

Survey using incognito standardized patients shows poor quality care in China's rural clinics

Sean Sylvia,¹ Yaojiang Shi,^{2*} Hao Xue,³ Xin Tian,³ Huan Wang,³ Qingmei Liu,⁴ Alexis Medina⁵ and Scott Rozelle⁵

¹School of Economics, Renmin University of China, 59 Zhongguancun Avenue, Beijing, 100872 China, ²Center for Experimental Economics in Education (CEEE), Shaanxi Normal University, 199 South Chang'an Road, Xi'an, Shaanxi, 710062 China, ³School of Economics and Management, Northwest University, No. 1 Xuefu Da Road, Xi'an, Shaanxi, 710127 China, ⁴Shaanxi No. 4 Provincial People's Hospital, 512 Xianning East Rd, Xi'an, Shaanxi, 710043 China and ⁵Freeman Spogli Institute for International Studies, Stanford University, 616 Serra Street, Stanford, CA 94305, USA

*Corresponding author. Center for Experimental Economics in Education (CEEE), Shaanxi Normal University, 199 South Chang'an Road, Xi'an, Shaanxi, China, 710062. E-mail: shiyaojiang7@gmail.com

Accepted 6 February 2014

Over the past decade, China has implemented reforms designed to expand access to health care in rural areas. Little objective evidence exists, however, on the quality of that care. This study reports results from a standardized patient study designed to assess the quality of care delivered by village clinicians in rural China. To measure quality, we recruited individuals from the local community to serve as undercover patients and trained them to present consistent symptoms of two common illnesses (dysentery and angina). Based on 82 covert interactions between the standardized patients and local clinicians, we find that the quality of care is low as measured by adherence to clinical checklists and the rates of correct diagnoses and treatments. Further analysis suggests that quality is most strongly correlated with provider qualifications. Our results highlight the need for policy action to address the low quality of care delivered by grassroots providers.

Keywords Physician quality, China, standardized patients

KEY MESSAGES

- We used standardized patients to measure the quality of health care in rural China.
- Clinicians completed only 18% of items on a checklist of recommended questions and exams.
- Only 26% of clinician diagnoses were correct.
- Clinicians provided medicine that was unnecessary/harmful in 64% of the interactions.

Introduction

Over the past decade, China has implemented comprehensive health system reforms and invested heavily in providing health care for the more than 600 million people who live in rural areas. Between 2000 and 2010 total health expenditure per capita in rural areas more than doubled, from 116 yuan (\$18.4) per capita to 326 yuan (\$51.7) per capita, mostly due to increased government investment (Yip *et al.* 2012). The majority of this investment has been aimed at expanding access to

health care by increasing coverage under new public health insurance programmes (Eggleston 2012). By the end of 2010 around 96% of rural residents had insurance coverage (Liang *et al.* 2012).

Although China has undoubtedly made great strides in improving access to health care in rural areas since the start of reforms, a fundamental question that remains is whether reforms have led to actual gains in population health outcomes (Wagstaff *et al.* 2009; Babiarz *et al.* 2010; Liang *et al.* 2012). Even with greater access to health care, health outcomes may not

improve if the quality of care is poor. Government attention has recently shifted towards improving health care services; however, little evidence exists on the quality of care currently delivered by providers in China's rural areas (Eggleston *et al.* 2008; Yip *et al.* 2012).

The primary goal of this study is to provide objective evidence on the quality of primary care delivered by village clinicians in China's rural areas. We provide estimates of the quality of care based on interactions between village clinicians and incognito standardized patients (SPs). The SP approach has long been considered the gold standard for assessing clinical practice in developed settings (Glassman *et al.* 2000; Peabody *et al.* 2000; Rethans *et al.* 2007; Mohanan *et al.* 2011; Das *et al.* 2012), and has recently begun to be applied in developing settings as well (Das *et al.* 2012). In our study, SPs were individuals recruited from the local community and trained to consistently and covertly present cases of common illnesses to clinicians. We use data collected during SP-provider interactions to evaluate providers on three main dimensions of care quality: (1) process quality (i.e. how closely clinicians adhere to disease-specific lists of recommended questions and physical exams), (2) the accuracy of diagnoses and (3) the appropriateness of treatment. In addition to describing the quality of care provided by village clinicians, we compare these results to a smaller sample of SP interactions with township health centre (THC) clinicians (the next level up in the tiered rural health care system). We also use additional data collected on providers prior to SP visits to assess how provider characteristics are correlated with different dimensions of quality as well as with fees charged for drugs and services. Finally, to provide additional perspective, we compare our results with results from a similar SP study of the quality of health care providers in rural and urban India.

