Part 4

Health System Topics from Disease Control Priorities, Third Edition
INTRODUCTION

Just after dawn, Vivej arrives at the hospital with her newborn under her arm to see you. She is 21 years old, two days postpartum, and exhausted after 36 hours of protracted labor. She is worried because she cannot get her firstborn, Esmile, to breastfeed. You learn that she delivered at a birthing clinic near her home and tells you that, even after her water broke, it took more than a day before the birth attendant could deliver her son. Your examination reveals a dire clinical picture: Esmile is lethargic and hypotonic, he has a poor suck reflex, his temperature is 39.8°C, his pulse is 180, and his breathing is labored. You check his white blood count, confirming leukocytosis. A spinal tap shows pleocytosis. You start him on fluids and antibiotics for neonatal sepsis with likely meningitis and quickly turn your attention to Vivej. Her situation is easier to diagnose but no less urgent: she is febrile and tachycardic, her blood pressure is 85/50. You give her fluids and start her on antibiotics. Ultimately, despite your efforts, both mother and child die.

What went wrong? This chapter looks narrowly at these situations—the critical points after access and availability (including affordability) are already accomplished, when patients are in health care facilities that are staffed and equipped with appropriate technology. These are the situations in which the inputs are brought together and it is up to the provider to improve the health of the patient. Simply put, this chapter looks at the decisions and actions of the provider when seeing a patient. It is at this critical moment when we expect the doctor or nurse, or whoever is caring for the patient, to provide the best possible care by skillfully combining the available resources and technologies with the best clinical evidence and professional judgment.

Esmile and Vivej received poor-quality care at the time of delivery. Several clinical steps were not taken. The prolonged rupture of membranes was not diagnosed in a timely manner. Vivej needed either to have her labor induced or, failing that, to be referred for a cesarean section. Prophylactic antibiotics should have been administered. Just as important, the provider at the birthing center needed support and professional oversight, with guidelines, supervision, or default referral systems in place to provide a path to the best care possible. The multiple failures in this case led to puerperal and neonatal sepsis. At worst, these conditions have a fatality rate greater than one in four; at best, they lead to protracted care, recovery, and clinical expense that could have been avoided. It is possible, however, to imagine providers in a different setting, with the same physical resources, giving better care and avoiding this tragic scenario.

In the next section, we answer the questions raised in this scenario and in countless clinics and hospitals around the world. How much variation is there in the
quality of care? How do we measure clinical practice? How and where has quality been systematically improved and practice variation reduced? What elements of care variation can be addressed by policy and what are the costs? Most important, what can be done to elevate the care given by providers in developing country settings? Our focus, therefore, is on the steps that can be taken to optimize the quality of care for patients like Esmile in pediatrics, Vivej in obstetrics, and other patients receiving care for the clinical conditions considered throughout the nine volumes of the third edition of Disease Control Priorities (DCP3).

PROBLEM OF VARIATIONS IN QUALITY OF CARE

Health policy makers, researchers, and clinicians recognize the wide variations in access to care (Peabody and others, forthcoming). However, once individuals and populations avail themselves of health care services, variations in health outcomes raise disturbing questions about the quality of care delivered, defined as “the degree to which health services for individuals and populations increase the likelihood of desired outcomes and are consistent with professional knowledge” (IOM 2013, 21). Variations in care entail policy challenges similar to those associated with variations in access, including equity and efficiency (Saleh, Alameddine, and Natafgi 2014). In studies comparing clinical practice with evidence-based standards, researchers found that high-quality care is provided inconsistently to large segments of the population (McClynn and others 2003). For example, a landmark Institute of Medicine report found that, in the United States, medical errors kill more people than traffic accidents (Kohn, Corrigan, and Donaldson 2000).

Many subsequent studies have documented variations in quality of care in low- and middle-income countries (LMICs) (Barber, Bertozzi, and Gertler 2007; Barber, Gertler, and Harimurti 2007; Hansen and others 2008; Loevinsohn, Guerrero, and Gregorio 1995; National Academies of Sciences, Engineering, and Medicine 2015; Peabody, Nordyke, and others 2006; World Bank 2003). In India, studies have found alarmingly low rates of correct diagnosis, limited adherence to treatment guidelines, and frequent use of harmful or unnecessary drugs. In one study, only 31 percent of standardized patients who described symptoms of unstable angina and 48 percent who reported symptoms of asthma were given the correct drugs (Das and Hammer 2014). Even more worrying, providers prescribed an incorrect or harmful treatment to more than 60 percent of patients reporting asthma symptoms. Clinicians failed to provide even the most basic care—only 12 percent of standardized patients who reported a child with symptoms of dysentery were told to give their child oral rehydration therapy (Das and others 2012). A study of 296 providers in India found that a mere 6 percent followed the six diagnostic standards of the International Standards for Tuberculosis Care (Achanta and others 2013).

Such deficits in quality of care can come from many sources, including gaps in knowledge, inappropriate application of available technology, and inability of organizations to monitor and support care standardization. This striking variation in quality within countries occurs across facilities, among providers, and between specialists and nonspecialists (Beracochea and others 1995; Das and Hammer 2007; Das and others 2012; Dumont and others 2002; Nolan and others 2001; Peabody, Gertler, and Leibowitz 1998; Weinberg 2001; Xu and others 2015).

Some cross-national comparisons have reached the same conclusion. A 2007 DCP-sponsored study that evaluated quality for three common clinical conditions in five countries simultaneously found that the average quality of care was low in every country (61 percent) and the difference in average score between countries was small (ranging from 60.2 to 62.6 percent). However, the quality scores within every country varied widely, ranging from 30 to 93 percent (Peabody and Liu 2007). This wide variation was constant across type of facility, medical condition, and domain of care.

Poor health outcomes are the result of many factors, ranging from the nature and severity of disease to patient behavior and structural elements of care (IOM and National Academy of Engineering 2011; Steinwachs and Hughes 2008; Xu and others 2015). Some factors are not amenable to change (genetic predisposition), while others are slow to affect outcomes (changes in payment incentives). Discouragingly, better access, more infrastructure, and structural measures of quality do not always translate into better health outcomes. Indeed, some structural indexes can be inversely related to health (for example, number of hospital beds versus health status) (Ng and others 2014). Thus, improving the quality of care may well provide the greatest sectoral opportunity to improve health outcomes (Peabody and others 2017). Care can be improved quickly and, if based on best evidence, improved care will improve outcomes and lower costs (Scott and Jha 2014). Reducing unwarranted variation and addressing poor-quality provider practices deserve the most urgent attention possible from policy makers (Kirkpatrick and Burkman 2010; Ransom, Pinsky, and Tropman 2000).

Providers, health care systems, governments, and payers are beginning to recognize this urgency and are
seeking innovative, effective ways to improve the quality of care. Metrics and measurement, pathways, clinical checklists, educational interventions, and payment incentives all raise awareness and offer opportunities to provide accountability and improve care. These approaches have been tried in many LMICs, but their effectiveness varies. Changing practice at the system level is difficult and requires coordination, vision, planning, and consideration of how effective, high-impact interventions can be scaled up and applied across an entire system (Massoud and Abrampah 2015). At the level of individual providers, knowledge improvement and acquisition of new skills need to be motivated by both extrinsic and intrinsic factors, which are enabled through access to knowledge and measurement tools that change behavior and ideally are accompanied by peer support (Schuster, Terwoord, and Tasosa 2006; Woolf 2000). We have learned that improved clinical practice requires active participation (not passive learning), peer and leadership support, and communication of relevant feedback (Kantrowitz 2014; Mostofian and others 2015). Multifaceted interventions seem more successful than single interventions, underscoring the importance of practice-level change that focuses on supporting the individual provider (training) and creating a suitable environment for change (accountability).

Even more challenging than finding disease-level interventions for individual providers is identifying health care policies that improve the quality of care for populations. While clinical practice interventions, such as checklists, for acute and chronic diseases work at the provider-patient level, policies need to address group-level practice, for example, through incentives and indirect means. Preparing the deaths of Vivej and Esmile, for example, would have required the timely use of simple uterotonic commodities and prophylactic antibiotics, which might happen with better supervision. An effective policy, however, compels groups of providers to set up the supervision or the training that leads to the use of oxytocin or cephalosporins.

