April 29, 2014

Disease Control Priorities in Developing Countries, 3rd Edition
Working Paper #8

Title: **Policy instruments to improve intervention uptake and provider quality**

Author (1): Devra Barter
barter@cddep.org

Affiliation: Center for Disease Dynamics, Economics & Policy
1616 P St NW Suite 6000 Washington DC 20036

Author (2): Charles H. Brower
brower@cddep.org

Affiliation: Center for Disease Dynamics, Economics & Policy
1616 P St NW Suite 6000 Washington DC 20036

Author (3): Aditi Nigam
nigam@cddep.org

Affiliation: Center for Disease Dynamics, Economics & Policy
1616 P St NW Suite 6000 Washington DC 20036

Author (4): Patrick Renschler
renschler@cddep.org

Affiliation: Center for Disease Dynamics, Economics & Policy
1616 P St NW Suite 6000 Washington DC 20036

Author (5): Kelsey Walters
walters@cddep.org

Affiliation: Center for Disease Dynamics, Economics & Policy
1616 P St NW Suite 6000 Washington DC 20036

Author (6): Ramanan Laxminarayan
ramanan@cddep.org

Affiliation: Center for Disease Dynamics, Economics & Policy

April 29, 2014

1616 P St NW Suite 6000 Washington DC 20036

Correspondence to: Aditi Nigam nigam@cddep.org

Keywords: conditional cash transfers, payment for performance, vignette studies, direct observation, protocol based methods

Abstract:

Often overlooked, complementary interventions allow technical interventions to achieve impact by enhancing access, uptake and quality. This paper discusses the availability and effectiveness of policy instruments by reviewing the evidence on five categories of policies to improve the uptake of intervention (demand side) and the quality of provision (supply side). Recent studies are reviewed to consider variations in provider quality and effort, interventions and incentives to improve health providers' performance, and protocol-based methods to improve provider quality. Financial interventions to improve patient uptake are also discussed.

1. Introduction

The vast literature on what should be purchased by public finances for population health avoids an explicit discussion of the complementary interventions that enable technical interventions to achieve impact. Interventions to improve the quality of technical interventions include financing tools on supply and demand side, operate at the individual and population level, and are applicable to both public and private sectors. Many of these complementary interventions enhance access, uptake, and quality and work both independently and in concert with other interventions. In this paper we discuss the availability and effectiveness of policy instruments to improve access, uptake, and quality of healthcare. We review the evidence on five categories of policies to improve the uptake of intervention (demand side) and the quality of provision (supply side). Policies are categorized as financial or procedural. Interventions such as health worker training are examined elsewhere and are outside the scope of this review. In section 2, we review the evidence on variations in provider quality and effort from recent studies. Next, we look at interventions and incentives to improve health providers' performance, and protocol-based methods to improve provider quality. Financial interventions to improve patient uptake are also discussed here. Section 3 concludes the paper.

2. Literature review

2.1 What do we know about providers' competence and effort?

Over the past 25 years countries across the world have improved access to healthcare providers and facilities.[1] Whether this increase in accessibility translates to an increase in quality of care is not clear.[2] The vast differences in providers' knowledge, effort, and setting can lead to drastic differences in quality of care. Improvements in measurements, through vignette studies and direct observations, have allowed researchers to measure quality of care across the world.[3] Vignette studies are hypothetical scenarios in which a researcher poses as a patient to present a hypothetical case to the provider with the provider's knowledge. By contrast, a direct observation occurs when the researcher observes patient-provider interactions without the provider's knowledge. Vignettes are used to measure what providers "know" (competence), and direct observations show what providers actually "do" (effort).[3] Together, these tools are useful for researchers to determine a provider's knowledge and effort. When vignettes and direct observations are combined to study quality of care, researchers can identify any "know-do" gaps.[3]

An overview by Das et al. (2008)[2] draws measurements of provider competence and effort from studies in four low-income countries—Tanzania[4], India[5], Paraguay[6], and Indonesia[7]—to determine quality of healthcare. Methods and samples used in these studies are in Table 4 from Das et al. (2008).[2] The review exposes the low quality of care as measured by provider competence in these countries as well as the compounded low quality of care due to the "know-do" gap.[2] Evidence from vignette results shows the low level of care quality due to provider competence in the countries studied. Multiple vignettes were independently created in each study, and each provider was presented with multiple cases.[2] The questions asked, examinations performed, diagnosis made, and treatment given by each provider were compared with checklists or protocols created by experts.[2] Results from the vignette responses (Figure 1 from Das et al. 2008[2]) show that a majority of providers in India, Indonesia, and Tanzania do not know basic essential procedures for common diseases in their countries.[2] In the India sample, a provider with average competence is more likely to harm than treat the patient.[2] The vignette results also show that in poorer areas, providers are less competent and therefore provide a lower quality of care than those in richer areas.[2]

In India and Tanzania, vignettes combined with direct observation show a positive correlation between competence and effort.[2] The data also expose a gap between what providers know and what they do.[69] Das et al. (2008) found that with low levels of competence, providers did what they knew. For example, in India's private sector, providers without an MBBS (equivalent to a medical qualification) knew about 20 percent of the essential tasks when tested and also performed about 20 percent of the essential tasks when interacting with a real patient.[2] The limiting factor for these providers was their competence. But as providers' competence increased, performance did not keep up.[3] Providers with an MBBS in the Indian private sector knew 40 percent of the essential tasks when tested but performed only 25 percent of them.[2] This gives clear evidence that the limitation in quality of care from more competent providers is a lack of effort.

Das and Hammer (2014) illustrate the "know-do" gap using data from Delhi, India (Figure 2).[3,5] If providers did what they knew, all three lines would be at a 45-degree angle to the axis, since their effort should match their competence.[3] However, both the public and private sector lines show very low effort for the highest level of knowledge. Figure 2 shows a larger gap in the public sector compared with the private sector. Differences in provider effort by affiliation can also be seen in Das et al. (2008).[2] In their India sample, public providers with an MBBS knew 30 percent of essential tasks but performed only 8 percent of them.[2] This "know-do" gap is larger in the public sector (knowledge: 30 percent, effort: 8 percent) than in the private sector (knowledge: 40, effort: 25 percent).[2] In the Paraguay sample, there was no difference in competence between providers working for NGOs and providers working for the public sector (both knew 50 percent of the tasks), but providers in the public sector put in significantly less effort (36 percent of tasks were performed) than those in the NGOs (44 percent).[2]

Variation in providers' effort across affiliations can also be seen through the average measurements of patient-provider interaction time, number of questions asked, and number of exams performed.[2] On average in public clinics in the India sample, providers spent less than two minutes with a patient, asked one question, and performed less than one examination per patient.[2] A comparison with the average values for the private sector given in Table 5 shows that providers in the public sector put in less effort. In Tanzania, providers working for NGOs are found to put in more effort than providers in the public sector.[2] In Paraguay, providers with

permanent contracts put in more effort than those with temporary contracts.[2] Clearly, public sector providers make less effort in providing quality healthcare.

Variation in provider effort also exists across countries.[2] Providers in Paraguay spent more time, with patients (8.33 minutes) on average, than providers in high-income countries such as Germany and Spain (Table 5).[2] This is also seen between Paraguay and middle-income countries such as India. On average, Paraguayan providers asked 8.23 questions and performed 2.65 physical exams while providers in Delhi spent 3.8 minutes with patients, asked 3.2 questions, and performed 1.09 exams.[2]

Motivational differences are potential causes of the varying effort across countries and affiliations. Brock et al.[8] found that peer observation and encouragement in Tanzania led to increases (4 and 10 percent, respectively) in providers' quality of care. Also using data from a study in Tanzania,[9] Das et al. (2008) found that as incentives increased in a facility, providers' effort increased as well.[2]

The studies above indicate the generally low quality of care and large "know-do" gaps that exist in low-income countries. Since the knowledge, motivation, and effort of a healthcare provider are major determinants of health outcomes, efforts to improve the quality of care should go beyond just ensuring access to healthcare facilities.

2.2 Payment-for-performance to healthcare providers

Payment-for-performance schemes have been implemented in both developing and developed countries and used to improve overall healthcare quality, chronic disease management, maternal and child health, and care for infectious diseases. These schemes create incentives for healthcare providers to improve quality of care or meet certain quality indicators. The incentives can target improved processes or better outcomes, and they can be implemented at the provider, group, or hospital level.[10] Examples of studies analyzing effects of payment-for-performance schemes in various geographical and healthcare settings are detailed below.

Several studies have evaluated payment-for-performance programs that focus on maternal and child health outcomes. Basinga et al.[11] conducted an impact evaluation of a performance-based payment of healthcare providers' intervention on prenatal care visits and institutional deliveries, quality of prenatal care, and child preventive care visits and immunization. The research team randomized 166 health facilities to receive the intervention (n=80) or continue with input-based funding (control group, n=86). The intervention consisted of direct payments made to healthcare facilities by the Rwandan Ministry on Health based on submission of monthly activity reports of 14 quality of care indicators related to maternal and child health. Compared with the control facilities, the intervention facilities were more likely to have a higher standardized total quality score ($p= 0.02$), mothers attending intervention facilities were more likely to have an institutional delivery ($p= 0.017$), and children under 23 months and 24–59 months were more likely to have had a preventive care visit in the previous four weeks ($p=0.004$ and $p<0.001$, respectively). No significant difference was observed for prenatal care or tetanus vaccine delivery. Additionally, the authors found that the effects were larger for facilities that received larger financial incentives.

Priedman Skiles et al.[12] analyzed the effects of performance-based financing in Rwanda on the use of facilities for maternal deliveries, antenatal care, and contraceptive use using a difference-in-differences model that compares intervention and control districts. In 2005, the Rwandan government implemented the incentive program in several districts to motivate providers to improve the quality of care and increase service output. The program monitored 14 quality indicators related to maternal and child primary healthcare, through quarterly site visits. Subsequent payments were based on service outputs and a quality score. Results indicated no significant differences in the use of maternal health services between intervention and control sites except for one indicator, facility birth deliveries ($p=0.014$), in which the probability increased by 10 percentage points for the intervention sites. The authors also found no significant differences when the outcomes were stratified by patients' income quintile. The authors concluded that performance-based financing may be useful if targeted at specific services, such as facility deliveries, but only if service use was consistently low.

Peabody et al.[13] considered payment-for-performance incentives and child health outcomes in the Philippines. The study evaluated an incentives program that was randomly assigned to 10 district

hospitals and matched to 10 control sites. The authors used clinical performance vignettes among randomly chosen physicians every six months over a three-year period to assess physicians' quality indicators, which determined bonus payments if qualifying scores were met. Outcomes of interest included age-adjusted wasting, C-reactive protein, hemoglobin level, parental self-reported health of children, and children under age five hospitalized for diarrhea or pneumonia. None of the results showed significant improvement in health outcomes between the control and intervention sites from the beginning of the study to after the intervention period, except for parental self-reported health of children, which increased by 6.53 percentage points in the intervention sites compared with a decrease of 0.85 percentage points in the control sites ($p=0.001$). The authors also found that the number of children who were wasted increased in both sites, but the increase was greater in the control sites (9.77 percentage points) than in the intervention sites (0.51 percentage points) ($p<0.0001$); the increase was attributed to extreme weather disturbance in 2006. The authors concluded that the payment-for-performance incentive program led to improvements in two health outcomes.