Village clinicians and China's rural health care reforms

Village clinicians are the front line of the rural health system in China. At the bottom tier of a three-tiered hierarchy, village clinicians provide primary care and public health services to China's rural population under the management of THCs (the middle tier) (Babiarz *et al.* 2012). Although China does not impose gatekeeping requirements, the majority of rural patients is first seen in village clinics, often the only source of medical care in remote villages (Babiarz *et al.* 2010; Li *et al.* 2013). In fact, the rural insurance system is set up to encourage patients to first contact their village clinician for treatment for a specified set of basic health conditions. County hospitals at the top tier of the rural health system provide relatively expensive, specialized care (Babiarz *et al.* 2012).

Expanding the scope of public health service provision through village providers is a primary focus of China's newest wave of health care reforms (Chen 2009; Yip *et al.* 2012). As part of these reforms, the government provides funding of 25 yuan (\$3.97) per head for village providers to deliver a defined package of public health services to the population in their catchment areas. This payment is on top of all existing wage and salary payments, and is meant to cover the additional labour involved in providing vaccinations, time spent administering the rural health insurance system inside the village and other public health services. In terms of primary health care

service provision, in the long-run, reforms envision village clinicians serving as gatekeepers to upper tiers of the system (Yip and Hsiao 2009; Yip *et al.* 2012).

Although village clinicians are shouldering an increased role as part of China's rural health system reforms, there is currently little objective evidence on the quality of care they provide (Eggleston *et al.* 2008). Indeed, some scholars have voiced concerns that primary care providers in rural areas lack the education and skills necessary to address the country's future health challenges (Eggleston *et al.* 2008). Non-communicable diseases are already widespread in both rural and urban areas (Yang *et al.* 2008, 2010). Heart conditions, cancer and cerebrovascular disease are now the leading causes of death in rural China (He *et al.* 2005). At the same time, the burden of many infectious diseases—such as tuberculosis and hepatitis B—remains significant (Wang *et al.* 2008). It is unclear whether a minimally educated village clinician has the training required to tackle this range of sophisticated medical conditions.

Another often-voiced concern is that primary health care providers have strong incentives for unnecessary care and waste (Eggleston *et al.* 2008; Yip and Hsiao 2009; Yip *et al.* 2010; Bloom, 2011; Wang *et al.* 2011a,b). Until recently, village clinicians were fully reliant on sales of services and drugs to clients. In recent years, in most villages, village clinicians receive a nominal base salary from the government. China's government is looking to gradually change this (as evidenced by the new per capita public health payments to village clinicians described earlier), but for now it is safe to say that village clinicians are almost fully private practitioners. They are best thought of as self-employed clinicians who live mostly off of their earnings in the village. Village clinicians' heavy reliance on drug sales and user fees for income paired with a price schedule distorted away from basic care and prevention have led to escalating patient costs and inappropriate treatment (Eggleston *et al.* 2008; Yip and Hsiao 2008; Yip *et al.* 2010). Studies have shown, for example, that antibiotics are widely overprescribed (Reynolds and McKee 2009; Currie *et al.* 2011).

Are China's rural clinicians prepared for the unique and daunting epidemiological landscape the country faces? Will the patients that utilize the services of these providers end up with improved health outcomes? Although some have voiced concerns that the answer to these questions is no, others argue that the quality of rural clinicians is sufficiently high as not to be of first order importance, and that reforms should first focus on other aspects of the health care system.

