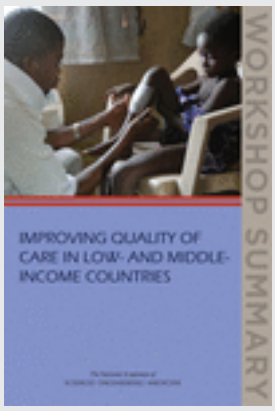

Improving Quality of Care in Low- and Middle-Income Countries: Workshop Summary

DETAILS

120 pages | 6 x 9 | PAPERBACK
ISBN 978-0-309-37342-5 | DOI: 10.17226/21736

AUTHORS

Gillian J. Buckley and Rachel E. Pittluck, Rapporteurs; Board on Global Health; Institute of Medicine; The National Academies of Sciences, Engineering, and Medicine

IMPROVING QUALITY OF CARE IN LOW- AND MIDDLE-INCOME COUNTRIES

WORKSHOP SUMMARY

Gillian J. Buckley and Rachel E. Pittluck, *Rapporteurs*

Board on Global Health

Institute of Medicine

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

This activity was supported by Contract/Grant No. APC-GM-0002 between the National Academy of Sciences and JSI Research and Training Institute/U.S. Agency for International Development. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

International Standard Book Number-13: 978-0-309-37342-5
International Standard Book Number-10: 0-309-37342-5

Additional copies of this workshop summary are available for sale from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Printed in the United States of America

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2015. *Improving quality of care in low- and middle-income countries: Workshop summary.* Washington, DC: The National Academies Press.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Ralph J. Cicerone is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.national-academies.org**.

PLANNING COMMITTEE FOR THE WORKSHOP ON IMPROVING QUALITY OF CARE IN LOW- AND MIDDLE-INCOME COUNTRIES[1]

SHEILA LEATHERMAN (*Chair*), Research Professor of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina at Chapel Hill
RIFAT ATUN, Professor of Global Health Systems; Head, Global Health Systems Cluster, Harvard T.H. Chan School of Public Health
KEDAR MATE, Senior Vice President for Innovation, Institute for Healthcare Improvement; Assistant Professor of Medicine, Weill Cornell Medical College

IOM Staff

GILLIAN J. BUCKLEY, Program Officer
RACHEL E. PITTLUCK, Research Associate
DAVID T. GARRISON, Senior Program Assistant (*from June 2015*)
FAYE HILLMAN, Financial Associate
PATRICK W. KELLEY, Director, Board on Global Health

[1] Institute of Medicine planning committees are solely responsible for organizing the workshop, identifying topics, and choosing speakers. The responsibility for the published workshop summary rests with the workshop rapporteurs and the institution.

Reviewers

This workshop summary has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published workshop summary as sound as possible and to ensure that the workshop summary meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the process. We wish to thank the following individuals for their review of this workshop summary:

A. Mushtaque R. Chowdhury, BRAC, Bangladesh; Columbia University Mailman School of Public Health
Corine Karema, Ministry of Health, Rwanda
Nachiket Mor, CARE India; Reserve Bank of India; CRISIL
Edward B. Perrin, University of Washington
Alexander Rowe, Centers for Disease Control and Prevention

Although the reviewers listed above have provided many constructive comments and suggestions, they did not see the final draft of the workshop summary before its release. The review of this workshop summary was overseen by **Robert S. Lawrence,** Professor of Environmental Health Sciences, Health Policy, and International Health, and Director, Center for a Livable Future, Johns Hopkins Bloomberg School of Public Health. He was responsible for making certain that an independent examination

of this workshop summary was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this workshop summary rests entirely with the rapporteurs and the institution.

Acknowledgments

This report is the product of the cooperation and contributions of numerous people. The planning committee and project staff are especially grateful to the presenters and moderators who participated in this workshop, lending their time and expertise to make the event a success. These individuals are listed in the workshop agenda in Appendix A. We would also like to thank the many participants who attended the workshop, both in person and via the live webcast.

Many staff of the Institute of Medicine and the National Academies of Sciences, Engineering, and Medicine have supported this activity from its inception to the workshop itself and through the publication of this workshop summary report. In particular, we wish to acknowledge the efforts of Daniel Bearss, Marton Cavani, Chelsea Frakes, Greta Gorman, Faye Hillman, Sarah Kelley, Ellen Kimmel, Rebecca Morgan, and Jose Portillo.

The planning committee and staff thank the U.S. Agency for International Development (USAID) for generously funding this activity.

Contents

Boxes, Figures, and Tables

BOXES

FIGURES

TABLES

Acronyms and Abbreviations

AIDS	acquired immune deficiency syndrome
AMTSL	active management of the third stage of labor
ART	antiretroviral therapy
ASSIST	Applying Science to Strengthen and Improve Systems
CDC	Centers for Disease Control and Prevention
COPE®	client-oriented, provider-efficient services
DALY	disability-adjusted life year
DHS	Demographic and Health Survey
Global Fund	Global Fund to Fight AIDS, Tuberculosis and Malaria
HIV	human immunodeficiency virus
IOM	Institute of Medicine
IQR	interquartile range
MES	median effect size(s)
MESI	Monitoring, Evaluation and Surveillance Interface
OECD	Organisation for Economic Co-operation and Development

PEPFAR	The President's Emergency Plan for AIDS Relief
SBM-R	Standards-Based Management and Recognition
URC	University Research Co., LLC
USAID	U.S. Agency for International Development
WHO	World Health Organization

1

Introduction[1]

A 2013 study on the global burden of medical error found that unsafe care causes 43 million injuries a year and the loss of 23 million disability-adjusted life years (DALYs), about two-thirds of them in low- and middle-income countries (Jha et al., 2013). By these calculations, adverse events, if they were a disease, would be the fifth leading cause of DALYs lost worldwide (Jha et al., 2013). Sobering though they are, such figures are likely underestimates, as the study included problems resulting from only seven common adverse events in inpatient hospitalization, which people in poor countries access at far lower rates than in rich ones. Furthermore, the data that inform these estimates come largely from medical records systems, which are inadequate in most low- and middle-income countries. Although the true scope of unsafe hospital care remains difficult to measure, the burden is clearly highest in the parts of the world with the least means to correct it (Adhikari, 2013).

There is reason to suspect quality problems with outpatient services as well. Studies employing standardized patient actors in India (both urban Delhi and rural Madhya Pradesh) found that only 4 percent of patients receive a correct diagnosis; 67 percent receive no diagnosis at all. When the researchers calculated the probability that the patient received treatment

[1] The planning committee's role was limited to planning the workshop, and the workshop summary has been prepared by the workshop rapporteurs as a factual summary of what occurred at the workshop. Statements, recommendations, and opinions expressed are those of individual presenters and participants, and are not necessarily endorsed or verified by the Institute of Medicine, and they should not be construed as reflecting any group consensus.

that does no harm, they found it ranged from .25 among diarrhea patients to .61 for pre-eclampsia. To put it another way, 75 percent of Indian children presenting at a clinic with diarrhea will receive a treatment that hurts them (Das et al., 2012).

Deficits in the quality of clinical care are only a small piece of the problem. The 2001 Institute of Medicine (IOM) report *Crossing the Quality Chasm* took a view of quality that incorporated efficiency, timeliness, patient-centeredness, and equity as important dimensions of care (IOM, 2001). This view of quality requires attention to human resources, infrastructure, communication, medicines safety, and logistics. Problems in these areas are no less detrimental to patient safety, but are harder to measure and rarely captured in morbidity and mortality data. In low- and middle-income countries, concern with access to services sometimes overshadows interest in the standard of the services provided (Berendes et al., 2011; Das, 2011).

The momentum for universal health coverage has underscored the problem of poor quality care in low- and middle-income countries. Universal coverage aims to make essential health services "of sufficient quality to be effective" available at a cost that does not expose the user to risk of financial hardship (WHO, 2012). It is an important piece of the post-2015 development agenda,[2] and many countries are making the free provision of a basic package of essential services a priority (IOM, 2014). As governments and donors spend more on health, they have greater concern that the services they pay for are safe and effective and, therefore, have more reason to invest in quality improvement.

Quality of care is a priority for the U.S. Agency for International Development (USAID). The agency's missions abroad and their host country partners work in quality improvement, but a lack of evidence about the best ways to facilitate such improvements has constrained their informed selection of interventions. In the absence of guidelines on how to invest, mission staff navigate an opaque market of quality improvement strategies. Six different methods—accreditation; client-oriented, provider-efficient services (COPE®); improvement collaborative; Standards-Based Management and Recognition (SBM-R); supervision; and clinical in-service training—currently make up the majority of this investment for USAID missions (see Box 1-1). The agency's Bureau for Global Health estimates that these six methods account for about 80 percent of the missions' spending on quality of care. As their already substantial investment in quality grows, there is demand for more scientific evidence on how to reliably improve quality of care in poor countries.

[2] The post-2015 development agenda is the set of targets for international development that will replace the Millennium Development Goals.

BOX 1-1
Six Quality Improvement Strategies Commonly Used in USAID Missions

Accreditation

Accreditation is an external quality evaluation through which an accrediting organization formally recognizes that an institution meets certain standards. Accreditation is usually a voluntary process, though some countries require accreditation for market entry or eligibility for government payment. The accrediting organization uses consensus standards to evaluate institutions, but there is wide variability in the evaluation process. Some accreditations rely largely on self-assessment; others require several weeks of on-site inspections from a team of accreditors.

Clinical In-Service Training

Clinical in-service training is a broad category of quality improvement strategies, including all training for health professionals who have already completed their formal credentialing process. In-service training is meant to either reinforce important concepts and practices or to introduce new knowledge about how a health professional should work.

COPE® (client-oriented, provider-efficient services)

COPE® is EngenderHealth's proprietary quality improvement method; it was designed for quality improvement in family planning and is now also used in maternal, child, and reproductive health. The strategy uses group problem solving and self-assessment to identify problems and set priorities for quality improvement. COPE® starts with an orientation for managers at the worksite, followed by a self-assessment where participants identify and rank their main problems. Facilitators help determine the root causes of these problems and develop action plans to fix them; the facilitator also helps select a COPE® committee, which is responsible for implementing and monitoring the action plan that staff develop. Three to 4 months after the first self-assessment, facilitators re-visit the implementing staff to review their progress and start the self-assessment process again. COPE® is meant to be implemented with other tools for continuous quality improvement, such as supervision and training.

Improvement Collaborative (also called collaborative improvement or, simply, collaboratives)

Improvement collaborative is a method for quality improvement developed in the 1990s at the Institute for Healthcare Improvement. Collaboratives aim to integrate into routine practice the best scientific evidence on how to improve outcomes and contain costs. To this end, they bring teams together with technical experts and process improvement coaches who use a continuous quality improvement process to make changes. The process uses iterative problem solving, encourages prompt

continued

BOX 1-1 Continued

process improvements, and emphasizes ongoing measurement and monitoring. Collaboratives usually last about 9–24 months, during which time the participating teams analyze a problem and its causes; plan changes and test the results of these changes; work with the coaches to set performance targets and measure progress toward meeting them. Collaboratives can be used to improve processes for patients and providers, teams, organizations, or systems.

SBM-R (Standards-Based Management and Recognition)

SBM-R is a management method developed by Jhpiego that aims to improve quality of care by improving health worker performance. It adapts the four main elements of the continuous quality improvement cycle (plan, do, study, act) to standardize, do, study, and reward. In the first step, participants are made aware of the national standard and trained on that standard. Next, the participants put the standard into practice, taking an initial assessment and correcting areas where their practice falls short of the standard. The study step involves repeated measurement of progress, including provider self-assessments, internal assessment by facility staff and managers, and external assessments. In the final step, the health workers are recognized for their efforts; rewards, such as feedback, praise, and social recognition, are seen as important in maintaining motivation.

Supportive Supervision

Supportive supervision refers to a process of working with staff to set goals, identify and correct problems, and monitor staff performance. It generally takes one of three forms: managerial, clinical, or educational.

SOURCE: Workshop read-ahead memo.

USAID missions, and many other organizations spending on quality improvement, would welcome more information about how different strategies work to improve quality, when and where certain tools are most effective, and the best ways to measure success and shortcomings. Such evidence would allow funders to make better-informed decisions about quality improvement programming. A better understanding of the evidence supporting different quality improvement tools and clarity on the causal pathways through which they work would, in turn, help advance the global quality improvement agenda.

On January 28–29, 2015, the IOM convened a 2-day workshop on improving quality of care in low- and middle-income countries. (See Box 1-2

BOX 1-2
Statement of Task

An ad hoc committee will plan a 2-day, public workshop on approaches to improving quality of care in low- and middle-income countries. A number of approaches to improving quality of health care are currently in use in low- and middle-income countries. These approaches are distinguished by distinct names, many of which are associated with specific organizations. The range of choices available has contributed to lack of clarity among U.S. Agency for International Development (USAID) missions and their counterparts. USAID requests that this workshop illuminate six different methods currently being used to improve quality of care in low- and middle-income countries, discussing their pros and cons. These methods are clinical in-service training; supervision; standards-based management and recognition; the client-oriented, provider-efficient method; improvement collaborative; and facility accreditation based on external evaluation. Evidence supporting these models might pertain to their cost-effectiveness, sustainability after donor support ends, and the degree to which these models have been made part of regular operations of the health system.

The public workshop will feature invited presentations and panel discussions. The planning committee will organize the workshop, select speakers and panelists, and serve as discussion moderators. Commissioned papers may be required to inform workshop discussions. A designated rapporteur or rapporteurs will prepare the workshop summary in accordance with institutional guidelines.

for the statement of task.) This workshop grew out of discussions between the IOM Standing Committee to Support USAID's Engagement in Health Systems Strengthening in Response to the Economic Transition of Health (hereafter, the standing committee) and the agency's Office of Health Systems. The workshop planning committee arranged the workshop agenda and invited a range of speakers from various organizations to respond to the statement of task.

This workshop summary is a description of the presentations and discussions as they occurred at the January workshop. The material is presented in roughly the order in which it was discussed, and the report is organized into sections corresponding to the sessions on the meeting agenda (see Appendix A). Views and opinions presented are those of individual speakers and do not reflect the consensus of the group, the planning committee, the IOM, or the workshop sponsor.

2

An Overview of Quality of Care in Low- and Middle-Income Countries

Key Points Made by Individual Speakers

- Unsafe medical care is a leading cause of death and disability around the world. It will not be possible to improve health in low- and middle-income countries without improving quality, but the topic gets considerably less attention than improving access. (Jha)
- The six strategies under consideration at the workshop are more similar than different. All can work in some contexts, and understanding the contextual factors that favor one over another is part of the challenge of implementation. (Jha)
- Like most other quality improvement methods, the six strategies put an emphasis on changing provider behavior, which is still several steps removed from changing patient outcomes. (Jha)
- USAID is a major supporter of quality improvement work. To do its work well, the agency needs to understand where different strategies are most suitable, how they work, and what the main gaps in the evidence are. (Heiby)
- It is difficult to glean impartial evidence about the different methods as long as the people closest to the work are asked to evaluate its effectiveness and perceive that they must minimize their failures and promote their successes to the wider audience. (Rowe)

After brief welcoming remarks from Victor Dzau, the IOM president, and Sheila Leatherman of the University of North Carolina at Chapel Hill, Ashish Jha of Harvard University gave a keynote address describ-

ing the current research on quality of care in low- and middle-income countries.

KEYNOTE REMARKS

Quality is a topic that has been neglected in global health, but that is changing. As countries grow wealthier, universal health coverage has become a common policy goal. The combination of growing wealth and increased access to services translates into more demand for health care. The number of doctor's visits, nurse's appointments, and hospital stays is going to grow rapidly in the next decade, and most of this growth will come from in low- and middle-income countries. At the same time, Jha pointed out, simply increasing use of services, or even demand for them, is not the goal of expanding coverage. The goal is to improve health, and that will not happen without changes to the quality of the services offered.

Jha then summarized his 2013 study that found the vast majority, about two-thirds, of the world's 43 million adverse events[1] occur in low- and middle-income countries (Jha et al., 2013). He qualified these numbers further, explaining that the calculations were based on a small group of in-hospital adverse events. Given what a small piece of health care hospitalization accounts for, Jha reckoned that unsafe care is probably 1 of the top 10 global causes of death and disability.

Further understanding of the problem comes from reviewing the work of Jishnu Das and similar researchers. Their work suggests that in low- and middle-income countries the probability of a patient receiving the correct diagnosis is, depending on other factors, in the range of 30 to 50 percent. Similar studies that have attempted to estimate the probability of a patient receiving non-harmful treatment found a likelihood of about 45 percent. While the probabilities vary widely depending on country and setting, Jha concluded that quality of care is a serious obstacle between expanded access and improved health.

Another important barrier to good quality care is lack of trust. The recent Ebola outbreak in West Africa has highlighted problems with public trust in the health system. Jha recounted a December meeting he attended on Ebola response. The participants included the health ministers of Guinea, Liberia, and Sierra Leone and many other global health leaders. One major theme from their discussion was that a lack of trust in the health system prevents people from using services. As long as people feel that they cannot trust their health system or that they are not treated respectfully,

[1] An adverse event refers to injury or harm to a patient as a result of medical care, rather than the underlying disease or condition.

the best quality services will still be useless, and the population's health will not improve.

Jha then briefly introduced the six strategies being discussed at the workshop (see Box 1-1). He explained that the literature suggests that each of these strategies works sometimes, in some contexts, but no strategy seems to work consistently in all contexts. He also observed that the six strategies have far more in common than they have differences. One example is a shared emphasis on measurement and iterative feedback.

All of the strategies rely on buy-in from leaders and adapting services to the local context. Quality improvement is generally a complex social intervention, and even seemingly universal tools like surgical checklists need to be tailored to the local environment. Adapting a tool requires consideration of the social relationships in the clinic. For example, if nurses are not empowered to correct doctors when they breach hygienic protocol, the organization will not be able to improve.

Another feature common to all six methods is a reliance on changing provider behavior, which is still several steps removed from the final goal of changing outcomes for patients. Measuring changes in the way care affects patients is complicated. Program evaluators usually have to be content with surrogate measures, such as changes in provider behavior. Often the success or failure of the program hinges on the choice of these surrogates. Successful programs rely on meaningful targets and set compelling goals. Jha felt that a goal such as eliminating hospital-based infection would motivate people more than a target of reducing hospital-based infections by 10 percentage points. Pay for performance programs may be particularly vulnerable to problems in this area. He explained that these programs, which have the power to improve efficiency, often choose targets that are not meaningful to clinicians or patients.

Quality programs that focus on provider behavior risk losing sight of the larger systemic obstacles that prevent change from taking root. Jha emphasized the need for commitment to change from the top levels of an organization. For example, a common goal of quality improvement programs in the United States is to reduce waste. To this end, doctors are encouraged to order fewer tests. But the organization's payments depend on the number of tests they order. This kind of obstacle makes it difficult to sustain any reduction in the number of tests ordered.

Jha concluded that improving quality is as important to global health as increasing coverage, although it gets much less attention. He suggested that quality improvement works best when the program prescribes a clear, meaningful goal, but not how to get there, which will vary in different settings. At the same time, the methods for quality improvement do matter. Accreditation, for example, may work better in places where patients use insurance or prepayment plans; in places where the workforce capability is

low, strategies such as supervision or collaboratives might not work well. Unfortunately, the literature is not clear on what the facilitating factors are for different strategies.

USAID'S WORK IN QUALITY OF CARE

James Heiby of USAID built on this point in his presentation about the agency's involvement in quality improvement. USAID has been active in quality improvement since around 1990, but initially most of that activity centered around the agency's headquarters. Over the past 10 years, the missions have expanded their efforts in this area, and the agency is now a major supporter of quality improvement.

Heiby framed his comments on quality in relation to the Donabedian model of a health system with inputs, processes, and outcomes. Quality improvement, especially at USAID, has attempted to change health care processes, the neglected piece of the triad. Quality improvement is essentially process improvement, and the agency has a stake in identifying the best strategies for process improvement. There are a range of strategies in use, exemplified by the six methods that dominate USAID's quality investments (shown in Box 1-1). Some strategies, like training and supervision, are so widely used that spending on them dwarfs any other global investment in quality, though many training and supervision programs are not primarily intended to improve quality. Other strategies, accreditation for example, are discrete interventions with the aim of changing the quality of services. These are different still from a set of discrete interventions that use management and incentive techniques to change provider behavior.

For USAID, it is important to understand how and where these different strategies are most suitable and what the main gaps in the evidence are. There are many possible strategies beyond the six discussed at the workshop, and Heiby encouraged the participants to think about the larger context of quality improvement. The first contextual question is whether or not everyone in the field understands quality to mean the same thing. Quality improvement is usually a complicated social change. It is hard to articulate the piece of the program that is the quality improvement, so people are more comfortable talking about measurable indicators. Researching and measuring quality tends to accompany a whole range of research projects—projects that produce a great deal of data and knowledge. Managing and synthesizing this information is a daunting responsibility. To complicate the matter, most data come from self-report and are difficult to validate. In discussing his concerns about the evidence base, Heiby reminded the audience that finding weak evidence for any or all of the target methods should not be seen as a failure in the discussion: understanding where the data are weak helps scientists form a clearer research agenda.

As an added benefit, increasing attention to quality will likely drive improvements in health information technology, an area of health systems strengthening that is often neglected. In a larger sense, all quality improvement can effect meaningful changes in low- and middle-income country health systems. Progress depends, however, on transferring the skills to the regular employees already working in country. When foreign consultants run these programs, the likelihood of sustaining changes becomes very low. Therefore, Heiby cautioned against blindly supporting quality improvement as a movement, saying that the programs should be held accountable for how they spend and how they influence their partner countries. The best way to do that may be by demanding rigorous evaluations of quality improvement programs.

The subsequent discussion illuminated some of the barriers to open sharing of program data and impact evaluations. Alexander Rowe of the Centers for Disease Control and Prevention (CDC) pointed out that the organizations charged with doing projects generally feel that they will be penalized if they report anything less than success. It will be difficult to glean impartial evidence about the pros and cons of different methods as long as the people closest to the work have the perception that they should be promoting their successes and downplaying their failures in the wider field.

3

Six Widely Used Methods to Improve Quality

Key Points Made by Individual Speakers

- The six strategies under consideration have much in common, including an interdisciplinary grounding, the use of assessments to identify problems, and benchmarking to measure progress. (Atun)
- The quality movement has grown out of the attempt to take the vast amount of evidence and guidelines afforded by modern science and capture them in practical ways. (Barker)
- Quality assurance tools are those that calibrate the performance of a system to certain standards. Quality improvement tools are more concerned with changing one area of a system, less with routine measuring against normative standards. The six methods discussed at the workshop fall across the continuum from quality assurance to quality improvement. (Barker)
- There is a need for rigorous evaluations on quality of care programs, but evaluating complex social interventions presents challenges, and there is limited funding for such work. (Cordero, Necochea, vanOstenberg)
- Implementers and funders of global quality programs would benefit from clear guidelines about what portion of spending to direct to programming and what portion to amassing evidence. (Necochea)

The next session introduced the six methods listed in Box 1-1: accreditation, clinical in-service training, COPE®, improvement collaborative, SBM-R, and supervision. Rifat Atun of the Harvard School of Public Health

gave an orientation to these six methods based on the standing committee's November 2014 meeting. The methods have two distinct components: the hardware, meaning the pieces of the intervention and the way those pieces fit together, and the software, meaning each intervention's unique design and the style with which it is implemented in different places.

Atun stressed some common elements of the six methods under question. All are multi-disciplinary, drawing on processes from social science, medicine, and management. They all rely on the use of assessments to identify problems and benchmarking to measure progress. Each one has an element of ongoing problem solving. In some of the methods, this means consciously employing a plan-do-study-act cycle; in others, the recursive element is the re-training or re-accreditation process. Facilitators or external advisors have a place in all the methods, though the prominence of the local team in solving problems varies.

Atun identified six criteria on which the methods could be evaluated. First is the cost-effectiveness, the results the intervention elicits relative to its price. The method's affordability is a similar concern, referring to the overall costs incurred by the host country. The feasibility of the method—whether it is realistic to implement the program in a range of settings where there are different resources—is another important factor to discuss. Along the same lines, the replicability of results in new settings and the scalability, or ease of expansion, are two features that funders need to understand. Lastly, the sustainability of the method, or the extent to which a program can be integrated into a health system, is of particular value both to external aid agencies and to the countries that host them.