In the second edition of DCP, the chapter on quality of care largely summarized the emerging policy evidence that better quality could lead to better outcomes (Peabody and others 2004, 771). Just a decade later, every volume in this edition discusses quality of care. We consider in this chapter the different policy interventions that have been tried around the world. We begin with the quality infrastructure that is required for every policy intervention, then expand on the policy framework for changing clinical practice, and use this expanded framework to discuss the challenges, returns, and costs of improved quality.

**QUALITY IMPROVEMENT INFRASTRUCTURE REQUIREMENTS**

Clinical solutions are typically not generalizable because they are disease-specific, vary by clinical condition, and rely on the training of health care providers and the context of the health care system (Dayal and Hort 2015). Policy, however, is designed to work at the group level—that is, at scales larger than the individual level. Effective quality improvement policies that work at the group level have several common features, specifically the means to collect information and synthesize it and the means to encourage skills and technologies to be applied in a timely fashion. The following four common policy attributes, detailed below, improve quality:

- Measurement of the clinical activity (including measurement tied to feedback)
- Standards for those measurements (based on scientific evidence for standardizing care)
- Training of providers (including supervision)
- Incentives that align and motivate providers (including financial incentives, but also incentives of professionalism and reputation).

**Measurement**

Accurate, affordable, and valid measurements “are the basis for quality of care assessments” (Peabody and others 2004, 771). For too long, routine measures of quality in LMICs relied on structural elements (rosters, catalogs, and inventories of coverage and access), giving little thought to how these elements improve health. Such elements are relatively easy to count and measure, but are only remotely linked to better outcomes. Improving quality requires measurement of the care process—that is, what providers do when they see patients (Ansong-Tornui and others 2007; Peabody, Taguiwalo, and others 2006; Peabody and others 2011).

Measurement of the care process is critical, creating awareness of deficits in practice, gaps in care, and accountability at the individual and system levels, which improves focus and motivation. To serve as an instrument of change and accountability, provider-level measurement needs to be ongoing and cyclical. Transparency of results can increase knowledge and change intentions, but requires a supportive context to be effective (National Patient Safety Foundation 2015).

When coupled with useful feedback and done in a timely manner, measurement is the foundation for improving quality. If the measures are reliable, affordable, and anchored in valid, evidence-based criteria,
quality of care can be followed over time and the impact of policy interventions can be assessed (Felt-Lisk and others 2012). Various quality measures have been developed, each with its own set of advantages and disadvantages. Although no measure is perfect, adequate measures exist, and every health system—from small clinics to national governments—can benefit from measurement. Feedback has the potential to promote improvement, but studies are limited, tending to focus on health care report cards (Baker and Cebul 2002; Dranove and others 2003; Kolstad 2013; Shaller and others 2003), which include public disclosure of quality scores that may not provide the same motivation to improve scores as when feedback is provided privately.

The available methods for measuring performance include provider self-reports, patient vignette simulations, patient self-reports, and reviews of medical records. These methods vary in their ability to capture improvement and account for differences in the type of patients treated (case-mix adjustment). They also vary in their economic feasibility (Epstein 2006; Spertus and others 2003), reliability (repeated measures), validity (against a gold standard), and ability to be “gamed” (Petersen and others 2006). The policy challenge is that performance-measurement methods may need to be developed and adapted to low-resource settings (Engelgau and others 2010). Table 10.1 lists available methods for measuring quality of the care process.

### Table 10.1 Methods for Measuring Quality of the Care Process

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantage</th>
<th>Issues</th>
</tr>
</thead>
</table>
| Chart abstraction or review of medical record | • Nearly ubiquitous and theoretically could be obtained after the patient-provider encounter, in practice, record keeping in most LMICs is inadequate  
• Electronic medical record technology: improved uniformity, legibility, communication  
• Records of clinical events | • May lack relevant clinical details, especially when written for other purposes, such as legal protection  
• Poor record keeping and documentation lead to incomplete and inaccurate content  
• Illegibility of handwritten notes  
• Inaccuracies in the process of abstracting to produce data suitable for analysis  
• High costs involved in training medical abstractors  
• Variation in documentation practices across providers, facilities, and countries |
| Direct observation and recording of visits | • Records of clinical events  
• First-hand observation of actual encounters | • Ethical considerations  
• Need to inform providers and patients, which can induce the Hawthorne effect (bias when participant changes his or her behavior as a result of being evaluated)  
• High cost of training observers  
• Variations across observers |
| Administrative data                   | • Available in most facilities  
• Ubiquitous and inexpensive to collect when data collection system is in place | • Lack sufficient clinical detail  
• Inaccuracies in content  
• Poor data collection or management systems, especially in LMICs |
| Standardized patients                 | • The gold standard for process measurement  
• Captures technical and interpersonal elements of process  
• Reliable over a range of conditions, providing valid measurements that accurately capture variation in clinical practice among providers across patients | • Expensive  
• Not practical for routinely evaluating quality  
• Limited range of applicability (works best for adult conditions and conditions that can be simulated) |
| Clinical vignettes                    | • Can measure quality within a group of providers and evaluate quality at the population level  
• Responsive to variations in quality  
• Cases simulate actual patient visit and evaluate physician’s knowledge  
• Validated against other methods and criteria for standard-of-quality measurement  
• Useful for comparison studies  
• Easy and inexpensive to administer  
• Ability to collect data independently | • Potential resistance of providers to complete the vignettes  
• Different methods for administering vignettes  
• Instrument validation  
• Link to patient-level data |

Sources: Bertelsen 1981; Peabody and others 2004; Peabody and others 2011; Peabody, Nordyke, and others 2006.

Note: LMICs = low- and middle-income countries.
The usefulness of any method for measuring process depends on the completeness and accuracy of the data collected—a ubiquitous problem with charts, medical records, and administrative data. Another significant concern is patient case mix, given that different patient characteristics may affect quality (Zaslavsky 2001). Validity and comparability of results across measurement units (individual patients, providers, facilities, and countries) are questionable unless these differences are controlled for through complex instrument design and statistical techniques (Peabody and others 2004). Operational concerns, such as the need for highly trained staff, can increase the cost and complexity of implementing some methods.

Data Derived from Medical Charts

Chart abstraction, or review of the medical record, has long been used to measure quality of care. Clinical audits, physician report cards, and profiles are based on chart abstraction. Reliable health records can provide credible evidence of the health status of patients and assist policy makers with developing plans and making decisions to improve health care delivery (Haux 2006). The core strength of the medical record is that it is ubiquitous and could potentially be obtained after each encounter.

Chart reviews, however, suffer from many problems. First, the medical chart must be completed (and found) to proceed with an abstraction. Handwritten notes on paper charts may be illegible. Medical charts may be generated for reasons other than documenting the key clinical events of the visit (for legal protection or obtaining payment) and thus may lack crucial clinical details. Luck and others (2000) found that charts identified only 70 percent of activities performed during the clinical encounter. Even abstracting measures of quality from electronic medical records is challenging given the heterogeneity in record-keeping practices (Ali, Shah, and Tandon 2011; Parsons and others 2012). The costs and logistical challenges of securing medical records, training medical abstractors, and reviewing records can be significant. Throughout acquisition, verification, and abstraction, a process is needed to ensure that the data collected are reliable (Koh and Tan 2005). Beyond these costs and challenges, chart review also suffers from the inability to control for patient case mix and difficulty of comparing physicians caring for different patient populations.

Direct Observation and Recording of Visits

Direct observation and recording of visits are common practices in LMICs (Nolan and others 2001). Some of the most obvious challenges to using direct observation are the need to staff projects and train evaluators, which can be difficult to scale up. Ethical challenges must be addressed, and both providers and patients must be informed of the observation or recording. Although research performed in Tanzania showed that the Hawthorne effect can disappear after 10 to 15 observations, this notification introduces participation bias when providers change their behavior as a result of being evaluated (Leonard and Masatu 2006). Perhaps a more salient problem is that trained observers are costly, and variation between observers is difficult to remedy. These challenges have stimulated the search for other ways to measure and record what happens in clinical visits.