Materials and methods

Survey sample and facility surveys

Our study was conducted in six counties located in southern Shaanxi province. Shaanxi province is a western inland province with a high number of poor, rural counties. The per capita rural income and average life expectancy in Shaanxi are among the lowest in all of China, both well below the national average (CNBS 2011). From each of our six sample counties, we randomly selected two townships (the middle level of administration between the county and the village) and three villages from each township. Within each village a census was taken of all clinics designated for reimbursement under the New

Cooperative Medical Scheme, the major public insurance programme in rural areas. From this list, one clinic per village was selected at random. [Most villages (67%) had only one eligible village clinic, in which case that clinic was automatically selected for participation in the study. In the remaining villages with two or more eligible clinics (only 8% of villages had more than two), we randomly selected one clinic for the study.] In all, the study sample consisted of 36 village clinics and 12 THCs and was representative of a catchment area serving approximately 2.35 million individuals.

Following sample selection, an initial facility survey was conducted in November 2012. The facility survey included questions on clinic infrastructure, patient load and disease frequency. A separate module collected information on clinicians employed at each clinic, including demographic characteristics, educational attainment, medical qualifications and training, and experience.

This study received ethical approval from the Stanford University Institutional Review Board (IRB) on 19 November 2012 (Protocol ID 25904). The IRB protocol allowed us to record SP interactions using a concealed device (discussed below), provided that we contacted providers in the sample at some point prior to SP visits and obtained their verbal consent to participate in the study. To obtain consent, clinicians were called approximately 4 months prior to the actual SP visits and asked if they agreed to be recorded by an individual posing as a patient. All of the clinicians asked gave their consent.

Standardized patients

The quality of care provided by rural clinicians was assessed based on interactions between rural clinicians and incognito SPs. SPs were individuals recruited from the local community and trained to present consistent disease cases. Using these interactions, we assessed clinicians on process quality, correctness of diagnoses and appropriateness of treatment.

The SP approach (also commonly referred to as ‘simulated’ patients) is considered the gold standard for assessing clinical practice (Glassman *et al.* 2000; Peabody *et al.* 2000; Rethans *et al.* 2007; Mohanan *et al.* 2011; Das *et al.* 2012). This approach has distinct advantages over other methods commonly used to measure provider quality (Woodward *et al.* 1985; Peabody *et al.* 2000; Das *et al.* 2012). First, because clinicians are unaware that they are being evaluated, assessments using SPs are not biased by changes in clinician behaviour due to observation alone (or ‘Hawthorne Effects’) (Woodward *et al.* 1985; Glassman *et al.* 2000; Peabody *et al.* 2000; Leonard and Masatu 2010). Second, this method measures ‘actual’ clinical practice as opposed to clinician knowledge (as is measured using clinical vignettes). Third, the SP approach (particularly when interactions are recorded) is less subject to recall bias than are assessments based on patient exit interviews, for example. Finally, because patients and cases are common and standardized across providers, the SP approach allows for comparisons across providers of different types and in different locations. Although SPs have long been used in western medical schools and in evaluating providers in developed countries, we are aware of only one other study that has used the SP approach to measure the quality of primary care providers in a developing country (India—Das *et al.* 2012).

Recruitment and training of SPs

To ensure that SPs were similar to patients typically seen by clinicians (in language, mannerisms, dress, etc.), SPs were recruited from the sample counties (though not from the specific villages included in the sample). A total of four SPs, two male and two female, were selected to participate in the study based on initial interviews and demographic match to the disease cases used in the study (discussed below).

Selected SPs were trained in a classroom setting for 4 days (32 h) by a team consisting of researchers and consulting medical professionals. Classroom training focused on preparing SPs to represent their assigned disease cases to providers in a consistent and unsuspecting manner. Medical professionals discussed the symptoms of diseases to be portrayed and the typical behaviour and presentation of real patients afflicted with the disease. Following classroom instruction, the SPs went through extensive field rehearsals in rural areas with clinicians who volunteered to assist with the study.

Cases presented by SPs

Each SP was trained to present a case of either dysentery or unstable angina. (For the dysentery case, female SPs presented the case of a child who was not present.) These diseases were selected to match those diseases used in the Das *et al.* (2012) study of rural India. Using the same diseases in each study allows for comparisons between rural China and rural India. These diseases are appropriate for the SP methodology as (1) there are no obvious physiological symptoms and (2) there is low risk that SPs would be exposed to invasive procedures or tests.