Managers often struggle to identify what tools work in which contexts, a difficult task when the tool in question is a complicated social intervention. Randomized trials have recently gained prominence outside of health, with Jeffrey Sachs and others undertaking innovative cluster randomized designs to test development projects. Although randomized trials are still the gold standard evaluation technique, they are a bit reductionist in relation to complex quality improvement programs. Even if cluster randomized trials of the different interventions show success, it is not always possible to say which process or tool drove the success, or to isolate effect sizes for different pieces of the intervention. A careful step-wedge design, with elements of the intervention introduced in sequence, would be a suitable design, but in practice may be difficult to achieve given the interdependence of the constituent elements of quality improvement interventions. Further, the people who work most closely with the methods do not see the components as different pieces that can be broken apart, but rather as a whole program. Identifying the best methods to evaluate quality improvement and other complex interventions is a challenge facing the field.

THEORIES OF CHANGE

Evaluating interventions requires first understanding how they work, specifically the theories of change that explain how different actions cause behavior change. Pierre Barker of the Institute for Healthcare Improvement gave the audience a theoretical grounding on the topic. Quality improvement aims to change the performance of a system not by adding or reducing the resources directed to it, but by rearranging them, much the way roads and traffic laws rearrange the motor traffic in a city. The quality movement, as Barker described it, has grown out of an attempt to take the vast amount of evidence and guidelines afforded by modern medicine and capture them in practical ways.

Taking guidelines and implementing them, particularly at large scale, was a concern of W. Edwards Deming, a founder of modern quality improvement. Barker explained that scale and feasibility depend on what Deming described as the psychology of change, the way of marrying the knowledge of the evidence with the knowledge of management. In Deming's model, which was not designed for health, the first step in quality improvement is to define the system and its limits—the specifications within which the system works. The quality controllers are responsible for making constant, small changes to keep the system working within defined parameters.

A complementary influence on quality improvement came from Joseph Juran, who described a trilogy of quality control, quality planning, and quality improvement. Barker placed these terms in the health context, explaining that quality planning refers to policy decisions affecting the way resources are coordinated and the checks in place to ensure accountability, while quality control refers to national guidelines and systems for professional oversight and accreditation and uses tools such as checklists and standards. Figure 3-1 shows how all three pieces of the Juran trilogy drive changes in the way systems perform.

Barker then explained the difference between quality assurance and quality improvement methods (see Table 3-1). Quality assurance has an overarching goal of calibrating the performance of a system to certain standards; managers drive this change and enforce compliance targets as much as possible. Quality improvement tends to give attention to change in one area of the system; it is less about routine measuring against normative standards and more about all participants teaching and learning from each other.

Quality assurance and quality improvement techniques differ in how they take shape on the front line, Barker continued. He described how managers of quality assurance programs often hit systemic barriers when implementing programs. The barriers are then reviewed in a plan-do-study-act cycle. The information from this cycle influences future iterations of the

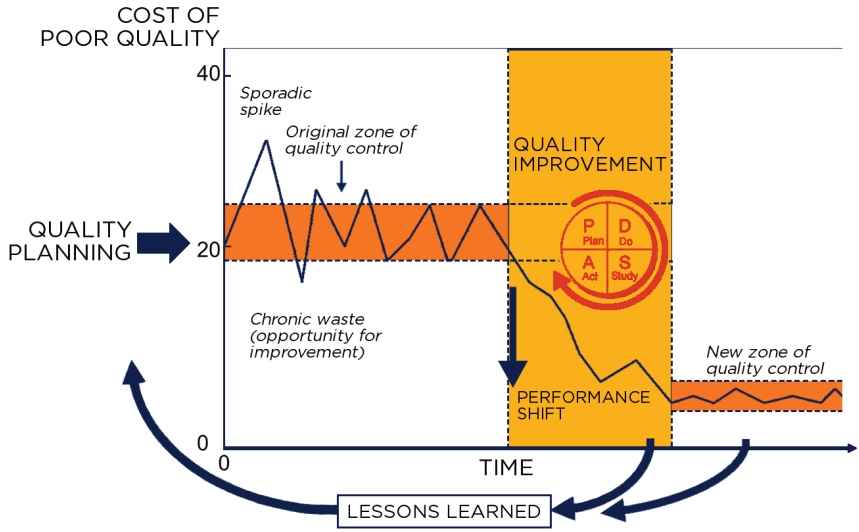

FIGURE 3-1 All three elements of the Juran trilogy are needed to improve outcomes.

SOURCE: Juran, J. M., and A. B. Godfrey. 1999. *Juran's quality handbook*. New York: McGraw-Hill. © McGraw-Hill Education. As presented by Barker on January 28, 2015.

TABLE 3-1 Differences Between Quality Assurance and Quality Improvement Methods

	Quality Assurance (QA)	Quality Improvement (QI)
Performance goal	Perform to standards (controls) across multiple parts of the system	Aspire to a best performance goal for a focused improvement area
Measurement	Periodic inspection of past events (large set of measures of inputs and/or processes)	Continuous tracking of current activity (few key processes linked to outcome)
Data tracking	Before and after change	Continuous (e.g., run charts)
Data system	External (e.g., inspection tools)	Internal (e.g., registers and tally sheets)
Changes	Standards driven; normative; can be linked to frontline system analysis	Theory driven; adaptive; always linked to frontline system analysis
Motivation for change	Management-led; compliance; incentives; competition	Shared governance; internal motivation; "all teach all learn"
Plan-Do-Study-Act (PDSA)	Management planning with single "slow" (months) intervention cycle; can use frontline rapid cycle to respond to defects	Frontline planning; rapid cycle (days/weeks) is core activity

SOURCE: Barker, 2015. Reproduced by permission of Pierre Barker, Institute for Healthcare Improvement (unpublished).

program, as well as changes to the normative standards. In quality improvement programs, the problem is taken straight to the front lines, where staff are asked to plan the necessary changes and there is great emphasis put on rapid plan-do-study-act cycles. Staff implement the proposed changes, and the success or failure is shared among all parties at the same time. Almost all modern quality improvement programs draw on the concepts of the Juran trilogy. The methods under discussion at the workshop set clear aims, establish terms for what should be considered improvement, and determine actions that will elicit the desired improvements. All the methods are grounded to some extent in the plan-do-study-act cycle, although the length of the cycle may vary. The six methods from the workshop fall across the quality assurance and improvement spectrum (see Figure 3-2). Barker admitted that all the methods have aspects of what might be considered a quality assurance orientation and a quality improvement slant. Depending on how it is implemented, each method could be placed differently along the continuum. In general, however, he thought that COPE® and SBM-R would tend to fall in the middle, with a slight tendency toward quality assurance, while improvement collaborative is meant to be more of an improvement tool, though it does reach into quality assurance at times. Barker saw training and supervision as general support to the whole continuous process; it is not possible to say exactly where any particular training or supervision program might fall.

Barker used an example from his fieldwork in sub-Saharan Africa to illustrate these theories. One way to decrease neonatal mortality is to give women who go into preterm labor a dose of the corticosteroid dexamethasone to mature the baby's lungs, but this intervention was not happening at a hospital in Malawi in late 2013 (see Figure 3-3). After a quality improvement program, the percentage of eligible women receiving the corticosteroid increased rapidly and stayed that way, with a few lapses during manage-

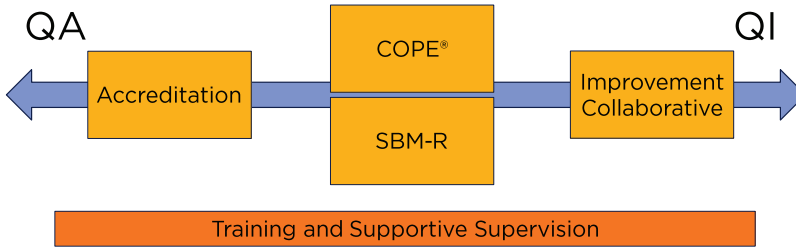

FIGURE 3-2 Putting the six methods on a continuum from quality assurance (QA) to quality improvement (QI).

SOURCE: Barker, 2015. Reproduced by permission of Pierre Barker, Institute for Healthcare Improvement (unpublished).

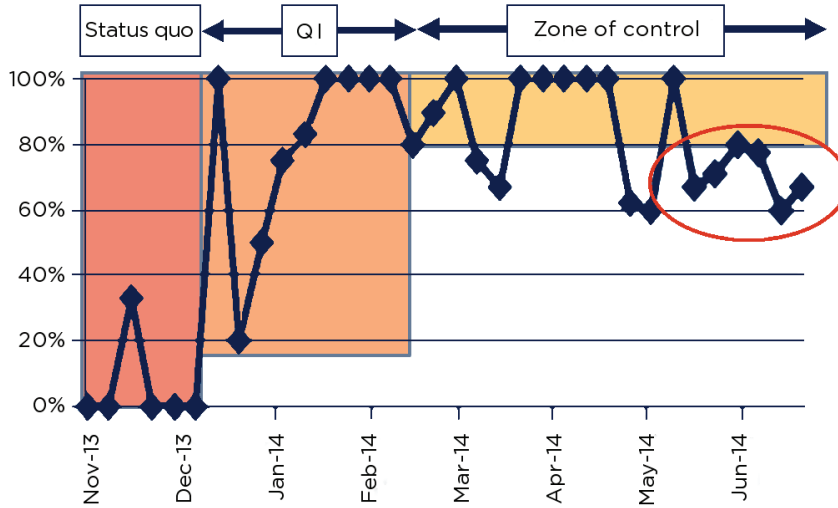

FIGURE 3-3 Percentage of eligible women receiving dexamethasone before, during, and after a quality improvement (QI) intervention. The red circle indicates a drop in compliance that should be seen as a warning signal.

SOURCE: Barker, 2015. Reproduced by permission of Pierre Barker, Institute for Healthcare Improvement (unpublished).

ment changes, until mid-2014, when compliance started to dip. This dip in compliance occurred at the flagship hospital as well as the other program hospitals. On a later visit to Malawi, Barker found the program almost abandoned. The government had, for many reasons, decided not to pursue the corticosteroid program, so women were no longer even being assessed for gestational age (necessary to determine if their labor is preterm), and supplies of the necessary drug had run out. The program's success could not be sustained because it depended on larger systemic factors, including the commitment of management up to the national level and the procurement of essential medicines. In this example, the hospital is something of a micro-health system, but sustaining success depends on factors beyond the hospital itself, such as the district policies, management's commitment, and a reliable medicines supply.

The sustainability of different quality improvement programs is a priority in global health, and some participants struggled with the extent to which successful techniques can be generalized. Barker concluded that, within a country, there is some generalizability of implementation approaches and that emphasis on context, though helpful, should not obscure these common threads.

TOOLS AND PROCESSES FOR QUALITY IMPROVEMENT

In the next session, experts in each of the six methods shown in Box 1-1 oriented the audience to their tool: the causal pathway through which it works, the key processes, and the relative strengths and weaknesses.

Client-Oriented, Provider-Efficient Services (COPE®)

Carmela Cordero of EngenderHealth started the session with a description of COPE®. COPE® is one part of EngenderHealth's suite of quality improvement interventions. It was developed in Kenya and Nigeria in the late 1980s, drawing on the work of Deming as well as other contemporary thinkers who were interested in clients' rights and providers' needs. Organizational psychology is important to COPE®; the method assumes that problems are more meaningful and solutions more effective if they come from facility staff.

The COPE® package has four main pieces: a handbook; a self-assessment guide that requires a systematic analysis of how services are provided; client interview guides that set out how staff members should talk to clients and identify what clients consider to be good quality care; and a client flow analysis to monitor how long clients are waiting and how long their contact with providers lasts. Figure 3-4 shows how the process and tools used in COPE® support continuous assessment of health services.

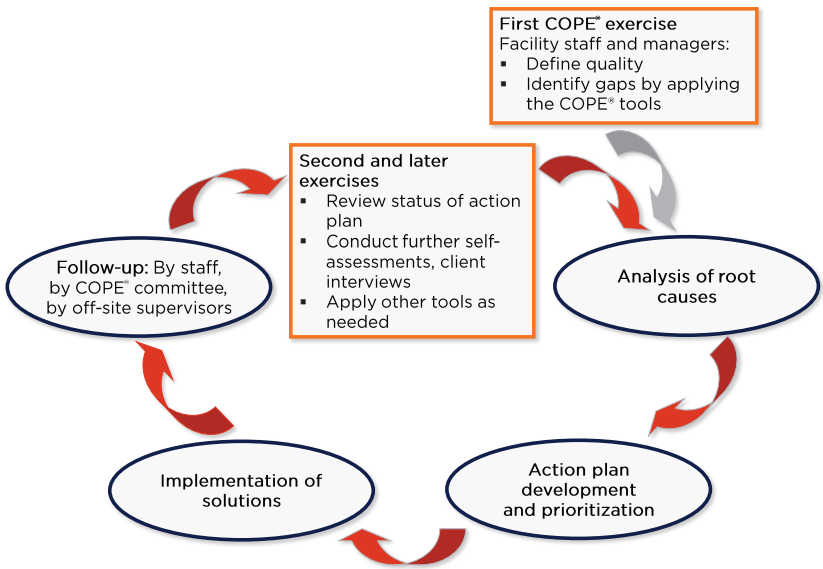

FIGURE 3-4 The processes and tools in COPE®.

SOURCE: Cordero, 2015. © Reproduced by permission of EngenderHealth.

COPE® exercises put a good deal of emphasis on local understanding, including how patients and providers perceive good quality care. External facilitators do not influence these discussions: their role is to help the local team analyze their service processes and identify the root causes of their problems. The COPE® facilitator is responsible for supporting local staff in developing an action plan. This includes establishing a local COPE® committee to start the process of implementing changes to the system. The process is meant to be highly participatory, based on the assumption that participating in the process gives workers a stake in the results (see Figure 3-5). The process also requires attention to recordkeeping and data analysis. Figure 3-6 shows the conceptual pathway through which COPE® elicits change.

Cordero described the emphasis on clients and communication as being among COPE®'s main strengths. Often the COPE® process is the staff's first introduction to the concept of quality, and it helps that a set of tools for problem solving accompanies the process, as does an introduction to relevant standards and guidelines. On the other hand, the process requires a lot of energy and commitment from both the organization's leaders and the local government, which can be seen as weaknesses. She closed her

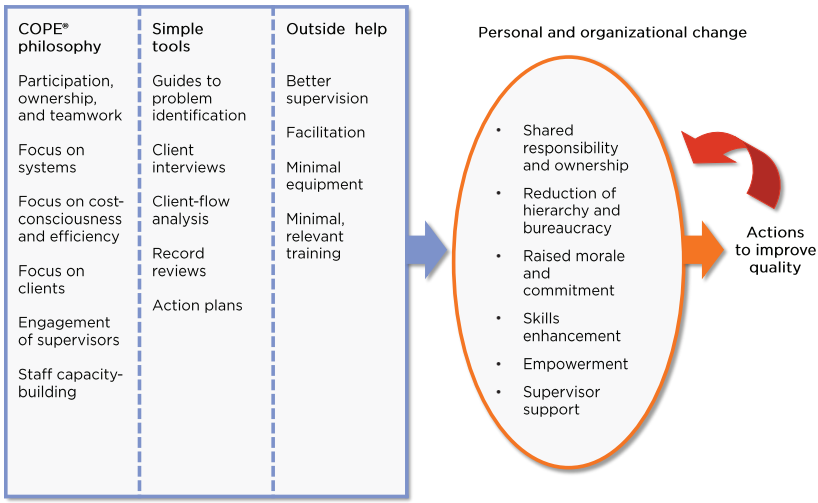

FIGURE 3-5 The COPE® model uses participatory methods that cultivate personal and organizational commitment to change and empower facility staff to make improvements in quality.

SOURCE: Bradley, J., and S. Igras. 2005. Improving the quality of child health services: Participatory action by providers. *International Journal for Quality in Health Care* 17(5):391–399. Reproduced by permission of Oxford University Press and the International Society for Quality in Health Care. As presented by Cordero on January 28, 2015.

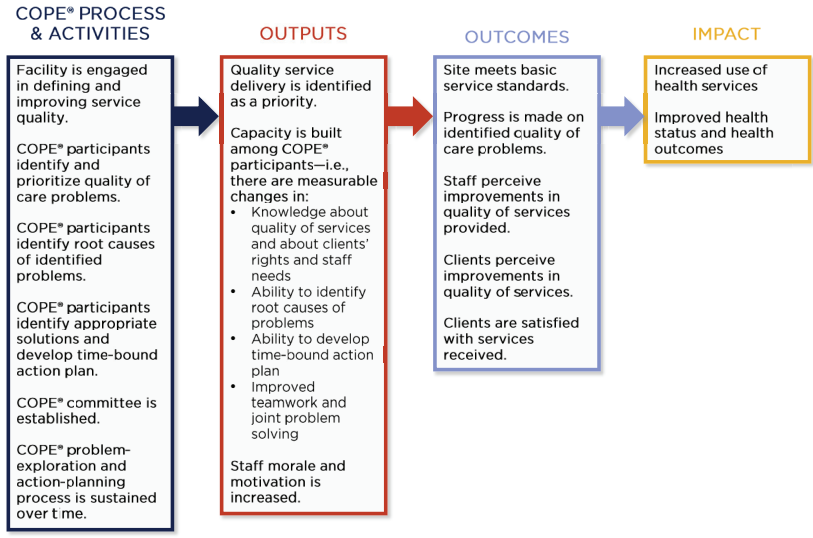

FIGURE 3-6 The logic model through which COPE® changes behavior and improves quality at the facility.

SOURCE: Cordero, 2015. © Reproduced by permission of EngenderHealth.

presentation by observing that COPE®, like all quality interventions, is not the only way to effect change and that the method works best when it is a piece of a larger quality strategy.

Standards-Based Management and Recognition (SBM-R)

Edgar Necochea of Jhpiego opened his presentation by reiterating Cordero's point that all of the methods being discussed can work under certain circumstances, and his presentation described what circumstances encourage success with SBM-R. He described SBM-R as a standardization approach, one that attempts to bridge a gap between the evidence and practice in low- and middle-income countries. In low- and middle-income countries, staff are typically overworked in poor conditions. They have to make the most of limited resources and often have had weak pre-service education. Management can be dysfunctional and morale low. In such environments, there is no shortage of evidence about what works or failure to translate evidence into concrete guidelines. Rather, the main problem is that guidelines are shelved by overworked staff and not translated into tools for daily use.

SBM-R builds on the Deming theory of quality improvement, modifying the plan-do-study-act cycle to standardize-do-study-reward as shown in Figure 3-7. The goal is to identify desired outcomes and processes that will lead to those outcomes. Clear written standards are key to SBM-R; these standards are developed with the host country counterparts and take considerable input from local stakeholder groups. Necochea shared an example of a standard (see Figure 3-8) from his fieldwork and called the audience's attention to the standard's emphasis on effective and respectful care, including details about both the substance of the intervention and the manner in which it should be done.

The first step in SBM-R is the training of supervisors and teams on the content of the standards and the quality improvement process. SBM-R facilitators then take a baseline assessment and work with teams, usually at health facilities, to analyze what obstacles are impeding implementation of standards. Interventions that correct these gaps must address all the factors that affect performance, including things like the provider's knowledge, the available resources, and the health workers' motivation. During the next stage in the process, internal and external supervisors monitor and measure the staff's progress, so there can be ongoing recognition of success. Necochea emphasized the importance of external recognition for both the workers and the supervisors, sharing examples of how ministers, managers, and other community leaders take part in the recognition process—a way of mobilizing the community for health.

Comparative studies on the effectiveness of SBM-R show some encouraging results. Data from 102 Mozambican hospitals showed substantial

FIGURE 3-7 The modified plan-do-study-act cycle that forms the basis of SBM-R.
SOURCE: Necochea, 2015. Reproduced by permission of Jhpiego Corporation.

Area: Antenatal Care			
Performance standard	Verification criteria	Y, N, N/A	Comments
1. The facility conducts a routine rapid assessment of pregnant women	Observe in the reception area or waiting room if the person who receives the pregnant woman: • Asks if she has or has had: ○ Vaginal bleeding ○ Headache or visual changes ○ Breathing difficulty ○ Severe abdominal pain ○ Fever • Immediately notifies the health provider if any of these conditions are present		

FIGURE 3-8 An SBM-R standard for antenatal care.
SOURCE: Necochea, 2015.

improvement in maternity practices, including treatment of eclampsia and pre-eclampsia, active management of the third stage of labor, and use of the partograph during labor, thereby reducing maternal mortality (see Figures 3-9 and 3-10). LiST[1] models confirm that the observed reduction in mortality at project hospitals is consistent with the measured improvement in practice.

Necochea reviewed SBM-R's strengths: it is a simple, intuitive, and systematic approach based on standardization of care. It also gives good attention to the human side of management: worker motivation, political will, and involvement of community leaders. Its emphasis on standards encourages the use of information technology. At the same time, the method does not work on all problems: some interventions do not lend themselves to standardization. SBM-R also takes time. Honing process efficiency and building support among local leaders can take months. This may be why SBM-R's sustainability and integration into national health systems remain a challenge.

Accreditation

Next, Paul vanOstenberg presented on accreditation. His remarks were not limited to the accreditation process of his organization, Joint Commission International, but he did draw some examples from their work. VanOstenberg defined accreditation as "a voluntary process by which a

[1] LiST, or the Lives Saved Tool, uses national demographic projections, burden of disease, and information about program effectiveness to estimate the effects of changes in the coverage of different maternal and child health interventions.

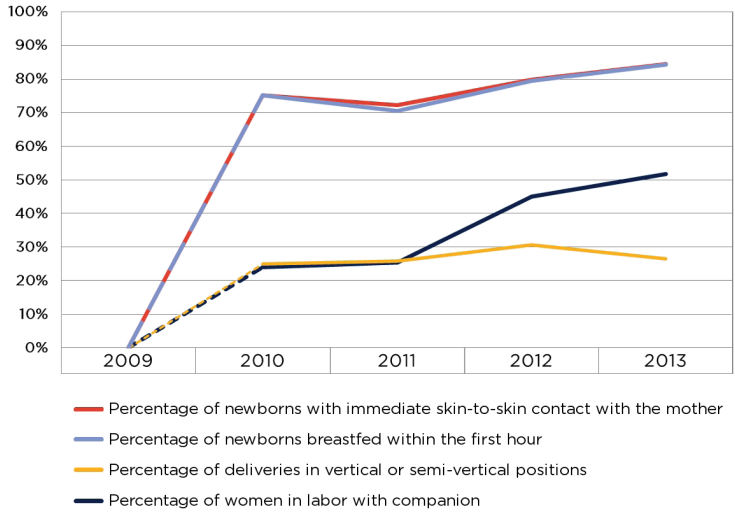

FIGURE 3-9 Results from a comparative study showing the trends in maternal and newborn health care practices after SBM-R was implemented in 102 facilities in Mozambique, 2009–2013.

SOURCE: Necochea, 2015. Reproduced by permission of Jhpiego Corporation.

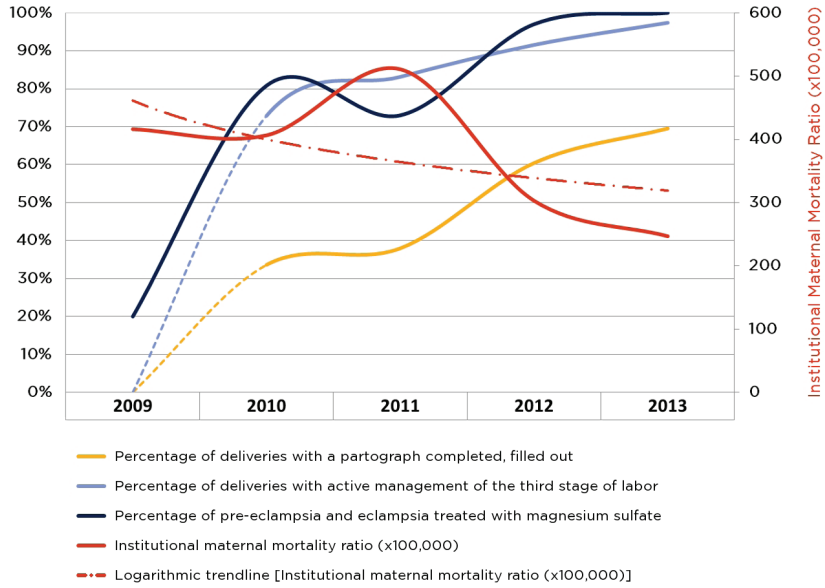

FIGURE 3-10 Results from a comparative study showing the trends in maternal health service delivery practices and health outcomes after SBM-R was implemented in 102 facilities in Mozambique, 2009–2013.