Administrative Data

Administrative data are available in all but the poorest settings. A data collection system, once established, can provide information on charges and many cost inputs. However, administrative data are assembled for purposes other than improving quality, such as documenting and processing medical claims (Calle and others 2000; Goeree and others 2009), and often lack sufficient clinical detail to be useful in evaluating clinical processes. In a 2004 study, an incorrect diagnosis was recorded 30 percent of the time (although the actual diagnosis was correct). The actual diagnosis was recorded only 57 percent of the time (Peabody and others 2004). As information systems advance, accuracy may improve, but the lack of adequate clinical detail will continue to limit the use of administrative data. Clinical databases such as registries may be helpful but are primarily available only in high-income countries (HICs) and for commercial interests.

Globally, both administrative and clinical health databases are of poor quality, and administrative databases are usually the only resource available in LMICs. Even when available, health information is underused for planning and decision making (Corrao and others 2009), especially in resource-constrained settings (Bosch-Capblanch and others 2009) and when data are paper based or decentralized to the district level (LaFond and Siddiqi 2003). District-level information systems often do not feed information back to the local level (Lippeveld, Sauerborn, and Bodart 2000). Paper-based information systems often generate poor-quality data (Lium, Tjora, and Faxvaag 2008), which weakens confidence in reported progress made toward health care system goals (Kerr and Fleming 2007) and toward the Sustainable Development Goals and the Millennium Development Goals (AbouZahr and Boerma 2005). In the absence of greater attention and resources from government or private
health insurance initiatives, using administrative data to measure and track clinical performance should be done cautiously.

**Standardized Patients**

Using standardized patients, when unannounced, is the gold standard for measuring process (Luck and Peabody 2002). Trained to simulate patients with a given illness, standardized patients present themselves in a clinical setting to providers who have given their consent to participate in the study. After the visit, the standardized patient reports on the technical and interpersonal elements of the care process. Interest in using standardized patients has been growing in LMICs, with most studies done in China and India (Das and others 2012; Das and others 2015; Mohanan and others 2015; Sylvia and others 2015). Well-trained standardized patients are not susceptible to observation bias (Glassman and others 2000) and, when rigorously monitored, enable comparisons of quality within and between facilities.

However, this method also has major drawbacks, including high costs of training, significant difficulties in large-scale application (consistent training), and limited conditions that actors can reliably portray, for example, excluding surgical and pediatric cases (Felt-Lisk and others 2012).

**Clinical Vignettes**

The shortcomings of the previous methods have spurred development of more facile methods. One of these, developed in work starting in 1999, is the use of validated clinical performance vignettes (Peabody and others 2000). Clinical performance vignettes use a full set of clinical care elements to assess the patient-provider interaction (Glassman and others 2000).

There are many types of vignettes from which to choose—for example, multiple choice versus open-ended, or short case versus full clinical care delivery scenarios—producing variable results at predicting actual practice. Clinical performance and value vignettes have been validated in randomized evaluations against standardized patients in two large trials (Peabody and others 2000; Peabody and others 2004). In these studies, vignette scores for clinical performance and value consistently reflected quality as measured by standardized patients better than abstracted medical records and worked across different health care systems, clinical conditions, and levels of training among randomly sampled physicians.

Various types of vignettes have been used in diverse settings around the world (Canchihuaman and others 2011; Das and Hammer 2005a, 2005b; Holm and Burkhardtzmeier 2015; Jörg and others 2006; Kaptanoğlu and Aktas 2013; Li and others 2007; Tiemeier and others 2002; Veloski and others 2005). Vignettes are particularly effective in comparative evaluations because the same case or type of case can be presented to many providers simultaneously, and the results can be examined over time. Vignettes have been used in large cross-national studies, such as a six-country policy evaluation in Central Asia and Eastern Europe (Peabody and others, forthcoming). This study, involving 1,039 facilities and 3,121 providers, evaluated quality of care in obstetrics, newborns, and chronic disease. Because vignettes are inexpensive to administer, they are especially well suited for use in resource-poor settings (Peabody, Luck, and others 2014; Peabody, Shimkhada, and others 2014; Peabody, Taguiwalo, and others 2006).

**Standards**

**Evidence-Based and Best-Practice Standards**

Much of the early disagreement about what to measure has given way to a consensus that performance should be measured against evidence-based criteria. The scientific literature is replete with evidence-based quality metrics that describe processes as varied as whether a patient’s blood pressure is under control, whether a patient is on the correct medication to slow down renal failure, whether the timing of a specific surgery is correct, or whether a diagnostic test is needed. Collectively, clinical care metrics are based on the evidence and the supposition that meeting these metrics results in better outcomes. Critics point out that evidence-based practice has only been established for a limited number of care elements (Contreras and others 2007; Karolinski and others 2009; Vogel and others 2014). However, clinicians routinely rely on best-practice standards, even as high-quality data from well-designed studies continue to emerge and evolve. In practical terms, there will never be a complete set of evidence-based standards, and quality of care will always rely on the best available evidence and local standards.

An important body of evidence-based, best-practice standards in LMICs comes from using surgical and childbirth safety checklists. Checklists have recently been rapidly introduced into LMIC settings, and the evidence indicates that using these evidence-based standards in checklist form improves health outcomes, primarily by setting a quality standard for treatment and facilitating communication within provider teams (Ergo and others 2012). An intervention in Michigan that used a surgical checklist to decrease catheter-related bloodstream infections in hospital intensive care units, for example, led the World Health Organization (WHO) to create the Surgical Safety
Clinical training starts in medical or other professional practice and has shown positive impacts on care. In Tanzania, training staff in the control of acute respiratory infections in young children reduced under-five mortality within two years (Mtango and Neuvians 1986). Physician-reported continuing medical education has been linked to better quality and health status when accountability is included using clinical performance vignettes (Luck and others 2014). A six-nation study linked continuing education to evidence-based practice as measured with simulated patients (Peabody and others, forthcoming). Using a systematic database of quality improvement studies, Rowe and colleagues at the U.S. Centers for Disease Control and Prevention (National Academies of Sciences, Engineering, and Medicine 2015) found that, in LMICs, training and supervision have modest positive effects on provider performance and that strategies may work better when used in combination than when used by themselves. Work by Das and others (2016) on providers in India suggests that better incentives can improve quality without any additional provider training.

Despite its ubiquity, continuing education will not greatly improve the quality of clinical practice or health outcomes (Davis and others 1999; Forsetlund and others 2009). An analysis of 62 studies and 20 systematic reviews found that the “continuing education ‘system,’ as it is structured today, is so deeply flawed that it cannot properly support the development of health professionals” (IOM, Committee on Planning a Continuing Health Professional Education Institute 2010, ix). Davis and others (2006) found that physicians cannot self-assess their skills accurately and suggested that external assessment, scoring, and feedback would drive more effective professional development. Moreover, physicians are often “not trained” to evaluate or use published guidelines and best practices. Passive dissemination of information (publishing guidelines, reading peer-reviewed articles) is generally ineffective at changing practice and is unlikely to change group-wide practice when used alone.

Newer educational techniques—targeted education, case-based learning, and interactive and multimodal teaching techniques—have had more success. Interventions that are multifaceted and include active participation and targeted feedback are much more likely to be effective than single interventions. Engaging clinicians is the key to translating training into improved quality (Mostofian and others 2015). Physicians engaged in hospital initiatives, for example, are much more likely to report successful experiences with quality improvement programs. Methods that require active physician learning (one-on-one meetings, small-group workshops, and programs tailored to a specific clinic) are effective at aligning patterns of physician practice with new clinical guidelines. In Guatemala, distance education that targeted diarrhea and cholera case...
Supervision

Supervision is an established method for assessing quality. The power and influence of peer review supervision, often conducted through professional societies, vary widely among countries (Heaton 2000). Large providers, such as hospitals or public health institutions, often have more resources for collecting information on provider practices and patient outcomes and for using those data to guide, educate, supervise, discipline, or recognize providers. Providers at clinics and primary care facilities also benefit from supervision (Loevinsohn, Guerrero, and Gregorio 1995). Other studies point to the benefits of quality review committees and standing groups that review all hospital deaths. However, oversight can also create an antagonistic relationship between workers and managers that may preclude cooperative problem solving and continuous improvement (Berwick 2002).