These diseases are also highly relevant to the current and future disease burden in rural China. Although deaths from dysentery have declined dramatically in recent years, it remains prevalent in many rural areas (Wang *et al.* 2006). In 2006, dysentery was the third highest notified communicable disease after tuberculosis and hepatitis B (Wang *et al.* 2008). The burden of heart disease is on the rise and is now a leading cause of death in rural China (He *et al.* 2005). The ability of rural clinicians to correctly diagnose and treat cases of dysentery and angina, therefore, provides an indication of how well rural providers are prepared to address both China’s ‘future’ prevalence of non-communicable disease and the infectious diseases of China’s ‘past’.

The SPs were trained on and followed scripts for each disease. Each script included disease symptoms and history as well as a detailed background story for each case. These scripts were based on those used in a previous study in India and were adapted for use in China by the research team, which included consulting doctors from local hospitals (Das *et al.* 2012). In designing case scripts, the intention was to make diagnosis relatively easy but to not lead clinicians or reveal information that could be used in assessment before being prompted by the clinician himself.

SP visits

SPs visited sample providers in late January 2013. Two SPs independently visited each provider in the sample (one presenting each disease case). One of the two SPs trained in each case was randomly assigned to each provider to reduce the potential for any remaining differences in individual SP

presentations to bias comparisons across providers. Upon entering each clinic, the SPs were seen by whoever would have seen them had they been a regular patient (they made no attempt to be seen by specific clinicians). As, according to our data, only one clinician was employed in the large majority of clinics (78%), this typically was not an issue.

We used three methods to collect data from SP–clinician interactions. First, SPs wore a concealed recording device. This allowed us to accurately score interactions without relying on the SP's ability to recall details. Second, SPs were administered a case-specific 'debriefing survey' upon exiting clinics. This survey covered the interaction with the clinician as well as the SP's own impressions of the providers and any additional observations made by the SP that they thought relevant but not captured on the audio recording. Finally, to collect information on drugs dispensed and fees charged, SPs were directed to purchase any and all medications prescribed (which are sold by providers in China) and to pay any additional fees charged by the provider.

Quality measures

We evaluated clinicians on three dimensions of quality observable from their interaction with the SPs: process quality, the accuracy of diagnoses and the appropriateness of treatment. We assessed process quality by grading recorded interactions against 'clinical checklists' of recommended questions to be asked of the patient and any physical exams to be performed (full checklists for each case are given in the [Supplementary Appendix](#)). These clinical checklists were identical to those used in the [Das et al. \(2012\)](#) study of quality of care in rural India. In adapting them to the Chinese setting, we checked the checklists against the Chinese national treatment guidelines; finding no discrepancies, and finding the checklists to be generally applicable to the rural Chinese clinical setting, we left the checklists unchanged. The clinical checklist for the unstable angina case included 22 'question' items and 5 'exam' items. The dysentery list consisted of 19 'question' items. There were no 'exam' items on the dysentery list as the child was not present. A subset of items on each checklist are considered 'essential'. These are items that were determined to be necessary for any prudent clinician to be able to make a basic diagnosis and protect a patient from harm. Six items on the angina checklist were deemed essential; four items on the dysentery checklist were deemed essential.

Diagnoses and treatments were assessed based on predetermined standards of correctness (full standards given in the [Supplementary Appendix](#)). Diagnoses were classified as 'correct', 'partially correct' or 'incorrect'. To ensure that diagnoses were given for each interaction, SPs were instructed to ask clinicians directly at the conclusion of the visit if a diagnosis had not already been volunteered. Treatments were similarly judged against a predetermined definition of appropriate treatments. Treatments were deemed 'correct' if clinicians dispensed any one of the 'correct' medications. For the unstable angina case, treatments were also considered correct if the clinician referred the SP to an upper level provider.

Results

Village provider characteristics

As part of the facility survey we asked village clinicians how frequently they saw cases of diarrhoea and heart problems ([Table 1](#)). On average, they reported seeing 1.17 cases of diarrhoea and 0.83 cases of heart problems in the previous 2 weeks. We also asked village clinicians whether they felt their clinic had the necessary equipment to treat 'moderate' and 'serious' cases of diarrhoea and heart problems. For diarrhoea, most providers (88.9%) said that they had the necessary equipment to treat moderate cases and 5.6% said that they could treat serious cases. Another 5.6% said that they did not have the necessary equipment to treat the disease. For heart problems, slightly under half (47.2%) said that they had the necessary equipment to treat moderate cases. None thought that they had the necessary equipment to treat severe cases. As providers at the lowest tier of China's health system, if village clinicians encounter diseases that they cannot treat, they are supposed to refer patients to the THCs (the next tier).