Fairbrother et al.[14] randomly assigned physicians in New York City to one of three intervention groups or a control group to assess physician-level interventions on rates of childhood immunization for DTP, Hib, OPV, and MMR. Physicians received (1) payment bonuses and feedback (information about immunization of their patients), (2) an enhanced fee for service and feedback, or (3) feedback only. Physicians in the first group received financial bonuses if their patients' up-to-date immunization status showed improvement from a baseline or if the physicians achieved 80 percent immunization coverage levels. Coverage rates increased over time for all physician groups, including control groups, but only the first group (bonus and feedback) had a statistically significant change ($p<0.01$) compared with the control group, with an increase in up-to-date immunization coverage of 25.3 percentage points. The authors also found that the percentage of immunizations received outside the participating practice increased significantly in the bonus and feedback group ($p<0.01$). The authors concluded that bonuses increased immunization coverage in the medial record, but this was due to better documentation instead of better performance or quality indicators.

Huntington and colleagues[15] conducted a case control study to analyze the effects of payment-for-performance schemes on maternal and child health care services in Egypt. The schemes were

designed to encourage facilities and providers to meet certain standards and performance targets for preventive and curative services and quality indicators. Some primary health care unit providers received performance incentives; others received equivalent amounts as salary "top-offs." Statistically significant ($p<0.05$) differences were observed between the two groups for the proportion of patients receiving laboratory investigation (11.2 percent for the incentive group vs. 26.2 percent for the top-off group), follow-up for family planning (73.9 vs. 59.0 percent), childcare services including treatment prescription (92.2 vs. 98.2 percent), childcare follow-up (54.9 vs. 30.7 percent), administration of medicine (20.0 vs. 9.5 percent), and some antenatal care indicators, such as urine tests (83.5 vs. 63.5 percent) and blood pressure tests (96.2 vs. 87.8 percent). Statistically significant differences were not seen with regard to history taking, medical examinations, or management of tetanus toxoid, iron, vitamins, or other treatment.

Payment-for-performance programs have been used to improve TB care. Li and colleagues[16] conducted a retrospective analysis of the effect of payment-for-performance incentive programs on cure rates and treatment length for TB patients in Taiwan between 2002 and 2005. Payment incentives were provided based on treatment categories and patient identification and cure. The authors found statistically significant differences ($p<0.01$) in the number of cases cured within nine months and the average length of treatment before and after implementation of the payment-for-performance program. Additionally, there was a significant difference in the number of patients cured in intervention vs. control hospitals ($p<0.01$). The 12-month cure rate was statistically higher in intervention vs. control hospitals (68.1 vs. 42.4 percent, $p<0.01$). The authors concluded that after the introduction of the payment incentives program, both the cure rate and the length of treatment for TB patients improved. Additionally, intervention hospitals had significantly better treatment outcomes than nonintervention hospitals.

Other studies have evaluated the effects of payment-for-performance schemes on chronic disease and primary healthcare outcomes, including diabetes, hypertension, and smoking cessation. Bardach et al.[17] conducted a cluster-randomized controlled trial of primary health care clinics in New York to evaluate the effects of a payment-for-performance incentives scheme on pre-specified quality performance criteria. Eighty-four clinics were randomized to receive either financial incentives and benchmarked quarterly reports of their performance (intervention, n=42) or quarterly reports only (control, n=42). Quality performance criteria included differences in

performance improvement from the start to the end of the study, prescriptions for aspirin and antithrombotic medication, blood pressure and cholesterol control, and smoking cessation interventions. Compared with control clinics, intervention clinics showed greater performance improvement from the start to the end of the study, for aspirin therapy (OR 1.28; 95 percent CI, 1.10–1.50), for smoking cessation intervention (OR 1.30; 95 percent CI, 1.04–1.63), and for blood pressure control in patients without diabetes mellitus or ischemic vascular disease, with only diabetes mellitus, and with both diabetes mellitus and ischemic vascular disease ($p<0.05$). No significant differences were observed for cholesterol control or blood pressure control for those with ischemic vascular disease only. The authors of this study concluded that payment-for-performance incentive schemes resulted in modest improvements in cardiovascular care in small primary care practices in New York.

Glickman et al.[18] conducted an analysis comparing acute myocardial infarction care outcomes for hospitals (n=54) participating in payment-for-performance programs initiated by the U.S. Centers for Medicare and Medicaid Services (CMS), launched in 2003, and control hospitals not participating in the program (n=446). Compared with control hospitals, intervention hospitals showed more improvement in quality measures for patients receiving aspirin at discharge ($p=0.04$), smoking cessation counseling ($p=0.05$), and a lipid-lowering agent at discharge ($p=0.02$) in 2006 (after introduction of the scheme) compared with 2003 (before). However, no significant difference was found in mortality or other quality-related indicators between intervention and control hospitals; both groups of hospitals showed improvement in most quality-related indicators assessed in the study from 2003 to 2006.

Rosenthal and colleagues[19] conducted a retrospective analysis comparing clinical quality outcomes of cervical cancer screening, mammography, and hemoglobin A_{1c} testing from large health plans that began paying one type of medical group bonuses if they met 10 clinical and service quality targets (n=134) compared with medical groups that did not receive bonuses for quality targets (n=33). Differences in improvement in clinical quality scores were significant for cervical cancer screening between intervention and control groups, with an improvement difference of 3.6 percent ($p=0.02$). However, there were no significant differences in improvement for mammography or hemoglobin A_{1c} testing ($p>0.05$) between the intervention and control groups. The authors concluded that there was little gain in paying medical groups to

reach a fixed performance target and that this type of incentive system might only reward medical groups with higher performances at baseline.

Yip and colleagues[20] conducted a matched-pair cluster-randomized controlled trial evaluating the effects of capitation with pay-per-performance on providers' antibiotic prescribing practices, healthcare spending, number of outpatient visits, and patient satisfaction in Ningexia Province in China. The authors compared the intervention with standard fee-for-service payments established by the government in China in the 1980s. Payment was given based on performance scores of village posts aggregated into an overall score of the township health centers and compared with average scores in the county. Results showed that in the intervention township health centers and village posts, antibiotic prescriptions fell by 6.6 percentage points ($p<0.005$) (adjusted relative risk, 15 percent) and 6.0 percentage points ($p<0.005$) (adjusted RR 16 percent), respectively, compared with control health centers. The authors found no other significant difference on spending, outpatient visits, or patient satisfaction. They concluded that capitation combined with payment-for-performance programs could improve prescribing practices.

A study by Peterson and colleagues[21] considered payment-for-performance incentive schemes for hypertension care. The authors evaluated payment-for-performance on outcomes, including appropriate response to uncontrolled blood pressure, medication prescription, and number of patients who developed hypotension at the individual physician level, practice level, both, and neither. Results showed that all incentive groups had a change in the percentage of physicians meeting the quality measures for blood pressure control and use of antihypertensive medications during the study period. Compared with the control group, only the individual incentives group had a significant difference in blood pressure control, 8.36 percent (95 percent CI, 2.4–13 percent). The authors concluded that only individual provider-level incentives resulted in increased control of blood pressure, but no incentive scheme led to increased use of antihypertensive medication or a greater incidence of hypotension compared with control groups.

Millet et al.[22] conducted a longitudinal study in 36 primary care practices of a payment-for-performance incentive program for providers in the United Kingdom, examining its effect on support for smoking cessation among patients with diabetes. In 2004, the country introduced quality targets related to smoking cessation and provided incentives to general practitioners to

achieve them. The authors found that smoking status was more likely to be recorded in the patient's medical record in 2005 (98.8 percent) compared with 2003 (90.0 percent) ($p<0.001$). Additionally, the proportion of patients with documented smoking cessation support increases from 2003 (48.0 percent) to 2005 (83.5 percent) ($p<0.001$). The authors concluded that the payment-for-performance scheme introduced in 2004 increased general providers' support for smoking cessation among their patients with diabetes and was associated with a documented decrease in smoking prevalence among these patients from 2003 to 2005.

Cheng et al.[10] analyzed the effects of a payment-for-performance program on diabetes care in Taiwan, in which physicians receive financial incentives if their patients have follow-up visits, the providers perform diabetes-specific tests and examinations, and the providers meet predetermined treatment goals. The authors used a natural experiment design with propensity score matching to analyze difference-in-differences to compare outcomes between the intervention and control groups before and after implementation of the program. Compared with the control group, the intervention group had a higher likelihood of completing the seven essential tests and examinations during the course of the study ($p<0.001$), as well as diabetes-related physician visits during the first year of the study ($p<0.001$), but the differences were not significant during the rest of the study period. The authors concluded that the program was successful in improving diabetes care.

Cutler and colleagues[23] evaluated a payment-for-performance program in California on chronic disease care management for enrollees in a medical group. The program pays medical groups who meet nine clinical quality measures established by the National Committee for Quality Assurance. This study evaluated a subset of these quality measures, specifically looking at cholesterol management, including an LDL-C test, and compared outcomes of cholesterol management for providers in participating and nonparticipating medical groups (with 165 and 1,694 patients, respectively). Testing for patients in the intervention group was higher than for the control group (91.5 vs. 67.8 percent) ($p<0.001$). The LDL-C goal attainment rate was also higher for the intervention group than for the control group (78.2 vs. 55.7 percent) ($p<0.001$). The authors concluded that patients managed under the payment-for-performance program had higher rates of LDL-C testing and higher rates of goal attainment than patients managed under routine care.

Other studies have evaluated payment-for-performance programs on overall healthcare quality. Lindenauer et al.[24] compared hospital healthcare quality voluntarily reported by acute-care hospitals participating in a payment-for-performance program funded by CMS (n=207) with healthcare quality reported by a group of nonparticipating hospitals (n=406). All hospitals reported on 10 quality measures related to heart failure, myocardial infarction, and pneumonia. Hospitals in the intervention group received bonus payments if they performed in the top 20 percent of composite measures of care quality. The authors estimated the incremental effect of payment-for-performance on quality measures. When matched for hospital characteristics, the incremental effect of the program on acute myocardial infarction was 4.3 percent (95 percent CI, 2.5–6.1 percent) ($p<0.001$); for heart failure, 5.2 percent (95 percent CI, 2.8–7.7 percent) ($p<0.001$); for pneumonia, 4.1 percent (95 percent CI, 2.3–5.9 percent) ($p<0.001$); and for the 10 overall composite measures, 4.3 percent (95 percent CI, 3.0–5.7 percent) ($p<0.001$). The authors concluded that hospitals participating in financial incentives programs achieved modest improvements in care quality over nonparticipating hospitals.

The studies summarized above give a varied picture of the effectiveness of payment-for-performance interventions on health outcomes. Several systematic reviews enumerate the effectiveness and breadth of payment-for-performance schemes, including a systematic review of systematic reviews.[25–31] The reviews note that the payment-for-performance effects range from having extremely positive effects to modest or no effects in addition to unintended consequences.[25,26,31,32] Houle and colleagues[31] detail the effect sizes of payment-for-performance interventions on various health outcomes in each of the papers included in the review (see Tables 1 and 2, below).