Incentives

Demand Incentives

Demand-side interventions, such as conditional cash transfer (CCT) and voucher programs, pay participants (not providers) a stipend for specific behaviors, for example, attending school, having up-to-date vaccinations, or visiting a health center for prenatal care (box 10.1). Although CCTs do not directly provide incentives to health care providers, they require quality health services, adding a supply- or provider-side component to demand-side interventions. There is also an indirect supply-side incentive when consumers use cash incentives to pay for services. A systematic review of the evidence suggests that CCTs improve the uptake of preventive services by children and pregnant women (Lagarde, Haines, and Palmer 2009).

However, in shorter time frames of months to a year, CCTs have difficulty driving lasting effects and affecting health (Beegle, Frankenberg, and Thomas 2001; World Bank 2003). From a policy perspective, it is also difficult to distinguish the effects of the CCT incentive from the impact of the cash itself, that is, it is unclear whether the behavioral change is associated with the conditional incentive or with an income effect (Fernald, Gertler, and Neufeld 2008). A systematic review of the impact of vouchers found modest evidence that the vouchers improved quality of care (Brody and others 2013). The question that remains is whether there are long-term effects because clinical practice was not improved.

Provider Payment

In the past two decades, health care administrators and policy makers in both LMICs and HICs have been using pay for performance (P4P) as a means to improve clinical practice. Although the details of programs vary, health care P4P programs link physician compensation to measures of clinical quality (Epstein, Lee, and Hamel 2004). P4P and other forms of results-based compensation have been used routinely in business settings. The challenge in health, however, is to identify suitable metrics that are under the control of the provider (Werner and Asch 2007). For example, care providers are hard pressed to reduce infant mortality rates that are driven primarily by poverty and nutrition, but they can readily change the frequency of unnecessary cesarean sections.

Even with suitable metrics, isolating and linking P4P changes in practice to better health has been challenging (Atkinson and others 2000; Derose and Petitti 2003). P4P might be linked, at best, to modest improvements in quality of care (Epstein 2007; Lindenauer and others 2007; Petersen and others 2006; Rosenthal and others 2005). However, most studies are not experimentally designed, and participation in P4P programs is voluntary, leading to selection bias. Although much of the literature on the equivocal benefit of provider incentive systems comes from HICs, the Quality Improvement Demonstration Study (QIDS), carried out in the Philippines as a social policy experiment, provides.
strong experimental evidence that P4P can be effective in an LMIC (Quimbo and others 2008) (see box 10.2). Similar results were found in the work by Gertler and Vermeersch (2013).

The large QIDS randomized community-level experiment found greater improvement in health outcomes than previous P4P studies (Peabody and others 2017). This finding may have occurred because most other studies providing incentives to doctors have been conducted in wealthier countries and been nonrandomized, which introduces the possibility of selection bias wherein providers who adopt the incentives may be the most likely to respond and improve their clinical practice anyway (Petersen and others 2006). Three randomized P4P studies conducted in the United States found that rewarding physicians improved outpatient care, such as immunization rates (Fairbrother and others 1997; Fairbrother and others 2001; Kouides and others 1998). However, other randomized studies found that physician P4P had no effect on mammography, other cancer screening, or adherence to pediatric preventive guidelines (Grady and others 1997; Hillman and others 1998; Hillman and others 1999). Three hospital-based studies examining inpatient P4P programs in the United States also included control hospitals. These studies, which focused on adult care in cardiovascular disease, community-acquired pneumonia, and joint replacement, found modest improvements of 2 to 4 percentage points in outcomes beyond the performance seen in controls (Glickman, Boulding, and others 2007; Grossbart 2006; Lindenauer and others 2007). Although these studies had controls, the interventions were not randomly assigned.

**Results- and Performance-Based Financing**

Results-based financing (RBF) encompasses various types of interventions that provide demand-side incentives (for example, CCTs), refine provider payments (for example, P4P), and trigger government reforms.

The RBF lending projects financed by the Health Results Innovation Trust Fund and World Bank credits or loans (World Bank 2014) operationalized the concept of RBF at a large scale in many LMICs and intended to provide incentives to policy makers to build and leverage their quality infrastructure as a condition for financing. Since 2008, RBF projects like these have been widely adopted in more than 30 countries, with interventions at the national, subnational, district, facility, and community levels. Operationally, funds are provided to governments at the national and subnational level based on agreed-on disbursement-linked indicators and their established targets (often nation- and state-wide estimates). At the facility level, payments to individual facilities are based on their contracts with fund holders (often district or provincial health authorities). And, increasingly used at the community level, payments are

**Box 10.2**

**Impact of P4P on Quality: Results of the Quality Improvement Demonstration Study**

The Quality Improvement Demonstration Study (QIDS) is unique in that it was an explicit policy experiment that randomized communities into pay for performance (P4P) versus universal health coverage versus a true control. P4P improved both quality and outcomes.

QIDS was a large policy experiment conducted in the Philippines among 119 doctors, 3,162 children, and 30 communities, covering about one-third of the country. The communities were randomized into an incentives-based policy program rewarding physicians financially for providing higher-quality care to children than provided by universal health coverage and controls (Quimbo and others 2008). In the communities where doctors were eligible for the bonus payments, the number of children who were not wasted (underweight for height) increased 9 percentage points relative to control sites. The share of parents reporting at least good health for their children was 7 percentage points higher in P4P sites than in controls (Peabody, Shimkhada, and others 2014).

The introduction of P4P led to improvements in quality of care as measured by clinical case vignettes (Peabody and others 2011). Difference-in-differences model estimations indicated that P4P improved not only the measured quality of physician practice but also health outcomes. The impact of policy can be measured in a relatively short (two-year) time frame when evaluation is integrated into policy making and planning before implementation (Peabody and others 2017), making it possible to measure policy effectiveness and to identify ineffective polices early on.
provided to community organizations or community health workers based on RBF contracts with fund holders (often districts or facilities).

A flexible approach, RBF focuses on results:

- Payments linked to results (both demand and supply side) based on context-specific health priorities
- Contracts or agreements that clarify the responsibilities of all stakeholders
- Autonomy for those contracted to be able to use RBF funds to attain the agreed-on results most effectively
- Verification of results to ensure that they are accurate and reliable
- Data sharing to enhance the results, which can be used for planning, design, and implementation
- Community involvement to enhance accountability

RBF operational data show improvement of quality (especially structural quality) in the RBF programs. Facilities’ quarterly quality scores, calculated based on a supervisory checklist, improved in almost all of these programs. In Burundi, for example, quality scores improved significantly during the first two years following rollout of a national RBF program (figure 10.1). In Ethiopia, where RBF was implemented at the national government level, the Ministry of Health undertook Service Availability and Readiness Assessment (SARA) of its primary care facilities on an annual basis to achieve targets associated with disbursement-linked indicators and developed action plans to address weaknesses identified through SARA.

Impact evaluation studies show positive evidence about the impact of RBF programs on certain dimensions of quality. Several countries, including Argentina, Rwanda, and Zimbabwe, report improvement in quality of prenatal care (Basinga and others 2010; Gertler and Vermeersch 2013; World Bank 2014). Afghanistan demonstrated substantial improvements in quality of examinations and counseling, as well as time spent with patients (Engineer and others 2016). Under Argentina’s Plan Nacer incentives-based program, the estimated probability of low birthweight was reduced by 19 percent among beneficiaries, and in-hospital neonatal mortality for babies of enrolled mothers was reduced by 74 percent (Gertler, Giovagnoli, and Martinez 2014).

RBF programs exercise interventions beyond provider performance incentives, such as policy reform, system strengthening, transparency improvement, and management and accountability enhancement. Because of this, establishing the effectiveness of clinical interventions through randomized controlled trials becomes a challenge. How best to use operational data and experiences remains important in disentangling the effects of incentives and the key bottlenecks addressed by RBF.

**LINKING POLICY AND PRACTICE AT THE PLATFORM LEVEL**

How do quality infrastructure policies at the government level translate into improved clinical care at the patient level? At its heart, quality improves only when providers deliver the right care to the patient at the right time, do so efficiently, and focus on the patient. However, less variation among a group of providers depends on individual providers treating their patients and their diseases the same way. This section examines how policy and practice come together at the platform level. Specifically, we review the policy elements described above that would be implemented for 11 clinical conditions across four platforms.

We start by looking at where care services are delivered. Delivery occurs through various platforms, from community and public health settings to primary care clinics, first-level hospitals, and the most advanced facilities in every country.