Village clinicians had low levels of general education and formal medical qualifications ([Table 2](#)). Only 20% completed academic high school or vocational college, and none completed academic college. The highest degree completed by the majority of clinicians (60.5%) was vocational high school. The majority (84.3%) of village clinicians held only the most basic certificate required to practice medicine in rural areas: a 'Rural Physician' certificate. The remaining 15.7% had higher certifications: either an 'Assistant Practising Physician' certificate or a 'Practising Physician' certificate. This pattern of education and medical qualifications is roughly similar to that found in larger surveys of village clinicians in China's western regions ([Li et al. 2013](#)).

Also of note is that only a fraction of the village clinicians' total income is derived from his/her base salary. On average, only 351 yuan (\$55) per month of a total income of 1355 yuan (\$215) was not from the sales of drugs or services rendered. Only 59% of village clinicians had any base salary at all; among these, the average base salary was 603 yuan (\$96) out of a total income of 1601 yuan (\$254). As noted earlier, this is the prevailing pattern of compensation among medical providers in China's villages and is the source of incentives to over-prescribe drugs and not refer patients to other tiers of the health system ([Eggleston et al. 2008](#); [Yip et al. 2010](#)).

Our survey data show that village clinics have the medical equipment needed for basic examinations. All of the clinics were stocked with at least one stethoscope, sphygmomanometer (blood pressure meter) and thermometer—three pieces of equipment that are needed for the recommended exams on the clinical checklist we used to score clinicians on process quality for angina. Due to low ownership rates of electrocardiogram (EKG) recording devices, the EKG item on the checklist was scored based on whether clinicians referred patients to another provider for the test.

SP interactions and provider quality

Given our sample of 36 village clinics, there were a total of 72 potential interactions with village clinicians across the two diseases. At the time of the SP visits, village clinicians were

Table 1 Clinic statistics on diseases used for SP cases

	Diarrhoea ^a Mean (SD) ^b or %	Heart problems Mean (SD) ^b or %
Number of patients seen in past 2 weeks with diarrhoea and heart problems	1.17 (1.84)	0.83 (1.4)
Clinic has the necessary equipment to treat		
Serious cases	5.6%	0%
Moderate cases	88.9%	47.2%
Does not have the necessary equipment	5.6%	52.8%

Source: Survey of providers conducted approximately 6 months before SP visit.

^aProviders were asked about frequency of 'diarrhoea' and 'heart problems', not dysentery and angina specifically.

^bSD: Standard deviation.

Table 2 Medical provider characteristics

	Village clinics (<i>n</i> = 36) Mean (SD) or %	Township health centres ^a (<i>n</i> = 12) Mean (SD) or %
Clinician characteristics		
<i>N</i>	45	24
Age (years)	45.85 (10.87)	43.78 (8.565)
Male	73.3%	83.3%
Experience (years)	23.33 (9.523)	20.89 (8.575)
Education		
Middle school	19.5%	0%
Vocational high school	60.5%	41.7%
Vocational college	10%	40.3%
General high school	10%	4.2%
College	0%	13.9%
Qualification		
Practising Physician	10%	66.7%
Assistant Practising Physician	5.7%	33.3%
Rural Physician	84.3%	0%
Hours worked per day	9.943 (4.311)	—
Base salary (yuan)	351.2 (377.7)	—
Total income (yuan)	1355.2 (1400.7)	—
Facility characteristics		
<i>N</i>	36	12
Patients per week	43.69 (31.69)	152.6 (141.5)
Value of medical instruments (10 000 yuan)	1.097 (1.265)	34.29 (31.59)
Stethoscope	100%	—
Sphygmomanometer	100%	—
Sterilization pot	97.2%	—
Thermometer	100%	—
Ultraviolet (UV) sterilizer	100%	—
Debridement kit	97.2%	—
Tongue depressor	88.9%	—
Suction device	16.7%	—
Oxygen bag	52.8%	—

Source: Survey of providers conducted approximately 6 months before SP visit.