A few studies noted the lack of research on the effect of payment-for-performance programs on healthcare quality, costs, or cost-effectiveness.[32,33] Several authors comment on the difficulty of drawing general conclusions about the effectiveness of interventions from the literature, since payment-for-performance evaluations vary in study design, intervention design (who receives payments, how frequently they are administered, the magnitude of incentives, etc.), and externalities (funding, setting, organization context, etc.).[27,30] Additionally, Eijkenaar and colleagues[27] point out that many studies have failed to disentangle the effect of the payment-for-performance interventions from the effect of other simultaneous improvements. The authors

concluded that overall, the effects of payment-for-performance on physicians' performance were estimated at 5 percent improvement. More rigorous evaluations with robust study designs will be useful in evaluating the true effect of payment-for-performance schemes on health outcomes.

2.3 Protocols to improve provider's quality of care

Protocol-based methods can improve quality of care at low cost.[34] The principal interventions in this domain have been surgical safety checklists and, more recently, childbirth safety checklists. Checklists improve health outcomes primarily by setting treatment quality standards and facilitating communication within provider teams. Current safety checklists have drawn on implementation science pioneered in other industries—from aviation[35] to Formula One racing[36] and nuclear power[37]—in hopes of injecting consistency and accountability into healthcare operations. Assuming high levels of compliance, carefully designed safety checklists can help teams manage crises efficiently,[38] eliminate "never events" (wrong patient, wrong procedure, wrong surgical site, retained surgical items, etc.),[39] and minimize additional health risks for patients, such as infections.[40] In this section we review research on the role of surgical safety checklists and childbirth safety checklists in improving the quality of healthcare delivery.

Complications arising from surgical care are a major cause of disability and death worldwide. The World Health Organization (WHO) cites surgical complication rates of 3 to 16 percent and surgical mortality rates of 0.4 to 0.8 percent for developed nations.[41] For developing nations, the surgical mortality rates are 5 to 10 percent.[41] Under these assumptions, the WHO estimates that each year 7 million patients suffer from significant surgical complications.[41] Despite the magnitude of this health problem, many of these adverse events are procedure related[42] and preventable within current resource constraints.[43–45] A surgical safety study performed across three U.S. hospitals reported that 43 percent of adverse events were associated with communication failures within the operating team.[46]

Following the success of a surgical checklist intervention that decreased catheter-related bloodstream infections in Michigan hospital ICUs,[47] the WHO Surgical Safety Checklist was designed[48] and has been adopted by 1,790 healthcare facilities worldwide since 2012.[49] The checklist addresses avoidable surgical complications and specifically targets never-events by

including steps such as marking the patient's operation site.[37] It also accounts for necessary equipment, details proper administration of anesthesia and antibiotic prophylaxis, and provides additional preventive measures by ensuring that patients who are at risk for complications (surgical site infections, blood loss, etc.) are identified.[39] Most importantly, the checklist (and accompanying briefings) promotes effective communication among the operating team, so that all individuals feel comfortable voicing their concerns.[37] Currently, the WHO Surgical Safety Checklist applies only to operating room procedures, but adaptations like the SURgical PAtient Safety System (SURPASS) checklist consider the entire surgical process, from intake to discharge.[50]

Given the theoretical benefits of surgical safety checklists, many studies have assessed their quantitative effect on health outcomes.[48,51–56] Bergs et al.[57] (2014) provide a systematic review and meta-analysis of seven studies that report quantitative effects of the WHO Surgical Safety Checklist on postoperative complications. The meta-analysis was performed for three health outcomes: any postoperative complication, mortality, and surgical site infections. The results were that all three negative health outcomes were reduced by the use of surgical safety checklists: any postoperative complication (Risk Ratio = 0.59; 95%CI = 0.47 – 0.74), mortality (Risk Ratio = 0.77; 95%CI = 0.60 – 0.98), and surgical site infections (Risk Ratio = 0.57; 95%CI = 0.41 – 0.79).[57] See Table 3 for the full results from the systematic review. Briefly, five (of six reporting) studies found statistically significant decreasing rates of postoperative complications: 11 vs. 7 percent (P <0.001),[48] 18.4 vs. 11.7 percent (P =0.001),[52] 22.9 vs. 10.0 percent (P =0.03),[54] 23.6 vs. 8 percent (P <0.001),[55] and 21.5 vs. 8.8 percent (P <0.001).[56] Two (of five reporting) studies found statistically significant decreasing mortality rates: 1.5 vs. 0.8 percent (P =0.003)[48] and 3.7 vs. 1.4 percent (P =0.007).[52] Three (of six reporting) studies reported statistically significant decreasing rates of surgical site infections: 6.2 to 3.4 percent (P<0.001),[48] 11.2 to 6.6 percent (P<0.001),[52] and 14.9 to 4.7 percent (P <0.001).[56]

Unlike surgical safety checklists, the literature on childbirth safety checklists is scant. To date, the one pilot study on the benefits of childbirth safety checklists suggests that of the 350,000 maternal deaths, 1.2 million intrapartum-related stillbirths, and 3.1 million neonatal deaths globally each year, many could be avoided within the current standards of care.[58] The WHO childbirth safety checklist was developed to help reduce the major causes of these deaths

(hemorrhage, infection, obstructed labor, etc.).[58] Since most deaths associated with childbirth occur within a 24-hour window and the major causes are well described, checklists pose a promising role for improving health delivery.[58] However, Spector et al.'s[58] (2012) pilot study did not quantify health outcomes and examined only adherence to accepted clinical practices. The authors report that 28 of 29 essential practices were delivered with significantly greater frequency when the childbirth safety checklist was used.[58] Follow-up studies intend to focus directly on health outcomes attributable to the increase in these practices.

Despite their benefits, checklists for surgery and childbirth safety should not be considered magic bullets: without the full engagement of healthcare providers, checklists will prove futile.[59] Thanks to increasing global awareness,[60–62] more healthcare providers have begun to voluntarily explore protocol-based methods for improving health outcomes.

2.4 Incentives to improve households' health behaviors

Conditional cash transfers (CCTs) are innovative social safety net programs that aim to break the cycle of poverty and promote intergenerational mobility.[63] Similar to welfare programs, CCTs give cash transfers to impoverished families that meet certain criteria (e.g., income cutoffs or residence in targeted geographical areas).[64] However, CCTs are distinguished from traditional social welfare programs in that payments are conditional on the recipients' compliance with certain requirements.[65,66]

Generally, the conditions for beneficiary status involve actions that invest in the human capital of the recipient family. For instance, Fernald et al.[64] note that the majority of CCT programs make transfers contingent on families' attendance at educational sessions on health and nutrition, regular visits to primary healthcare facilities for preventive health services, and school attendance rates of children. In this way, CCT programs aim to reduce current poverty immediately while reducing future poverty through investments in human capital.[67]

CCT programs are widespread—as of 2009, some 29 developing countries had CCT programs[68]—and often are a country's largest social assistance program.[68] The review below summarizes the results of 11 studies on various outcome measures (use of health services, health

outcomes, nutrition and development status, and school attendance rates) from CCT programs in eight countries as well as a systematic review of CCTs worldwide. For the purpose of this review, it is important to note that recent work in this field (e.g., Baird et al. 2012 and Robertson et al. 2013)[67,69] has investigated and compared the relative effectiveness of CCTs and unconditional cash transfer (UCT) programs; the latter eliminate operational costs required to ensure that conditions for payments are met.

Barber[7] investigated the effect of Mexico's large-scale CCT program, *Oportunidades* (formerly known as PROGRESA), on caesarean section rates using retrospective reports from a government-sponsored randomized evaluation of the program. The *Oportunidades* program provides impoverished families with cash transfers conditional on family members' use of certain health and education services. Barber[7] found that beneficiary status in the CCT program was associated with a 5.1 percent increase in caesarean rates (p=0.05). If families were enrolled in the CCT program for more than six months, the effect on caesarean rates increased to 7.5 percent. In terms of delivery location, differences in caesarean rates between beneficiaries and nonbeneficiaries were significant only in social security and government health facilities, not in private facilities. Therefore, Barber[7]concluded that the CCTs were correlated with higher caesarean section rates in these types of healthcare facilities and that this effect was likely precipitated by the additional disposable income generated from the CCT program.

Salinas-Rodriguez and Manrique-Espinoza[70] conducted a cross-sectional study investigating the effect of the *Oportunidades* program on vaccination coverage in older people in Mexico. They used propensity score matching and a linear probability model to estimate the effect of being a beneficiary of the CCT program on vaccination coverage. Beneficiary status had a significant effect, increasing the vaccination proportions of older people by 5.5 percent (95 percent CI, 2.8–8.3) for complete schedule coverage (i.e., all vaccines studied), 6.9 percent (95 percent CI, 3.8–9.6) for influenza coverage, 7.2 (95 percent CI, 4.3–10.2) for pneumococcal coverage, and 6.6 percent (95 percent CI, 4.1–9.2) for tetanus coverage. The authors concluded that CCT programs such as *Oportunidades* could significantly increase vaccination coverage in the elderly—likely because the program requires that recipients have regular health checkups, and because it also increases awareness of health services.

Like Mexico's *Oportunidades*, the *Bolsa Familia* Programme (BFP), launched in Brazil in 2003,[71] transfers payments to families on the condition that beneficiaries obtain health services (such as vaccinations and prenatal care for pregnant women) and that children maintain a minimum daily attendance rate at school. Rasella et al.[71] investigated the effect of BFP on overall under-five mortality and under-five mortality from causes associated with poverty, such as malnutrition, diarrhea, and lower respiratory infections. Using a mixed ecological design over 2004–2009, Rasella et al. [71] employed multivariable regression analyses of panel data with BFP coverage classified as low (0–17.1 percent), intermediate (17.2–32.0 percent), high (>32.0 percent) or consolidated (>32 percent with 100 percent target population coverage for at least four years). Using the mortality rate ratio as their primary outcome variable, the authors determined rate ratios for the effect of BFP on overall under-five mortality of 0.94 (95 percent CI, 0.92–0.96) for intermediate coverage, 0.88 (95 percent CI, 0.85–0.91) for high coverage, and 0.83 (95 percent CI, 0.79–0.88) for consolidated coverage. Their findings demonstrate that CCT programs (especially in conjunction with primary healthcare programs, such as Brazil's Family Health Program) have the potential to greatly reduce childhood mortality.

Using different methods, Shei[63] examined the effect of the expansion of the *Bolsa Família* CCT program on infant, neonatal, and postneonatal infant mortality in Brazil. Using a pooled time-series, cross-sectional design in conjunction with a fixed-effects approach, Shei examined infant mortality rates by municipality as the BFP expanded heterogeneously across municipalities. Results from this study suggest that the "treatment effect" of the program precipitated a 9.3 percent ($p<0.01$) decline in the infant mortality rate and a 24.3 percent ($p<0.01$) decrease in the postneonatal mortality rate. Decreases in the neonatal mortality rate were not significantly associated with increases in program coverage. Additionally, Shei[63] found that declines in mortality rates were greatest in areas with the highest baseline rates and where Brazil's primary health care program was well-established. In conclusion, this study indicated that CCT programs like BFP could significantly improve population health.