SOURCE: Necochea, 2015. Reproduced by permission of Jhpiego Corporation.

government or nongovernment agency grants recognition to health care institutions which meet certain standards that require continuous improvement in structures, processes, and outcomes." In English, the terms accreditation, certification, and licensure are often mistakenly used as synonyms. Although the meaning of each term varies in different countries, accreditation often differs from certification in that the latter establishes an organization's (or a person's) competence in a particular procedure (e.g., a certified mammography center).

Accreditation is essentially a risk-reduction strategy, one that works by applying standards and evaluating adherence to them. Each institution chooses its own path to meet the accreditation standard; in that way, the process is a vehicle for different quality improvement methods. Because accreditation requires regular re-review, it encourages a culture of continuous quality improvement. Table 3-2 shows the steps in the accreditation process and estimates the amount of time they take in developed countries; in low- and middle-income countries, certain steps can take much longer.

VanOstenberg explained that accreditation's main strengths lie in its emphasis on the whole health system: accreditors look at all the processes, materials, and staff involved with an institution and evaluate them against consensus standards. The process can encourage new working relationships across different parts of the system and in different provider networks. It also lends itself to public–private partnerships: the government often sets the standards, but leaves the evaluation and survey of them to a private accrediting body.

The emphasis on standards can be a double-edged sword, especially in low- and middle-income countries, he continued. Meeting hundreds of standards can seem impossible and overwhelming in places where resources are limited. To complicate the matter, accreditation standards can vary in their content and the way they are described, even when the standard comes from the same accrediting body. The evaluation process can also vary

TABLE 3-2 The Accreditation Process, Including the Typical Timeline in Developed Countries

Step in the process	Amount of time to complete
Obtain and study standards	6 months
Implement standards and prepare for the evaluation	12–18 months
External evaluation	3–5 days
Decision and recognition	1 month
Re-review	2–4 years

SOURCE: vanOstenberg, 2015.

widely: some accrediting bodies allow a brief self-evaluation, while others require weeks of staff time with the external reviewers.

VanOstenberg gave an example of how the accreditation process can be adapted for very basic clinics. The SafeCare Initiative, a partnership among the South African accreditor, Joint Commission International, and PharmAccess Foundation, aims to distill the essence of accreditation—the external, standards-based quality evaluation—to more than 1,500 clinics in rural South Africa. The program recognizes clinics for reaching different levels of compliance with standards, providing loans and incentives to help the clinic management meet each level. By the time a clinic has reached the highest intermediary level, it is in a good position to try for formal accreditation.

Most of the research about the effectiveness of accreditation as a quality improvement tool comes from developed countries. VanOstenberg mentioned the dearth of literature from the rest of the world and cautioned that accreditation is a tool for continuous quality improvement, not a substitute for a national licensure system that codifies the national minimum standards for health.

Discussion

In the discussion following their presentations, the panelists gave their views on what makes a quality program sustainable. Cordero and Necochea agreed that integrating quality assurance into routine management is the essence of sustainability. Regardless of the method used, they saw continual, deliberate monitoring as essential for sustainability, as are communication and involvement of all local stakeholders. VanOstenberg noted that, when community members understand the value that quality control adds to health services, they will advocate for more attention to the problem.

Some participants discussed their experiences with implementing programs, and one person questioned whether ministries of health ever ask to see evidence that the program should work. In the early days, counterparts might ask for an explanation of how the process would work and, if it made good sense, they agreed to try it. Lately there is more demand for evidence, which the organizations are starting to amass, but the panelists agreed that the available evidence is generally weak. They have common problems identifying the best methods to evaluate their work. As a result, the evidence base is limited to small studies, studies without a comparison group, and studies demonstrating non-causal associations.

According to the panelists, one barrier to more rigorous evaluations is the limited funding for such research. The need for evidence poses difficult questions for implementing organizations and their funders. If there were clear consensus about what portion of spending goes to programming

and what portion goes to building the evidence base, some of the problem could be solved. There is also the risk that opportunities for establishing the effectiveness of quality programs are lost through insufficient attention to evaluation in the program's planning stage. Implementing and funding agencies need to identify the intended outcomes of their program and concrete measures of its success before the project is under way. Too often they wait until the program implementation phase, at which point it is too late.

The other three methods under consideration (improvement collaborative, clinical in-service training, and supervision) were discussed in the next panel.

Collaborative Improvement

Rashad Massoud of University Research Co., LLC (URC) gave the first presentation on collaborative improvement. Although the knowledge exists, life-saving services still fail to reach patients in low- and middle-income countries. Deficits in the clinical processes and organization of the health system are at the root of the problem that collaboratives aim to correct. In the collaborative model, multiple sites work on the same problem at the same time. The method encourages learning from peers, who are all testing different ways to improve on common indicators. Collaboratives support analysis of the process for delivering care and data collection on the outcomes being studied. They use plan-do-study-act cycles and measure the effects of changes to the procedures. Collaborative improvement can address clinical or managerial topics, such as recordkeeping and waiting times. As URC implements them, the collaboratives work on problems the host country and the USAID mission have identified as priorities.

Massoud illustrated this point with an example from Niger of a collaborative that aimed to reduce post-partum hemorrhage, a major cause of maternal death worldwide, with active management of the third stage of labor (see Figure 3-11). Thirty-three participating hospitals formed quality improvement teams. URC worked with these teams, training them on technical material and on ways to make the process run more smoothly. One team promptly identified a problem with the availability of oxytocin, a uterotonic drug used to treat post-partum hemorrhage. Oxytocin is not stable at room temperature; for this reason, it is always stored in the dispensary refrigerator. Deliveries happen around the clock, but the dispensary is usually locked when the pharmacist is not on duty. After identifying this problem, the teams found creative ways to make the drug available during deliveries, such as storing a single dose on ice or putting some aside in a cooler.

The dramatic reduction in post-partum hemorrhage, shown in Figure 3-11, persisted even after the program ended in 2008. External evalu-

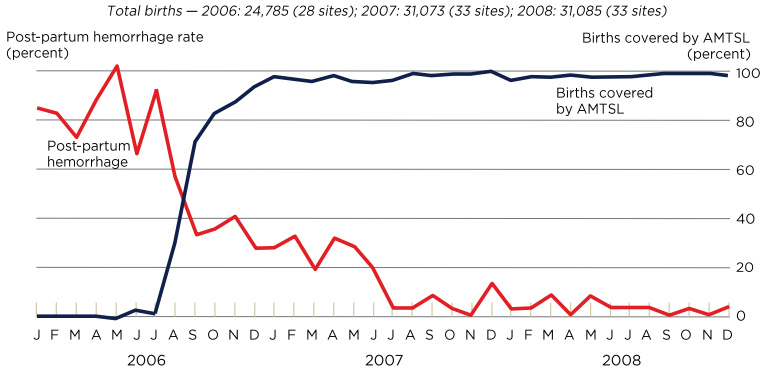

FIGURE 3-11 Active management of the third stage of labor (AMTSL) and post-partum hemorrhage rates, in participating Nigerien hospitals, January 2006 to December 2008.
SOURCE: Data from Maina Boucar, URC, as presented by Massoud on January 28, 2015.

ators and a records audit both showed adherence to active management of the third stage of labor, and the consequent reduction in hemorrhage up to 2 years after the end of the collaborative. A cost-effectiveness analysis of the project found that the cost per delivery fell from $35 to $28, with an incremental cost-effectiveness of $147 per DALY averted (Broughton et al., 2013). As further evidence of the program's sustainability, Massoud discussed how the Nigerien team replicated their work in neighboring Mali without technical assistance from URC headquarters (Boucar et al., 2014).

Other collaborative improvement programs have also shown good sustainability over time. A collaborative to improve HIV care in Uganda continued to improve adherence to treatment and clinical outcomes 2 years after the intervention phase. Similarly, the evaluation of a collaborative on compliance with best practices for management of noncommunicable diseases in the Republic of Georgia found that 40 to 91 percent of the improvement in practice could be attributed to the collaborative (see Table 3-3).

In describing the pros and cons of collaboratives, Massoud mentioned the breadth of expertise necessary to run them: experts in medicine, implementation science, and group process need to work together at the patient care, management, and policy levels. The process is time consuming, and some would say that it takes away from delivering services. At the same time, the intensity of local involvement can be seen as one of the method's strengths. Solutions identified through collaboratives are already suited to the local context; the process of identifying and implementing these solutions builds local capacity and empowers managers to pursue con-

TABLE 3-3 Results from a Controlled Evaluation of Noncommunicable Disease Compliance in the Republic of Georgia, Showing the Percent Difference Attributable to the Collaborative Intervention

Indicator	Attributable difference	p-value
% of chronic obstructive pulmonary disease (COPD) patients given evidence-based medications for management on discharge	40%	<0.001
% of coronary artery disease patients put on secondary prevention (aspirin, beta-blocker, angiotensin-converting enzyme inhibitor/angiotensin receptor blocker, statin) with all four medications	56%	<0.001
% of acute coronary syndrome patients with initial treatment (morphine, oxygen, nitrate, aspirin) recorded	44%	<0.001
% of COPD patients where all risk factors recorded (smoking, body mass index, physical activity)	91%	<0.001
% of charts of patients with COPD where all risk factors recorded anywhere in the chart (smoking, body mass index, physical activity)	60%	<0.001
% of pneumonia patients assessed for respiratory status severity	43%	<0.001

SOURCE: Data from Tamar Chitashvili, URC, as presented by Massoud on January 28, 2015.

tinuous quality improvement. Other strengths of the method are its cost-effectiveness and emphasis on data management.

Clinical In-Service Training

Mike English of KEMRI Wellcome Trust gave the next presentation on clinical in-service training. He asked the group whether or not training works and then cited a systematic review from the early 2000s that showed a 10 percent median effect of training on changing provider behavior (Forsetlund et al., 2009). English conceded that training clinicians is not a particularly effective way to improve health, or at least it is far less powerful than one might assume. He concluded that training may be necessary when providers' knowledge is not up to par, but it is almost never sufficient to sustain changes in care, especially given how quickly the training material is forgotten.

In general, training is understood as a tool that works when providers are not well informed: once they have the knowledge, they will change their practice. Many of the other presenters showed through their examples, however, that provider knowledge may have a relatively small influ-

ence on the standard of care in low- and middle-income countries. When a country runs out of dexamethasone, for example, women in preterm labor cannot be treated regardless of the provider's knowledge. Training as a way to reliably improve practice may be an oversimplified understanding of the relationship between knowledge and behavior. English compared it to alcohol use: we all know that drinking alcohol is unhealthy, but people drink it anyway. Training can give clinicians the competence to do a task, but it cannot change larger problems with their motivation or work environment.

It is not inexpensive to run in-service training programs. In English's experience in Kenya, a training program with qualified teachers costs about $100 per person per day. When foreign governments run the programs, the costs only increase. This analysis does not even account for the opportunity costs associated with keeping both providers and their trainers out of work for the duration of the program. Aggravating the problem is the authority of managers to choose training participants. English's experience, seconded by many in the room, was that the same people are re-trained multiple times per year while most of their co-workers are never trained once. Students earn generous per diems at trainings, an incentive that donors have encouraged.

English explained that, in an attempt to minimize the costs of training, global health programs often use a cascading model for in-service training in which the student who attends a training is responsible for educating a larger group at his or her home office. Educational theory offers little evidence to support this model. It is extremely difficult for one person to change the behavior of a group of 25 or 30 peers. This might be contrasted with pre-service training, wherein one skilled professor, working with the right curriculum, can have a powerful influence on a large group of students.

At the same time, training can be an important part of a larger management intervention. Figure 3-12 shows the relationship between behavior change and the provider's capability, motivation, and opportunities to practice. Clear thinking about exactly which behaviors the training intervention is meant to change and how it should change them can allow for a more efficient program evaluation. At the moment, there is not enough data about how training influences the relationships shown in Figure 3-12. English recommended more attention to the related conceptual problem of measuring quality. He cautioned that, by not investing in data, we perpetuate the lack of understanding about how management and quality interventions work.

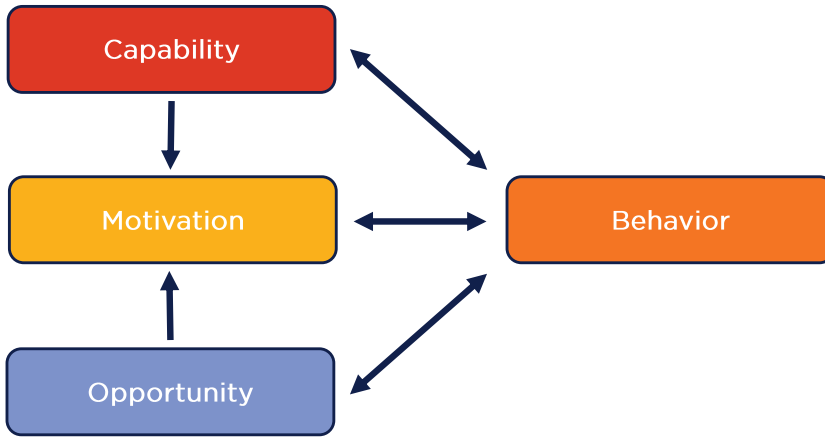

FIGURE 3-12 The capability, opportunity, motivation model of behavior change.
SOURCE: Michie et al., 2011. As presented by English on January 28, 2015.

Supervision

Xavier Bosch-Capblanch of the Swiss Tropical and Public Health Institute gave the final presentation of the session, examining the role of supervision in quality improvement. His comments drew heavily from a 2008 literature review and a 2011 Cochrane review on the effects of supervision on quality of care, as well as from a more recent summary paper on improving the quality of integrated community case management programs (Bosch-Capblanch and Garner, 2008; Bosch-Capblanch and Marceau, 2014; Bosch-Capblanch et al., 2011). These papers drew on evidence published from observational and experimental studies as well as programmatic literature.

Supervision can take many forms. In health care, the term most often refers to the managerial duties of a district health officer or head of a hospital. Supervisors deal with information, such as the volume of activity in the system and caseloads reported from the periphery to the headquarters. Usually, supervision involves visits from the central office to the field offices. The frequency of such visits varies widely but is generally about every other month. Logistics and costs pose serious obstacles to more frequent supervision. In the 2008 and 2011 reviews, Bosch-Capblanch and his colleagues found that quality changes are more meaningful and sustainable when supervision is part of a bigger program that includes health worker training, incentives, and improved supply chain management.

The 2011 Cochrane review found nine studies of suitable rigor for in-

clusion: five cluster randomized trials and four controlled before-and-after designs. Bosch-Capblanch conceded that the quality of the evidence was poor; most studies had a high risk of bias. It was also essentially impossible to summarize the comparisons across the nine studies because the control and intervention groups differed in every study (see Table 3-4). The outcomes chosen also varied widely from health worker performance to the quality of chart documentation in hospitals. The review concluded that supervision elicits no certain, long-term effect on quality of care. Ultimately, supervision is a complicated intervention; it is implemented in many different ways and usually accompanied by other quality improvement measures. The literature is not always clear about separating the supervision intervention from the incentives that go along with it, such as bonus pay or supply chain reform.

Bosch-Capblanch echoed English's sentiment that it is difficult to articulate the causal pathway by which supervision changes health outcomes, a problem his team has been struggling with as they update the Cochrane review on supervision and quality of care. The health systems framework that some experts refer to is descriptive: it identifies the components of a health system. It is not an analytical framework that would allow researchers to estimate the effects of submitting the system to different stresses. The

TABLE 3-4 Summary of the Nine Types of Comparisons Examined in a Cochrane Review on the Effects of Supervision on Quality of Care

#	Intervention	Control
1	Supervision and monitoring with feedback	No supervision
2	Training of providers, cascade training package, one supervisory visit	Same, without supervisory visit
3	Two supervisory visits on adherence to guidelines and stock management	No supervision
4	Training on Integrated Management of Childhood Illness (IMCI), enhanced supervision, enhanced package of support	Training on IMCI, routine supervision, usual support package
5	Training of supervisors with Modified Matrix model or the Centre for Health and Social Studies (CHESS) model	Routine training of supervisors
6	Community leaders involvement in supervision	Routine supervision
7	Intensive monthly supervision	Routine supervision
8	Supervision with training and checklists	Routine supervision only
9	Quarterly supervision	Monthly supervision

SOURCES: Bosch-Capblanch, 2015; Bosch-Capblanch et al., 2011.

lack of a clear analytic framework prevents any meaningful understanding of supervision or its effects. Bosch-Capblanch cited the development of such a framework as his group's top priority; otherwise there is the risk of doing pointless research on outcomes that turn out to be irrelevant to the intervention.

4

Reviewing the Evidence for Different Quality Improvement Methods

Key Points Made by Individual Speakers

- An exhaustive systematic review on the effect of different quality improvement strategies to change provider performance was unable to make any conclusions about the effectiveness of COPE®, SBM-R, or accreditation because of insufficient information. (Rowe)
- Training and supervision have modest positive effects on provider performance. The two strategies in combination may work better than either one alone. (Rowe)
- The data suggested a somewhat larger effect for improvement collaboratives, but high risk of bias in these studies prevents any firm conclusion. (Rowe)
- There is a need for more head-to-head comparisons of quality improvement tools, but it is not clear who should be charged with such research. Implementing organizations may have a conflict of interest in evaluating their own work. (Broughton, Mate, Rowe)
- Electronic data management systems can improve the efficiency of the health system and, when properly managed, can provide the data necessary to establish the link between quality programs and improved health. (Agins)
- Monitoring quality of care in a country depends on accurate clinical and death registries and the ability to link patient data across registries. (Klazinga)
- Cost-effectiveness analysis requires information on the costs and health consequences of every outcome a quality of care program could have on a system, and it depends on epidemiological modeling,

the assumptions of which are always debatable. But the complexity of such research is no excuse not to do it. (Broughton)
- Cost-effectiveness research is neglected, an omission that could undermine the field of quality improvement. (Broughton)
- More randomized trials of improvement interventions could yield a wealth of information for policy makers. (Broughton)
- Early planning for program evaluations could help make them more of a priority. (Barker, Broughton, English)

A consistent thread in the January 28 discussion was the need for more evidence, but the workshop aimed to glean as much insight as possible from the evidence that already exists. In the next session, participants considered an exhaustive systematic review on the effects of different quality improvement strategies on provider behavior. They also discussed challenges and opportunities in measuring quality of care and how to establish the cost-effectiveness and feasibility of quality improvement programs.

WHAT WORKS?: THE RESULTS OF A SYSTEMATIC REVIEW

Alexander Rowe of the CDC presented preliminary results of the Health Care Provider Performance Review, a systematic review of quality improvement methods to change the behavior of health care providers (hereafter, providers), funded by The Bill & Melinda Gates Foundation, the CDC, and the World Bank.

The systematic review included any strategy for improving provider performance, with providers being defined rather broadly to include private health workers, pharmacists, and drug shopkeepers. Both published and unpublished studies were included, and there were no language restrictions on inclusion. Pre- and post-intervention studies with a comparison group were eligible, as were post-intervention only studies with randomized controls, and interrupted time series studies with at least three data points before and after the intervention. Rowe's team identified studies from 15 electronic databases; this search was finished in 2006. Next, they reviewed personal libraries, searched document inventories of 30 organizations, and asked colleagues for references and unpublished studies, a phase of the review that ended in 2008. They also conducted a hand search of bibliographies from 510 previous reviews and other studies. Over the project's last years, 17 investigators sent in new reports from their research. Two people reviewed each study report, corresponding with authors when details of their study or strategy were not clear. Through the extensive literature review and verification process, the final database came to contain more information than the published reports.

In preparation for the IOM workshop, Rowe and his colleagues reviewed their database for information about the effectiveness of the six quality improvement methods being discussed. The database has about 150 variables that separately code each component of a strategy. For example, training and supervision are separate components of quality improvement strategies. Because there is no universal taxonomy of strategies to improve provider performance, all the definitions in the database are working definitions created for the analysis. Table 4-1 shows how each strategy was defined. Rowe conceded that some experts might object to the analytic organization. The accreditation category, for example, includes licensing, certification, accreditation, or registration programs. The proprietary methods of COPE® and SBM-R also posed an analytic challenge. As the database had no studies specifically involving COPE® or SBM-R, Rowe

TABLE 4-1 Definitions of the Six Strategies Using Component Variables from the Health Care Provider Performance Review (HCPPR) Database

Strategy	Definition
High-intensity training only	Training >5 days (or ongoing training) with ≥1 interactive method (i.e., clinical practice, interactive sessions, or role play). No other components.[a]
Low-intensity training only	Any training that is not high-intensity training. No other components.[a]
Supervision only	Supervision. Excludes strategies that resemble supervision (e.g., audit and feedback). No other components.
Accreditation only	A strategy with only the component: "licensing, certification, accreditation, or registration."
Improvement collaborative only	Improvement collaborative, as defined by authors of the report. No other components.
Client-oriented, provider-efficient (COPE)® method[b] only	A "COPE®-like" strategy was defined as having all the following components: provider self-assessment, continuous quality improvement (includes team-based problem solving), and peer review. No other components.
Standards-based management and recognition (SBM-R)[b] only	An "SBM-R-like" strategy was defined as having all the following components: standard health facility specifications were introduced, health facility received recognition after meeting certain criteria, health care provider self-assessment, team-based problem solving, supervision, and low-intensity training (according to HCPPR's definition, which also includes informal education by a peer).

[a] Excludes academic detailing and informal education by a peer. Also, training is allowed to have job aids or printed educational materials for health care providers.

[b] No studies involving COPE® or SBM-R were included in the HCPPR database. Using variables from the database, COPE®- and SBM-R-like strategies were constructed for the analysis.

SOURCE: Rowe, 2015.

constructed COPE®-like and SBM-R-like variables based on consultation with Carmela Cordero and Edgar Necochea.

Studies on provider performance use a wide range of outcomes. Some researchers look at the effect of their program on mortality rates, others on taking a patient history; such disparate outcomes cannot be compared in the same analysis. Therefore, the analysis presented included only outcomes on the process of care expressed as a percentage (e.g., the proportion of patients correctly assessed, diagnosed, or treated). The analysis was also stratified by two main groups of providers. Rowe reasoned that lay health workers are different from health professionals (e.g., doctors, nurses, pharmacists) in important ways, and studies on these populations should be dealt with separately.

Rowe used methods based on the Cochrane Effective Practice and Organisation of Care recommendations for determining risk of bias (Higgins et al., 2008). In this system, risk of bias is a function of study design, the number of clusters in each arm, data completeness, between-group comparability at baseline, the outcome's reliability, concealment of allocation for studies randomized at the patient level, the likelihood that the intervention could change data collection, and having fewer than six data points before or after an intervention for interrupted time series. Studies were coded as having low, moderate, high, or very high risk of bias.

Analytic Strategy

Estimates of effect size were expressed in terms of absolute percentage point change using the equation:

$$size = (Follow\ Up - Baseline)_{intervention} - (Follow\ Up - Baseline)_{control}.$$

Interpreting the difference of differences calculation is straightforward: if the intervention group sees the proportion of patients correctly treated rise to 50 percent from 20 percent, that is a 30 percentage point improvement. If the control group sees an increase of 5 percentage points, then the difference of differences is 25 percentage points; so, for every 100 patients seen, 25 are correctly treated in a way attributable to the intervention. This calculation has an added intuitive appeal because positive values indicate an improvement. (For studies that were designed to show a decrease in certain outcomes, the calculation was flipped.) For interrupted time series data, the investigators used a similar approach, but with values derived from segmented regression modeling.

The analysis mainly considered comparisons between a particular strategy and some kind of control; head-to-head comparisons of different methods were not included. If a study reported more than one primary

outcome, the median effect size was the statistic of interest. Investigators compared median effect size distributions using interquartile ranges and a weight calculated as:

$$(1 + ln\ (\textit{the number of providers or sites})).$$

For strategies with fewer than five studies, investigators used the unweighted median.

Exploratory analysis in the larger database and a priori knowledge of the topic suggested possible confounding effects, so the investigators attempted to adjust all the results to a partly standard context, the result that might have been observed if all studies had a similar baseline. To this end, they adjusted the analysis for two main effect modifiers: baseline performance and a public health facility setting. Rowe explained that, regardless of the strategy, better providers or clinics have less room for improvement; for every 10 percentage point increase in baseline performance, the observed effect size decreased by 2 percentage points on average. Therefore, for every 10 percentage points over the average baseline performance, investigators added 2 percentage points to the effect size estimate.[1] Similarly, the mean effect was 8 percentage points higher in a public facility than in any other setting, regardless of strategy. As about half of the effect sizes were from studies with a public facility–only setting, the adjustment subtracted about 4 percentage points from effect sizes of studies with a public facility–only setting and added 4 percentage points to effect sizes from other settings.