The quality of care will vary in each setting, which means that the policy elements discussed above are relevant to each setting. These policy elements are categorized as quality measurement, practice standards, training management, and incentives.

Table 10.2 shows how each policy element might be implemented across the four delivery platforms.
Table 10.2 Infrastructure Elements for Improving Quality for 4 Delivery Platforms and 11 Representative Clinical Conditions

<table>
<thead>
<tr>
<th>Infrastructure elements</th>
<th>Disease or condition</th>
<th>Community-based services</th>
<th>Primary health centers</th>
<th>First-level hospitals</th>
<th>Referral and specialized hospitals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Outcomes</td>
<td>Metrics</td>
<td>Outcomes</td>
<td>Metrics</td>
</tr>
<tr>
<td>Quality measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurement</td>
<td>Reproductive health</td>
<td>Coverage rates or service use, provider knowledge, patient behavior (condom use); unintended pregnancy rate</td>
<td>Prenatal, perinatal care; recognition of high-risk pregnancies</td>
<td>Referral rates; folic acid coverage; ability to recognize high risk</td>
<td>Management of labor; vaginal delivery or cesarean section</td>
</tr>
<tr>
<td>Feedback and accountability</td>
<td>Cardiovascular disease</td>
<td>Use of nutritional and exercise programs</td>
<td>Blood pressure, lipid, diabetes screening; management</td>
<td>Screening at the population level; screening with patient data on blood pressure, lipids</td>
<td>Triage of acute myocardial infarction and treatment of congestive heart failure</td>
</tr>
<tr>
<td>Practice standards</td>
<td>Sexually transmitted infections</td>
<td>Patient knowledge, safe sex practices</td>
<td>Patient surveys of knowledge, behaviors; provider surveys of clinical knowledge regarding sexually transmitted infections; cultural competency in communication</td>
<td>Syphilis screening, treatment of gonorrhea</td>
<td>Direct observation of successful management and treatment</td>
</tr>
</tbody>
</table>

*table continues next page*
<table>
<thead>
<tr>
<th>Infrastructure elements</th>
<th>Disease or condition</th>
<th>Community-based services</th>
<th>Primary health centers</th>
<th>First-level hospitals</th>
<th>Referral and specialized hospitals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Outcomes</td>
<td>Metrics</td>
<td>Outcomes</td>
<td>Metrics</td>
</tr>
<tr>
<td>Checklists, clinical guidelines</td>
<td>Pediatric infectious diseases</td>
<td>Preventive, evidence-based measures to prevent disease</td>
<td>Immunization rates; incidence of disease</td>
<td>Diarrhea treatment, referral</td>
<td>Provider compliance with guidelines (charts, vignettes); provider's ability to diagnose, referral rates</td>
</tr>
<tr>
<td>Licensing, certification, accreditation</td>
<td>Infectious disease</td>
<td>Provider hygiene, handwashing; proper disposal of needles</td>
<td>Direct observation of program implementation</td>
<td>Wound care with suturing; aseptic technique; instrument sterilization</td>
<td>Direct observation of explicit management or treatment criteria</td>
</tr>
</tbody>
</table>

**Training, management**

| Training | Mental health | Provider and community awareness of mental health; destigmatization of mental health illness | Provider and community attitudes, knowledge using surveys; incidence surveys of mental illness by socioeconomic status; destigmatization of mental health at community level | Acute mental health first aid and triage (suicide prevention, crisis intervention, disaster counseling); screening for ASD | Institutional training outcomes (provider knowledge) for diagnosis and counseling; provider's use of screening for ASD | Emergency care and hospitalization for acute psychosis; treatment and detoxification of substance abuse | Presence of care coordination and team practice with counseling and drug therapies available; provider's ability to diagnose; referral rates; treatment per institutional guidelines | Long-term care for dementia, chronic affective disorders, schizophrenia | Provider-level compliance according to evidence-based care; patient-level data: use of procedures, complications |

*Table 10.2 Infrastructure Elements for Improving Quality for 4 Delivery Platforms and 11 Representative Clinical Conditions (continued)*
<table>
<thead>
<tr>
<th>Infrastructure elements</th>
<th>Disease or condition</th>
<th>Community-based services</th>
<th>Primary health centers</th>
<th>First-level hospitals</th>
<th>Referral and specialized hospitals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuing medical education</td>
<td>Diabetes</td>
<td>Patient preventive behaviors: physical activity, healthy eating</td>
<td>Whether continuing medical education requirements are being met; provider knowledge regarding patient programs and ways to engage patient in behavioral change; patient knowledge</td>
<td>Diabetes management with behavioral interventions and medication</td>
<td>Knowledge-based testing</td>
</tr>
<tr>
<td>Management</td>
<td>Accidents, injury, trauma</td>
<td>Provider's and patient’s knowledge and use of preventive measures for injury (child safety, car seats, water safety; elder safety)</td>
<td>Provider's and patient knowledge surveys</td>
<td>Provider's ability to recognize and assess severity of injury or complications</td>
<td>Provider-level data on ability to make correct diagnosis (vignettes), time to treatment, referral rates</td>
</tr>
<tr>
<td>Professional oversight</td>
<td>Cancer</td>
<td>Smoking cessation, hepatitis B immunization rates, school-based human papillomavirus vaccination</td>
<td>Patient-level data on immunization rates, cancer incidence; hospital-based cancer registries</td>
<td>Screening for breast, colon, cervical, lung, and skin cancer</td>
<td>Assessment of provider’s knowledge of risk and referral standards (set and disseminated); patient screening rates</td>
</tr>
</tbody>
</table>
### Table 10.2 Infrastructure Elements for Improving Quality for 4 Delivery Platforms and 11 Representative Clinical Conditions (continued)

<table>
<thead>
<tr>
<th>Infrastructure elements</th>
<th>Disease or condition</th>
<th>Community-based services</th>
<th>Primary health centers</th>
<th>First-level hospitals</th>
<th>Referral and specialized hospitals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Outcomes</td>
<td>Metrics</td>
<td>Outcomes</td>
<td>Metrics</td>
</tr>
<tr>
<td><strong>Incentives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance-based remuneration</td>
<td>Orthopedics</td>
<td>Provider’s knowledge and communication of preventive behaviors (healthy eating, immunization, healthy lifting); patient’s behavior of the same</td>
<td>Provider’s knowledge and behavior surveys; patient participation in preventive health programs; monitoring of physical activity of patients</td>
<td>Care for low-level trauma (simple broken bones), management of low-back pain</td>
<td>Provider’s ability to diagnose low-back pain (vignettes); use of certain prescription drugs for pain management</td>
</tr>
<tr>
<td>Team-based, multidisciplinary care (global payment)</td>
<td>Malaria</td>
<td>Provider’s knowledge and communication of vector control to patient (use of insecticides)</td>
<td>Provider’s knowledge and behavior surveys; patient’s participation in preventive health programs; community malaria rates</td>
<td>Provider’s ability to recognize malaria rapidly; use of ACT</td>
<td>Provider’s ability to diagnose type of malaria (drug resistant or not); proper treatment, long-term follow-up, management</td>
</tr>
</tbody>
</table>

Note: Quality interventions provided at lower-level platforms are also provided at higher-level facilities. ACT = artemisinin-based combination therapy; ASD = autism spectrum disorder; HIV/AIDS = human immunodeficiency virus/acquired immune deficiency syndrome; SARS = severe acute respiratory syndrome.
For each element, the table details how quality outcomes and metrics could be operationalized for a given disease or clinical condition. For example, community-based services in reproductive health (a condition) would focus on family planning and fertility management, which can be assessed by metrics of patient behavior (condom use); primary clinics would focus on high-risk pregnancies, which can be assessed using referral rates for women at risk. Outcomes and metrics tend to become more concrete as care progresses across platforms. Primary clinics and first-level hospitals, for example, might require chart-level data or provider-level assessments of skill, knowledge, and practice. Specialized hospitals, where care is more complex (treatment of birth complications) and outcome metrics are more serious (mortality rates), are likely to have more readily available data and better outcomes. A key element of quality improvement, whether at the specialized hospital or community clinic level, is that the effectiveness of the improvement strategy must be assessed regularly. Recommendations published by the WHO and the International Association for Trauma Surgery and Intensive Care on quality improvement strategies are broadly applicable to all levels of care and types of settings and include strategies such as morbidity and mortality conferences to review errors occurring during the care of patients, panel reviews of preventable deaths, and tracking of complications, adverse events, sentinel events, and errors (WHO, Association for Trauma Surgery and Intensive Care, and International Society of Surgery 2009).