Studies on CCT programs have also demonstrated significant effects on education. Barrera-Osorio et al.[72] investigated the effects of CCT program design on school attendance, re-enrollment, and graduation or tertiary enrollment using a randomized control study in Colombia. This study divided participants into a control group and three treatment groups: standard,

"savings," and "tertiary." The standard CCT design is based on Mexico's *Oportunidades* program, wherein households are paid bimonthly on the condition that their children meet a specified attendance target and certain health criteria. In the savings CCT design, families receive only part of the standard monthly transfer; the remainder is postponed until students re-enroll in school. In the tertiary treatment, bimonthly payments for good attendance are lower but the family is guaranteed a large payment upon graduation; however, this transfer can be accessed earlier if the student matriculates into a tertiary institution. In terms of attendance, all three treatment designs generated increases in attendance rates of 3 to 5 percent ($p<0.05$) compared with the control group. The standard, savings, and tertiary designs increased re-enrollment by 1.7 ($p<0.1$), 4.0 ($p<0.01$), and 3.7 ($p<0.1$) percent, respectively. The savings and tertiary designs significantly increased matriculation into institutions of higher education, by 9.4 percent ($p<0.05$) and 48.9 percent ($p<0.01$), respectively. Barrera-Osorio et al. concluded that altering the standard CCT design to incentivize graduation or re-enrollment could improve both daily attendance and attainment of higher levels of education.

Baird et al.[67] investigated the indirect effect of a cash transfer program (divided into conditional and unconditional treatment groups) on the prevalence of HIV and herpes simplex type 2 (HSV-2) in women aged 13 to 22 years in Malawi using a cluster-randomized trial. At the end of the study period, weighted HIV prevalence was 1.2 percent for those receiving cash transfers compared with 3.0 percent for those in the control group (odds ratio 0.36, 95 percent CI, 0.14–0.91). Additionally, the weighted prevalence of HSV-2 was 1.3 percent higher in the control group than in the group receiving cash transfers (0.7 vs. 3.0 percent, OR 0.24, 95 percent CI, 0.09–0.65). Baird et al. concluded that cash transfer programs (both conditional upon school attendance and unconditional) decreased the prevalence of HIV and HSV-2 infections in adolescent schoolgirls in low-income settings.

De Walque et al.[66] investigated the potential of CCTs to incentivize safe sex for the prevention of HIV and sexually transmitted infections (STIs), using an unblinded, individually randomized and controlled trial in rural Tanzania. Participants were randomly assigned to a high-value intervention ($20 per testing round), a low-value intervention ($10), and a control group. All were tested at four-month intervals for four STIs that serve as reasonable proxies for risky sexual behavior, such as unprotected sex. Adjusted relative risk (RR) for the high-value CCT treatment

group was 0.73 (95 percent CI, 0.47– 0.99) compared with the control group, indicating a significant reduction in STI incidence as a result of the financial incentive. However, results indicated no significant reduction in STI prevalence for the low-value CCT group compared with the control group. De Walque et al. concluded that CCTs were a promising tool to prevent the spread of HIV and other STIs but called for a larger study to confirm these results and determine the appropriate cash transfer amount.

Robertson et al.[69] studied the effects of unconditional and conditional cash transfers on child health and development in Zimbabwe, with indicators such as birth registration, vaccination uptake, and school attendance**.** They used a matched, cluster-randomized controlled trial in which each study site was divided into three clusters: UCT, CCT, or control. Compared with control clusters, the proportion of children under four years of age with birth certificates had increased by 1.5 percent (95 percent CI, –7.1–10.1) in groups receiving UCTs and by 16.4 percent (95 percent CI, 7.8–25.0) in groups receiving CCTs. Moreover, the proportion of children under four with complete vaccination records increased by 3.1 percent (95 percent CI, –3.8–9.9) and 1.8 percent (95 percent CI, –5.0–8.7) in the UCT and CCT groups, respectively, compared with the control group. Finally, school attendance (i.e., at least 80 percent attendance) of children aged six to 12 years was 7.2 percent (95 percent CI, 0.8–13.7) and 7.6 percent (95 percent CI, 1.2–14.1) greater in the UCT and CCT groups, respectively, than in the control group. All endpoint indicators had significant increases in the UCT and CCT groups compared with the control group except for vaccination uptake. Ultimately, Robertson et al.[69] concluded that UCT and CCT programs could be valuable tools for improving child health and development in sub-Saharan Africa, but more research was needed to understand whether UCTs or CCTs would be more effective.

Randive et al.[73] investigated the effect of India's *Janani Suraksha Yojana* (JSY) CCT program, which provides cash incentives to women who deliver their babies in formal health institutions, on the proportion of institutional childbirths and the maternal mortality ratio. Using correlation and multivariate regression models, they found that JSY prompted a 29 percent increase in the proportion of institutional births from the preprogram average over a five-year period (from 20 to 49 percent, $p<0.05$). However, using bivariate analysis, Randive et al.[73] found that the increase in the proportion of institutional births did not significantly reduce maternal mortality. Randive et

al. concluded that because of JSY, more women gave birth in healthcare facilities, but other factors—poor-quality facilities, poor-quality care, nonfinancial access barriers, and negligence in antenatal and postnatal care—may have prevented the program from reducing maternal mortality.

Fernald et al.[64] set out to disaggregate the effects of the cash component of CCTs from that of the "conditionalities" for payments using data from the initial two-stage introduction of *Oportunidades* in Mexico. Their aim was to determine whether receiving higher cumulative transfers in the CCT program was correlated with improvements in various indicators of child growth, health, and development. They selected a study population in which total cumulative transfers varied based on how long families had been in the *Oportunidades* program and on family demographics. Using regression analysis to account for a wide range of covariates (many of which were measures of socioeconomic status), Fernald et al.[64] found that a doubling of cumulative cash transfers resulted in a 9 percent decrease in the prevalence of stunting ($p<0.0001$), a 0.16 standard deviation increase in height-for-age z-score ($p<0.0001$), a 6 percent lower prevalence of overweight children ($p=0.0001$), and higher hemoglobin concentrations ($p=0.03$). Moreover, a doubling of the cash component of the CCT was significantly associated with improvements in several measures of cognitive and motor development. Fernald et al.'s [64]results indicated that improvements in child health, growth, and development were associated with the cash component of the CCT, rather than its various conditions.

Filmer and Schady[68] investigated the effect of CCTs of different magnitude on school attendance in the CESSP Scholarship Program in Cambodia to determine whether there were diminishing returns to the cash amount of the transfer. They used regression discontinuity (RD) analysis to compare attendance rates of students who received no transfer payments, small transfers (US$45 per year), and large transfers (US$60 per year). RD analysis enabled them to estimate CCT program effects for different transfer amounts. The various models indicated that compared with no transfer at all, the small transfer delivered a large increase in school attendance, with effects ranging from 18 to 28 percent. However, they found very little difference in attendance rates between children receiving large and small transfer payments. Emphasizing this result, the models estimated that the average per dollar percentage point increase in attendance resulting from the large (versus the small) transfer was –0.07 to 0.31 percent, whereas the increase

resulting from receiving the first small payment (versus no payment at all) was 0.40 to 0.62 percent. Filmer and Schady[68] therefore concluded there was "clear evidence of diminishing marginal returns to program size," and that enrollment response to the program was identical for both transfer amounts.

Lagarde et al.[65] conducted a systematic review of CCTs in low- and middle-income countries to see whether CCTs improve access to and use of healthcare services as well as health outcomes. A total of 28 papers were assessed, of which 10 papers from six CCT interventions were included in the review. The results indicated that CCTs were generally effective in increasing access to and use of preventive health services. However, evidence also suggested that the effect of CCTs on health outcomes and health status was less clear. In a more in-depth version of the same review, Lagarde et al.[74] noted that some of the CCT programs reviewed had clear benefits for nutritional status and health outcomes, but it remained difficult to ascribe these effects to cash incentives because so many other factors contribute to health outcomes. Lagarde et al.[65] also emphasized that the success of CCT strategies seemed to hinge on the existence of primary healthcare systems and infrastructure (such as the Family Health Program in Brazil): beneficiaries of CCT programs must be able to use the additional resources effectively. Finally, Lagarde et al.[74] question whether CCT programs will prove as successful in settings with fewer resources, such as sub-Saharan Africa.

Ten of the 11 CCT studies reviewed above found significant positive effects on the outcome variable being examined. Only the JSY program from India had no significant benefit, but its failure to lower the maternal mortality rate likely stems from beneficiaries' lack of access to quality healthcare facilities. Despite the substantial evidence that CCTs can significantly raise social outcomes, such as school attendance and health status, more research into the design of CCT programs and the actual cash value transferred is needed if countries are to realize the full potential of these innovative social welfare programs.

3. Conclusions and research directions

There has been a surge of interest in implementing and evaluating innovations to improve intervention uptake and provider quality. Progress has been made in high-income countries and

in LMICs with significant opportunities for cross-learning but there remains significant opportunity and the need for additional research to inform the design of interventions.

Payment-for-performance schemes have been widely implemented in developed countries such as the United States and Britain to improve chronic disease management and overall healthcare quality, and they have also been implemented in many developing countries to improve maternal and child health and infectious disease outcomes. Of the fifteen studies reviewed in this paper, only four report significant improvement in healthcare outcomes from various payment-for-performance programs.[16,22–24] Overall, results on the effectiveness of these programs on health outcomes are mixed; however, there is room for a wider experimentation with payment-for-performance schemes in many LMICs where they are currently absent. A major challenge with program evaluation and drawing general conclusions on the effectiveness of the interventions is that payment-for-performance evaluation schemes are heterogeneous in terms of study design, intervention design, and externalities.[27,30] Moreover, many studies have failed to disentangle the effect of the payment-for-performance interventions on the effect of other simultaneous improvements.[27] More rigorous evaluations with robust study designs will be useful in evaluating the true effect of payment-for-performance schemes on health outcomes.

We also looked at protocol-based methods, such as safety checklists, that can improve patient health outcomes by setting treatment quality standards and by facilitating improved communication within provider teams. Following the success of a surgical checklist intervention that decreased catheter-related bloodstream infections in Michigan hospital ICUs,[47], the WHO Surgical Safety Checklist[48] was designed and has been adopted by 1,790 healthcare facilities worldwide since 2012[49] to address avoidable surgical complications. There have been many studies assessing the quantitative impact of the checklist interventions on health outcomes.[48,51–56] Bergs et al.'s [57] (2014) systematic review and meta-analysis of seven studies investigates the quantitative effect of the WHO Surgical Safety Checklist on postoperative complications. We conclude that more research that separates out the Hawthorne effect of checklist implementation from the true effect of the checklists is an important next step for scale up of these potentially useful interventions.

On the demand side, we reviewed a number of conditional cash transfers (CCTs), where payments are made to families conditional on whether they comply with certain pre-defined requirements (such as maintaining certain attendance rates in school-age children or utilizing certain health services).[65,66] Of the eleven studies reviewed here, ten of them reveal significant positive effects on the outcome variable (such as the use of health services, health outcomes, nutrition/development status and attendance rates) being examined. For instance, Shei[63] examines the effect of Brazil's *Bolsa Família* CCT program on infant, neonatal and postneonatal infant mortality and found that the CCT program precipitated a 9.3 percent ($p<0.01$) decline in the infant mortality rate and a 24.3 percent ($p<0.01$) decrease in the postneonatal mortality rate. Only India's *Janani Suraksha Yojana* (JSY) CCT program had no significant effect on decreasing the maternal mortality rate and this result likely stems from a lack of quality healthcare that beneficiaries can access[73].