For strategies that appeared to have the greatest effectiveness, the analysis checked for confounding by limited variability, the chance that the observed effect came from an idiosyncrasy of study design—that is, a setting unusually well suited to a strategy. When this confounder was a concern, the investigators broadened the definition of a strategy to include more studies with the same basic strategy components. If the adjustment brought about a large decrease in the estimate of effect size, confounding is likely.[2]

Rowe also briefly discussed his plans for secondary analyses. In the future, the group will attempt different adjustments for sample size and ways to summarize effect estimates. Network meta-analysis, a relatively

[1] For further clarity, the average baseline performance in the database is 41 percent, meaning that for every 100 patients supposed to be treated or diagnosed, 41 get the service as it is meant to be done. If the baseline in a particular study were 10 percentage points above the mean, or 51 percent, then the adjustment increased the effect size by 2 percentage points.

[2] For instance, Rowe gave a hypothetical example of three studies of licensing suggesting an effect size of 50 percentage points, but, by broadening the definition of the licensing strategy to include studies of licensing with other components, the median effect size would fall to only 17 percentage points. In such a case, it would be prudent to conclude the effect of licensing is about 17 percentage points.

TABLE 4-2 Number of Studies in the Health Care Provider Performance Review Database Comparing an Active-Strategy Group with a No-Intervention or Historical Control Group

Strategy	Number of comparisons
High-intensity training only	9
Low-intensity training only	36
Supervision only	7
Accreditation only	0
Improvement collaborative only	7
COPE®-like strategy only	0
SBM-R-like strategy only	0
High-intensity training + supervision	4
Low-intensity training + supervision	5
Low-intensity training + improvement collaborative	3

SOURCE: Rowe, 2015.

new analytic technique, will allow for the inclusion of head-to-head comparisons. Explaining why this was not the primary analytic strategy, Rowe raised concerns with the validity of the sample sizes and, therefore, with analytic weighting, especially in the older studies. CONSORT[3] guidelines require investigators to report a fair amount of detail on the design and conduct of randomized trials, such as the number of subjects per cluster, the number of clusters per study arm, the unit of randomization, and the inter-cluster correlation coefficient. Many of the studies in Rowe's database pre-date such reporting requirements, making the more conservative approach preferable.

Preliminary Results

From the more than 105,000 citations screened and 824 reports included in the database, 66 were eligible for the analysis presented at the IOM workshop. Table 4-2 shows the number of comparisons in Rowe's database on the six strategies discussed at the workshop. After removing the strategy groups not mentioned in the database, only 11 percent (n = 7) of the studies had low risk of bias, 23 percent (n = 15) had moderate risk

[3] CONSORT, or Consolidated Standards of Reporting Trials, is a set of guidelines for reporting the results of trials that came into common use in the mid-1990s, with an extension for cluster randomized trials coming into use in the early 2000s.

of bias, 36 percent had high risk of bias (n = 24), and 30 percent (n = 20) had very high risk of bias. Table 4-3 shows the breakdown of risk of bias by study strategy.

Thirty-three countries were represented in this analysis, 56 percent of which were low-income countries. Figure 4-1 gives more detail on the geographic breakdown of the data. Almost half of the studies were randomized

TABLE 4-3 Breakdown of the Risk of Bias in the Strategy Studies

Strategy	Number of comparisons	Risk of bias: Low/ Moderate	Risk of bias: High/ Very high
High-intensity training only	9	4	5
Low-intensity training only	36	15	21
Supervision only	7	3	4
Improvement collaborative only	7	0	7
High-intensity training + supervision	4	0	4
Low-intensity training + supervision	5	2	3
Low-intensity training + improvement collaborative	3	0	3

SOURCE: Rowe, 2015.

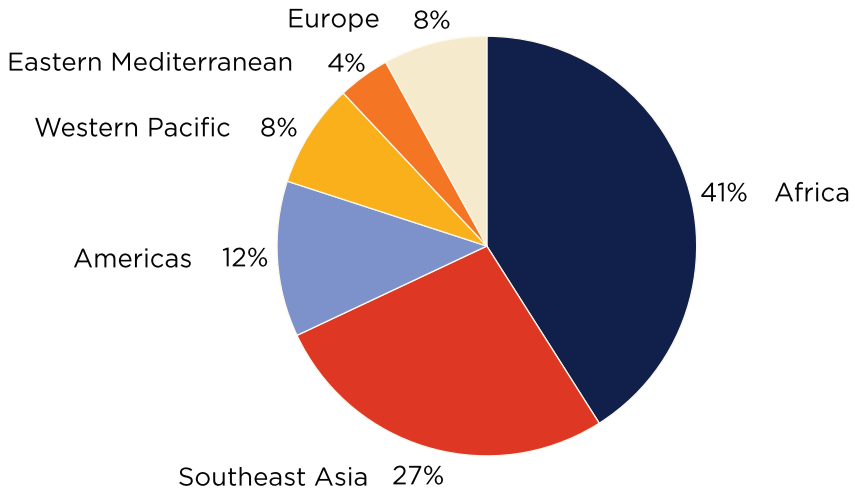

FIGURE 4-1 Breakdown of the strategy studies by region.
SOURCE: Rowe, 2015.

controlled trials (see Table 4-4); most studies were conducted in the 1990s and 2000s.

Rural, urban, peri-urban, and mixed settings accounted for similar proportions of the 66 studies, and there was fairly wide representation of different types of facilities (see Table 4-5), providers (see Table 4-6), health conditions (see Table 4-7), and methods of data collection (see Table 4-8).

Figure 4-2 shows the weighted median adjusted median effect sizes (MES) and interquartile range (IQR) for the strategies being discussed. The combination of improvement collaborative and low-intensity training showed unusually high effectiveness. The effectiveness, and the small number of studies informing the comparison, caused the investigators to suspect confounding by limited variability. After broadening the strategy inclusion criteria somewhat, the effectiveness estimate for the strategy declined from

TABLE 4-4 Breakdown of the Strategy Comparisons by Study Design

Study design	Number of studies (%)
Pre-post study with randomized controls	27 (41)
Pre-post study with non-randomized controls	22 (33)
Interrupted time series	13 (20)
Post-only with randomized controls	4 (6)

SOURCE: Rowe, 2015.

TABLE 4-5 Breakdown of the Strategy Studies by Setting

Setting	Number of studies (%)
Urban or peri-urban area only	20 (30)
Rural area only	15 (23)
Mixed setting	19 (29)
Public or governmental only	40 (61)
Any private sector	14 (21)
Other (e.g., household)	12 (18)
Outpatient health facilities	39 (59)
Hospital outpatient departments	20 (30)
Hospital or health facility inpatient wards	15 (23)
Community settings	8 (12)

SOURCE: Rowe, 2015.

TABLE 4-6 Breakdown of the Strategy Studies by Health Care Provider Type

Type of provider	Number of studies (%)
Nurse or midwife	36 (55)
Physician	35 (53)
Nurse or midwife aide	26 (39)
Pharmacist/lab worker	10 (15)
Paramedic	9 (14)
Lay health worker	8 (12)
Health educator	6 (9)

SOURCE: Rowe, 2015.

TABLE 4-7 Breakdown of the Strategy Studies by Health Condition

Health topic	Number of studies (%)
Multiple (or all) health conditions	27 (41)
Acute respiratory infections	10 (15)
Diarrhea	10 (15)
Pregnancy	8 (12)
HIV/AIDS +/− other sexually transmitted diseases	8 (12)
Newborn health	4 (6)
Malaria	3 (5)
Malnutrition	3 (5)
Reproductive health (not pregnancy)	2 (3)
Tuberculosis	1 (2)

SOURCE: Rowe, 2015.

60 percent (IQR: 30 to 76 percent) to 11 percent (IQR: 6 to 60 percent) (see Table 4-9). The sharp drop in the effectiveness estimate suggests the studies informing the 60 percent estimate have low generalizability.[4] Rowe also pointed out the risk of bias in the data. The strategies with the highest

[4] After the discussion, participants asked if broadening the definition of a strategy changed the effect estimates for other strategies. Rowe said they have observed something similar, though not on the same scale, in the larger database when considering the group-based problem solving and training with supervision strategies.

TABLE 4-8 Breakdown of the Strategy Studies by Method of Data Collection

Method	Number of studies (%)
Record review	46 (70)
Interview with patient or caretaker	17 (26)
Interview with health care providers	14 (21)
Observation of health care provider–patient interaction	13 (20)
Questionnaire for health care provider	10 (15)
Simulated client	4 (6)
Physical exam of patient	2 (3)

SOURCE: Rowe, 2015.

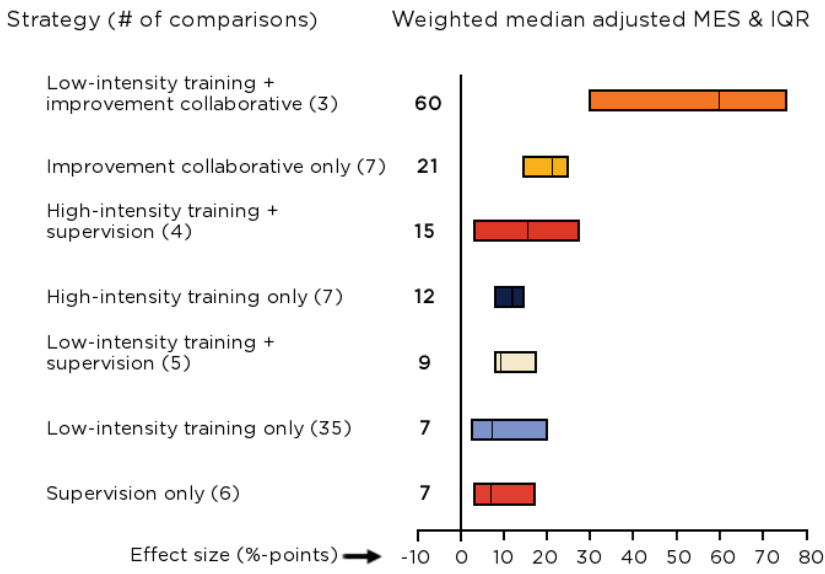

FIGURE 4-2 Weighted median adjusted median effect sizes (MES) and interquartile range (IQR) for selected strategies. For the three strategies with greatest MES, all studies had high or very high risk of bias (see Table 4-3).

SOURCE: Rowe, 2015.

effect estimates come from 14 studies whose risk of bias was classified as high or very high (see Figure 4-2).

In an effort to determine where different strategies work best, Rowe stratified the results by country income level. This analysis, though con-

TABLE 4-9 Broadening the Definition of Improvement Collaborative with Low-Intensity Training to Adjust for the Probable Confounding of Limited Variability Sharply Reduces Estimates of Effectiveness

Strategy group	Number of comparisons	Median adjusted MES (IQR)
Original definition Improvement collaborative + low-intensity training	3	**60** (30, 76)
Broadened definition Group problem solving + low-intensity training +/− other components	6	**11** (6, 60)

SOURCE: Rowe, 2015.

ceptually useful, was practically prevented by small sample sizes. Similarly, in attempting to determine the strategies that work best with lay health workers, Rowe identified only four comparisons, all from low-income countries and with a high or very high risk of bias.

A post-hoc analysis aimed to identify factors associated with the effectiveness of training and supervision. The analysis used mixed linear regression modeling and drew on all training and supervision studies in the database (not just the 66 studies identified as relevant to the IOM workshop). The data indicate that, for trainings on one topic, the length of training is not associated with its effectiveness. When several topics are covered in the same training, however, the training's effectiveness increases by 1 or 2 percentage points for each added day. Only after 5 days do the trainings on multiple topics reach the same effectiveness as a single-day training on one topic. Regarding supervision, strategies emphasizing feedback to the provider appeared to be about 11 percentage points more effective than other types of supervision. Rowe cautioned against over-interpreting these results, however. The models had serious problems with missing data and, therefore, substantial risk of bias and confounding.

Study Limitations

In discussing the limitations of the study, Rowe mentioned the lack of studies on COPE®, SBM-R, and accreditation in the database and how attempts to make COPE®-like study and SBM-R-like study variables failed. He also cited limitations in the studies themselves, such as lack of detail on strategy and context, lack of standardization, difficulty in assessing study precision and strength of implementation, and high risk of bias.

Discussion

Because of insufficient information about COPE®, SBM-R, and accreditation, it was not possible to compare the effectiveness of the six target strategies. Rowe expressed interest in including gray literature from EngenderHealth and Jhpiego in future additions to the database. As for accreditation, it might simply be a strategy that is not implemented without other components, so its effectiveness should be studied this way.

Results indicated modest effectiveness for both training and supervision. For facilities with a baseline performance of about 40 percent, training and supervision of providers can boost performance to about 50 percent. Combining the two strategies seems to work better than employing either alone, though conclusions about the combination of high-intensity training and supervision should be guarded because of the high risk of bias in these studies. Similarly, improvement collaboratives showed somewhat larger effect estimates, but such results should be interpreted cautiously because of the high risk of bias and limited variability in the data.

Rowe concluded his comments with a request for more studies of rigorous design. He recognized USAID's need for more information about which strategy is the best investment and encouraged the agency to fund the research that could help determine that, particularly head-to-head comparisons of different strategies. Still, it is not clear who should be responsible for such research. Implementing organizations may have a conflict of interest in evaluating their own programs. Rowe expressed some regret that the database does not include information on the study funder, as people often ask how that might influence the results.

In the session that followed the presentation, participants discussed the study methods. Concerning the adjustment for baseline performance, one participant asked if the higher effectiveness in places with poor baseline performance might be a reflection of the effectiveness of the strategy rather than the low baseline. Rowe agreed, pointing out that this raises a larger question of how to handle contextual factors in analysis. One approach would be to adjust for them until all the comparisons are between similar strategies. Another would be to stratify the analysis by important contextual factors and see how the data vary within strata. These are the kinds of analyses the team plans to undertake in the future.

Rowe was also asked whether his team had considered the scale of the intervention, if the effectiveness changes when 10 providers are involved versus 100 or 1,000. He explained some possible analyses that could address this question, using the number of providers or facilities as a proxy for study scale. The data indicate that, as the study scale gets larger, effectiveness gets smaller. This analysis is subject to risk of confounding, however, as certain strategies tend to be implemented at larger scale than others.

Several speakers asked Rowe how to distinguish between bias and suitability of strategy to setting. He explained that working in a place friendly to a particular strategy is not itself a bias, but it does limit the generalizability of the results. The contextual factors that improve a place's receptivity to quality improvement are difficult to measure. Identifying these factors is another important question for implementation research.

HOW DO WE KNOW IT WORKS?: MEASURING CHANGES IN QUALITY

Bruce Agins of New York State Department of Health AIDS Institute started the next session with a brief review of the role for information systems in measuring performance. He described a goal of using real-time data in clinics for continuous improvement. This would allow managers to link process improvements at clinics or in certain regions with corresponding changes in health. Better attention to data collection can also build the health system's capacity for quality improvement. To illustrate these points and to give an example of the feasibility even in a poor country, Agins shared a case study of a program to improve information technology in the Haitian ministry of health. The President's Emergency Plan for AIDS Relief (PEPFAR) and the Global Fund[5] support the program, both with direct contributions to the ministry and through local provider networks.

The CDC funded two Web-based electronic records systems for the Haitian PEPFAR program: the Monitoring, Evaluation and Surveillance Interface, or MESI, and an electronic medical records system used in 80 percent (n = 144) of the Haitian antiretroviral clinics. In the early days of the program (2004–2005), the monitoring and evaluation system was used primarily for collecting data to report to donors. Then in 2005, I-TECH, a health systems development organization run by the University of Washington and the University of California, San Francisco, with support from PEPFAR and the CDC, developed iSanté, an electronic medical records system that supports both individual and population health. iSanté gave clinicians a way to manage longitudinal data and to make data easily accessible to the ministry.

iSanté is an open-source system developed using Linux OS, Apache Web server, MySQL database, and PHP scripting language. When the program was new, the main servers were kept in Seattle, but eventually they were transferred to Haiti. As of April 2015, more than 100 sites—including government clinics, teaching hospitals, mission hospitals, and nongovernmental organizations—use iSanté to manage more than 500,000 patient records. Eighty-seven of the clinics have local servers, eliminating

[5] Officially, the Global Fund to Fight AIDS, Tuberculosis and Malaria.

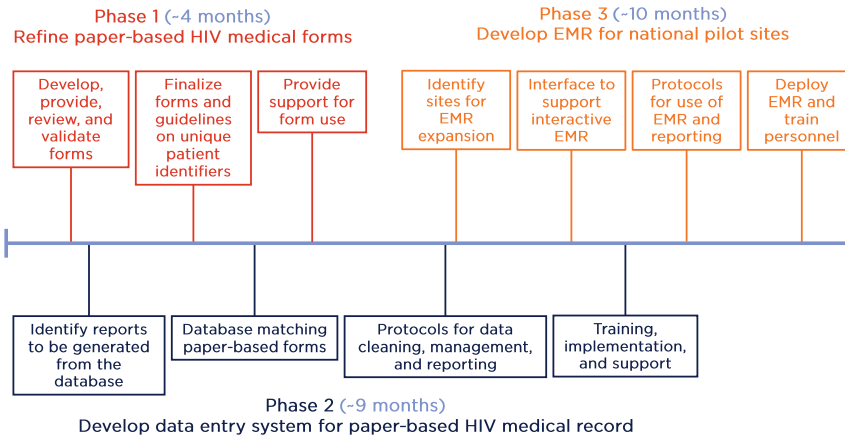

FIGURE 4-3 The timeline for developing the iSanté electronic medical records (EMRs) system.

SOURCE: Agins, 2015.

reliance on slow Internet connections and allowing for automatic backup to a central data repository.

Rolling out the iSanté system took about 1 year (see Figure 4-3) and required input from an electrical engineer, a database specialist, a programmer, an analyst, and a network specialist. Ten full-time staff operate the system, and health care workers need a 2-day orientation before they can use it. The cost of implementing the system was $400,000 plus an additional 10 percent annually for ongoing maintenance and technical assistance. Agins acknowledged that the low cost can be attributed in part to idiosyncrasies of the Haitian labor market: Haiti has a proportionately large number of workers qualified for technology jobs, and pay scales in Haiti are much lower than in most other PEPFAR countries. But, even considering the additional ongoing costs, iSanté was inexpensive and good value for the money.

The iSanté system allows the health worker to make different kinds of retrospective or prospective reports, generate case lists, or set reminders. This applies at the level of the patient in clinic, but also at the district or national level. Agins chose one indicator, the enrollment of eligible patients on antiretroviral therapy (ART), and shared it as an example of how electronic medical records can be used to improve quality.

Clinic managers used iSanté to generate their baseline data on ART enrollment. Then, using plan-do-study-act cycles, the clinic staff were able to test the effectiveness of three strategies to improve enrollment, with the goal of increasing enrollment of eligible patients by 20 percentage points. The iSanté system afforded the clinic managers confidence in their data and

the ability to see the effect of their changes almost instantly. When they found that reducing the amount of time patients spend waiting to start ART had the best effect on enrollment, they were able to give more attention to that strategy.

Haiti's national quality advisory board also made use of the electronic medical records system. The real-time data allowed them to set goals, measure their performance, and give feedback to the district and local health offices about their progress. So, in fiscal year 2012, when the national target was to add 10,000 patients to ART, every participating clinic was given a specific goal for enrollment. Analysts used iSanté data to identify the main problems preventing new enrollment and to test different solutions; there was no need to marshal a separate data collection and analysis program. As of September 2014, iSanté data indicated that 83 percent of eligible patients were on ART.

Agins then briefly introduced MESI, the Haitian electronic monitoring and evaluation system. Like iSanté, MESI was a relatively inexpensive program to set up: including recurring maintenance and technical assistance, the platform cost about $200,000. Both systems link local, district, and regional data and enable mandatory case reporting and mandatory reporting of aggregate epidemiological data. Because the value of the electronic systems was quickly evident to the Haitian government and donors, there has been enthusiasm for broadening the platforms to include primary and chronic disease care.

The Haitian example underscores the promise of electronic management systems to improve quality in low-income countries; its success might be replicated, assuming there is political commitment for change. Ministries of health are complicated organizations; in Haiti, the senior leadership was willing to cooperate across departments and allocate staff to the project. Though relatively modest, the additional staffing demands can surpass the workforce capacity in the least developed countries. Agins emphasized that the investment in electronic systems can contribute to sustainability because the platform helps improve efficiency and integrate different work streams, ultimately establishing the link between quality improvement programs and improved health.

In the next presentation, Niek Klazinga of the Organisation for Economic Co-operation and Development (OECD) built on Agins's closing remarks. He stressed the importance of an information infrastructure for health systems strengthening and reminded the audience to be mindful of how their programs contribute to developing this infrastructure.

Measuring the performance of health systems has been a priority for the OECD for many years. Since 2005, they have used a framework guided by the IOM report *Crossing the Quality Chasm* (see Figure 4-4). This framework recognizes four functions of the health system: effectiveness,

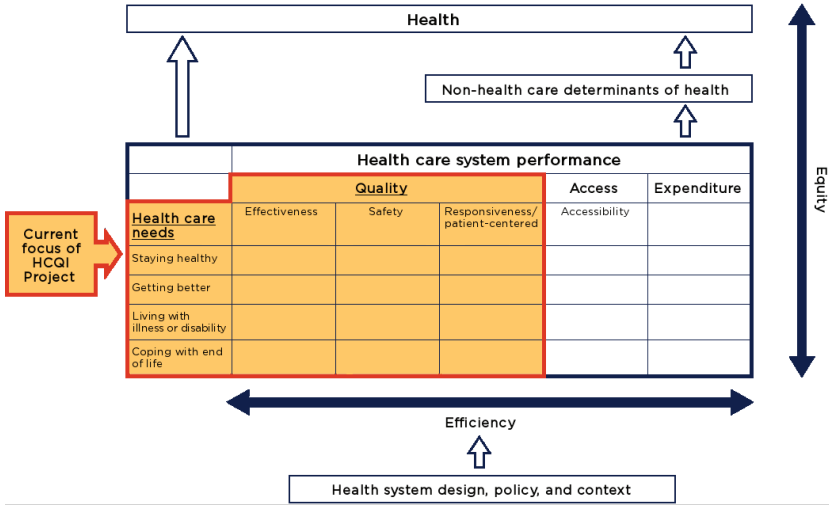

FIGURE 4-4 The OECD health system performance assessment framework.

NOTE: HCQI = Health Care Quality Indicators.

SOURCE: Arah, O. A., G. P. Westert, J. Hurst, and N. S. Klazinga. 2006. A conceptual framework for the OECD Health Care Quality Indicators project. *International Journal for Quality in Health Care* 18 (Suppl 1):5–13. Reproduced by permission of Oxford University Press and the International Society for Quality in Health Care. Adapted from Kelley and Hurst, 2006. As presented by Klazinga on January 28, 2015.

safety, patient-centeredness, and accessibility. The OECD tracks hundreds of indicators that give information about different functions of the health system; to simplify their work for the audience, Klazinga pulled out the set of key indicators most often used to track quality (see Table 4-10). By tracking these indicators over time, the OECD analysts have determined that, for example, survival of patients with acute myocardial infarctions is improving in all its member countries. While richer countries clearly have an advantage, the data show an association between better health outcomes and the amount of policy attention a country gives to quality improvement.

The same information about health outcomes is not readily available in low- and middle-income countries. The joint OECD/World Health Organization (WHO) report *Health at a Glance: Asia/Pacific* aims to present similar data for all countries in the Asia-Pacific region. Many countries in this region are trying to measure more than the minimum (e.g., infant mortality and vaccination rates) and therefore are investing in improved cancer registries, death registries, and electronic administrative databases; the growing momentum for universal coverage has driven some of these improvements. But tracking the 30-day case fatality rates of chronic disease remains challenging. The ability to link patient data across registries is an essential prerequisite to monitoring most of the OECD's key indicators.