**UPDATED QUALITY OF CARE FRAMEWORK**

As shown in figure 10.2, health actions take place in the context of and are influenced by political (laws, governmental stability), cultural and social (societal norms, norms, policies), environmental (natural, political, economic, social, and cultural factors), and institutional (policy levers, payment systems, provider, patient, platform, and access) factors. The framework illustrates how these factors interact and influence each other, ultimately affecting health outcomes. The policy levers include health care access, structure, systems, and human resources; clinical process and care; and provider skill, performance, and adherence to guidelines. The payment systems include incentives and the provider, patient, platform, and access levels. The social and cultural factors include norms and values that influence health behaviors and outcomes.

**Figure 10.2 DCP3 Approaches to Improving Quality of Care Framework**

Note: Blue indicates items discussed in this chapter; DCP3 = Disease Control Priorities, third edition.
practices), environmental (natural disasters), and institutional (functioning health departments, corruption) factors. Demographic and socioeconomic makeup, including genetics and personal resources, also affect the health status of individuals seeking care.

The classic construct of structure, process, and outcome is at the core of the framework (Brook, McGlynn, and Cleary 1996; De Geyndt 1995; Donabedian 1980; McGlynn 1997). These three elements are described in table 10.3.

Structure refers to stable, material health care assets (infrastructure, tools, technology, implements), the resources of the organizations providing care, and the financing of that care (levels of funding, staffing, payment schemes, incentives). These factors can be measured inexpensively and data are typically readily available (De Geyndt 1995).

Process captures the interaction between caregivers and patients, including appropriateness of the care delivered, cognitive skill of the provider, and communication (Murray, Gakidou, and Frenk 1999). The private nature of the doctor-patient consultation, lack of measurement criteria, and absence of reliable measurement tools make it difficult to assess process, especially in LMICs (Peabody and others 2004). However, new approaches to measuring process have come a long way toward capturing process measures across settings.

Outcome includes direct measures of health status, death, or disability-adjusted life years as well as patient satisfaction or patient responsiveness to the health care system. Outcome measurement has matured in the past decade with the use of electronic medical records and data registries.

The updated framework in figure 10.2 adds policy levers for improving quality of care and showcases the provider’s practice and behavior as well as the unique perspectives of policy makers, physicians, and patients, which are essential to establishing accountability. The early frameworks focused on the lack of structural inputs, whereas recent frameworks look at care processes (Kruk and others 2009). The Institute of Medicine was the first to include additional elements of care regarding safety and efficacy, patient focus, affordability and timeliness, and effectiveness (Berwick 2002; IOM, Committee on Quality Health Care in America 2001). The remainder of this section discusses these elements.

### Safety and Efficacy
Patient safety has not received enough attention in LMICs. Globally, up to 1 in 10 patients is harmed by an adverse incident in a hospital not directly related to his or her clinical care, with approximately US$6 billion in costs per year (WHO 2008). Even procedures that are not considered high risk in HICs have the potential to lead to harm or poor outcomes in LMICs. For example, up to 1 in 4 cataract surgeries in India results in poor visual acuity (Lindfield and others 2012).

A study on patient safety practices in low-income countries suggests that improved staff-patient communication, use of protocols, control of infections, and standardization between providers can improve overall safety (Lindfield, Knight, and Bwonya 2015). Efficacy of care has an ascendant role in clinical practice as the practice of evidence-based medicine continues to expand. Many new and exciting studies of clinical efficacy are

### Table 10.3 Quality-of-Care Framework

<table>
<thead>
<tr>
<th>Elements of quality</th>
<th>Description</th>
<th>Subcomponents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure</strong></td>
<td>Stable, material characteristics (infrastructure, tools, technology) and resources of the organizations that provide care and the financing of care (levels of funding, staffing, payment schemes, incentives)</td>
<td>• Physical characteristics&lt;br&gt;• Management (executive leadership, board responsibilities)&lt;br&gt;• Culture&lt;br&gt;• Organizational design&lt;br&gt;• Information management&lt;br&gt;• Incentives</td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td>The interaction between caregivers and patients during which structural inputs from the health care system are transformed into health outcomes</td>
<td>• Making the diagnosis&lt;br&gt;• Providing evidence-based treatment</td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td>Measures of health status, deaths, or disability-adjusted life years (a measure that encompasses the morbidity and mortality of patients or groups of patients); outcomes such as patient satisfaction or patient responsiveness to the health care system</td>
<td>• Morbidity&lt;br&gt;• Mortality&lt;br&gt;• Patient satisfaction</td>
</tr>
</tbody>
</table>

Sources: Glickman, Boulding, and others 2007; Peabody, Taguiwalo, and others 2006.
driving better care, including the use of antibiotic prophylaxis before surgery and the elimination of antibiotics for otitis media.

**Patient Focus**

As with efficacy, focus on the patient and his or her perspective has become more prominent, leading evaluations of performance to include satisfaction as a necessary outcome. The availability and growing acceptance of patient satisfaction surveys are striking given that these tools were almost unheard of 20 years ago.

The focus on the patient is important because patients’ or users’ perspectives determine whether they seek care and where they obtain services (demand). This perspective is based on the individual’s own opinions, previous experiences with the health system, and input received from others.

Perception of low quality has been reported as a major factor in the decision not to use or to bypass health services. For example, in a study in Tanzania, 42 percent of women who delivered children in a health care facility in rural parts of the country bypassed the local primary care clinic and delivered directly in a hospital or health center (Kruk and others 2014). This finding is striking because all of them lived near a functioning clinic with delivery services and the sample excluded women referred to a hospital. Primary care clinics tend to have poor infrastructure, lack equipment, and are understaffed, and women may choose care based on their perception of specific factors, such as a provider’s attitude or competency and the availability of drugs and medical equipment.

**Affordability and Timeliness**

Determining affordability is challenging given that there is no recognized, consistent association between affordability and quality. High-quality care is often assumed to mean more expensive care (Starfield and others 1994). Indeed, early quality improvement efforts were often costly because the quality interventions themselves had to be paid for, and new measures of performance had to be introduced to calibrate the baseline quality and detect subsequent change (U.S. Congress, Office of Technology Assessment 1994).

However, high-quality care is potentially more affordable care because consistent, high-quality, standardized care entails fewer unnecessary tests, less time spent in the hospital, and fewer complications. In the United States, as much as one-third of health care costs are unnecessary, and as much as US$799 billion in costs is due to unexplained variation in practice and quality (Health Affairs 2014). Estimates are not available for LMICs, but as much as one-third of health care costs may be due to unexplained variation in quality and unnecessary care in practice. A study in eight countries found that the introduction of surgical guidelines in hospitals led to less variation in quality, better health outcomes, and lower costs (Haynes and others 2009).

**Effectiveness**

Effectiveness refers to how well evidence-based practices are followed. Translating promising research findings and evidence, especially results that improve health or lower health care costs, into scalable interventions is challenging. The high stakes—and rare successes—have led to increasing calls for evidence-based policy making. Ideally, evidence-based policy making is based on evaluations of real-world economic effectiveness, allowing a determination to be made of value as well as efficacy.

With this effort has come interest in determining the comparative cost-effectiveness of clinical interventions. Few studies compare policy approaches to quality improvement. Peabody and others (2017) compared a demand-side intervention (universal health coverage for children under age five years) with a supply-side intervention (P4P scheme for physicians) and found that both interventions were effective, reducing wasting by about 9 percent (relative to controls). Costs were notably lower in the supply-side intervention than in the demand-side intervention, suggesting that increasing quality is more cost-effective than expanding insurance benefits in resource-constrained settings.

**CHALLENGES FOR ASSESSMENT**

The conversation on quality needs to include issues related to equity, misdiagnosis, perceptions, accountability, and learning from patients, all of which are challenging to assess.