Despite this review presenting substantial evidence that CCTs can significantly benefit social outcomes such as school attendance and health status, more research on the design of CCT programs and the actual cash value transferred is needed to realize the full potential of these innovative social welfare programs. For instance, Filmer & Schady[68] found explicit evidence of diminishing marginal returns to the amount of cash being transferred in a CCT program aiming to promote school attendance in Cambodia. Additionally, Barrera-Osorio et al.[72] found that alternative designs to traditional CCT programs could produce greater effects on the outcome the program aims to addresses. Thus, more research in program size and design is necessary to increase the efficiency and effectiveness of CCTs.

4. References

1. Das J, Gertler PJ. Variations in practice quality in five low-income countries: a conceptual overview. *Health Aff.* 2007;26(3):w296–w309.
2. Das J, Hammer J, Leonard K. The quality of medical advice in low-income countries. *J. Econ. Perspect.* 2008;22(2):93–114.
3. Das J, Hammer J. Quality of Primary Care in Low-Income Countries: Facts and Economics. 2014.
4. Kenneth L. Leonard, Melklory C. Masatu. Variations In The Quality Of Care Accessible To Rural Communities In Tanzania. *Health Aff.* 2007;26(3):w380–w392.
5. Das J, Hammer J. Money for nothing: The dire straits of medical practice in Delhi, India. *J. Dev. Econ.* 2007;83(1):1–36.
6. Das J, Sohnesen TP. Variations in doctor effort: evidence from Paraguay. *Health Aff.* 2007;26(3):w324–w337.
7. Barber SL, Gertler PJ, Harimurti P. The contribution of human resources for health to the quality of care in Indonesia. *Health Aff.* 2007;26(3):w367–w379.
8. Brock JM, Lange A, Leonard KL. Generosity norms and intrinsic motivation in health care provision: evidence from the laboratory and the field. Working papers 147, European Bank for Reconstruction and Development, 2012.
9. Kenneth L. Leonard, Melklory C. Masatu, Alexandre Vialou. Getting Doctors to Do Their Best: The Roles of Ability and Motivation in Health Care Quality. *J. Hum. Resour.* 2007;42(3):682–700.
10. Cheng S, Lee T, Chen C. A Longitudinal Examination of a Pay-for-Performance Program for Diabetes Care. *Med. Care*. 2012;50(2):109–116.
11. Basinga P, Gertler PJ, Binagwaho A, Soucat ALB, Sturdy J, Vermeersch CMJ. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *Lancet*. 2011;377(9775):1421–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21515164. Accessed March 24, 2014.
12. Priedeman Skiles M, Curtis SL, Basinga P, Angeles G. An equity analysis of performance-based financing in Rwanda: are services reaching the poorest women? *Health Policy Plan.* 2013;28(8):825–37.
13. Peabody JW, Shimkhada R, Quimbo S, Solon O, Javier X, McCulloch C. The impact of performance incentives on child health outcomes: results from a cluster randomized controlled trial in the Philippines. *Health Policy Plan.* 2013:1–7.
14. Fairbrother G, Hanson KL, Friedman S, Butts GC. The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. *Am. J. Public Health*. 1999;89(2):171–5.
15. Huntington D, Zaky HHM, Shawky S, Fattah FA, El-Hadary E. Impact of a service provider incentive payment scheme on quality of reproductive and child-health services in Egypt. *J. Health. Popul. Nutr.* 2010;28(3):273–80.

16. Li Y-H, Tsai W-C, Khan M, et al. The effects of pay-for-performance on tuberculosis treatment in Taiwan. *Health Policy Plan.* 2010;25(4):334–41.

17. Bardach NS, Wang JJ, De Leon SF, et al. Effect of pay-for-performance incentives on quality of care in small practices with electronic health records: a randomized trial. *JAMA.* 2013;310(10):1051–9.

18. Glickman SW, Delong ER, Roe MT, Lytle BL, Schulman KA, Peterson ED. Pay for Performance , Quality of Care , and Outcomes in Acute Myocardial Infarction. *JAMA.* 2014;297(21):2373–2380.

19. Rosenthal MB, Frank RG, Epstein AM. Early Experience With Pay-for-Performance. *JAMA.* 2014;294(14):1788–1793.

20. Yip W, Powell-jackson T, Chen W, et al. Capitation Combined With Pay-For-Performance Improves Antibiotic Prescribing Practices In Rural China. *Health Aff.* 2014;33(3):502–510.

21. Petersen L a, Simpson K, Pietz K, et al. Effects of individual physician-level and practice-level financial incentives on hypertension care: a randomized trial. *JAMA.* 2013;310(10):1042–50.

22. Millett C, Gray J, Saxena S, Netuveli G, Majeed A. Impact of a pay-for-performance incentive on support for smoking cessation and on smoking prevalence among people with diabetes. *Can. Med. Assoc. J.* 2007;176(12):1705–1710.

23. Cutler TW. Evaluation of the Relationship Between a Chronic Disease Care Management Program and California Pay-for-Performance Diabetes Care Cholesterol Measures in One Medical Group. *J. Manag. Care Pharm.* 2007;13(7):578–588.

24. Lindenauer P, Remus D, Roman S, et al. Public Reporting and Pay for Performance in Hospital Quality Improvement. *N. Engl. J. Med.* 2007;356(5):486–496.

25. Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: Effects, design choices, and context of pay-for-performance in health care. *BMC Health Serv. Res.* 2010;10:247.

26. Gillam SJ. Pay-for-Performance in the United Kingdom : Impact of the Quality and Outcomes Framework — A Systematic Review. *Ann. Fam. Med.* 2012;10(5):461–468.

27. Eijkenaar F, Emmert M, Scheppach M, Schöffski O. Effects of pay for performance in health care: a systematic review of systematic reviews. *Health Policy.* 2013;110(2-3):115–30.

28. Mehrotra A, Damberg CL, Sorbero MES, Teleki SS. Pay for performance in the hospital setting: what is the state of the evidence? *Am. J. Med. Qual.* 2009;24(1):19–28.

29. Flodgren G, Mp E, Shepperd S, Scott A, Parmelli E, Fr B. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes (Review). *Cochrane Database Syst. Rev.* 2011;(7).

30. Witter S, Fretheim A, Fl K, Ak L. Paying for performance to improve the delivery of health interventions in low- and middle-income countries (Review). *Cochrane Database Syst. Rev.* 2012;(2).

31. Houle SKD, McAlister F a, Jackevicius C a, Chuck AW, Tsuyuki RT. Does performance-based remuneration for individual health care practitioners affect patient care?: a systematic review. *Ann. Intern. Med.* 2012;157(12):889–99.
32. Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S. Improving Patient Care Does Pay-for-Performance Improve the Quality of Health Care? *Ann. Intern. Med.* 2006;145(4):265–272.
33. De Bruin SR, Baan C a, Struijs JN. Pay-for-performance in disease management: a systematic review of the literature. *BMC Health Serv. Res.* 2011;11(1):272.
34. Semel ME, Resch S, Haynes AB, et al. Adopting A Surgical Safety Checklist Could Save Money And Improve The Quality Of Care In U.S. Hospitals. *Health Aff.* 2010;29(9):1593–1599.
35. McCulloch P, Mishra A, Handa A, Dale T, Hirst G, Catchpole K. The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. *Qual. Saf. Health Care*. 2009;18(2):109–115.
36. Catchpole K, Sellers R, Goldman A, McCulloch P, Hignett S. Patient handovers within the hospital: translating knowledge from motor racing to healthcare. *Qual. Saf. Health Care*. 2010;19(4):318–322.
37. Fowler AJ. A review of recent advances in perioperative patient safety. *Ann. Med. Surg.* 2013;2(1):10–14.
38. Arriaga AF, Bader AM, Wong JM, et al. Simulation-Based Trial of Surgical-Crisis Checklists. *N. Engl. J. Med.* 2013;368(3):246–253.
39. McDowell DS, McComb SA. Safety Checklist Briefings: A Systematic Review of the Literature. *AORN J.* 2014;99(1):125–137.e13.
40. David Classen, R. Scott Evans, Stanley Pestotnik, Susan Horn, Ronald Menlove, John P. Burke. The timing of prophylactic administration of antibiotics and the risk of surgical-wound infection. *N. Engl. J. Med.* 1992;326(5).
41. World Health Organization. Safe Surgery Saves Lives. Patient Safety, WHO Guidelines. 2009. Available at: http://www.who.int/patientsafety/safesurgery/tools_resources/9789241598552/en/. Accessed on: March 26, 2014.
42. Classen DC, Resar R, Griffin F, et al. "Global trigger tool" shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff. (Millwood).* 2011;30(4):581–589.
43. De Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. *Qual. Saf. Heal. Care*. 2008;17(3):216–223.
44. Pearse RM, Moreno RP, Bauer P, et al. Mortality after surgery in Europe: a 7 day cohort study. *Lancet*. 2012;380(9847):1059–1065.
45. Zegers M, de Bruijne MC, de Keizer B, et al. The incidence, root-causes, and outcomes of adverse events in surgical units: implication for potential prevention strategies. *Patient Saf. Surg.* 2011;5:13.

46. Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery*. 2003;133(6):614–621.

47. Pronovost P, Needham D, Berenholtz S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. *N. Engl. J. Med.* 2006;355(26):2725–2732.

48. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N. Engl. J. Med.* 2009;360(5):491–499.

49. Surgical Safety Web Map. Patient Safety A World Alliance for Safer Health Care. The World Health Organization. Available at: http://maps.cga.harvard.edu:8080/Hospital/ Accessed on: February 18, 2014

50. De Vries EN, Prins HA, Crolla RM, den Outer AJ, van Andel G, van Helden SH, et al. Effect of a comprehensive surgical safety system on patient outcomes. *N. Engl. J. Med.* 2010;363(20):1928–1937.

51. Van Klei WA, Hoff RG, van Aarnhem EEHL, Simmermacher RK, Regli LP, et al. Effects of the introduction of the WHO "Surgical Safety Checklist" on in-hospital mortality: a cohort study. *Ann. Surg.* 2012;255(1):44–49.

52. Weiser TG, Haynes AB, Dziekan G, Berry WR, Lipsitz SR, Gawande AA, et al. Effect of a 19-item surgical safety checklist during urgent operations in a global patient population. *Ann. Surg.* 2010;251(5):976–980.

53. Sewell M, Adebibe M, Jayakumar P, Jowett C, Kong K, Levack B. Use of the WHO surgical safety checklist in trauma and orthopaedic patients. *Int. Orthop.* 2011;35(6):897–901.

54. Askarian M, Kouchak F, Palenik CJ. Effect of surgical safety checklists on postoperative morbidity and mortality rates, Shiraz, Faghihy Hospital, a 1-year study. *Qual. Manag. Health Care*. 2011;20(4):293–297.