TABLE 4-10 Types of OECD Health Systems Indicators

Infectious diseases	• Vaccination among children • Flu vaccination among the elderly
Mortality	• Infant mortality • Maternal mortality
Acute care	• 30-day case fatality following acute myocardial infarction • 30-day case fatality following stroke
Primary care	• Hospital admission for chronic conditions (diabetes, asthma/chronic obstructive pulmonary disease, chronic heart failure) • Prescribing of antibiotics
Cancer care	• Screening • Mortality • 5-year survival
Mental health	• Excess mortality in persons with severe mental health problems
Patient safety	• Post-operative complications (sepsis, deep vein thrombosis/pulmonary embolism) • Obstetric complications
Patient experiences	• Respect • Autonomy • Communication

SOURCES: Klazinga, 2015. Drawn from OECD, 2013, and Carinci et al., 2015.

Klazinga recommended that the first priority for low- and middle-income countries be developing clinical registries and administrative databases.

Over the past 5 years, the OECD quality experts have been asked to analyze the national quality strategies for a growing list of the non-OECD countries. When they give input, the OECD experts look at the way the health system is governed—for example, the methods used to ensure health worker competence, hospital quality, and the safety of medicines and equipment. The evaluation gives some attention to the country's efforts at standardization of processes, the use of guidelines, consideration for patients' rights, and the prominence of a safety strategy. Klazinga reminded the audience that all quality improvement strategies, including the six being discussed at the workshop, complement the national quality improvement architecture. What he described as the three management strategies (COPE®, improvement collaborative, and SBM-R) help develop the overall management culture in a country. Training and supervision can accompany the continuing education and professional development plans. Accreditation is a part of the safety strategy and a way to promote standard guidelines. The OECD's initial survey data have established that many of

the 26 countries in the Asia-Pacific region have a basic quality architecture in place (see Table 4-11); the task is to build on it.

Klazinga closed his presentation with points on ensuring successful project implementation. Management interventions, for example, generally have a certain overall philosophy; it is important to determine if that philosophy meshes well with local management styles. He encouraged training and supervision programs designed to complement local incentive and education structure. Involving policy makers and local experts, even in the program's pilot stages, helps ensure long-term success.

Following the presentations, participants shared their own experiences with medical information systems. William Tierney of the Indiana University School of Medicine mentioned national programs that are now under way in Bangladesh, Kenya, Mozambique, the Philippines, Rwanda, and Tanzania. There is an opportunity to use these systems—all at different stages of development—to build national quality improvement programs, though Tierney expressed concern that connections between the new electronic systems and quality managers were somewhat haphazard. Mike English built on this point, saying that the electronic medical records system he worked with in Kenya is not as sophisticated as the example Agins shared from Haiti. He found that the HIV care system was far superior, but parallel, to the national health information systems. Problems of interoperability, which Western countries are struggling with too, are vastly aggravated in Kenya, where various small entrepreneurs develop software packages and sell them to hospitals.

The participants also discussed the promise of electronic medical records systems, as well as some risks to control. Klazinga had concerns that most of the attention to information technology was in hospitals. Many OECD

TABLE 4-11 Number and Percentage of 26 Asia-Pacific Countries Reporting Selected Components of a Quality Improvement System

Component of national quality improvement system	Number of countries (%)
Mandatory continuing medical education or continuing professional development	16 (61.5)
Mandatory hospital accreditation	8 (30.8)
Voluntary hospital accreditation	13 (50.0)
Technology assessment studies on medicines	15 (57.7)
Standards on safe blood use	23 (88.5)
Pharmacovigilance system	21 (80.8)

SOURCES: Klazinga, 2015; OECD and WHO, 2014.

countries are now at a point where poor information about primary care is limiting their ability to monitor quality of care. It is not possible to know, for example, if someone is hospitalized for a chronic condition without integrating primary care and hospital databases. The problem could be avoided in low- and middle-income countries with attention to computerized information systems in primary care. DHIS 2, a health information system used in 47 low- and middle-income countries and 23 international organizations, has the ability to link hospital and primary care data and to control the burden of data collection on health workers (DHIS 2, n.d.).

There was also discussion of the best ways to measure quality in the poorest countries. Information technology, though full of promise, is going to remain out of reach for the poorest patients, many of whom are treated in their homes or in rural clinics. A participant pointed out that only half of the world's newborns are even weighed at birth and asked how, in such settings, we can ensure reliable information on chronic disease. Klazinga agreed and suggested that the most realistic first steps would be developing basic death registries and then cancer registries. He cited success from South America where cancer registries have developed rapidly over the past 5 years.

AT WHAT COST?: WHAT MAKES A PROGRAM COST-EFFECTIVE

In the last panel discussion of the day, Edward Broughton of URC and Dinesh Nair of the World Bank discussed cost-effectiveness and feasibility in quality improvement. Broughton opened his presentation with a simple example of cost-effectiveness in health. If a person knows his or her likelihood of getting sick, weighing the costs of treatment against foregone wages is relatively simple. But in quality improvement, although the basic principle is the same, the analysis gets complicated quickly. For example, if the USAID mission to Liberia did a quality improvement program in connection to the Ebola response, the epidemiology of Ebola would not be the only thing changing in the health system. Clinics that treat Ebola patients presumably treat others as well; attention to these conditions would improve, suffer, or stay the same (see Figure 4-5). Cost-effectiveness analysis needs information on the cost and health consequences of all those variables. Usually, as Figure 4-5 shows, the analysis accounts for costs and consequences associated with improving or not improving treatment. Accounting for changes in mortality requires sophisticated epidemiological modeling. Modeling depends on assumptions that are invariably debatable, leaving the final conclusions open to criticism.

For one thing, it is difficult to say how long the effects of a quality improvement intervention last. The initial intervention might go on for 6 months to 2 years. Modeling must make assumptions about how the im-

FIGURE 4-5 The possible outcomes accounted for in a hypothetical cost-effectiveness analysis of a quality improvement program in a Liberian district health center.
SOURCE: Broughton, 2015.

provements will continue after the technical experts leave. While it is probably not sensible to assume processes return to baseline after the project, it is also imprudent to suppose that all gains are sustained. Sometimes the local host agency expands the program after the USAID implementing partner leaves. Estimating cost-effectiveness requires making assumptions about how effective the implementation will be with different technical staff in charge. Finally, Broughton stressed that cost-effectiveness analysis generally takes the perspective of the costs to the funder—in the case of his organization that means the cost to USAID. A stronger model would consider the costs to society and thereby give the host country ministry the information it needs to determine cost-effectiveness.

After describing the factors that make health systems research difficult, Broughton advised that the only justifiable course of action was to accept the inherent challenges of such research and get on with it. The complexity of the research is not, he maintained, an excuse for not doing it. A quick review of cost-effectiveness studies indexed to PubMed found 2,962 articles about the cost-effectiveness of quality interventions, but only 10 (0.34 percent) of those papers actually published results for cost-effectiveness analysis (see Figure 4-6).

Of the 10 papers that published economic analysis of quality interventions, half were from the USAID Applying Science to Strengthen and Improve Systems (ASSIST) project. None of the analyses, including those from the ASSIST collaboratives, looked explicitly at the effectiveness of the different brand names being discussed at the workshop. In seven of the studies, researchers found the quality changes to be cost-effective; in the

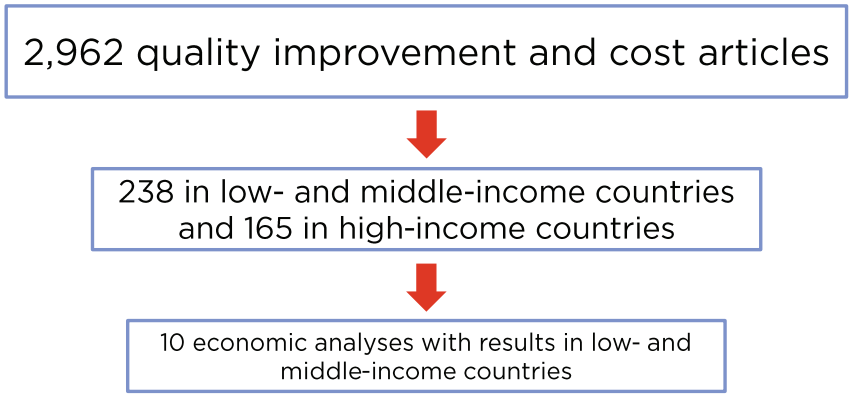

FIGURE 4-6 A quick literature review of cost-effectiveness studies indexed to PubMed found only 10 papers included cost-effectiveness analyses.
SOURCE: Broughton, 2015.

remaining three, the interventions were found to be cost-saving, meaning the change not only improved health but also saved money.

Broughton was frank about the limitations of these studies. None of the USAID-funded research used control groups. There were also problems with generalizability and short time frames. *BMJ* guidelines for cost-effectiveness analysis ask investigators to report certain information about the method of modeling, confidence intervals on key estimates, and sensitivity analyses; the guidelines give considerable attention to describing uncertainty in the data, threats to the study validity, and the assessment of outcomes. The 10 studies of cost-effectiveness of quality interventions did not rank highly on these criteria and would probably not have met publication standards for *BMJ* and similar journals. The fact that all 10 studies found positive results raises questions about publication bias. Because all the studies were evaluated by people close to the implementation, the risk of confirmation bias also seems high.

Broughton concluded his talk by observing that cost-effectiveness research is a serious weakness in quality improvement. He sees this weakness as threatening to undermine the whole quality improvement field and echoed a sentiment expressed in earlier presentations about the need for more randomized trials of improvement interventions.

Nair built on similar themes in his presentation on feasibility and cost-effectiveness in quality strategies. He agreed that economic evaluation of quality strategies—research weighing the costs and benefits of correcting inefficiency—is neglected. The few studies published in peer-reviewed journals tend to look at a narrow group of maternal and child health outcomes. More importantly, cost-effectiveness depends on the comparisons shown in Figure 4-7. This analysis can fail to balance costs against feasibility.

Nair felt each of the six methods being discussed at the workshop had feasibility issues. For example, the standards-based methods can face problems with inconsistencies in the national guidelines. Many of the methods have requirements for foreign technical support, decreasing the feasibility. Nair emphasized the value of evaluating the methods under discussion, and all quality improvement strategies, against a strong counterfactual.

In the final discussion of the day, the participants reflected on the limitations of the data. One participant brought up the possibility that the problem of poor quality data is not likely to change and suggested that global health researchers might do well to anticipate this, moving to Bayesian approaches. Classical statistics need a kind of data that quality improvement simply may not be able to provide. On the other hand, classical statistics command an authority in medicine where there is wide acceptance that decisions should be evidence based. Although health systems researchers in developing countries do not have the same resources available to them as

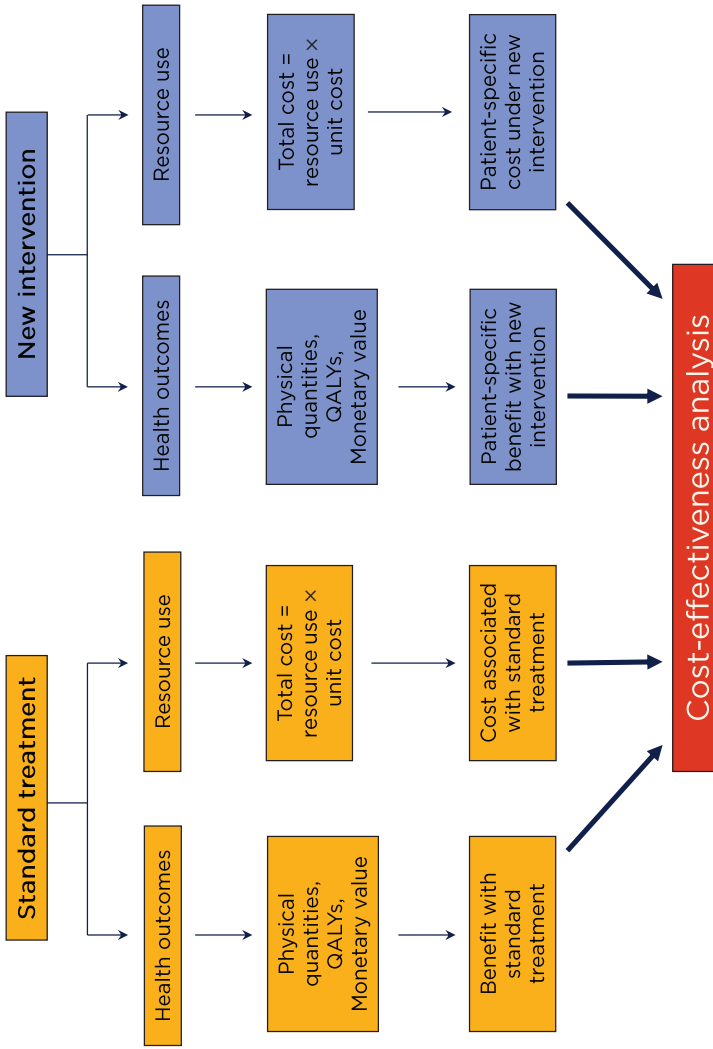

FIGURE 4-7 General structure of economic evaluations.

NOTE: QALY = quality-adjusted life year.

SOURCE: Favato, 2008, as presented by Nair on January 28, 2015.

other kinds of medical research, there is still a need for convincing evidence and an obligation to provide it. If the case can be made, with objective data, that quality improvement can actually save a country money, then ministers of health and finance will be eager to support it.

The discussion also touched upon implementation research with some participants acknowledging the constraints of working with foreign hosts. A host country agency may request, for example, that a project take place in 100 clinics in 2 districts; they often choose the districts and set some of the terms of the project. Some guests suggested that the implementers could try to persuade their hosts to implement the project in two phases, randomly assigning clinics to begin the intervention at different times in a step-wedge design. While not the strongest design, such models would be an improvement on what is often done now.

Participants highlighted opportunities for future quality improvement work to confront the weaknesses discussed at the workshop and improve on them. Studies can be made stronger by articulating clear aims and tying those aims to the improvement strategy being used. Project evaluations can be given better priority by planning early, allocating time and money for evaluation from the start. Other participants questioned the relevance of the experimental research design for quality improvement given the importance of context. Some of their colleagues responded that one way to address this concern is for researchers to describe the context and limitations of their work, so that future work can weigh the likely relevance of a given study to their setting.

A slight contradiction was identified between the first sessions of the day, where experts on the six strategies emphasized the importance of data collection to each of their methods, and the last sessions, where there was a consistent concern that no one has data about quality improvement. Although the morning speakers explained how quality improvement programs collect considerable data, Broughton made the point that quality improvement research requires information about a comparable control group. Jishnu Das of the World Bank pointed out that publicly available data, such as the Demographic and Health Surveys (DHS), can be valuable in closing the data gap. Triangulating controls off DHS data could help solve the problem of the poor control group, though DHS data is sometimes not de-aggregated below the district level.

CLOSING REMARKS

After the discussion, Kedar Mate of the Institute for Healthcare Improvement gave closing remarks on the day's proceedings. He described the workshop's main theme as improving the science of quality in low- and middle-income countries. Implementation experts, policy makers, academics,

and researchers may come at the problem from different perspectives, but the goal is the same.

Over the course of the day, participants heard a good deal about accreditation, COPE®, collaboratives, SBM-R, supervision, and training, but less about the wider quality improvement movement. Mate echoed Ashish Jha's judgment that the six methods are more similar than different, and that the ultimate success or failure of a strategy has as much to do with its suitability to a particular environment as its technical merits. At the same time, it is imprudent to be too confident about the value of any method, as some of the most promising results come from the weakest study designs. Several presenters had brought up the need for more head-to-head comparisons of different strategies. There was similar demand for independent evaluation of quality improvement work, with the distance between evaluator and implementer allowing for more confidence in the results.

5

Synthesizing Evidence, Identifying Gaps

Key Points Made by Individual Speakers

- Outstanding clinicians are often tapped for executive positions in hospitals and provider networks, though little in their formal training qualifies them for management. Training health executives in management might make for leaders who are more receptive to quality improvement in health. (Ruelas)
- There is a pronounced gap between what providers know and how they practice. Closing this gap probably depends on training and on involving the private sector. (Das)
- The vast majority of strategies in the systematic review were tested only once or twice, often in studies with high risk of bias. Without more consistent attention to research questions, it will not be possible to build a convincing body of evidence on any strategy. (Rowe)
- In the absence of clear treatment guidelines, market forces and patient demand will have a strong influence on provider behavior. Therefore, emphasis on the frontline health worker to improve quality might be misplaced. (Das, Heiby, Laxminarayan, Mor, Ruelas, audience member)
- Discussions about quality would be eased by a common vocabulary of terms used to describe different strategies. (Rowe)

The second day of the workshop opened with brief welcoming comments from Sheila Leatherman. Then Enrique Ruelas of the Monterrey Institute of Technology and Higher Education gave the opening address. His comments broadened the scope of the discussion beyond the six methods,

setting the topic squarely within current policy debates about health equity and human rights. He stressed the importance of treating patients with dignity as a fundamental human right. Respectful treatment builds trust between patients and providers. Over time, this translates into wider use of services and better health.

The concept of health as a human right has been central to the recent discussion about universal coverage, and improving quality protects that right. Ruelas took issue with a point in the universal coverage discussion: the policy documents, particularly the ones released by the WHO and the Pan American Health Organization, do not reference quality beyond a cursory mention of "quality services." While considerable attention is paid to questions of access, funding, professional training, and infrastructure, quality of care is neglected.

Ruelas also cautioned the audience against lumping low- and middle-income countries together, describing the differences among low-, lower-middle-, and upper-middle-income countries (see Table 5-1). Given the variability in the baseline health indicators and available funding, it seems wise to consider quality improvement strategies differently for each category.

In reflecting on the previous day's discussion, Ruelas shared some of his experiences from working in quality improvement in Mexico. When he started in the mid-1980s, the country was in a financial crisis. There was some resistance to the idea of investing in quality when so many people had no access to services. Ruelas argued that the need for quality was greater precisely because so few people had access to services. He understood the concern that accreditation standards, for example, may seem unrealistically high in poor countries; he had felt the same way when responsible for oversight of accreditation in Mexico. Nevertheless, he saw a vicious cycle in lowering the standards or removing the items that seem too difficult to attain. Once a standard is taken out of consideration, the health workers

TABLE 5-1 Variability in Key Health Indicators by World Bank Income Group in 2013

	World Bank income group		
Health indicator	Low	Lower-middle	Upper-middle
Life expectancy at birth (years)	62	66	74
Child mortality rate (per 1,000 live births)	76	59	20
Per capita spending on health (current US$)	36	88	479

SOURCE: World Bank, 2015.

will never reach it. He counseled the workshop audience to keep this in mind as they try to strike the difficult balance between setting achievable goals and encouraging progress.

Making progress at a large scale—moving beyond a single clinic or village to an entire district or hospital system—often requires the cooperation of the central bureaucracy, and good programs can help ensure this by empowering staff to work with the government. Ruelas acknowledged that sometimes this process is beyond the control of health staff. In government, central ministry staff may change with the administration, and the new political appointees might not share an interest in quality improvement. Even the best quality of care program can fail because of the larger political context.

Ruelas concluded with comments on the current gaps in quality of care programming. He discussed how to create an environment conducive to triggering change. Building such an environment takes years, he cautioned; it is not a matter of a few conversations with political leaders. In Mexico, they learned from a 1994 survey that people were unhappy with the quality of health services, but it was not for another few years that the national quality assessment data were available to substantiate these concerns. The national quality strategy came into place in 2000, in part because President Vicente Fox, the former Coca-Cola chairman for Latin America, understood the importance of quality control.

Data and political will are important for expanding quality programs, but Ruelas advised that these are not the only factors driving sustainability and scale. One interpretation of Rowe's analysis is that combining certain strategies is more effective than using the same tools singly. This was consistent with Ruelas's experience in Mexico, where the national quality strategy used 10 interventions. Using only one, he felt, would have been far less successful.

Along the same lines, Ruelas saw a gap between improving quality for the individual patient and improving the quality and efficiency of the overall system. Improving the systems means working with more than 10 clinics or 20 hospitals; the units of the intervention are an order of magnitude larger. Quality improvement at large scale requires clarity on the central message and a consistent incentive system. Scale also means involving the different networks. Diabetes patients, for example, are treated in primary care and occasionally hospitalized. Integrating quality work in hospital and primary care systems protects these patients from falling into gaps in the system.

Quality improvement, Ruelas explained, is ultimately about management, and a final gap he discussed was the training of managers. He observed that, in low- and middle-income countries, outstanding clinicians are often made chief executives of hospitals or put in charge of provider networks. The training and skills that qualify a clinician are different,

however, from those that make a good manager. It is not reasonable to expect that a retired clinician will have the same understanding of quality improvement and how to make changes as someone trained in management. Ruelas encouraged the audience to think more about how to provide essential management training to health executives.

The next session built on Ruelas's comments, as the speakers continued discussing the evidence and gaps in it. Jishnu Das of the World Bank gave the first presentation. He described the overwhelming similarities among the six strategies being discussed: all are based around health services, depend greatly on context, and are labor-intensive. There is no randomized controlled trial establishing the effectiveness of any one of them, and the analysis of these methods raises questions about what constitutes quality of care.

Das encouraged the group to first clarify if quality of care is a policy or a product and, if policy is the question of interest, whether or not the policy is explicitly a health policy. The World Bank research suggests that changes outside of health—improving roads, for example—can drive more improvement in infant mortality than health interventions. Das then shared some research on improving the quality of primary care in India, summarizing work his group has been doing for the past decade.[1]

On average, India has about 4.4 health providers in each village. The vast majority (77 percent) have no formal medical training. Although unlicensed practice is illegal in India, 2011 estimates indicate that it accounted for 70 percent of the first contact with primary care in the country.

Research using standardized patients found that about 4 percent of patients get the correct treatment and no incorrect treatment; 40 percent of the time patients get the correct treatment plus another treatment, and 75 percent of the time patients receive at least one incorrect treatment (Das et al., 2012). Other research indicates that ignorance of proper treatment protocol is not the problem. Das described this tension as "the know-do gap," a problem that has been documented in Canada, the Netherlands, Rwanda, and Tanzania, as well as India. When providers are presented with vignettes describing common problems such as diarrhea, unstable angina, and asthma, they respond with the correct treatment more than 80 percent of the time (see Figure 5-1). This is true regardless of whether they work in the public or private sectors, or even hold a qualifying medical degree. But when actual standardized patients present in their clinics with identical problems, 50 to 70 percent of survey providers give some other, incorrect treatment. The know-do gap shown in Figure 5-1 is most pronounced for the treatment of diarrhea: although on average 88 percent of providers

[1] Das presented ongoing work that he is doing in collaboration with Karthik Muralidharan and others.

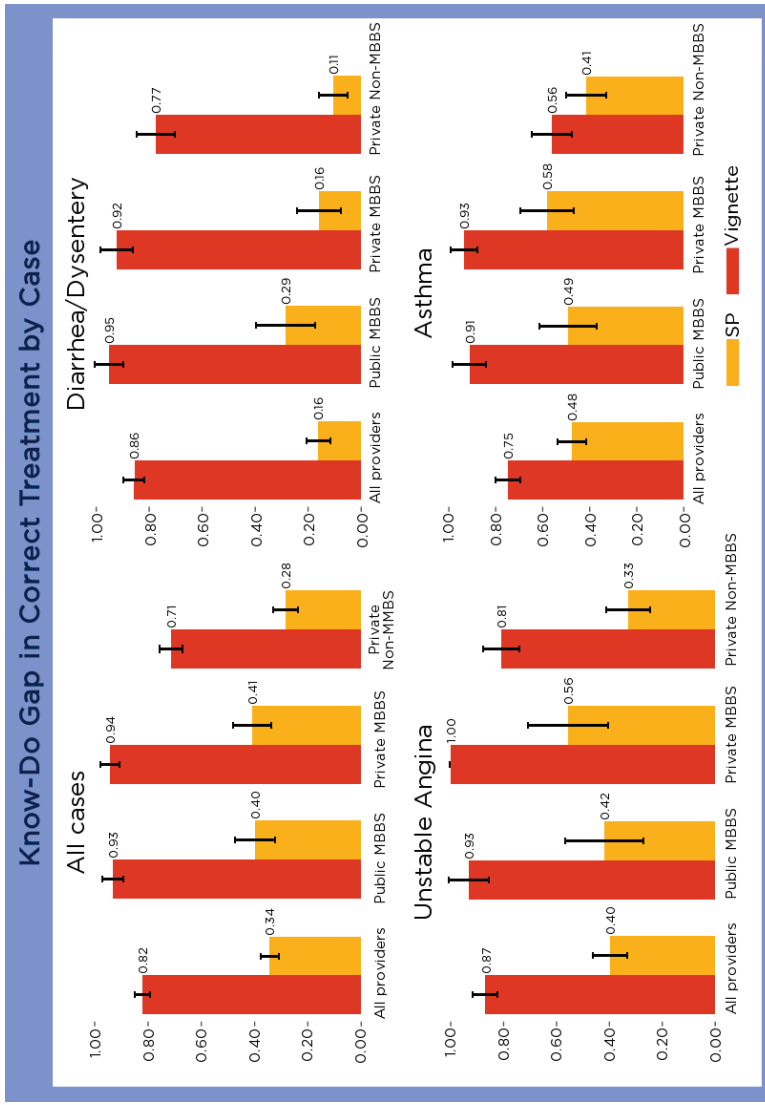

FIGURE 5-1 The know-do gap in correct treatment by case among different types of providers in India.