**Equity**

Equity is an increasingly recognized part of the quality equation. Inequality—a situation in which poor-quality care is disproportionately provided to people from a particular disadvantaged group—is rampant worldwide (Barber, Bertozzi, and Gertler 2007; Barber, Gertler, and Harimurti 2007; Hansen and others 2008). Socioeconomically disadvantaged groups have poorer access to services and, once they have access, are less likely to receive effective treatment (Garrido-Cumbera and others 2010; Health Affairs 2011; Rogers 2004).
If they are lucky enough to obtain treatment, they receive poorer-quality care than people from other groups. The impact of quality interventions on equity has not received enough attention in the literature.

**Misdiagnosis**

Misdiagnosis, also referred to as diagnostic error, is a significant shortcoming, with worrisome, albeit poorly understood, consequences (box 10.3). For example, a study reported that 5 percent of adults are misdiagnosed during outpatient visits, and about 50 percent of these errors could harm the patient (WHO 2000). Misdiagnosis in breast cancer is as high as 20 percent in some cases (Lozano and others 2006).

Misdiagnosis is likely to be especially high in LMICs (Galaktionova and others 2015). A study in India found that only one-third of primary care providers articulated a diagnosis, either correct or incorrect, and when a diagnosis was given, close to 50 percent were wrong (Marchant and others 2015). In an observational study of primary care providers in rural China, the misdiagnosis rate was 74 percent, and clinicians provided medicine that was unnecessary or harmful to 64 percent of their patients (WHO and World Bank 2014). Diagnostic errors occur around the world and in all types of settings, suggesting a need to include misdiagnosis in conceptualizing quality-of-care deficiencies.

**Real-world practicalities make investigating misdiagnoses a substantial challenge. Methodological problems include the difficulty of aggregating patients with the same diagnosis to overcome the unobserved (and unrecorded) case-mix variation, legitimate disagreements on reference standards for practice, reliance on recorded retrospective data, and challenges of measuring a clinician's cognitive thought processes. Perhaps the biggest methodological challenge is to reach some agreement regarding the correct diagnosis. Short of having a group of experts reexamine the patient, the correctness of diagnoses is difficult to evaluate.**

**Perceptions of Quality**

Identifying a perspective—or multiple perspectives—from which to assess quality is difficult (Strauss and Corbin 1998; Tafreshi, Pazargadi, and Abed Saeedi 2007; Van der Bij, Vollmar, and Weggeman 1998; Wisniewski and Wisniewski 2005). Judging quality requires balancing the competing viewpoints of many players in the system. For example, payers and purchasers typically judge quality by how well insurance premium dollars are spent for each covered life; patients typically judge quality by how well their individual needs are addressed; and physicians assess quality by their own clinical judgment or training, patient demands, available resources, and cost-controlling mechanisms (Luck and others 2014).

---

**Box 10.3**

**Misdiagnosis as a Core Element of Poor Quality**

Diagnosis is a key determinant of a successful outcome (Freedman and Kruk 2014). Yet the extent of misdiagnosis has not been fully recognized (Jamison and others 2013; Ng and others 2014; OECD 2015; Rockers, Kruk, and Laugesen 2012; WHO 2000). A wrong diagnosis will lead, at best, to unnecessary evaluations and treatment and, at worst, to harmful tests and toxic treatment. Diagnostic errors result in potential delays in treatment, putting the patient at risk (WHO 2000) and leading to severe complications and overtreatment. They are an important cause of preventable morbidity and mortality (Freedman and Kruk 2014; Jamison and others 2013; Ng and others 2014; Rockers, Kruk, and Laugesen 2012).

The field of obstetrics provides a rich opportunity to study misdiagnosis in LMICs. A study examined the prevalence and consequences of misdiagnosis among 103 obstetrical providers in an urban area of the Philippines using identical vignettes and reviewing each provider’s clinical records (Shimkhada and others 2016). The misdiagnosis of three common obstetric conditions—obstructed labor, postpartum hemorrhage, and preeclampsia—was almost 30 percent overall. Providers who misdiagnosed these conditions were more likely to have patients with a complication. Patients with a complication were significantly less likely to be referred to a hospital immediately and were more likely to be readmitted to a hospital after delivery, to have significantly higher medical costs, and to lose more income than patients without a complication.
When different perspectives collide—for example, when physician performance metrics (penalties for high surgical complication rates) are not in the best interest of the patient (a diabetic who is a higher surgical risk and may be turned down for surgery to keep complication rates low)—the patient’s outcomes, including satisfaction, should be given the greatest weight.

Accountability
Establishing accountability is challenging. It can be difficult to determine which platform is responsible for achieving certain measurement goals and which individuals within each level should be held accountable for those measures (Emanuel and Emanuel 1996; Wachter 2013). The challenge of establishing accountability is tied to the larger challenge of convincing all players that poor quality should not be attributed to an individual clinician. Poor quality cuts across all types of care, facilities, providers, health insurance offerings, geographic areas, and patient populations. Accountability must be established at all levels (Brinkerhoff 2003). Holding physicians accountable may be especially difficult in a fee-for-service environment where individuals are used to being independent, and there are significant methodological, political, and legal obstacles to measuring accountability (Quimbo and others 2008).

A common trap is to let the availability of data determine which system-level metrics are tracked. System accountability is analogous to provider accountability, and metrics must be relevant, reliable, valid, comprehensive, and financially achievable; data availability should not drive the selection of metrics (Hsia 2003). Accountability also means that those who judge quality have the opportunity to go beyond explicit, evidence-based measures of practice or even structure. Recent work points to system- and platform-level accountability for collaboration, local ownership, and shared learning (Boucar and others 2014).

Learning from Patients
A final, neglected area of quality assessment is health system responsiveness to patients, specifically data on the patient’s experience and satisfaction with care (Bernhart and others 1999). Therefore, improving the patient experience is a stand-alone goal of health systems in the updated framework (Rockers, Kruk, and Laugesen 2012; WHO 2000).

Initiatives such as the current push for universal health coverage assume that people will value and want to fund health benefits, whether through taxes or premiums. Public support, however, is shaped in important ways by an individual’s health system experiences. For example, in addition to health outcome data, the Organisation for Economic Co-Operation and Development now measures the patient experience, including metrics on wait times, communication, and costs of care.

Methods of obtaining data on the patient experience include exit surveys (in person or anonymous), mailed or online questionnaires, and, increasingly, phone surveys. The large and growing penetration of mobile phones makes it more and more feasible to collect short telephone or mobile Web assessments of the patient experience in LMICs (Solon and others 2009).

IMPACT OF QUALITY IMPROVEMENT
Global health goals and projections are predicated on assumptions about achieving high coverage and improving the quality of care in high-mortality countries (Jamison and others 2013). Given the lack of high-quality data from LMICs, data from high-income settings are used to predict health gains from expanded coverage in LMICs. These extrapolations do not reflect the real-life impact of quality on use and eventual outcomes in LMICs. Diagnosis and treatment, for example, are often egregiously poor in understaffed, underresourced and underregulated health systems. Yet it is critical to understand whether health care visits translate into quality health care—both for projecting better health and for estimating the health returns on initiatives such as universal health coverage.

Influence on Demand for Services and Outcomes
Quality of care is a major driver of use. Various studies have shown that perceived quality of care influences patterns of use—for example, perceptions of poor quality can motivate patients to stay at home or to choose far-away providers perceived to be more competent (Bohren and others 2014; Kruk and others 2009; Leonard 2014). Perceptions of poor quality are a strong factor pushing patients to bypass care, as are users’ assessments of the complexity of their health needs (Akin and Hutchinson 1999; Kruk and others 2014; Leonard, Mliga, and Mariam 2002). In sum, patients in low-income settings increasingly behave like their rich-country counterparts: as active consumers making rational choices about their care rather than as passive beneficiaries of health care.

The demand for quality is likely to grow as coverage expands. Kruk and others (2015) found that, when childbirth at a health facility (that is, in-facility delivery)
exceeds 80 percent of all births in a community, proximity to hospitals, not primary care clinics, matters in predicting delivery of care, potentially because of growing demand for high-quality care that is difficult for low-volume clinics to deliver.