55. Bliss LA, Ross-Richardson CB, Sanzari LJ, Shapiro DS, Lukianoff AE, Bernstein BA, et al. Thirty-day outcomes support implementation of a surgical safety checklist. *J. Am. Coll. Surg.* 2012;215(6):766–776.

56. Kwok AC, Funk LM, Baltaga R, Lipsitz SR, Merry AF, Dziekan G, et al. Implementation of the World Health Organization surgical safety checklist, including introduction of pulse oximetry, in a resource-limited setting. *Ann. Surg.* 2013;257(4):633–639.

57. Bergs J, Hellings J, Cleemput I, Surel O, De Troyer V, Van Hiel M, et al. Systematic review and meta-analysis of the effect of the World Health Organization surgical safety checklist on postoperative complications. *Br. J. Surg.* 2014;101(3):150–158.

58. Spector JM, Agrawal P, Kodkany B, Lipsitz S, Lashoher A, Dziekan G, et al. Improving Quality of Care for Maternal and Newborn Health: Prospective Pilot Study of the WHO Safe Childbirth Checklist Program. Middleton P, ed. *PLoS One*. 2012;7(5):e35151.

59. Bosk CL, Dixon-Woods M, Goeschel CA, Pronovost PJ. Reality check for checklists. *Lancet*. 2009;374(9688):444–445.

60. Peter Provonost. A doctor's crusade to improve patient safety. *Biomed. Instrum. Technol.* 2013;47(4):319–22.

61. Landro L. The Secret to Fighting Infections. March 28, 2011. Peter Pronovost and the Checklist. The Wall Street Journal Online. Last updated March 28, 2014. Available at: http://online.wsj.com/news/articles/SB10001424052748704364004576131963185893084 Accessed: March 30, 2014

62. Atul Gawande. *The checklist manifesto: how to get things right.* New York, NY: Metropolitan Books; 2009.

63. Shei A. Brazil's conditional cash transfer program associated with declines in infant mortality rates. *Health Aff. (Millwood).* 2013;32(7):1274–81.

64. Fernald LCH, Gertler PJ, Neufeld LM. Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades. *Lancet*. 2008;371(9615):828–837.

65. Lagarde M, Haines A, Palmer N. Conditional Cash Transfers for Improving Uptake of Health Interventions in Low- and Middle-Income Countries. *JAMA J. Am. Med. Assoc.* 2007;298(16):1900–1910.

66. De Walque D, Dow WH, Nathan R, et al. Incentivising safe sex: a randomised trial of conditional cash transfers for HIV and sexually transmitted infection prevention in rural Tanzania. *BMJ Open*. 2012;2:e000747.

67. Baird SJ, Garfein RS, McIntosh CT, Ozler B. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. *Lancet*. 2012;379(9823):1320–9.

68. Filmer D, Schady N. Are there diminishing returns to transfer size in conditional cash transfers? 2009:1–22.

69. Robertson L, Mushati P, Eaton JW, et al. Effects of unconditional and conditional cash transfers on child health and development in Zimbabwe: a cluster-randomised trial. *Lancet*. 2013;381(9874):1283–92.

70. Salinas-Rodríguez A, Manrique-Espinoza BS. Effect of the conditional cash transfer program Oportunidades on vaccination coverage in older Mexican people. *BMC Int. Health Hum. Rights*. 2013;13:30.

71. Rasella D, Aquino R, Santos C a T, Paes-Sousa R, Barreto ML. Effect of a conditional cash transfer programme on childhood mortality: a nationwide analysis of Brazilian municipalities. *Lancet*. 2013;382(9886):57–64.

72. Barrera-Osorio F, Bertrand M. Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia. *Am. Econ.* 2011;3:167–195.

73. Randive B, Diwan V, De Costa A. India's Conditional Cash Transfer Programme (the JSY) to Promote Institutional Birth: Is There an Association between Institutional Birth Proportion and Maternal Mortality? Baradaran HR, ed. *PLoS One*. 2013;8(6):e67452.

74. Lagarde M, Haines A, Palmer N. The impact of conditional cash transfers on health outcomes and use of health services in low and middle income countries. (The Cochrane Collaboration, ed.). Chichester, UK: John Wiley & Sons, Ltd; 2009.

5. Tables and Figures

Table 1. Effect of payment-for-performance on preventive care or screening

Results of Included Studies on the Effect of P4P on Preventive Care or Screening

Study, Year (Reference)	Outcome(s) Measured	Results*		
		Control	Intervention	Comparison
Randomized, controlled trials				
Fairbrother et al, 2001 (9)	Change in percentage of children receiving recommended vaccinations over 1 y	−2.5 percentage points	5.9 percentage points	$P < 0.05$
Kouides et al, 1998 (10)	Change in mean influenza immunization rates over 1 y	2.5 percentage points	10.3 percentage points	$P = 0.03$
Grady et al, 1997 (12)	Change in mean rate (per practice, over 1 y) of:			
	Mammography referral	25.0 percentage points	26.0 percentage points	$P = 0.46$
	Mammography completion	30.2 percentage points	28.2 percentage points	$P = 0.14$
Nonrandomized, controlled study		Control	Intervention	Interaction Term Comparing Trends
Gavagan et al, 2010 (15)	6-y linear trend models for achievement of performance thresholds for:			
	Papanicolaou smears	Slope, −0.004	Slope, 0.005	$P = 0.053$
	Mammography	Slope, 0.0015	Slope, 0.003	$P = 0.076$
Controlled before–after studies		Pre-P4P	Post-P4P	Between-Group Difference†
Rosenthal et al, 2005 (13)	Cervical cancer screening	Intervention: 39.2% Control: 55.4%	Intervention: 44.5% Control: 57.1%	3.6 ± 1.8 percentage points ($P = 0.02$)
	Mammography	Intervention: 66.1% Control: 72.4%	Intervention: 68.0% Control: 72.6%	1.7 ± 1.5 percentage points ($P = 0.13$)
	Hemoglobin A_{1c} testing	Intervention: 62.0% Control: 80.0%	Intervention: 64.1% Control: 82.1%	0.0 ± 3.5 percentage points ($P = 0.50$)
Fagan et al, 2010 (14)	Influenza vaccination	NR	NR	OR, 1.79 (CI, 1.37–2.35)
	Hemoglobin A_{1c} testing	NR	NR	OR, 0.44 (CI, 0.30–0.65)
	LDL cholesterol screening	NR	NR	OR, 0.62 (CI, 0.44–0.86)
	Retinopathy screening	NR	NR	OR, 0.98 (CI, 0.61–1.58)
	Nephropathy screening	NR	NR	OR, 0.95 (CI, 0.62–1.46)
Uncontrolled before–after studies		Pre-P4P	Post-P4P	Between-Group Difference
Morrow et al, 1995 (16)	Measles/mumps/rubella immunization in children	78.1% (CI, 73.9%–82.1%)	95.6% (CI, 93.5%–97.7%)	$P < 0.001$
	Cholesterol screening in adults	91.9% (CI, 87.8%–94.4%)	95.4% (CI, 91.7%–97.4%)	$P < 0.001$
Chung et al, 2003 (17)	Patients with diabetes receiving hemoglobin A_{1c} test annually	51.5%	79.6%	$P < 0.001$
	Children receiving measles/mumps/rubella vaccine	83.2%	87.3% (year 2); 81.8% (year 3)	$P = 0.061$ (year 2); $P < 0.001$ (year 3)
Armour et al, 2004 (11)	Colorectal cancer screening	23.4%	26.4%	$P < 0.01$
Multivariate analysis of cohort study		Adjusted Relative Risk (CI)‡		
Ettner et al, 2006 (18)	Proportion of patients receiving the following when P4P bonuses were based on quality or patient satisfaction scores:			
	Hemoglobin A_{1c}	0.99 (0.90–1.11)		
	Proteinuria assessment	0.86 (0.71–1.13)		
	Lipid panel assessment	1.05 (0.90–1.25)		
	Dilated eye examination	1.00 (0.89–1.14)		
	Foot examination	1.02 (0.91–1.18)		
	Advice to take daily aspirin	1.19 (0.99–1.48)		
	Influenza immunization	1.06 (0.90–1.29)		
	Proportion of patients receiving the following when P4P bonuses were based on patients' outpatient utilization or care costs:			
	Hemoglobin A_{1c}	0.99 (0.88–1.05)		
	Proteinuria assessment	1.13 (1.03–1.24)		
	Lipid panel assessment	1.01 (0.87–1.11)		
	Dilated eye examination	1.02 (0.93–1.09)		
	Foot examination	1.01 (0.88–1.08)		
	Advice to take daily aspirin	0.87 (0.71–1.04)		
	Influenza immunization	1.02 (0.89–1.14)		

LDL = low-density lipoprotein; NR= not reported; OR= odds ratio; P4P= pay-for-performance.
* All CIs are 95% CIs.
† Values expressed with a plus/minus sign are SEs.
‡ Adjusted for sociodemographic(age, sex, ethnicity, education, household income, source of insurance) and clinical (type of diabetes treatment [insulin, oral agents, diet only], years since diabetes diagnoses, Charlson comorbidity index, and Short-Form-12 physical and mental component summary score) characteristics of patient population.

Source: Table from Houle et al. (2012)[31]

Table 2. Effect of payment-for-performance on quality of care for chronic conditions

Results of Included Studies on the Effect of P4P on Quality of Care for Chronic Conditions