NOTE: MBBS = Bachelor of Medicine, Bachelor of Surgery; SP = standardized patient.

SOURCE: © 2015 Das, Hammer, and Mohpal (unpublished), as presented in Das, 2015.

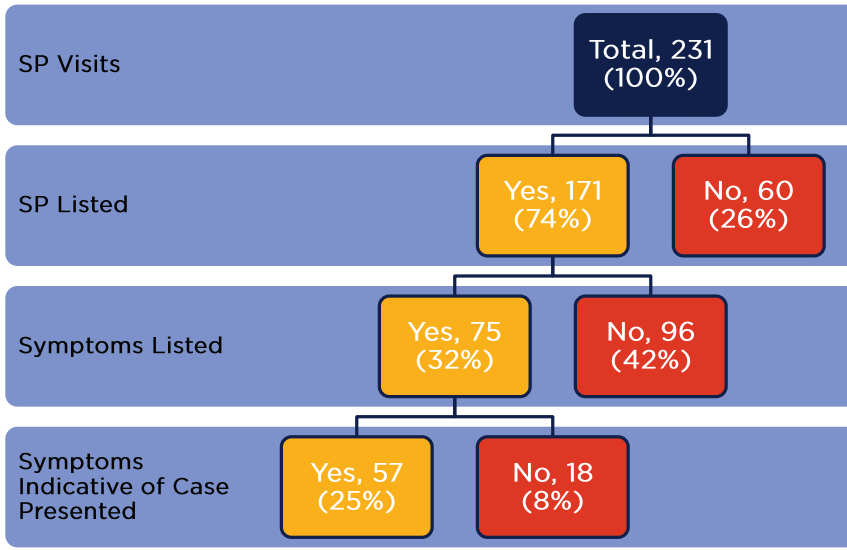

FIGURE 5-2 A review of clinic records 6 months after the standardized patient (SP) visits raised concerns about missing and erroneous information.

SOURCE: Analysis of MAQARI data by Aakash Mohpal and Jishnu Das (unpublished), as presented in Das, 2015.

know how to treat the condition properly, only 16 percent actually do.[2] Beyond knowing that the prescribed treatment is incorrect, it is not clear in what way it is incorrect—that is, if the provider's course of action would be properly classified as over- or under-treatment.

Auditing clinic records after presentation by standardized patients raises more concerns. Six months after the patient actors visited clinics, the investigators searched clinic records for the date of their visits using the patient's reported name (see Figure 5-2). While 71 percent of the standardized patients were listed in the clinic records, only 32 percent had any symptoms recorded, and of those only 25 percent of the records showed the symptoms the patient actually reported. In one case, a patient with chest pains and symptoms of angina was recorded as having a headache. Das concluded that the records are, for practical purposes, erroneous and not a valid data source for managers.

It is difficult to know how to fix a problem of such scope. Das outlined some solutions, starting with involving the private sector. In India, private sector clinicians provide 80 to 90 percent of care; some of these providers have long experience and apprenticeship-style training, despite having no

[2] These figures come from ongoing work by Das, Jeffrey Hammer, and Aakash Mohpal.

recognized degree. Others he described as "orphans of government and donor projects," often attempts at task-shifting programs. Targeting these providers is necessary, he said, but must be done for a reasonable cost. In Kenya, the government asked the World Bank to come up with a product that would cost $100–$125 per facility. In India, Das and his colleagues worked with a budget of $150 per provider.

Within this budget, the researchers randomly assigned 304 providers to receive an intensive training or to be in a control group (Das et al., 2015). The training, designed and implemented by the Liver Foundation of West Bengal, lasted 2 days per week, 4 hours per day, for 9 months; used multiple trainers and a variety of teaching methods; and covered multiple topics. Six months after the training ended, standardized patients were sent to all 304 providers to test the intervention's effectiveness. They found that the training increased the likelihood of correctly treating a standardized patient by 7–8 percentage points over a baseline of 52 percent. Trained providers improved adherence to a treatment checklist by 4–7 percentage points over the baseline adherence of 27 percent. At the same time, there was no difference between trained and untrained provider groups on measures of over-treatment (e.g., over-use of antibiotics, polypharmacy,[3] or injections). Das and his colleagues found that the improved services increased demand for care by two patients per day. At that rate, providers can recoup the cost of training in about 6 months.

The researchers also compared the trained private-sector providers to the public-sector, degree-holding providers in the same villages on various measures of quality (see Table 5-2). All informal private-sector providers spent more time with their patients, were less likely to prescribe antibiotics for asthma or myocardial infarctions, and were less likely to give unnecessary treatments or polypharmacy. However, trained public-sector providers were more likely to correctly diagnose and manage the patient's condition; training the private-sector providers reduced that difference by half.

Das concluded that randomized controlled trials of quality programs can yield a wealth of information for policy makers. He also emphasized the value of involving the private sector, given that private providers are responsible for more than 80 percent of medical care in India. He argued that the costs of such programs, between $100 and $150 per provider, are sufficiently modest to allow for large-scale, sustainable implementation.

Nynke van den Broek of the Liverpool School of Tropical Medicine then gave a presentation on the quality gaps in maternal and newborn health. Indicators of neonatal and maternal health are the most disparate in the world, making maternal and neonatal health a litmus test for the strength of the health system. Despite having ample information about

[3] The use of four or more medications by a patient.

TABLE 5-2 Proportion of Providers Adhering to Various Treatment Guidelines in the Public Sector (PHC), Private Sector Non-Intervention (control), and Private Sector Intervention (treatment) Groups

	PHC	Control	Treatment
Checklist – all	0.202	0.273	0.313
Correct treatment	0.667	0.520	0.594
Average quality treatment	0.182	0.114	0.174
Correct diagnosis	0.182	0.136	0.188
Consultation length (minutes)	1.735	3.252	3.495
Gave antibiotics	0.667	0.477	0.480
Antibiotics (asthma and myocardial infarction)	0.636	0.331	0.332
Offered injection	0.045	0.011	0.019
Treatment – polypharmacy	2.758	2.162	2.208

SOURCE: © 2015 Das, Chowdhury, Hussam, and Banerjee (unpublished), as presented in Das, 2015.

what works to protect mothers and newborns, there is little information about how interventions could be effectively bundled together to improve quality.

Effectively combining essential interventions for pregnancy and delivery is one goal of the Liverpool School of Tropical Medicine's Making It Happen program. The program aims to improve both the availability and the quality of obstetric and neonatal care. Quality improvement is supported using the audit method, which, though different from the six methods described at the workshop, has significant overlap with all of them. Van den Broek described audit as a way of asking questions. For example, it might ask, "why do mothers in this clinic die?" The audit method has a long history: a 1935 *BMJ* paper used the technique to estimate maternal mortality in Rochdale, near Liverpool, United Kingdom (Oxley et al., 1935). Audit has much in common with plan-do-study-act cycles; in audit, health workers gather information, review it, and make decisions based on it (see Figure 5-3).

Making It Happen makes use of audit and implementation research in the program's 11 target countries. Researchers answer questions such as whether emergency obstetric training is more effective when given alone or in combination with other quality improvement methods. Project staff first defined quality and came to the conclusion that, "Quality of care is the degree to which maternal health services for individuals and populations increase the likelihood of timely and appropriate treatment for the purpose of achieving

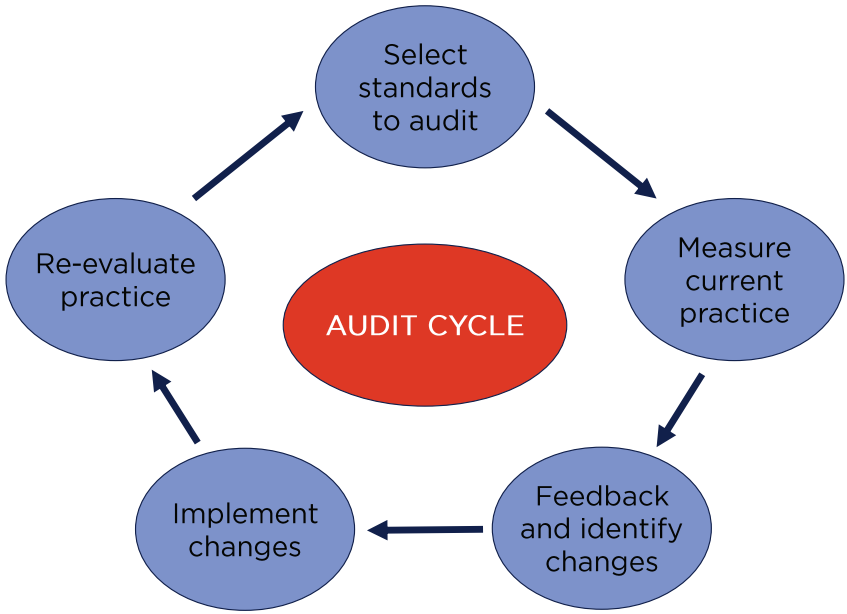

FIGURE 5-3 The audit cycle.
SOURCE: van den Broek, 2015.

desired outcomes that are both consistent with current professional knowledge and uphold basic reproductive rights" (Hulton et al., 2000).

Edward Kelley[4] of the WHO gave the last presentation of the session, joining via videoconference from Geneva. He echoed van den Broek's point about the challenges of measuring the effectiveness of bundles of interventions, as well as Ruelas's idea that the momentum for universal coverage will make quality questions even more important. Kelley emphasized the challenge of looking at quality in the context of health systems strengthening. Much as malaria and HIV programming has often taken too narrow a view of improving services, so has health systems research sometimes been too insular. He observed that health systems experts need to work closely with ministers of finance, a lesson that his work on Ebola has made imminently clear.

In the subsequent discussion, participants considered gaps in the research environment that might be holding back the field. Rowe pointed out that much of the information on quality improvement is hard to find,

[4] Much of Kelley's presentation was not comprehensible to the Washington audience because of a bad videoconferencing connection.

a problem that could be corrected with a public data clearinghouse. Such a clearinghouse would allow researchers to know if and when certain strategies have been tried. It could be made more useful by setting out a taxonomy of different strategies to improve provider performance or quality of care. Currently, there is no universal classification for these methods. Even at the workshop, the method that some participants described as audit and feedback others called the plan-do-study-act cycle. A common vocabulary could do much to simplify the discussion and facilitate comparisons of similar methods.

Another major challenge facing implementation researchers is the need for repetition. Rowe explained that his group had combed through almost 500 studies and found more than 80 distinct strategies to improve provider performance. The vast majority of the strategies were tested only once or twice. Without more consistent attention to a research question, it is difficult to build a convincing body of evidence on the effectiveness of any strategy. But correcting this problem may require explicit incentives to repeat and replicate other groups' work.

The relationship between frontline quality improvement and policy change is something many of the meeting guests had experience with. James Heiby explained that USAID's primary interest is improving care on the front lines, often focusing on providers' practices. Guests from the implementing organizations saw themselves as working both at the provider level (in clinics or hospitals, for example) and at the district and national levels, feeding policy makers with the information they need to change national guidelines.

In the absence of clear treatment guidelines, market forces and patient demand can have a strong influence on provider behavior. A few participants suggested that sometimes the provider tendency to give injections and prescribe unnecessary medicines is driven by the patient's expectations; financial incentives may also play a role. Research from China has used the audit method to make a compelling point about financial incentives to over-prescribe antibiotics (Currie et al., 2014). Given the important influence that contextual factors, such as prescribing regulations, have on quality, some participants questioned USAID's emphasis on the frontline health worker. Quality of care may depend on policy changes at a higher level.

Developing countries provide interesting examples of unregulated markets for health care. Das observed that the private, unregulated health market in India rewards perceived quality: providers who complete more of the clinical checklist charge significantly more than providers who do not. One problem is that, in many countries, there is no link between quality and wages in the public sector. The majority of public-sector health workers in India earn six times more than their counterparts in the private sector, a fact that is often obscured by high salaries at the top private hospitals.

Das mentioned an explosion in private medical schools in India, some of which are very poor quality and have instructors who rarely come to class. He described meeting students from these schools and asking them about their plans for the future. Many intend to open a private practice in a slum, then apply and re-apply for a government job. Working in an undesirable location makes them competitive for government work and, because the government cannot discriminate in wages, they would be guaranteed a good salary.

Some participants saw this as an example of how governments and donors need to look beyond the proximal good when setting policy. The initial effect of reimbursing public providers equally, without regard to the quality of their education, might have been positive, but over time, it has warped the market. Similarly, a quality improvement strategy put into place now, based on good proximal evidence, can change the equilibrium of a health system. Over years, after the system comes back to equilibrium, the changes may have different consequences than those immediately apparent. It might be beneficial for policy makers to try to anticipate these changes and consider how stress in one part of a system can affect other parts.

6

Cross-Cutting Approaches to Improve Quality

Key Points Made by Individual Speakers

- The evidence presented at the workshop was among the best in the field, but not as high-quality as anyone would like. (Heiby)
- Quality improvement work confronts tradeoffs among generalizability, simplicity, and local usefulness; though all three are desirable, it is not possible to have them all in one program. (Berwick)
- A "select and censure" approach to quality gives the field a bad name. A better method encourages innovation and continual improvement. (Berwick)
- Berwick saw evaluations and randomized trials as toxic and impeding progress. Randomization in particular he saw as a way to lose valuable contextual data. (Berwick)

At the start of the next session, meeting participants broke into three groups to discuss the practice, policy, and research of quality improvement. Afterward, a representative from each group summarized the main points of the discussions (see Boxes 6-1, 6-2, and 6-3).

PRACTICE TO IMPROVE QUALITY: HOW TO ACCELERATE PROGRESS?

Pierre Barker relayed the key themes from the discussion on practice. Members of the group were concerned that a failure to recognize successes

and learn from them has held back quality improvement as a field. They observed that people working in quality improvement lack a common forum in which to discuss their experiences. The paucity of collaboration can lead to practitioners wasting time and repeating mistakes. Many participants thought it would be valuable for quality improvement teams to have a way to learn from each other. One simple option would be writing up careful case studies of programs that work well. Another possibility, put forth by William Tierney, would be to establish a peer learning group to encourage collaboration. Setting up such a network would be a logistical challenge, but the challenge could be overcome with the proper support.

If there were a forum for collaboration, one of the first topics it might broach is motivation. Though almost everyone in the field uses plan-do-study-act cycles and collects data about their projects, understanding what motivates people to change is a serious question. There is growing interest among international organizations in giving incentives for quality improvement to providers and managers. Some participants welcomed this interest, but wondered how to select the best incentives. Without better understanding what motivates providers, it is hard to say what incentives to offer them.

Barker also described how the group had addressed questions of capacity building. As the meeting made clear, there are many different quality improvement strategies in use. Barker reported a common concern that quality improvement programs neglect patient-centeredness—respecting the patient's needs and supporting them to make good decisions regarding their health. This problem could be corrected with more explicit attention to patient-centeredness in capacity building programs.

BOX 6-1
Main Points of Practice Discussion[a]

- People working in quality improvement have trouble recognizing their successes and learning from them. There is no clear forum for people working at different organizations to collaborate. One simple way to advance the field would be to give some attention to sharing detailed case studies of successful programs.
- International organizations are increasingly interested in using incentives to improve quality of care. This is a promising strategy, but depends on a better understanding of what motivates providers.
- Patient-centeredness can be neglected in quality improvement work, partly because patient empowerment and respectful treatment is a complicated skill to build. Quality improvement programs are good venues for capacity building, and building the ability of providers to give patient-centered services is essential.

[a] As summarized by breakout group rapporteur.

POLICY TO IMPROVE QUALITY: HOW TO CREATE AN ENVIRONMENT TO IMPROVE QUALITY?

Kedar Mate summarized the policy group's discussion on how to create an environment conducive to improving quality of care. The group talked about policy implementation as a process that starts with evidence, which is then translated into a policy and implemented. Many participants thought that community stakeholders could be better included in this process. Quality improvement is often concerned with patient-centeredness, and Mate reported how some group members saw patient-centered policy as an outgrowth of involving patient groups and providers in policy making.

The group members recognized room for improvement in the way program staff interact with policy makers. Some suggested that, when involved with any quality program, whether implementing a collaborative or accrediting a clinic, the technical experts could include policy makers as part of the effort. This would ensure that policy makers are better informed of the quality strategies used in their countries and better able to translate quality improvement programs into policy.

Mate explained that policy makers need data to understand the effects of their efforts. He described this as a learning system, one that can be used to understand how well policies are working and how they might be adjusted. Members of the group acknowledged that it is difficult to say exactly what systems should be in place to support policy makers in translating evidence into policy; this could be an area for further research.

BOX 6-2
Main Points of Policy Discussion[a]

- Policy implementation is a cycle. Building evidence is the first step; evidence is then translated into policy.
- Policy should be made in close collaboration with communities. Patient groups are important stakeholders, as are providers.
- Policy makers need data and a good understanding of research to translate evidence to practice. This requires regular involvement with research and implementation staff.
- A learning system provides policy makers with regular feedback about quality improvement. Learning systems also hold promise for researchers as they attempt to understand why some policies succeed and others fail.

[a] As summarized by breakout group rapporteur.

RESEARCH TO IMPROVE QUALITY: HOW TO ADVANCE THE FIELD?

Margaret Kruk of the Harvard School of Public Health presented the main points from the research discussion. First, a number of group members were concerned that many thoughtful, evidence-based policies are never implemented. One useful goal for research would be to explain why some policies never come into force. Kruk advised that such research need not be heavily theoretical, but could be a practical comparison of settings where implementation succeeds with those where it fails.

The group members also discussed the importance of better measurement, both to the field of quality improvement and to health systems research more broadly. Many of the previous workshop speakers had emphasized the importance of process indicators in quality improvement. Kruk pointed out that there is good consensus on meaningful outcome measures in health, such as child mortality rates, but process indicators are less standardized. She explained how donor needs and competing project priorities have littered health systems research with hundreds, if not thousands, of indicators; researchers could do better work, and compare their results across studies, if there were a smaller, more tightly defined group of indicators.

Along the same lines, better guidelines on how to report the process of implementing quality improvement programs would help the researchers involved with program evaluation. Kruk described a common problem of evaluations in the peer-reviewed literature containing copious detail on everything except the meat of the project: the actual change implemented is not described. A good description of the process contributes to a better understanding of a program's generalizability and external validity. In general, clearer guidelines on the best methods for evaluating quality improvement programs would be beneficial. Many participants in the research discussion were positive about using standardized patients to test providers, but acknowledged that observation may work better in some settings. In any case, the lack of gold standard research methods is holding back the field.

Improving measurement is linked to improving study designs. Kruk emphasized the value of prospective studies with evaluations incorporated into the design of the study from the start. Likewise, there was support for involving the evaluation staff from the start of a project to set up a rigorous evaluation. Several group members observed that study designs would be improved by applying a longer time frame; 6 months is too little time to observe a system and determine the effects of a change. Researchers would also do well to replicate the same study in different settings to observe the effects of contextual variables on outcomes and establish whether findings are generalizable.

Research on quality improvement requires investigators to first diagnose the problem in the health system and only then identify solutions. This ensures that the problem and the solution work at the same level. So, if a quality of care problem is caused by a policy or a regulation, then the solution would aim to correct that policy, rather than change the management of a particular clinic. Similarly, if the improvement method relies on management, it should first be clear that there are enough capable managers in the system to support the change. Kruk reported that the group members discussed the importance of establishing baseline performance before rushing into any improvement program because baseline data has a strong influence on possible effect sizes.

Members of the group were interested in having better data, especially from low-income countries. Some suggested that USAID could play a meaningful role by investing in a common information infrastructure in the poorest countries. Kruk commended the agency for insisting that data paid for by the taxpayers are a public good and should be publicly available. Similarly, many of the research group participants argued for better data sharing across projects. This would include more openness about projects that do not work, thereby allowing researchers to avoid repeating mistakes.

Lastly, the research discussion considered some of the obstacles in the organization of research at national and international levels. Some participants were concerned that the increasing interest in quality puts stress on the grassroots staff to do more without proper support from the government. They suggested that important decisions about how research is organized should be confronted at higher levels. This would give governments clear ideas about the responsibility of the data collectors and their authority over the data. For example, if projects are to be peer-reviewed, that arrangement would be formally recognized before the program begins.

BOX 6-3
Main Points of Research Discussion[a]

- Researchers would do well to explore what makes policy implementation a success or failure. Illustrative case studies may be a constructive way to do this.
- Quality of care research, and health systems research more broadly, has no consensus on a set of meaningful process indicators that could be reported across studies, allowing more direct comparisons of results.

continued

BOX 6-3 Continued

- The rigor of quality research could be improved with more prospective studies, research over a longer time frame, and agreement on methodological gold standards.
- Clarity on the nature and potential causes of a quality problem should precede action on the solution. Establishing baseline performance is also an indispensable early step for researchers.
- A common data infrastructure would ease sharing of information across projects, enabling learning from other groups' successes and failures.
- There is a need for more attention to governance of research, including the role of national and global stakeholders, researcher independence, peer review, and mechanisms for policy uptake.

[a] As summarized by breakout group rapporteur.

REACTIONS TO THE WORKSHOP

Heiby then shared his reflections on the workshop. He opened by noting that quality improvement in low- and middle-income countries is no longer in its infancy. Quality experts have collected considerable experience over the past 10 to 15 years. Heiby thanked the workshop participants for summarizing the experience that brought them to that point and identifying the problems that confront the field. Though the discussions had mostly looked at the evidence on quality improvement outcomes, he asked them to also consider how those results come to pass. One of the goals of quality improvement work is to build understanding as to what makes a program successful and how to replicate that success.

Heiby reiterated the opinion of many participants that the quality of the evidence presented at the workshop, though the best available in the field, was not as high as it could be. Most data came from program evaluations, a relatively new field. Heiby encouraged more research on program implementation—that is, on how programs work on the ground and how they can be improved. As an example, he described a large-scale quality improvement program in East Africa, which was faced with a decision about how to structure the supervision and coaching of its teams. One option was to place the central office, with its highly trained and intelligent staff, in charge; the alternative was to delegate responsibility to the provincial offices, where staff would be more familiar with local conditions. Through a field test comparison, they determined that the central office and provin-

cial offices had the same effect. But using teams from the latter cost only one-fifth of what it would to use the central teams.

Heiby used this example to illustrate the importance of examining the design of quality improvement programs, an area of research that warrants further study. He stressed the value of the depth of knowledge in quality improvement programming and encouraged the audience to unlock more transferable lessons from this data, adding that it would be unfair to suggest the field of quality improvement does not have data or understand the effectiveness of its interventions. Rather, the challenge is to improve the usefulness of the available data and evaluations and enhance monitoring and reporting.

Quality improvement tends to produce a good amount of information about effectiveness, but weaker data in other areas. Future projects could benefit from more attention to cost-effectiveness, non-clinical processes (e.g., human resources management), scalability, and sustainability (i.e., how to foster a culture of quality improvement in health organizations). Heiby cited the role of external advisors in quality programs as an area of particular concern for USAID, as programs that depend on foreign support are not sustainable. Program evaluations often blur the work of host country nationals and outside technical experts, to the detriment of the analysis.

Heiby shared his vision for the workshop's outcomes, including briefing USAID's senior management on the key messages. Because quality improvement is a relatively young field, there are opportunities to influence program design and capture the knowledge generated by quality programs. Heiby agreed that the thousands of professionals working in the field would benefit from a public clearinghouse on quality improvement programming and data. He saw the failure to share this data with people who would benefit as a profound waste of both existing and future resources because more time and money will be spent duplicating work.

Heiby closed by affirming a point made in the policy breakout discussion, that there is considerable room for improvement in how quality experts communicate their work to the public. Calling the discussion on research refreshing and overdue, he underlined the need to examine failures, not only successes, and to build research into quality improvement programs. Finally, he welcomed the suggestion that donors fund more prospective studies, as well as longer studies, on quality of care.