^aDetailed questions on medical training and equipment were not asked of township providers.

absent in five instances for the dysentery case and six instances for the angina case. The rate of absenteeism was thus around 15%—lower than that found in many other developing countries (Chaudhury *et al.* 2006). It is also important to note

that village clinicians in China often make house calls and perform public health duties outside of the clinic, so absence at the time of the unannounced visit is not necessarily a sign of shirking.

The detection rate of the SPs was assessed in the post-interaction debriefing survey and found to be extremely low. In only two instances (two interactions for the angina case), was there any suspicion that the clinician might have detected the SP. We excluded these two instances from the analysis. To further test whether suspicion of the SPs may have influenced clinician performance, we examined the correlation between the probability that SPs were ‘correctly’ treated and (1) time spent on diagnosis and (2) the percentage of checklist items completed. We find both to be positive. Assuming that clinicians would be more likely to treat ‘incorrectly’ if they believed that the patient was faking their condition, this is the opposite of what one would expect if obtaining more information from an SP raised suspicion that the SPs were genuine patients.

After excluding the two cases where SPs were plausibly detected plus the 11 cases in which the village doctor was absent, the final sample contained 59 total interactions with village clinicians across the two diseases.

Process quality

Unannounced SP visits revealed the overall quality of care provided in village clinics to be poor. Although the average length of each SP interaction with providers was 7.2 min, the majority of this time (3.6 min) was spent filling prescriptions (Table 3). On average, clinicians only spent 1.6 min consulting with patients. During these consultations, village clinicians asked 18% of the recommended questions on average. For the unstable angina case, only 15% of the recommended exams were performed. Of checklist items deemed essential to make a proper diagnosis and protect patients from harm, clinicians addressed only 36% on average. No village clinicians addressed all essential items.

In looking at the frequency of individual questions asked and exams performed (Figures 1 and 2), the emphasis appears to be on collecting the information needed to provide medicine. In the dysentery case, 97% of clinicians asked the age of the child—the main basis for prescription of medication—whereas the remaining questions were asked around half of the time or less. Only 10% of clinicians asked about the frequency of urination, an indication of dehydration that is supposed to be one of the primary concerns of a provider in cases of dysentery.

Clinicians asked unstable angina SPs about pain location in almost all interactions yet asked other checklist items and conducted exams infrequently. Few clinicians (only 14%) asked about pain radiation, a clear symptom of stable or unstable angina. Further, although considered essential, and even though all providers had the necessary equipment, only 11% of clinicians took the blood pressure of the SPs.

Diagnoses

Given the low level of process quality, it is unsurprising that the quality of diagnoses and treatments was also poor. Diagnoses were volunteered in 71% of the angina cases and 32% of the dysentery cases. A fully correct diagnosis was made in only 26% of the interactions (Table 3). Diagnoses were completely incorrect in 41% of interactions. The most common incorrect diagnosis was ‘indigestion’ for the dysentery case and ‘sprain’ for the angina case.

Treatments

Across the two cases, treatments were deemed correct or partially correct in 53% of interactions (Table 3). Note, though, that treatments are judged correct or partially correct if they include any of the correct medications or if clinicians gave a referral in the angina case. This number, therefore, may be a poor indicator of treatment quality.

A better indication of appropriate treatment is perhaps given by the overall appropriateness of the medications dispensed. Village clinicians dispensed medications in 75% of interactions (61% of angina interactions and 87% of dysentery interactions). When medications were prescribed, on average, each village clinician prescribed 2.07 medications. In the cases where drugs were dispensed, consulting physicians determined that they were either unnecessary or harmful 64% of the time (see Supplementary Table 2 for standards for correct and incorrect treatments).

Failure of clinicians to refer patients was not entirely due to misdiagnosis. For the angina case (and heart problems more generally), for example, correct treatment dictates that patients are referred to higher levels of the rural health system. However, of clinicians who correctly or partially correctly diagnosed unstable angina, 25% failed to refer patients.

Patient costs

Total fees charged by village clinicians (including medication and clinic fees) were 11.5 yuan (\$1.85) on average. Angina SPs were charged 14.5 yuan (\$2.30) on average and dysentery SPs were charged 9 yuan (\$1.43). Village providers typically do not provide itemized receipts so obtaining a detailed breakdown of fees charged was not possible. The post-interaction debrief survey of SPs revealed that most of the fees charged were for drug sales.