Study, Year (Reference)	Outcome(s) Measured	Results*		
		Control	Intervention	Comparison
Randomized, controlled trial				
Twardella and Brenner, 2007 (19)	Proportion of patients smoke-free as validated by a negative blood cotinine level after 1 y	2.7%	3.5%	P = 0.75
Interrupted time series		Pre-P4P	Post-P4P	ITS Analyses
Campbell et al, 2009 (20)	Mean scores for processes of care quality indicators:			
	Coronary heart disease	Improved from 58.6% in 1998 to 76.2% in 2003	Improved to 85.0% in 2005 and 84.8% in 2007	P = 0.06 for greater than expected change in 2005 given secular trends before; P = 0.02 for less than expected improvement in 2005–2007 versus 1998–2003
	Asthma	Improved from 60.2% in 1998 to 70.3% in 2003	Improved to 84.3% in 2005 and 85.0% in 2007	P = 0.001 for greater than expected change in 2005 given secular trends before; P = 0.16 for less than expected improvement in 2005–2007 versus 1998–2003
	Diabetes	Improved from 61.6% in 1998 to 70.4% in 2003	Improved to 81.4% in 2005 and 83.7% in 2007	P < 0.001 for greater than expected change in 2005 given secular trends before; P = 0.91 for less than expected improvement in 2005–2007 versus 1998–2003
Serumaga et al, 2011 (21)	Proportion of patients with blood pressure controlled to ≤150/90 mm Hg	70%	67.3%	Level change: −1.19 (CI, −2.06 to 1.09) percentage points Trend change: −0.01 (CI, −0.06 to 0.03) percentage points
	Proportion of patients with blood pressure measured	47.7%	53.2%	Level change: 0.85 (CI, −3.04 to 4.74) percentage points Trend change: −0.01 (CI, −0.24 to 0.21) percentage points
	Proportion of patients who began receiving new drug treatment	0.05% per month	0.05% per month	Level change: 0.67 (CI, −2.37 to 3.81) percentage points Trend change: 0.02 (CI, −0.23 to 0.19) percentage points
	Occurrence of hypertension-related outcomes (myocardial infarction, stroke, heart failure, renal failure, death)	NR	NR	Level change: 0.07% (CI, −0.13 to 0.28) percentage points Trend change: 0.05% (CI, −0.02 to 0.07) percentage points
Vamos et al, 2011 (22)	Rate of improvement per year in achieving the following targets for diabetic patients:			
	Blood pressure	2.2% (CI, 1.9% to 2.6%)	3.8% (CI, 2.7% to 4.9%)	P < 0.001 for improvement after P4P
	Total cholesterol level ≤5 mmol/L (193.05 mg/dL)	4.9% (CI, 4.3% to 5.3%)	7.4% (CI, 6.0% to 8.8%)	P < 0.05 for improvement after P4P
	Hemoglobin A_{1c} value ≤7.0%	2.0% (CI, 1.3% to 2.7%)	−0.2% (−2.0% to 1.6%)	P < 0.01 for worsening after P4P
Alshamsan et al, 2012 (23)	Hemoglobin A_{1c}	NR	NR	Level change: 0.04 (CI, −0.04 to 0.12) percentage points Trend change: 0.19 (CI, 0.15 to 0.22) percentage points
	Total cholesterol	NR	NR	Level change: −0.12 (CI, −0.18 to −0.06) mmol/L Trend change: 0.03 (CI, 0.01 to 0.05) mmol/L
	Systolic blood pressure	NR	NR	Level change: −1.95 (CI, −2.87 to −1.02) mm Hg Trend change: −1.04 (CI, −1.42 to −0.64) mm Hg
MacBride-Stewart et al, 2008 (24)	Percentage increase in the defined daily dose prescribed per patient:			
	Drugs incentivized by the Quality and Outcomes Framework	1.3% (CI, 1.2% to 1.4%)	1.0% (CI, 0.9% to 1.1%)	P < 0.001 for less than expected change in prescribing between 2004–2006 versus 2002–2004, taking into account secular trends
	Drugs not incentivized by the Quality and Outcomes Framework	0.2% (CI, 0.2% to 0.3%)	0.3% (CI, 0.3% to 0.4%)	P = 0.09, meaning no significant difference in rate of change in prescribing for 2004–2006 versus 2002–2004

Continued on following page

Continued

Study, Year (Reference)	Outcome(s) Measured	Results*		
		Control	Intervention	Comparison
Controlled before–after study		Pre-P4P	Post-P4P	Between-Group Difference
Beaulieu and Horrigan, 2005 (25)	Performance of tests/examinations:			
	Hemoglobin A_{1c}	Intervention: 80.3% Control: 87.1%	Intervention: 83.4% Control: 86.5%	2.5 percentage points
	Lipid panel	Intervention: 68.5% Control: 88.8%	Intervention: 86.8% Control: 91.0%	16.1 percentage points
	Diabetic retinal examination	Intervention: 37.4% Control: 50.3%	Intervention: 63.0% Control: 51.5%	24.4 percentage points
	Nephropathy test	Intervention: 41.8% Control: 48.5%	Intervention: 78.8% Control: 51.6%	33.9 percentage points
	Outcome measures:			
	Hemoglobin A_{1c} value ≤9.5%	Intervention: 61.8% Control: 72.4%	Intervention: 75.6% Control: 74.2%	12.0 percentage points
	LDL cholesterol level ≤130 mg/dL (3.4 mmol/L)	Intervention: 46.0% Control: 60.4%	Intervention: 69.5% Control: 59.7%	24.2 percentage points
Uncontrolled before–after studies		Pre-P4P	Post-P4P	Between-Group Difference
Chung et al, 2003 (17)	Patients with heart failure prescribed ACE inhibitor or angiotensin-receptor blocker	40.8%	64.2%	$P < 0.001$
Coleman et al, 2007 (26)	Provision of brief smoking cessation advice to smokers			Rate ratio, 3.03 (CI, 2.98 to 3.09)
Doran et al, 2011 (27)	Percentage change in prescribing rates above projected trends:			
	For post-P4P year 2004–2005			Incentivized prescriptions: 4.3% (CI, 3.3% to 5.3%) Nonincentivized prescriptions: −0.9% (CI, −1.9% to 0.2%)
	For post-P4P year 2006–2007			Incentivized prescriptions: 2.9% (CI, 2.0% to 3.7%) Nonincentivized prescriptions: −1.7% (CI, −2.7% to −0.0%)
Greene et al, 2004 (28)	Number of exceptions (care decisions deviating from the recommended treatment algorithm) per 1000 episodes:			
	Of any type	326	261	$P < 0.005$
	For prescribing ofl ess appropriate or ineffective antibiotics	199	136	$P < 0.005$
	For inappropriate radiologic studies	15	12	$P < 0.005$
Kiran et al, 2012 (29)	Proportion of diabetic patients receiving 1 retinal eye examination, 4 hemoglobin A_{1c} tests, and 2 cholesterol tests in the previous 2 y	38%	45%	Improvement in 2 y after P4P similar to improvement in 2 y before P4P (relative risk, 1.22 [CI, 1.21 to 1.23] versus 1.31 [CI, 1.30 to 1.32])
McGovern et al, 2008 (30)	Percentage of diabetic patients with hemoglobin A_{1c} ≤7.4%	45.0%	52.7%	$P < 0.05$
	Percentage of diabetic patients with blood pressure ≤145/85 mmHg	63.2%	69.5%	$P < 0.05$
	Percentage of diabetic patients with total cholesterol ≤5 mmol/L (193.05 mg/dL)	67.5%	66.2%	$P < 0.05$ (reduction)
McGovern et al, 2008 (31)	Proportion of patients with coronary heart disease with:			
	Smoking cessation advice (if smoker)	81.0%	96.2%	$P < 0.05$
	Blood pressure controlled to ≤150/90 mm Hg	79.3%	80.0%	$P < 0.05$
	Total cholesterol level ≤5 mmol/L (193.05 mg/dL)	86.3%	75.5%	$P < 0.05$ (reduction)
	Antiplatelet or anticoagulant therapy	65.8%	90.3%	$P < 0.05$
	β-blocker therapy	42.6%	70.0%	$P < 0.05$
	ACE inhibitor therapy	66.4%	77.9%	$P < 0.05$
	Influenza vaccination up to date	57.4%	85.5%	$P < 0.05$
Millett et al, 2007 (32)	Percentage of diabetic smokers given cessation advice	48.0%	83.5%	$P < 0.001$

Continued on following page

Continued

Study, Year (Reference)	Outcome(s) Measured	Results*		
		Control	Intervention	Comparison
Simpson et al, 2011 (33)	Change (2006 versus 2001) in proportion of patients:			
	Blood pressure ≤140/90 mm Hg			Difference, 18.9% (CI, 18.5% to 19.4%)
	With hypertension who were treated with at least 1 antihypertensive drug			Difference, 9.2% (CI, 9.0% to 9.5%)
Simpson et al, 2006 (34)	Percentage of stroke patients with:			
	Advice on smoking cessation (if smokers)	79.0%	95.9%	Difference, 17.0% (CI, 15.7% to 18.3%)
	Total cholesterol level controlled to ≤5 mmol/L (193.05 mg/dL)	65.8%	66.1%	Difference, 2.3% (CI, 0.6% to 2.3%)
	Blood pressure controlled to ≤150/90 mm Hg	75.2%	76.5%	Difference, 1.3% (CI, 0.4% to 2.2%)
	Antiplatelet or anticoagulant therapy	55.9%	88.2%	Difference, 32.3% (CI, 31.5% to 33.1%)
	Influenza vaccination	47.1%	81.3%	Difference, 34.2% (CI, 33.4% to 35.0%)
St. Jacques et al, 2004 (35)	Mean number of cases per physician delayed by inappropriately excessive anesthesiology procedure, induction, or emergence time	14.9 (SD, 2.9)	3.34 (SD, 1.0)	P < 0.001
Steel et al, 2007 (36)	Percentage of quality indicators achieved:			
	Asthma (incentivized)	59% (SD, 24%)	73% (SD, 23%)	P < 0.001
	Hypertension (incentivized)	58% (SD, 17%)	70% (SD, 16%)	P < 0.001
	Depression (nonincentivized)	37% (SD, 14%)	38% (SD, 14%)	P = 0.22
	Osteoarthritis (nonincentivized)	36% (SD, 19%)	38% (SD, 22%)	P = 0.43
Young et al, 2007 (37)	Difference in rate of change for adherence to targets of:			
	2 hemoglobin A_{1c} measurements per year	0.0176	0.0262	P = NS
	Annual LDL cholesterol measurement	0.0439	0.0472	P = NS
	Annual urinalysis/microalbumin measurement	0.0278	0.0105	P = NS
	Annual eye examination	0.0195	0.0713	P < 0.001
Regression analysis of administrative data		Adjusted Odds Ratio (CI)†		
Pourat et al, 2005 (38)	Association between reimbursement method and self-reported:			
	Screening sexually active females aged 15–19 y for chlamydia annually:			
	FFS	0.96 (0.58 to 1.59)		
	Capitation with P4P	1.32 (0.65 to 2.72)		
	Salary with P4P	1.45 (0.66 to 3.15)		
	Screening sexually active females aged 20–25 y for chlamydia annually			
	FFS	1.32 (0.55 to 3.15)		
	Capitation with P4P	0.87 (0.43 to 1.75)		
	Salary with P4P	1.37 (0.50 to 3.74)		
	Provide chlamydia drugs to affected patients for treatment of partner			
	FFS	0.59 (0.30 to 1.15)		
	Capitation with P4P	0.93 (0.49 to 1.77)		
	Salary with P4P	0.78 (0.51 to 1.20)		

ACE = angiotensin-converting enzyme; FFS = fee-for-service; ITS = interrupted time series; LDL = low-density lipoprotein; NR = not reported; NS = not significant; P4P = pay-for-performance.
* All CIs are 95% CIs.
† Adjusted for practice characteristics (practice setting, volume of Medicaid patients, number of Medicaid HMO contracts, number of medical group contracts with Medicaid business), individual physician characteristics (sex, specialty, years in practice), having sexually transmitted disease guidelines from the Centers for Disease Control and Prevention and U.S. Preventive Services Task Force, having ever received feedback on sexually transmitted disease screening from the contracted Medicaid HMO or medical group, and the type of contracted Medicaid managed care health plan.