CLOSING REMARKS

Don Berwick, president emeritus and senior fellow at the Institute for Healthcare Improvement, gave closing remarks for the workshop. Patrick Kelley, who directs the IOM Board on Global Health, briefly introduced the session, telling the audience how Berwick, along with Heiby and Sheila

Leatherman, had begun encouraging the IOM to work on quality of care in developing countries 12 years ago. Kelley reflected on how global health had changed during that time, especially with the billions of dollars spent through U.S. programs like PEPFAR and the President's Malaria Initiative. The 2014 Ebola epidemic further underscored the value of developing strong health systems around the world. All of which, Kelley concluded, have shown the world that quality of care should be a priority for international action. But the question remains as to how quality can best be advanced in poor countries. The often-cited 2001 IOM report *Crossing the Quality Chasm* changed the way people think about quality of care in the United States, and Kelley described the challenges of adapting the earlier report to a global audience. Quality improvement in the United States makes certain assumptions—the stability and safety of the drug supply, for example—that do not always apply in low- and middle-income countries.

Kelley then introduced the IOM president Victor Dzau. Dzau added his reflections on the importance of strengthening systems for health, quality, and patient safety around the world. He observed that, for some time, the focus in global health has been to create better access to care and to keep services affordable. Only recently has the discussion turned to quality. It is not enough, he asserted, to invest money in access without paying attention to quality.

Dzau described how *Crossing the Quality Chasm* changed the practice of medicine in the United States, and he agreed that the dimensions of quality identified in that report (safety, effectiveness, timeliness, patient-centeredness, efficiency, equity) are universal. But he also echoed Kelley's point that a larger set of quality issues is at work in low- and middle-income countries. He commended the workshop participants for their commitment to the field and expressed his support for a global quality of care project.

After an introduction by Dzau, Berwick began his presentation by thanking the participants, reiterating that such a meeting would not have been held 5 or 10 years earlier. The interest that has grown in the meantime is driven, he felt, by a desire to identify simple, generalizable methods that can be used to improve quality of care locally. Berwick described the tradeoffs among the three qualities shown in Figure 6-1, saying that it is not possible to find methods for quality improvement that meet all three conditions. When a solution is generalizable and simple, it is usually not useful locally and must be adapted. If something is generalizable and locally useful, then it will not be simple. Berwick encouraged the audience to keep this tension in the back of their minds.

Like other aspects of modern public policy, quality improvement rests on the assumption that inspection and attention will drive improvements. Berwick compared it to a Ghanaian proverb, "Weighing a pig doesn't make it fatter." He proposed that not only does measurement not cause improve-

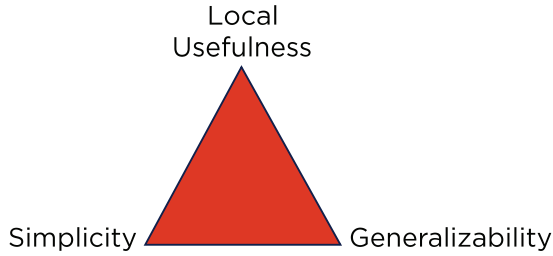

FIGURE 6-1 Tradeoffs in improving quality.
SOURCE: Berwick, 2015.

ment, it can have a demoralizing effect on workers, the people who most need encouragement. Bill Scherkenbach of General Motors identified this problem in the 1960s, calling it a "cycle of fear" (Scherkenbach, 1991). When workers in a hospital system know that an inspection is coming, the anxiety interferes with their performance because invariably inspection results in the culling, or at least retraining, of some workers. The concept of removing poor performers from a system goes back to the 1920s: it was Henry Ford's method of keeping production lines consistent. Berwick explained that modern quality improvement has moved beyond that in every field but health.

He conceded that systems do need to be accountable; especially in public or politically managed organizations, it is important to catch thieves and sack them. But he cautioned against a culture of inspection that poisons the well against quality improvement. At the same time, Berwick praised measurement as essential for learning. He explained that measurement can go down two paths. The first is the "select and censure" path that gives quality improvement a bad name; the second is measuring from a safe place, encouraging innovation and continual improvement while protecting employees and, in the case of health care, patients (Berwick et al., 2003).

In reflecting on the previous 2 days, Berwick concluded that quality defects are pervasive and problems with the health system drive these defects. As Barker had pointed out earlier, every system is perfectly designed to get the results it gets. So problems with quality of care are systems problems, particularly distressing ones in poor countries because that is where people can least afford poor quality. He asserted that improvement is often a nonlinear process in which effective learning requires both clear aims and iterative testing (see Figure 6-2). All six models that the workshop participants discussed depend on repeated testing to determine if the program is bringing the health system closer to a clearly articulated aim.

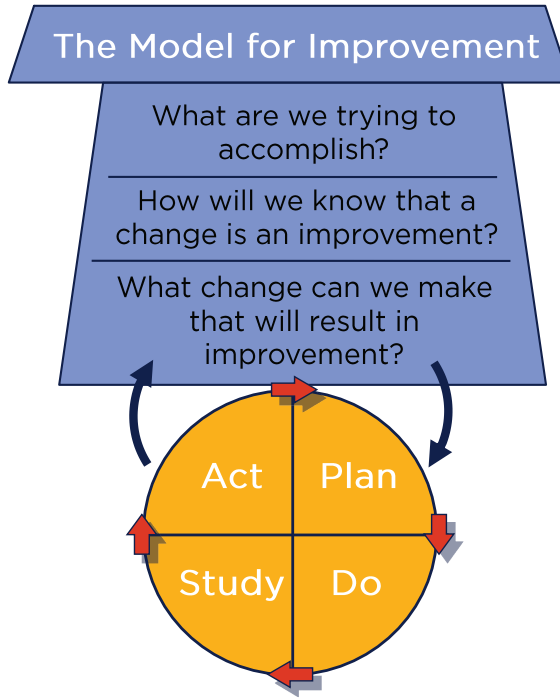

FIGURE 6-2 Combining three key questions with a plan-do-study-act cycle is a model for improvement.

SOURCE: Langley, G. J. 2009. *The improvement guide: A practical approach to enhancing organizational performance.* San Francisco: Jossey-Bass. © John Wiley & Sons, Inc. As presented by Berwick on January 29, 2015.

While Berwick saw repeated testing and the emphasis on goals as positive influences on quality improvement, he described evaluation as toxic and actively impeding progress. First, he lamented the amount of time evaluation cycles require. Clinic managers do not have several years to wait before knowing how their changes are affecting patients; they need fast results. Along the same lines, he saw evaluations, particularly randomized trials, as being insensitive to local context. The goal of randomization is to remove the influence of confounders, which Berwick saw as a loss of valuable data. In quality improvement, the process by which some changes take root and others do not is as important as the final effects on health. Berwick suggested that both methodological rigor and the utility of evaluation would be enhanced with deeper understandings of local contexts and how they influence processes and outcomes.

Some of the meeting participants had reservations about placing too much emphasis on local context, however. They brought up the cost and

staffing constraints that get in the way of doing detailed implementation research or qualitative analysis. Jishnu Das pointed out that researchers are not inclined to work on large, messy public datasets for which there is little publication market, and major funders, including USAID, are not interested in funding large research projects. Some alternatives to traditional evaluations might give more attention to real-time data and feedback or use anthropology and ethnographic methods to understand how a system works.

An interest in local learning and context could help build will for a new global initiative in quality of care. Berwick spoke of how the IOM reports *Crossing the Quality Chasm* and *To Err Is Human* changed the way Americans think about medical error and quality of care. Although progress on the reports' recommendations has been variable, the influence they had in terms of making managers aware of safety, timeliness, effectiveness, patient-centeredness, efficiency, and equity is undeniable. Outside of the United States, the reports' effects have been less pronounced. Although Berwick noted progress in Denmark, Scotland, and Singapore, very little attention has been given to the topic in low- and middle-income countries. He asked for a well-considered statement about the nature and urgency of the gaps in the performance of health systems. This kind of information would call attention to the waste of resources that health markets are capable of and would motivate public interest in change.

Crossing the Quality Chasm identified a chasm between the care people have and what they could have. Berwick acknowledged that, while this gap still exists in developed countries, it is far worse in developing ones. He thought the time was right for more attention to quality problems in the parts of the world least able afford them and to encourage governments and donor organizations to action.

References

Adhikari, N. K. 2013. Patient safety without borders: Measuring the global burden of adverse events. *BMJ Quality & Safety* 22(10):798-801.

Agins, B. D. 2015 (January 28). How to know if it works: Measuring changes in quality: Can we link improvement data to public health outcomes? Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Arah, O. A., G. P. Westert, J. Hurst, and N. S. Klazinga. 2006. A conceptual framework for the OECD Health Care Quality Indicators Project. *International Journal for Quality in Health Care* 18(Suppl 1):5-13.

Barker, P. 2015 (January 28). Improving quality: Theories of change. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Berendes, S., P. Heywood, S. Oliver, and P. Garner. 2011. Quality of private and public ambulatory health care in low and middle income countries: Systematic review of comparative studies. *PLoS Medicine* 8(4):e1000433.

Berwick, D. 2015 (January 29). Closing remarks. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Berwick, D. M., B. James, and M. J. Coye. 2003. Connections between quality measurement and improvement. *Medical Care* 41(1 Suppl):I30-I38.

Bosch-Capblanch, X. 2015 (January 28). Tools and processes for quality improvement: Supervision. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Washington, DC.

Bosch-Capblanch, X., and P. Garner. 2008. Primary health care supervision in developing countries. *Tropical Medicine and International Health* 13(3):369-383.

Bosch-Capblanch, X., and C. Marceau. 2014. Training, supervision and quality of care in selected integrated community case management (iCCM) programmes: A scoping review of programmatic evidence. *Journal of Global Health* 4(2):020403.

Bosch-Capblanch, X., S. Liaqat, and P. Garner. 2011. Managerial supervision to improve primary health care in low- and middle-income countries. *Cochrane Database of Systematic Reviews* (9):Cd006413.

Boucar, M., K. Hill, A. Coly, S. Djibrina, Z. Saley, K. Sangare, E. Kamgang, and S. Hiltebeitel. 2014. Improving postpartum care for mothers and newborns in Niger and Mali: A case study of an integrated maternal and newborn improvement programme. *BJOG: An International Journal of Obstetrics and Gynaecology* 121(Suppl 4):127-133.

Bradley, J., and S. Igras. 2005. Improving the quality of child health services: Participatory action by providers. *International Journal for Quality in Health Care* 17(5):391-399.

Broughton, E. 2015 (January 28). The cost-effectiveness of improving health care. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Broughton, E., Z. Saley, M. Boucar, D. Alagane, K. Hill, A. Marafa, Y. Asma, and K. Sani. 2013. Cost-effectiveness of a quality improvement collaborative for obstetric and newborn care in Niger. *International Journal of Health Care Quality Assurance* 26(3):250-261.

Carinci, F., K. Van Gool, J. Mainz, J. Veillard, E. C. Pichora, J. M. Januel, I. Arispe, S. M. Kim, and N. S. Klazinga. 2015. Towards actionable international comparisons of health system performance: Expert revision of the OECD framework and quality indicators. *International Journal for Quality in Health Care* 27(2):137-146.

Cordero, C. 2015 (January 28). COPE®: A process for improving quality in health services. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Currie, J., W. Lin, and J. Meng. 2014 (unpublished). Addressing antibiotic abuse in China: An experimental audit study.

Das, J. 2011. The quality of medical care in low-income countries: From providers to markets. *PLoS Medicine* 8(4):e1000432.

Das, J. 2015 (January 29). Improving quality of care. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Das, J., A. Holla, V. Das, M. Mohanan, D. Tabak, and B. Chan. 2012. In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. *Health Affairs* 31(12):2774-2784.

Das, J., A. Chowdhury, R. Hussam, and A. Banjerjee. 2015 (unpublished). The impact of training informal providers on clinical practice in West Bengal, India: A randomized controlled trial.

DHIS 2. n.d. https://www.dhis2.org (accessed April 6, 2015).

English, M. 2015 (January 28). In-service training to improve quality of care. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Favato, G. 2008 (June 26-28). Health care decision making. Presented at the University Program in Health Economics, Ragusa, Italy.

Forsetlund, L., A. Bjorndal, A. Rashidian, G. Jamtvedt, M. A. O'Brien, F. Wolf, D. Davis, J. Odgaard-Jensen, and A. D. Oxman. 2009. Continuing education meetings and workshops: Effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* (2):Cd003030.

Higgins, J. P., D. G. Altman, and J. A. Sterne (editors). 2008. Chapter 8: Assessing risk of bias in included studies. In *Cochrane handbook for systematic reviews of interventions* Version 5.0.1 (updated September 2008), edited by J. P. Higgins and S. Green. Chichester, West Sussex, UK: The Cochrane Collaboration.

Hulton, L. A., Z. Matthews, and R. W. Stones. 2000. *A framework for the evaluation of quality of care in maternity services*. Highfield, Southampton, UK: University of Southampton.

IOM (Institute of Medicine). 2001. *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.

IOM. 2014. *Investing in global health systems: Sustaining gains, transforming lives*. Edited by G. J. Buckley, J. E. Lange, and E. A. Peterson. Washington, DC: The National Academies Press.

Jha, A. K., I. Larizgoitia, C. Audera-Lopez, N. Prasopa-Plaizier, H. Waters, and D. W. Bates. 2013. The global burden of unsafe medical care: Analytic modelling of observational studies. *BMJ Quality & Safety* 22(10):809-815.

Juran, J. M., and A. B. Godfrey. 1999. *Juran's quality handbook*. New York: McGraw-Hill.

Kelley, E., and J. Hurst. 2006. Health Care Quality Indicators Project: Conceptual framework paper. *OECD Health Working Papers,* No. 23, OECD Publishing. http://dx.doi.org/10.1787/440134737301 (accessed July 20, 2015).

Klazinga, N. 2015 (January 28). How to know it works? Measuring and assuring changes in quality: An OECD perspective. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Washington, DC.

Langley, G. J. 2009. *The improvement guide: A practical approach to enhancing organizational performance*. San Francisco: Jossey-Bass.

Massoud, M. R. 2015 (January 28). What is collaborative improvement? Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Michie, S., M. M. van Stralen, and R. West. 2011. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science* 6:42.

Nair, D. 2015 (January 28). Quality of care from a cost-effectiveness and feasibility perspective. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Necochea, E. 2015 (January 28). Standards-based management and recognition. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

OECD (Organisation for Economic Co-operation and Development). 2013. *Health at a glance 2013: OECD indicators*. Paris: OECD Publishing.

OECD and WHO (World Health Organization). 2014. *Health at a glance: Asia/Pacific 2014: Measuring progress towards universal health coverage*. Paris, France: OECD.

Oxley, W. H., M. H. Phillips, and J. Young. 1935. Maternal mortality in Rochdale: An achievement in a black area. *British Medical Journal* 1(3867):304-307.

Rowe, A. 2015 (January 28). Health care provider performance review. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

Scherkenbach, W. W. 1991. *Deming's road to continual improvement*. Knoxville, TN: SPC Press.

van den Broek, N. 2015 (January 29). Quality of care in maternal and newborn health. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

vanOstenberg, P. 2015 (January 28). Accreditation. Presented at Workshop on Improving Quality of Care in Low- and Middle-Income Countries, Institute of Medicine, Washington, DC.

WHO (World Health Organization). 2012. Online Q&A: What is universal health coverage? http://www.who.int/features/qa/universal_health_coverage/en (accessed May 21, 2015).
World Bank. 2015. Data: Countries and economies. http://data.worldbank.org/country (accessed April 15, 2015).

Appendix A

Workshop Agenda

Workshop on Improving Quality of Care in Low- and Middle-Income Countries
January 28–29, 2015
The Keck Center, 500 Fifth Street NW
Washington, DC 20001

JANUARY 28, 2015
Room 100

8:30-9:00 Breakfast available

Session 1
Welcome and Overview

9:00-9:15 Welcome
Victor Dzau, *President*, Institute of Medicine
Sheila Leatherman, Planning Committee Chair
Research Professor, Gillings School of Global Public Health, University of North Carolina at Chapel Hill

9:15-9:45 Keynote
An Overview of Quality of Care in Low- and Middle-Income Countries
Ashish Jha, *Director*, Harvard Global Health Institute (by video)

9:45-10:15 USAID's Work in Quality of Care: Challenges and Common Themes
James Heiby, *Medical Officer*, USAID

10:15-10:30 Break

Session 2
Six Widely Used Methods to Improve Quality

10:30-11:00 A Brief Overview of the Six Methods in Question
Rifat Atun, Planning Committee Member
Professor of Global Health Systems, Harvard School of Public Health (by video)

11:00-11:30 Theories of Change
Pierre Barker, *Senior Vice President*, Institute for Healthcare Improvement

11:30-12:30 Panel Discussion: Tools and Processes for Quality Improvement
Carmela Cordero, *Senior Clinical Advisor*, EngenderHealth
Paul vanOstenberg, *Senior Advisor, Growth and Innovation*, Joint Commission International
Edgar Necochea, *Director, Health Systems Development*, Jhpiego
Moderator: Nachiket Mor, *Chairman of the Board*, CARE India

12:30-1:30 Lunch

1:30-2:30 Panel Discussion: Tools and Processes for Quality Improvement
Rashad Massoud, *Senior Vice President*, University Research Co., LLC
Mike English, *Head of Health Services*, KEMRI Wellcome Trust
Xavier Bosch-Capblanch, *Group Leader*, Swiss Tropical and Public Health Institute (by video)
Moderator: Kedar Mate, Planning Committee Member
Senior Vice President, Institute for Healthcare Improvement

2:30-2:45 Break

Session 3
Reviewing Evidence for Different Quality Improvement Methods

2:45-3:45 What Works?: Results of a Systematic Review
Alexander Rowe, *Medical Epidemiologist*, Malaria Branch, CDC

3:45-4:30 Panel Discussion: How to Know It Works? Measuring Changes in Quality
Bruce Agins, *Medical Director*, New York State Department of Health AIDS Institute
Director, HEALTHQUAL International
Niek Klazinga, *Coordinator, Health Care Quality Indicators program*, OECD
Moderator: Sheila Leatherman, *Research Professor*, Gillings School of Global Public Health, University of North Carolina at Chapel Hill

4:30-5:15 Panel Discussion: At What Cost?: What Makes a Program Cost-Effective and Feasible?
Edward Broughton, *Director of Research & Evaluation*, University Research Co., LLC
Dinesh Nair, *Senior Health Specialist*, The World Bank
Moderator: Kedar Mate, *Senior Vice President*, Institute for Healthcare Improvement

5:15-5:30 Closing Remarks
Kedar Mate, *Senior Vice President*, Institute for Healthcare Improvement

5:30 Adjourn

JANUARY 29, 2015
Room 100

8:30-9:00 Breakfast available

Session 4
Synthesizing Evidence, Identifying Gaps

9:00-9:15 Welcome and Overview
Sheila Leatherman, *Research Professor*, Gillings School of Global Public Health, University of North Carolina at Chapel Hill

9:15-9:45 Synthesizing the Evidence
Enrique Ruelas, *Professor and Director of Public Policy and Health Systems*, Monterrey Institute of Technology and Higher Education

9:45-11:00 Panel Discussion: Identifying Gaps and Solutions
Jishnu Das, *Lead Economist, Development Research Group*, The World Bank
Edward Kelley, *Director, Department of Service Delivery and Safety*, World Health Organization (by video)
Nynke van den Broek, *Professor*, Liverpool School of Tropical Medicine
Moderator: Ramanan Laxminarayan, *Director and Senior Fellow*, The Center for Disease Dynamics, Economics and Policy, and *Vice President, Research and Policy*, Public Health Foundation of India

11:00-11:15 Break

Session 5
Cross-Cutting Approaches to Improve Quality

11:15-12:30 Breakout Group Discussions

Policies to Improve Quality: How to Create an Environment Conducive to Improving Quality of Care?
Rapporteur: Kedar Mate, *Senior Vice President*, Institute for Healthcare Improvement

Practice to Improve Quality: How to Accelerate Progress?
Rapporteur: Pierre Barker, *Senior Vice President*, Institute for Healthcare Improvement

	Research to Improve Quality: How to Advance the Field? **Rapporteur: Margaret Kruk,** *Associate Professor of Global Health*, Harvard School of Public Health
12:30-1:15	Lunch
1:15-2:00	Reports from Breakout Groups
2:00-2:15	Reactions to Workshop **James Heiby,** *Medical Officer*, USAID
2:15-3:15	Closing Remarks **Victor Dzau,** *President*, Institute of Medicine **Don Berwick,** *Former Administrator*, Centers for Medicare & Medicaid Services
3:15	Adjourn

Appendix B

Participant Biographies

Bruce Agins, M.D., M.P.H., is the medical director of the New York State Department of Health AIDS Institute, where he oversees a staff of 40 individuals involved in guidelines development, quality management, and education programs. He is the principal architect of New York's HIV Quality of Care Program and has more than 20 years of HIV-specific quality improvement experience. He is principal investigator of the National Quality Center and the NYLINKS programs funded by the Health Resources and Services Administration (HRSA). Since 2002, Dr. Agins has applied his experience in the field of HIV quality improvement to the international setting, directing HEALTHQUAL International, which is supported through the President's Emergency Plan for AIDS Relief (PEPFAR). HEALTHQUAL partners with ministries of health in countries throughout Africa, Asia, and the Caribbean and in Papua New Guinea to provide technical support and coaching to build capacity for quality management as part of sustainable health systems. Dr. Agins has participated in consultations with the World Health Organization (WHO) and participated as faculty in the Salzburg Seminar devoted to the development of guidelines and models of quality management in resource-limited settings. Dr. Agins is an infectious disease specialist with extensive experience in the oversight of government-sponsored programs. Dr. Agins holds academic appointments as full clinical professor in the Department of Epidemiology and Biostatistics in Global Health Sciences (volunteer) at the University of California, San Francisco, School of Medicine and as adjunct professor in the Division of Infectious Diseases and Immunology at New York University School of Medicine. He is a graduate of Haverford College (1975) and Case Western Reserve School

of Medicine (1980) and received his M.P.H. from the Mailman School of Public Health at Columbia University in 1994.

Rifat Atun, M.B.B.S., M.B.A., FRCGP, FFPH, FRCP, is a professor of global health systems at the Harvard School of Public Health, Harvard University, and director of the Global Health Systems Cluster. He is an honorary professor at the London School of Hygiene and Tropical Medicine, and from 2006–2013, he was a professor of international health management at Imperial College London. Between 2008 and 2012, he was a member of the Executive Management Team of the Global Fund to Fight AIDS, Tuberculosis and Malaria in Switzerland as the director of the Strategy, Performance and Evaluation Cluster.

His research is empirically oriented and focuses on health systems reform, diffusion of innovations in health systems, and global health financing, including research and development. He has published extensively in these areas in *Lancet*, *PLoS Medicine*, *BMJ*, *Lancet Infectious Diseases*, and the *Cochrane Database of Systematic Reviews*.

Dr. Atun has worked at the UK Department for International Development (DFID) Health Systems Resource Centre and has acted as a consultant for the World Bank, the WHO, and a number of international agencies on the design, implementation, and evaluation of health system reforms. He has served as a member of the Advisory Committee for the WHO Research Centre for Health Development in Japan. He is a member of the PEPFAR Scientific Advisory Board, the UK Medical Research Council's Global Health Group, and the Advisory Board for the Norwegian Research Council's Programme for Global Health and Vaccination Research.

Dr. Atun studied medicine at University of London as a Commonwealth Scholar and subsequently completed his postgraduate medical studies and master of business administration at University of London and Imperial College London. He is a Fellow of the Faculty of Public Health of the Royal College of Physicians (UK), a Fellow of the Royal College of General Practitioners (UK), and a Fellow of the Royal College of Physicians (UK).

Pierre M. Barker, M.D., is senior vice president of the Institute for Healthcare Improvement (IHI), where he is responsible for IHI's large-scale health systems improvement initiatives outside the United States. Working closely with partners in government and the nongovernmental sector, IHI has an expanding portfolio of work to improve the reliability and scale-up of effective health programs in Africa, Australasia, Europe, Latin America, and the Middle East. In addition to advising governments and large organizations on quality strategies, IHI uses the science of improvement to promote improved outcomes in the areas of patient safety, population health, patient-centered care, and cost of care. Dr. Barker attended medical school

in South Africa and trained in pediatrics in the United Kingdom and the United States.