Correlates of village provider quality and patient costs

The summary of quality discussed earlier hides meaningful differences across individual providers in our sample. For instance, while on average village clinicians asked 18.2% of recommended questions on average, clinicians at the 10th percentile asked 5.3% of recommended questions whereas clinicians at the 90th percentile asked 36.8% of recommended questions. In order to ascertain what might account for these differences in provider quality, in this section we present analyses that consider correlations between provider quality and individual characteristics such as physician qualifications and financial incentives.

Figure 3 shows results from multiple linear regressions for the different dimensions of provider quality and for fees charged during SP visits (full specification details are in the Supplementary Appendix). By far, the factors most strongly correlated with dimensions of quality were clinician education and qualifications. Village clinicians with ‘Practising Physician’ qualifications, for example, addressed 12.3% more recommended checklist items than those with lower levels of qualification. Clinicians with an upper secondary degree were 45 percentage points more likely to give a correct diagnosis and nearly 30 percentage points more likely to offer an appropriate treatment.

Table 3 Standardized patient measures of provider quality^a

	Village clinics			Township health centres		
	All (<i>n</i> = 59) Mean (SE) ^c	Angina (<i>n</i> = 28) Mean (SE) ^c	Dysentery (<i>n</i> = 31) Mean (SE) ^c	All (<i>n</i> = 23) Mean (SE) ^c	Angina (<i>n</i> = 12) Mean (SE) ^c	Dysentery (<i>n</i> = 11) Mean (SE) ^c
Time						
Total time (min)	7.24 (1.05)	8.57 (1.83)	6.05 (1.12)	10.52 (2.29)	15.08 (3.77)	5.54 (1.55)
Wait time (min)	1.72 (0.67)	2.57 (1.34)	0.96 (0.36)	1.17 (0.44)	1.12 (0.40)	1.23 (0.83)
Diagnosis time (min)	1.61 (0.19)	2.18 (0.33)	1.09 (0.16)	5.06 (1.34)	8.56 (2.12)	1.25 (0.19)
Prescription time (min), if given	3.63 (0.73)	3.82 (0.85)	3.51 (1.07)	4.10 (0.97)	4.22 (0.80)	3.95 (2.05)
Process quality						
Number of recommended questions asked	3.73 (0.28)	4.11 (0.36)	3.39 (0.42)	4.78 (0.47)	6.00 (0.63)	3.45 (0.47)
% recommended questions asked	18.23 (1.38)	18.67 (1.66)	17.83 (2.19)	22.92 (2.10)	27.27 (2.85)	18.18 (2.50)
Number of recommended exams performed	0.89 (0.22)	0.89 (0.22)	—	1.67 (0.26)	1.67 (0.26)	—
% recommended exams performed ^b	14.88 (3.67)	14.88 (3.67)	—	27.78 (4.27)	27.78 (4.27)	—
% recommended questions and exams	17.84 (1.41)	17.86 (1.77)	17.83 (2.19)	22.98 (2.05)	27.38 (2.69)	18.18 (2.50)
% essential questions and exams	36.44 (2.40)	32.14 (2.96)	40.32 (3.61)	43.48 (3.18)	37.50 (3.63)	50.00 (4.77)
Diagnoses^c						
Diagnosis correct (0/1)	0.26 (0.06)	0.25 (0.08)	0.27 (0.08)	0.52 (0.11)	0.58 (0.15)	0.45 (0.16)
Diagnosis partially correct (0/1)	0.33 (0.06)	0.32 (0.09)	0.33 (0.09)	0.13 (0.07)	0.25 (0.13)	0.00 (0.00)
Treatment						
Treatment correct or partially correct ^c (0/1)	0.53 (0.07)	0.61 (0.09)	0.45 (0.09)	0.52 (0.11)	0.67 (0.14)	0.36 (0.15)
Medications dispensed (0/1)	0.75 (0.06)	0.61 (0.09)	0.87 (0.06)	0.48 (0.11)	0.50 (0.15)	0.45 (0.16)
Number of medications dispensed, if any	2.07 (0.14)	2.35 (0.24)	1.89 (0.16)	2.55 (0.37)	2.17 (0.60)	3.00 (0.32)
Medication correct, if dispensed (0/1)	0.36 (0.07)	0.12 (0.08)	0.52 (0.10)	0.45 (0.16)	0.17 (0.17)	0.80 (0.20)
Referral to other provider ^d (0/1)	0.37 (0.06)	0.61 (0.09)	0.16 (0.07)	0.48 (0.11)	0.67 (0.14)	0.27 (0.14)
Referred to county provider, if referral (0/1)	0.41 (0.11)	0.47 (0.12)	0.20 (0.20)	0.91 (0.09)	0.88 (0.13)	1.00 (0.00)
Referred to township provider, if referral (0/1)	0.27 (0.10)	0.24 (0.11)	0.40 (0.24)	—	—	—
Suggested follow up visit (0/1)	0.19 (0.05)	0.04 (0.04)	0.32 (0.09)	0.61 (0.10)	0.33 (0.14)	0.91 (0.09)
Fees						
Total fees charged (yuan)	11.57 (1.83)	14.46 (3.60)	8.96 (1.16)	29.56 (7.16)	32.02 (10.24)	25.12 (9.02)