Source: Table from Houle et al. (2012) [31]

Table 3. Summary of major clinical outcomes before and after checklist implementation

	Any complication (%)			Mortality (%)			Surgical-site infection (%)		
	Before	After	*P*	Before	After	*P*	Before	After	*P*
Haynes *et al.*[18]	11·0	7·0	< 0·001	1·5	0·8	0·003	6·2	3·4	< 0·001
Site 1	11·6	7·0	< 0·050	1·0	0	< 0·050	4·0	2·0	< 0·050
Site 2	7·8	6·3	> 0·050	1·1	0·3	> 0·050	2·0	1·7	> 0·050
Site 3	13·5	9·7	> 0·050	0·8	1·4	> 0·050	5·8	4·3	> 0·050
Site 4	7·5	5·5	> 0·050	1·0	0·6	> 0·050	3·1	2·6	> 0·050
Site 5	21·4	5·5	< 0·050	1·4	0·0	< 0·050	20·5	3·6	< 0·050
Site 6	10·1	9·7	> 0·050	3·6	1·7	> 0·050	4·0	4·0	> 0·050
Site 7	12·4	8·0	< 0·050	2·1	1·7	> 0·050	9·5	5·8	> 0·050
Site 8	6·1	3·6	> 0·050	1·4	0·3	> 0·050	4·1	2·4	> 0·050
Weiser *et al.*[19]	18·4	11·7	0·001	3·7	1·4	0·007	11·2	6·6	< 0·001
Sewell *et al.*[20]	8·5	7·6	RR 0·89 (0·58,1·37)	1·9	1·6	RR 0·88 (0·34, 2·26)	4·4	3·5	n.a.
Askarian *et al.*[21]	22·9	10·0	0·03	n.a.	n.a.	n.a.	10·4	5·3	0·1
van Klei *et al.*[11]	n.a.	n.a.	n.a.	3·1	2·9	OR 0·91 (0·78, 1·05)	n.a.	n.a.	n.a.
Bliss *et al.*[22]	23·6	8	< 0·001	n.a.	n.a.	n.a.	6·2	5	0·845
Kwok *et al.*[23]	21·5	8·8	< 0·001	4·0	3·1	0·151	14·9	4·7	< 0·001

Values in parentheses are 95 per cent confidence intervals. RR, risk ratio; n.a., not available; OR, odds ratio.

Source: Table from Bergs et al. (2014)[57]

Table 4. Survey instruments used in four countries

Survey Instruments Used in the Four Countries

Country	*Methodology*	*Observations*	*Sample*	*Year of survey*
Tanzania, Arusha region	Vignettes and direct observation	111 doctors, 1178 provider–patient interactions	Rural/urban doctors practicing in Western-style health facilities	2002–03
India	Vignettes and direct observation	215 providers, 4108 provider–patient interactions	Delhi-based public and private providers spanning a wide range of treatment styles	2002–03
Paraguay	Direct observation and exit surveys	286 providers, 2200 provider–patient interactions	Public providers practicing "Western" forms of medicine sampled from 4 "departments" in the country.	2004
Indonesia	Vignettes	992 Facilities	Providers practicing "Western" forms of medicine sampled from the entire country.	1993 Indonesia Family Life Survey

Notes: Kenneth Leonard was the principal investigator on the Tanzanian survey; Jishnu Das and Jeffrey Hammer on the Indian survey; Jishnu Das and Daniel Dulitzky on the Paraguay survey. The Indonesian results are drawn from the Indonesia Family Life Survey, data for which were provided by Sarah Barber.

Source: Table from Das et al. (2008) [2]

Table 5. International comparisons of effort

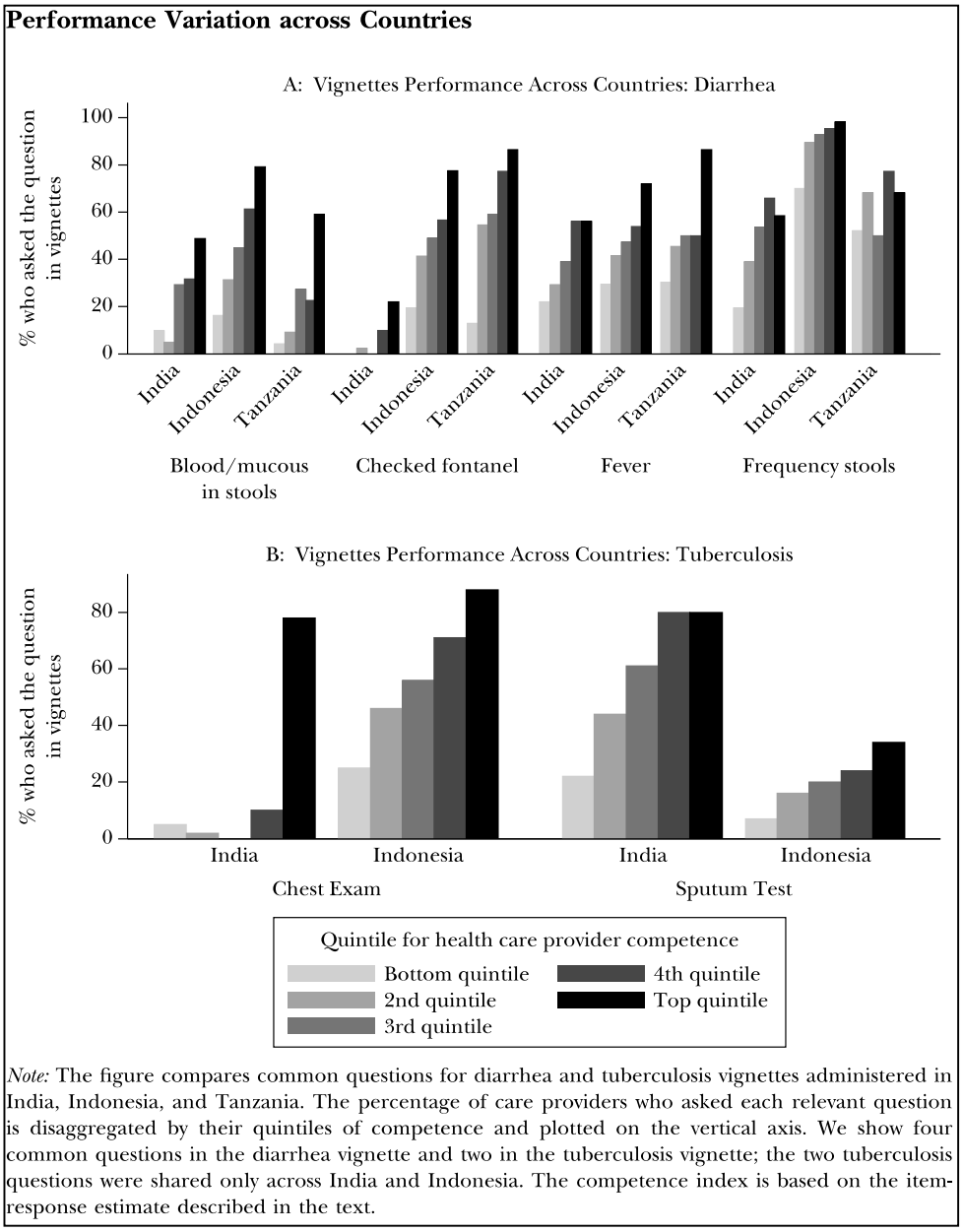
International Comparisons of Effort

Country/Effort category	*Time spent*	*Questions asked of patient*	*Number of physical exams*	*(Total number of medicines given)*
Dehli				
Doctors who exert low effort	1.9	1.36	0.97	2.13
Doctors who exert medium effort	3.36	2.94	1.0	2.72
Doctors who exert high effort	6.15	5.32	1.37	3.05
All doctors	**3.80**	**3.20**	**1.09**	**2.63**
Paraguay				
Doctors who exert low effort	5.79	5.33	1.38	1.36
Doctors who exert medium effort	7.90	7.50	2.93	1.55
Doctors who exert high effort	11.34	11.91	3.64	1.65
All doctors	**8.33**	**8.23**	**2.65**	**1.52**
Tanzania				
Doctors who exert low effort (25th Percentile)	3	2	0	N/A
All doctors	**6.32**	**3.96**	**1.51**	**N/A**
Germany	7.6	N/A	N/A	N/A
Spain	7.8	N/A	N/A	N/A
Belgium	15.0	N/A	N/A	N/A
United Kingdom	9.4	N/A	N/A	N/A

Notes: We divide doctors by terciles of effort in India and Paraguay, and the 25th percentile versus all doctors for Tanzania. The data are based on the following sources: India—Das and Hammer (2007); Paraguay—Das and Sohnesen (2007); Tanzania—based on calculations by Kenneth Leonard; International Comparisons—Hogelzeir et al. (1993) and Deveugele, Derese, Brink-Muinen, Bensing, and De Maeseneer (2003).

Source: Table from Das et al. (2008)[2]

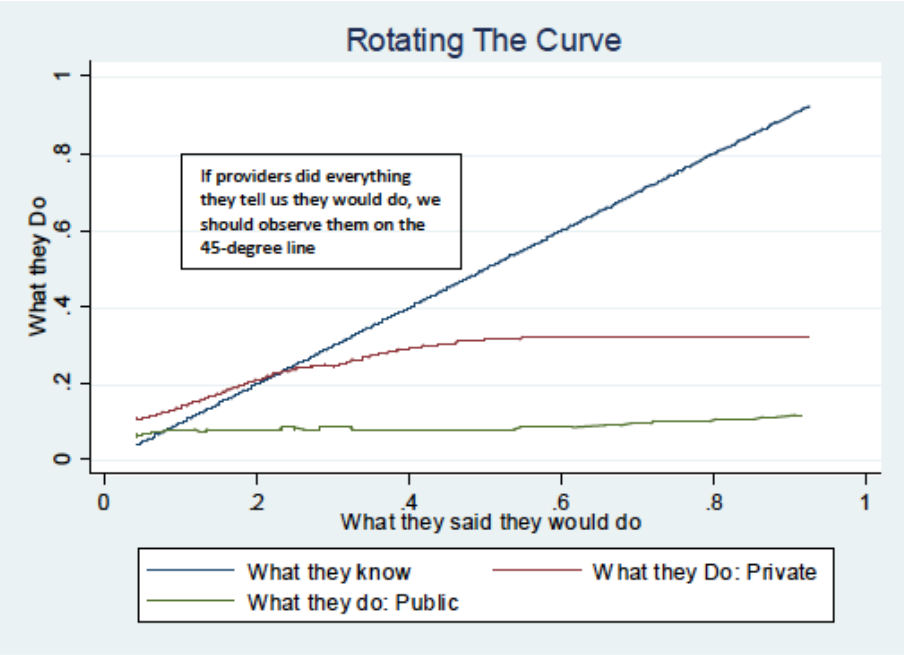
Figure 1. Performance variation across countries



Note: The figure compares common questions for diarrhea and tuberculosis vignettes administered in India, Indonesia, and Tanzania. The percentage of care providers who asked each relevant question is disaggregated by their quintiles of competence and plotted on the vertical axis. We show four common questions in the diarrhea vignette and two in the tuberculosis vignette; the two tuberculosis questions were shared only across India and Indonesia. The competence index is based on the item-response estimate described in the text.

Source: Figure from Das et al. (2008)[2]

Figure 2: The know-do gap



Note: This figure shows the "know-do" gap in medical care. On the horizontal axis we plot what a provider knows, as measured by medical vignettes, using compliance with a medically necessary case-specific checklist of history questions and examinations. On the vertical axis, we plot what the provider actually did with a similar patient, observed in practice. Every history question and examination can be compared in a pairwise comparison. The figure shows (A) at very low levels of knowledge, practice is constrained by knowledge; (B) at higher levels of knowledge, there is a significant gap between knowledge and practice and (C) the know-do gap is larger in the public sector (where there is no correlation between practice and knowledge), but even in the private sector, there is a significant gap at higher levels of knowledge.

Source: Figure from Das and Hammer (2014)[3]