Before joining IHI, he was medical director of University of North Carolina (UNC) Children's Hospital clinics and was responsible for leading health system–wide initiatives on improving access to care and chronic disease management. A renowned authority on improving health systems, Dr. Barker initially served at IHI as in-country director of IHI's South Africa projects and then as head of IHI programs in low- and middle-income countries. He retains a position of clinical professor of pediatrics in the Maternal and Child Health Department at the UNC Gillings School of Global Public Health. He also advises the WHO on health systems strengthening, redesign of HIV care, and infant feeding guidelines.

Donald M. Berwick, M.D., M.P.P., FRCP, is president emeritus and senior fellow at IHI, an organization that Dr. Berwick co-founded and led as president and CEO for 18 years. He is one of the nation's leading authorities on health care quality and improvement. In July 2010, President Obama appointed Dr. Berwick to the position of administrator of the Centers for Medicare & Medicaid Services (CMS), which he held until December 2011. A pediatrician by background, Dr. Berwick has served as clinical professor of pediatrics and health care policy at the Harvard Medical School, professor of health policy and management at the Harvard School of Public Health and member of the staffs of Boston Children's Hospital Medical Center, Massachusetts General Hospital, and the Brigham and Women's Hospital. He has also served as vice chair of the U.S. Preventive Services Task Force, the first Independent Member of the Board of Trustees of the American Hospital Association, and chair of the National Advisory Council of the Agency for Healthcare Research and Quality. An elected member of the National Academy of Medicine (NAM), Dr. Berwick served two terms on the Institute of Medicine's (IOM's) governing council and was a member of the IOM's Board on Global Health. He served on President Clinton's Advisory Commission on Consumer Protection and Quality in the Healthcare Industry.

Dr. Berwick is a recipient of numerous awards, including the 1999 Joint Commission's Ernest Amory Codman Award, the 2002 American Hospital Association's Award of Honor, the 2006 John M. Eisenberg Patient Safety and Quality Award for Individual Achievement from the National Quality Forum and the Joint Commission on Accreditation of Healthcare Organizations, the 2007 William B. Graham Prize for Health Services Research, the 2007 Heinz Award for Public Policy from the Heinz Family Foundation, the 2012 Gustav O. Lienhard Award from the IOM, and the 2013 Nathan Davis Award from the American Medical Association. In 2005, he was appointed Honorary Knight Commander of the British Empire by the Queen

of England, the highest honor awarded by the United Kingdom to non-British subjects, in recognition of his work with the British National Health Service. Dr. Berwick is the author or co-author of more than 160 scientific articles and 4 books. He also serves now as lecturer in the Department of Health Care Policy at Harvard Medical School.

Xavier Bosch-Capblanch, M.D., M.Sc., Ph.D., is a medical doctor, with an official medical specialty in public health (Spain), M.Sc. in tropical medicine and hygiene (Spain), and Ph.D. on evidence and policy of vaccination programs in low- and middle-income countries (University of Amsterdam). He is Group Leader (Data-Evidence-Evaluation-Policy) at the Swiss Tropical and Public Health Institute (Basel, Switzerland). He is also honorary lecturer at the Liverpool School of Tropical Medicine.

Dr. Bosch-Capblanch started his career working as a clinician in Spain, and he has been living and working in rural settings of sub-Saharan Africa for 10 years. He progressively moved from clinical work to health care management, public health, project management and research, including 4 years coordinating a rural health research center in Mozambique. He has worked in more than 20 countries.

Dr. Bosch-Capblanch has undertaken formal lecturing and on-the-job training of government officials and consultants involved in the design and/or implementation of public health programs in Saudi Arabia, Sudan, and Syria, as well as in Basel and Liverpool.

His areas of expertise include information and evidence methods on health systems interventions and immunization. In the former, Dr. Bosch-Capblanch has been involved in numerous initiatives to assess the quality of routine administrative and surveys data, has led several surveys in different countries, and has lectured on data and information for decision making. A significant part of this work is directly related to health systems evidence and vaccination coverage data. He has been involved in several Cochrane and non-Cochrane systematic reviews and has led initiatives to bridge the gap between research evidence and management and policy making.

Edward Broughton, Ph.D., M.P.H., PT, is director of research and evaluation for the U.S. Agency for International Development (USAID) Applying Science to Strengthen and Improve Systems (ASSIST) Project at University Research Co., LLC (URC), where he leads a portfolio of more than 20 evaluations and cost-effectiveness studies on health care improvement methods in low- and middle-income countries. He has published several research papers in peer-reviewed journals and presented at many international health conferences. He received his Ph.D. in international health from Johns Hopkins University and formerly was adjunct faculty at Columbia University.

Carmela Cordero, M.D., is an obstetrician-gynecologist with 30 years of experience as a service provider, master trainer, technical advisor, manager, and program leader in both the public and private sectors of reproductive health and family planning programs. She is currently a senior clinical advisor at EngenderHealth, where she leads the Clinical Support Team, a multiple-country group of clinicians providing support to global and country programs and projects. Dr. Cordero is a highly experienced trainer and has worked in more than 37 countries facilitating the introduction and dissemination of new family planning technologies, especially long-acting reversible methods and permanent methods; the development of training manuals, job aides and curricula; the development and implementation of national norms and standards in contraception and reproductive health; and the introduction of maternity care technologies and approaches that improve service quality.

Dr. Cordero is widely recognized for stimulating the development, introduction, and adoption of clinical safety and quality assurance and improvement systems in health care facilities around the world. She has 9 years of experience as a hospital gynecologist and obstetrician in Maternidad Nuestra Señora de la Altagracia, the main maternity care and university hospital in Santo Domingo, Dominican Republic. Dr. Cordero is an active member of the Dominican Society of Obstetrics and Gynecology and of the International Federation of Gynecology and Obstetrics. She is fluent in English, French, and Spanish.

Jishnu Das, Ph.D., is a lead economist in the Development Research Group (Human Development and Public Services Team) at the World Bank and a visiting fellow at the Center for Policy Research, New Delhi. Dr. Das's work focuses on the delivery of basic services, particularly health and education. He has worked on the quality of health care, mental health, information in health and education markets, child learning and test-scores and the determinants of trust. His work has been published in leading economics, health, and education journals and widely covered in the media and policy forums. In 2011, he was part of the core team on the *World Development Report* on gender and development. He received the George Bereday Award from the Comparative and International Education Society and the Stockholm Challenge Award for the best information and communications technology project in the public administration category in 2006, and the Research Academy award from the World Bank in 2013.

Michael English, M.B.B.Chir., worked in Kilifi from 1992 on malaria, in the early years of the Kilifi program, and returned to the United Kingdom in 1996 to complete specialist training as a general pediatrician in 1998. He returned to Kilifi in 1999 to work on neonatal illnesses as part of a

Wellcome Trust Career Development Fellowship while also working as a pediatrician in Kilifi District Hospital. In 2004, after some work at the national level on quality of pediatric care, he moved to Nairobi where he continues to work with the KEMRI Wellcome Trust Research Programme as a Wellcome Trust Senior Research Fellow. He was made Professor of International Child Health in Oxford in 2010. His work has included developing national, evidence-based guidelines for care of severely ill children and newborns, at first in 2005 and then updated in 2010 and 2013. To complement these, Dr. English and colleagues developed the ETAT+ course, adapting the WHO's ETAT course and expanding its scope to include evidence-based case management of serious illness in the child and newborn periods. The ETAT+ course is now provided with the help of multiple colleagues and the Kenya Paediatric Association with training conducted across Kenya and for Kenyan medical students. Others have taken the course to Rwanda, Somaliland, and Uganda. The Health Services Unit he leads has undertaken long-term studies by a multidisciplinary team on initiating and establishing best practices within rural government hospitals. This has resulted in a Kenyan team working with the support of international collaborators on hospital performance measurement, cost-effectiveness, motivation, task-shifting, and barriers to implementation. More recently, work has started on governance, leadership, human resources for health and knowledge translation. The group is well known for their work on measuring and testing interventions to improve pediatric and neonatal quality of care. Dr. English and the group work closely with the Kenyan ministry of health, and he provides technical advice to the WHO on a range of issues related to child and newborn survival.

James Heiby, M.D., M.P.H., is a medical officer in the USAID Office of Health Systems. Since 1985, Dr. Heiby's work has focused on adapting modern quality improvement approaches from industrialized countries for use in the health systems of low- and middle-income countries. He has developed and managed a series of 5-year projects, including the current USAID ASSIST Project, which works in more than 20 countries. His work on quality improvement was recognized with the USAID Science and Technology for Development Award and the Distinguished Honor Award, and he has published several papers related to this field. He lectures on quality improvement at the schools of public health of Johns Hopkins, Harvard, George Washington, and Columbia Universities and serves as a reviewer for several journals. Prior to joining USAID, he worked in the Bureau of Epidemiology at the Centers for Disease Control and Prevention (CDC). Dr. Heiby has a degree in medicine from Johns Hopkins and in public health from Harvard and completed clinical training in internal medicine at New York Hospital-Cornell Medical Center.

Ashish K. Jha, M.D., M.P.H., is director for the Harvard Global Health Institute, K.T. Li Professor of International Health & Health Policy at the Harvard School of Public Health, professor of medicine at Harvard Medical School, and a practicing internal medicine physician at the Veterans Affairs Boston Healthcare System.

Dr. Jha received his M.D. from Harvard Medical School and trained in internal medicine at the University of California, San Francisco, where he also served as chief medical resident. He completed his general medicine fellowship at Brigham and Women's Hospital and Harvard Medical School and received his M.P.H. from the Harvard School of Public Health.

Dr. Jha's major research interests lie in improving the quality and costs of health care with a specific focus on the impact of policy efforts. His work has focused on a broad set of issues, including transparency and public reporting of provider performance, financial incentives, health information technology, and leadership, and the roles they play in fixing health care delivery systems.

Edward Kelley, Ph.D., is director for the Department of Service Delivery and Safety at the World Health Organization. In this role, he leads the WHO's efforts at strengthening the safety, quality, integration, and people-centeredness of health services globally and manages the WHO's work in a wide range of programs, including health services integration and regulation, patient safety and quality, blood safety, injection safety, transplantation, traditional medicine, essential and safe surgery, and emerging areas such as mHealth for health services and genomics. Very recently, Dr. Kelley was asked to lead the Infection Prevention and Safety and the Health System Recovery teams for the WHO's Ebola response effort.

Prior to joining the WHO, he served as director of the U.S. National Healthcare Reports for the U.S. Department of Health and Human Services in the Agency for Healthcare Research and Quality. These reports track levels and changes in the quality of care for the American health care system at the national and state level, as well as disparities in quality and access across priority populations. Dr. Kelley also directed the 28-country Health Care Quality Indicators (HCQI) Project of the Organisation for Economic Co-operation and Development (OECD). Formerly, Dr. Kelley served as a senior researcher and quality assurance advisor for the USAID-sponsored Quality Assurance Project (QAP) and Partnerships for Health Reform Project Plus (PHRPlus). In these capacities, he worked for 10 years in West and North Africa and Latin America, directing research on the Integrated Management of Childhood Illness in Niger. Prior to this, Dr. Kelley directed the international division of a large U.S.-based hospital consulting firm, the Advisory Board Company. His research focused on patient safety, quality and organization of health services, metrics and measurement in health services, and health systems improvement approaches and policies.

Niek Klazinga, M.D., Ph.D., is presently the coordinator of the Health Care Quality Indicators (HCQI) project at the OECD in Paris. He combines this work with a professorship in social medicine at the Academic Medical Centre at the University of Amsterdam. Dr. Klazinga has been involved over the past 25 years in numerous health services research projects and policy debates on quality of care and has published widely on the subject. His present commitments include a visiting professorship at the Corvinus University in Budapest, advisor to the WHO/Europe, advisor to the Canadian Institute for Health Informatics, member of the Outcome Framework Technical Advisory Committee of the English Department of Health, member of the advisory committee on Hospital Groups of the Irish Ministry of Health, and member of the board of trustees of the Isala Clinics (Zwolle, a large teaching hospital in the Netherlands) and Arkin (Amsterdam, one of the largest mental health care institutes in the Netherlands). Dr. Klazinga has (co)authored around 200 articles in peer-reviewed journals and to date completed the supervision of 30 Ph.D. trajectories.

Margaret E. Kruk, M.D., M.P.H., is associate professor of global health at the Harvard T.H. Chan School of Public Health. Dr. Kruk's research emphasizes health care utilization, health financing, quality of care, and population preferences for health services in low-income countries. Dr. Kruk is interested in the development of novel evaluation methods for assessing the effectiveness of complex interventions and health system reforms. She collaborates with governments and academics in several African countries, most recently Ethiopia, Ghana, Liberia, Mozambique, and Tanzania. She has published more than 50 papers in peer-reviewed journals and was a commissioner on the *Lancet* Global Health 2035 Commission. Prior to joining Harvard, Dr. Kruk was assistant professor of health policy and management at the Columbia University Mailman School of Public Health where she also directed the Global Health Systems, Coverage, and Quality Program. Before that, she was an assistant professor in health management and policy at the University of Michigan School of Public Health and policy advisor for health at the Millennium Project, an advisory body to the United Nations Secretary-General on the Millennium Development Goals. She holds an M.D. degree from McMaster University and an M.P.H. from Harvard University.

Ramanan Laxminarayan, Ph.D., M.P.H., directs the Center for Disease Dynamics, Economics & Policy (CDDEP). He is also a research scholar and lecturer at Princeton University. His research deals with the integration of epidemiological models of infectious diseases and drug resistance into the economic analysis of public health problems. He has worked to improve

understanding of drug resistance as a problem of managing a shared global resource. Dr. Laxminarayan has worked with the WHO and the World Bank on evaluating malaria treatment policy, vaccination strategies, the economic burden of tuberculosis, and control of noncommunicable diseases. He has served on a number of advisory committees at the WHO, the CDC, and the IOM. In 2003–2004, he served on the IOM Committee on the Economics of Antimalarial Drugs and subsequently helped create the Affordable Medicines Facility for malaria, a novel financing mechanism for antimalarials. His work has been covered in major media outlets, including the Associated Press, BBC, CNN, the *Economist*, the *LA Times*, NBC, NPR, Reuters, *Science*, *Wall Street Journal*, and *National Journal*.

Sheila Leatherman, M.S.W., is a research professor at the UNC Gillings School of Global Public Health. She conducts research and policy analysis internationally focusing on quality of care, health systems reform, methodologies for evaluating the performance of health care systems, and integrating microfinance and community health interventions. She was elected to the U.S. National Academy of Sciences in 2002 as a member of the IOM and made an Honorary Fellow of the Royal College of Physicians (2006).

In the international field of health care quality and health systems strengthening, Ms. Leatherman worked from 1997–2008 as an independent evaluator of the impact of government reforms on quality of care in the National Health Service of the United Kingdom (resulting in three books). In 2007, she was awarded the honor of Commander of the British Empire (CBE) by Queen Elizabeth for her work over the past decade in the National Health Service. In the United States, she has authored a series of books on quality of health care: general (2002), child and adolescent health (2004), Medicare population (2005). She co-authored the first national report for Canada on quality, commissioned by the Canadian Health Services Research Foundation, which was published in 2010. She works with multiple low-, middle-, and upper-income countries advising and assisting in the development of national quality agendas.

Her second area of research and practice is in the emerging field of integrating microfinance and income generation with community health interventions for poverty reduction and health promotion; projects have been conducted in Benin, Bolivia, Burkina Faso, Cambodia, India, Nigeria, Peru, the Philippines, and Tanzania.

She has a broad background in health care management in state and federal health agencies, as chief executive of a health maintenance organization (HMO), and as a senior executive of United Health Group in the United States, where she founded and directed a research center for more than 10 years.

M. Rashad Massoud, M.D., M.P.H., FACP, is a physician and public health specialist internationally recognized for his leadership in global health care improvement. He is the director of the USAID ASSIST Project. He is senior vice president of the Quality and Performance Institute at URC, leading URC's quality improvement efforts in more than 30 countries, applying improvement science to deliver better results in global health priority areas. He has a proven record of strong leadership and management.

Previously, he was senior vice president at IHI in Cambridge, Massachusetts, responsible for its Strategic Partners—IHI's key customers working on innovation, transformation, and large-scale spread. Dr. Massoud pioneered the application of collaborative improvement methodology in several low- and middle-income countries. He helped develop the WHO strategy for design and scale-up of antiretroviral therapy to meet the 3×5 target and large-scale improvement in the Russian Federation. He founded and, for several years, led the Palestinian health care quality improvement effort. He was a founding member and chairman of the Quality Management Program for Health Care Organizations in the Middle East and North Africa, which helped improve health care in five participating Middle East countries. He has worked on health care quality improvement for the Harvard Institute for International Development and the Palestine Council of Health.

Kedar Mate, M.D., is an internal medicine physician and an assistant professor of medicine at Weill Cornell Medical College and a research fellow at Harvard Medical School's Division of Global Health Equity. In addition, he serves as the senior vice president for innovation at IHI and the regional senior vice president for the Middle East and Asia-Pacific. Previously, he worked with Partners In Health, served as a special assistant to the director of the HIV/AIDS Department at the WHO, and led IHI's national program in South Africa.

In addition to his clinical expertise in hospital-based medicine, Dr. Mate has developed broad expertise in health systems improvement, innovation, and implementation science. He advises initiatives in multiple countries on developing and applying novel strategies to strengthen health systems to improve delivery of critical health services. In his leadership role at IHI, Dr. Mate has overseen the developments of innovative new systems designs to implement high-quality, low-cost health care both in the United States and in international settings.

Dr. Mate has published numerous peer-reviewed articles, book chapters, and white papers and delivered keynote speeches in forums all over the world. He teaches undergraduate and graduate-level courses in Haiti, New York, South Africa, and Tanzania. He graduated from Brown University with a degree in American history and from Harvard Medical School with

his medical degree. He trained in internal medicine at the Brigham and Women's Hospital in Boston.

Nachiket Mor, M.B.A., Ph.D., is the chairman of the board of CARE India, a board member of the Reserve Bank of India, and a board member of CRISIL. He has a background in finance and economics with a specific interest in financial access and health care. Dr. Mor worked with ICICI Bank, India's second largest bank, from 1987 to 2007 and was a member of its board of directors from 2001 to 2007. From 2007 to 2011, he served as the founding president of the ICICI Foundation for Inclusive Growth and, during this period, was also the chair of the Governing Council of IFMR Trust and board chair of FINO, both leading participants in the field of financial inclusion in India. While at ICICI, he also served as a board member of Wipro for 5 years and board chair of the Fixed Income Money Market and Derivatives Association of India for 2 years. During 2011–2012, he served as a member of the High Level Expert Group on Universal Health Coverage for India appointed by the Planning Commission of India and, during 2012–2013, as a member of the health subcommittee of the National Advisory Council of the government of India. Dr. Mor is currently also a member of the board of directors of the IKP Centre for Technologies in Public Health and Sughavazhvu Healthcare. Dr. Mor is a Yale World Fellow, has a Ph.D. in economics from the University of Pennsylvania with a specialization in finance from the Wharton School, an M.B.A. from the Indian Institute of Management, Ahmadabad, and an undergraduate degree in physics from the Mumbai University.

Dinesh Nair, M.B.B.S., AFIH, M.H.A., is a core member of the Results Based Financing (RBF) team at the Health, Nutrition and Population (HNP) Global Practice in the World Bank. His main areas of work include providing operational support to countries for the implementation of RBF and drawing together lessons from implementation, especially building the knowledge base for RBF. As a senior health specialist in the Bank, Dr. Nair brings a wide range of experiences from working as a primary care physician in a tribal coal mining area in Central India to leading some of the Bank's pioneering work in Africa. Dr. Nair has been the technical quality adviser for several West African countries and team leader of three of Nigeria's large health projects: Results Based Financing Project, Malaria Control Project, and the Polio Eradication Project. He has earlier worked with the World Bank South Asia HNP team, leading the multi-donor-financed Bangladesh health sector–wide program (Bangladesh HNPSP) and coordinating for the human development team in Bangladesh.

Prior to joining the World Bank, Dr. Nair was a health adviser with DFID, the United Kingdom's aid program, covering India and then

Bangladesh. He has rich experience in working with the nongovernment and the public sectors in India. Dr. Nair has a master's degree in health administration from the Tata Institute of Social Sciences, India, and graduated in medicine from Calcutta, India. He has had higher training in industrial health and epidemiology and surveillance.

Edgar Necochea, M.D., M.P.H., is the director for health systems development at Jhpiego. He has more than 20 years of experience in health programs and services management in developing countries, with a special focus on clinical services quality management. He co-developed the Standards-Based Management and Recognition (SBM-R) approach for performance and quality improvement in health service delivery and supported its implementation and scale-up in many countries of Africa, Asia, and Latin America. Dr. Necochea also works on health systems and human capacity development, including the elaboration of a manual for the systemic management of human resources for health, the development of human resources for health information systems, performance support systems, leadership, deployment and incentives, and workplace safety and health.

Alexander Rowe, M.D., M.P.H., is a medical epidemiologist with the Malaria Branch of the CDC. He received an M.D. from Cornell University and an M.P.H. from Emory University. He has worked at the CDC since 1994 in several areas: the chronic diseases center, an international child survival unit, and malaria. His key interests include improving health worker performance in developing countries (for all health conditions—not just malaria), strengthening health systems, monitoring and evaluation methods, and systematic reviews. He is the author or co-author of more than 50 scientific publications.

Enrique Ruelas, M.D., M.H.Sc., M.P.A., is a physician trained in public and health administration in Mexico and in Canada. He has accumulated extensive experience in academic, consulting, government, and philanthropic organizations. He was the dean of the National School of Public Health of Mexico; program director of the W.K. Kellogg Foundation for Latin America; founding president and CEO of QualiMed (the leading consulting firm on quality improvement in Latin America); vice minister of health; and secretary of the General Health Council of Mexico (a position similar to Surgeon General in the United States).

As vice minister of health, he was responsible for the design and conduction of the first national strategy for quality improvement in health care in Mexico, as a component of a major health reform. As secretary of the General Health Council, he chaired the Mexican Commission on

Accreditation of Health Care Facilities. He was founding president of the Mexican Society for Quality in Health Care; president of the Mexican Hospital Association; and president of ISQua. He was also president of the Latin American Society for Quality in Health Care and president of the National Academy of Medicine of Mexico, the most prestigious medical organization in this country. He has published extensively on quality in health care and health systems and has lectured on these topics in more than 20 countries.

He is now immediate past-president of the National Academy of Medicine of Mexico; a member of the board of directors and senior fellow of IHI; and professor and director of public policy and health systems of the Monterrey Institute of Technology and Higher Education, a leading private university in Mexico.

Nynke van den Broek, Ph.D., FRCOG, DFFP, DTM&H, is a professor in maternal and newborn health at the Liverpool School of Tropical Medicine (LSTM). She is a recognized international expert in global maternal and newborn health. At LSTM, she established and leads the Centre for Maternal and Newborn Health (CMNH), a WHO Collaborating Centre. Dr. van den Broek has designed and conducted large population-based randomized controlled trials of single interventions for improved maternal and newborn outcomes. She has used this experience to develop complex packages of interventions and to design and conduct operational research programs in multi-country settings in both Asia and sub-Saharan Africa. Impact has been ascertained through the development and application of new monitoring and evaluation frameworks and indicators to measure quality of care and maternal morbidity. Dr. van den Broek enjoys the challenge of bringing the discipline of good research methodology to the planning and evaluation of complex development programs that aim to strengthen health systems where this is expected to directly benefit maternal and newborn health.

Paul R. vanOstenberg, D.D.S., M.S., serves as the senior advisor for global growth and innovation for the Joint Commission Enterprise, including Joint Commission International (JCI). Until January 2014, he was vice president for international accreditation, standards, and measurement. Prior to returning to the JCI headquarters in November 2007, he served as the first managing director for the JCI Asia-Pacific office in Singapore and as the first managing director for the JCI European office. Dr. vanOstenberg was appointed the first executive director of international accreditation in 1998 and charged with the development of international standards and survey methods and the promotion of accreditation around the world. At that time, he was director of the Department of Standards at the Joint Commission USA.

Dr. vanOstenberg earned a B.A. from the University of South Florida (Tampa). He also received a D.D.S. (doctor of dental surgery) from the Medical College of Virginia (Richmond) and an M.S. (master in gerontology and health administration) from the Virginia Commonwealth University (Richmond).