How accurately do patients assess quality? Although patients are well positioned to report on interpersonal or nontechnical quality-of-care issues, such as clarity of communication, respect, confidentiality, and waiting times, they do not have full information with which to gauge the technical quality of care. Doyle, Lennox, and Bell (2013) found that the patient experience of care was positively associated with clinical effectiveness and safety in more than 75 percent of studies. For example, Glickman and others (2010) found that higher patient satisfaction was linked to lower mortality among patients with acute myocardial infarction. Similarly, more satisfied patients had lower 30-day hospital readmission rates and higher adherence to physician recommendations (Boulding and others 2011; Fenton and others 2012). Other research found little correlation between patient ratings of care and chart-measured adherence to standards of care, use of inpatient care, or mortality (Chang and others 2006).

Whether accurate or not, perceptions drive behavior. Patient ratings of quality and satisfaction are also associated with future care seeking, an important consideration given the rise of chronic diseases requiring ongoing contact with the health system (Bohren and others 2014; Groene 2011; Kruk and others 2014; Sun and others 2000). More work is needed to understand which patient assessments are most reliable and the best ways to collect these data.

Patient-reported quality and satisfaction are important indicators of the responsiveness and accountability of health systems (Thaddeus and Maine 1994). Responsiveness, defined as meeting patients’ nonhealth expectations, should be a goal of every health system (WHO 2000). Yet recent research has documented disrespectful and abusive treatment of patients in health facilities. For example, nearly 20 percent of women in two districts of Tanzania reported harsh treatment by health workers, including yelling and slapping (Freedman and Kruk 2014). Such treatment leads to a loss of confidence (Kujawski and others 2015). Abusive treatment is distressingly common in other settings as well (Asefa and Bekele 2015; Gourlay and others 2014; Okafor, Ugwu, and Obi 2015; Sando and others 2014).

**Fit between Services and Patient Needs**

One promising strategy is to improve the fit between people’s expectations and health needs and the health services available to them. This tailoring of care is an example of patient-centered reform (Groene 2011). For example, when the quality of obstetric care provided at first-level, low-volume facilities is of poor quality, referrals to higher levels of emergency care is inefficient, resulting in excessively high maternal and newborn mortality (Hsia and others 2012; Thorsen and others 2014). Women who deliver in the health system clearly prefer higher-volume, higher-quality facilities, as evidenced by choice of provider. Thus, the answer to improved quality and outcomes may be to establish high-volume maternity health centers or hospital units and provide support for travel to these facilities, rather than to invest more in primary care obstetrics or low-volume, first-level facilities. Focusing on customer service and respect requires paying attention to staffing, training, and supervision.

Health systems that can satisfy people’s expectations may experience a double benefit: better health outcomes and greater support for the health system. For example, women who bypassed their first-level clinic and delivered in hospitals rated quality of care more highly than women who delivered in first-level clinics across a wide range of indicators (Kruk and others 2014). Experiencing responsive health services may enhance confidence in government. A multicountry study of LMICs found that a combination of high-quality care and financial risk protection raised the probability of having trust in government by 13 percent (Rockers, Kruk, and Laugesen 2012). More responsive, patient-centered health systems should be a health and political priority.

**COSTS OF IMPROVING QUALITY**

Almost all deficits in the quality of care can be addressed if enough resources are made available for the purpose. The question is not, “Can we improve the quality of health care services?” Instead, it is, “How can we use the resources available to achieve that improvement?” Thus, when resource constraints are considered, policy makers will have to choose from a range of interventions, and the question becomes, “What are the most efficient and feasible ways to improve health outcomes?” For example, nosocomial infections could be treated with costly antibiotics, new facilities, and equipment. However, it is likely to be far more efficient to introduce a handwashing protocol, to ensure that providers comply with it, and to develop a rapid response team that can be deployed when infections occur.

The costs of improving quality are different from the costs of the intervention itself. For example, the cost of delivering care to patients with closed fractures requiring internal fixation includes facility costs (patient room,
equipment, sterile supplies), personnel costs (clinicians, support staff), and patient costs (transportation to the facility, time costs). If a high proportion of patients develop nosocomial infections, the cost of quality would be the costs incurred to reduce the risk of facility-associated infection through strategies such as providing training, supervising staff, procuring new cleaning and sterilization equipment, and developing care pathways or checklists.

Cost-effectiveness analysis (CEA) can be used to determine how cost-efficient a quality improvement intervention is. CEAs compare the resources consumed and the effects on the desired outcome of an intervention to improve the quality of care against a valid comparison, which is either business-as-usual or a different intervention. Three results are possible. First, the intervention may fail to improve the outcome of interest and is not cost-effective at any price. Second, the intervention may achieve the intended improvements, but require additional resources, in which case implementation is a matter of willingness to pay for the level of improvement achieved. Third, the intervention may improve health outcomes as a result of better quality while also reducing overall expenditure. Lower cost comes from spending a lesser amount on care or avoiding an expensive complication or an adverse event. Economically, it is best to implement all interventions matching the third result.

There is a dearth of literature on the cost-efficiency of quality of care interventions (IOM, Committee on Quality of Health Care in America 2001). Several difficulties are involved in determining efficiency:

- Inaccurate, incomplete, or unavailable routinely collected data
- Fidelity of the intervention to the outcome stated in research design
- The challenge of choosing comparison groups to isolate the variable of interest
- The difficulty of capturing all of the effects of the intervention to account for positive or negative spillover effects
- The challenge of calibrating the extent to which the quality improvement can be attributed to the intervention
- The perceived costs and economic consequences meaningful to different audiences
- The difficulty of valuing in-kind contributions
- The difficulty of capturing complexity of a system and the implications for economic evaluation.

Nevertheless, CEA can show substantive returns from better quality. In one study in Niger, high quality from a quality improvement collaborative conducted in childbirth facilities reduced the overall costs per birth an average of 20 percent (from US$35 to US$28); when accounting for the decrease in average clinical costs due to improved efficiency and the reductions in postpartum hemorrhage, the authors determined that the incremental cost of the improvement collaborative was US$2.43 (Broughton, Boucar, and Alagane 2012). The incremental cost-effectiveness was an impressive US$147 per disability-adjusted life year averted, compared with US$870 for the rotavirus vaccine, US$135 for hypertension treatment, and US$1,480 for a tobacco tax (Tran and others 2014). Interventions to improve health care quality can also save money as shown in the example of improving uptake of Kangaroo Mother Care for premature and low birthweight infants in Nicaragua (Broughton and others 2013). In this case, the cost of the improvement intervention was less than the cost savings realized from decreased treatment costs resulting from improved adherence to evidence-based standards of care.

Despite the many difficulties in determining efficient ways to address deficits in the quality of health care, it is important to include these cost analyses in every quality improvement intervention. Systematic accounting for the resources and rigorous evaluation of the effects on the outcomes of interest are essential for prioritizing decision making. Basic guidance on what costs to include in economic evaluations and how to analyze cost and effectiveness data is needed to move the field of health care quality forward.

CONCLUSIONS

In LMICs, quality of care is an emerging conversation. Mostly ignored a few decades ago, studies are now examining health system priorities once access to care has been addressed. Conversations over the past 10 years have largely acknowledged the importance of quality of care in resource-constrained LMICs. Quality of care is discussed in all volumes of *DCP3*.

Quality of care matters because it relates directly to outcomes and can be addressed in a shorter time frame than other policy interventions. The updated quality framework presented in this chapter describes the urgency, connections, and responsibilities for creating quality infrastructure that ties this responsibility to individual providers through the diseases they address and the patients who access care via various health care platforms. The framework is applicable across country settings, emphasizing the fundamental role that providers and patients play in
determining quality. With the growing evidence base on quality improvement efforts around the globe, there is reason for renewed hope that quality can be improved and done so rapidly. Successful policies will always be linked to practice on a disease-by-disease basis and will only occur where access to health care is not in question.

NOTES

World Bank Income Classifications as of July 2014 are as follows, based on estimates of gross national income (GNI) per capita for 2013:

- Low-income countries (LICs) = US$1,045 or less
- Middle-income countries (MICs) are subdivided:
  - (a) lower-middle-income = US$1,046 to US$4,125
  - (b) upper-middle-income (UMICs) = US$4,126 to US$12,745
- High-income countries (HICs) = US$12,746 or more.

1. A program that delivers insurance for maternal and child health services to uninsured families.

REFERENCES


of Routine Health Information at District Level.” RHINO (Routine Health Information Network), Mpokweni Sun, Eastern Cape, South Africa, September 29–October 4.


