in order to answer your questions?	Examine number of HL-7 (OBX) elements exchanged	Look at time-date stamps of the data throughout the implementation	Look at usage statistics: how often to clinicians access the system? What is the number of patients for which data was used?	Satisfaction surveys

How will you make your measurements?		Review time-stamps for different result types generated by different laboratories for different types of providers.	Denominator = number of clinicians in the catchment area Numerator = number of discrete clinicians accessing the system Denominator = number of patients in the catchment area with results captured by the data exchange network. Numerator = number of patients for whom data was accessed	Develop clinician satisfaction survey. Administer pre-implementation, then 6 and 12 months post-implementation
How will you design your study? What comparison group will you use?	Will not use comparison group as we started from zero exchange of data-will look at trends over time	Monitor this time throughout the implementation process		Pre-implementation versus post-implementation comparison
For quantitative measures only: What types of statistical analysis will you to perform on your measurements?	Graph on-going trends	Graph ongoing trends	Graph trends over time, for different provider types at different locations	Graph trends. T-test comparison for satisfaction levels (analyzed as continuous variable) across different time points
How would the answers to your questions change future decision–making and/or implementation?	Look at what was done to bring the system from zero exchange up to 100% exchange	Pinpoint trouble spots in the data exchange network and use the data to drive improvement.	If clinicians were not using the system would want to consider how to increase that participation. Might interview clinicians to see what the barriers are to usage	Want to understand how the ability to better locate data on a patient impacts professional satisfaction.

Appendix A

Following is a simple, hypothetical example to illustrate the importance of sample size:

Before implementation of an e-prescribing tool in the outpatient setting, 5 prescribing errors per 100 prescriptions written are noted. After implementation of the e-prescribing tool, the rate drops to only 2.5 errors per 100 prescriptions. If you select 100 prescriptions at random for review both before and after the implementation of e-prescribing, you might observe the following:

	BEFORE	AFTER
Number of Errors in 100 sampled prescriptions	5	3
Observed Error Rate	5%	3%

Would you feel confident concluding that the error rate actually fell? Most people would answer "no". Statistics show us that repeated samples of 100 would reveal slightly different rates. Since the number of observed events (prescription errors) is so small, the errors may have shown up in the sampled prescriptions by chance. If you are particularly unlucky, chance may lead you to observe three fewer errors in the review of the 100 prescriptions before implementation of e-prescribing, creating the appearance that e-prescribing was causing *errors rather than preventing them.*

The picture changes, however, if you could afford to examine 100,000 prescriptions before and after implementation of the e-prescribing system. Instead, you might observe:

	BEFORE	AFTER
Number of Errors in 100,000 Sampled Prescriptions	4,932	2,592
Observed Error Rate	4.9%	2.6%

Looking at the observed data now, would you feel more confident that the drop in the error rate is real and not due to a random phenomenon? Most people would say "yes". Even if, by chance, the observed data are a few errors off from the "true" error rate, you still would conclude that the prescribing error rate was very different after implementation of e-prescribing.

The actual number of observations required in this example (i.e., the minimal sample size), falls somewhere between 100 and 100,000. To determine the exact number required, you need to do a "sample size calculation". A full discussion of sample size calculations is beyond the scope of

this toolkit, but resources are readily available to you to help you carry out a sample size calculation. Statistics textbooks cover this topic when they discuss statistical power. Many free tools are available on the Internet and may be found through a simple search. You may consult a statistician, either locally or through the AHRQ National Resource Center; or you may use one of the many software programs available to do these calculations.

No matter how you perform the sample size calculation, it is important to do it before you embark on an evaluation. Many evaluation projects have failed after the investigators found that insufficient data were collected to show a statistically significant difference. A sample size calculation can be a sobering experience: You may learn that your team cannot answer the desired question because the required sample size is too large. In that case, you may need to address a question that is less interesting but feasible to answer.