Source: Authors' analysis.

^aUnit of observation is the SP-provider interaction. Doctors were absent in five instances for the dysentery case and six instances for the angina case. Two angina interactions with village doctors were excluded due to suspicion that the doctor detected the SP.

^bNo exams recommended for dysentery case as child was not present.

^cSee [Supplementary Appendix](#) for classification of correct, partially correct, and wrong diagnoses and treatment.

^dReferrals to places other than the county or township provider were typically to a 'larger hospital'.

^eSE: Standard error in parentheses.

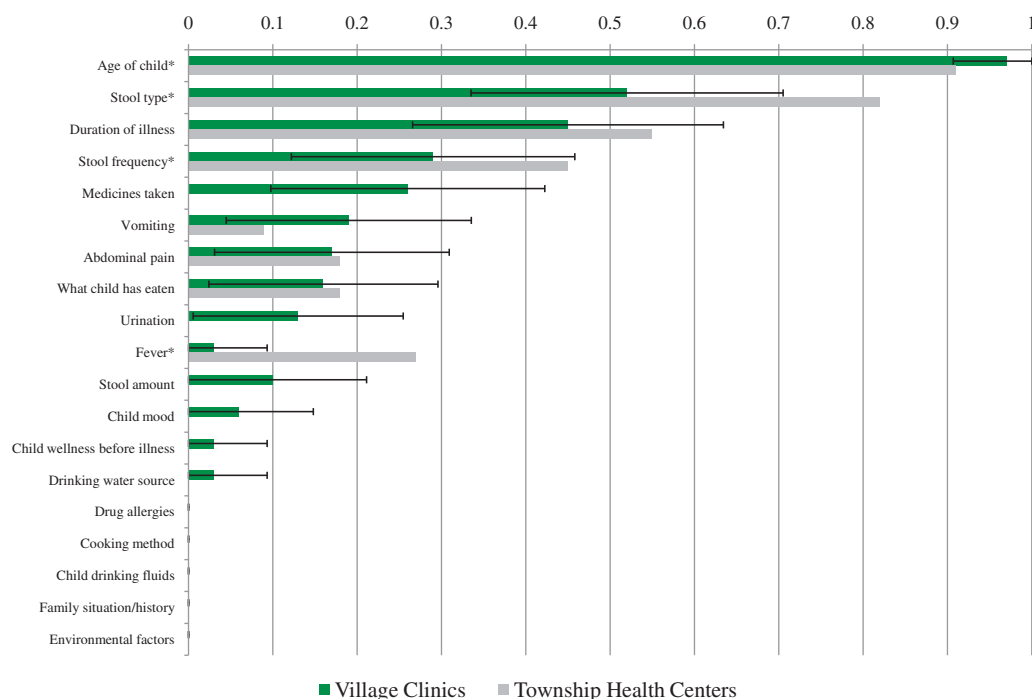


Figure 1 Adherence to clinical checklist, dysentery case

Source: Authors' analysis.

Notes: Figure shows means and 95% confidence intervals for clinician adherence to dysentery case clinical checklist items. Items with asterisks (*) are considered essential.

Apart from clinician education and qualifications, few other factors were correlated significantly with provider quality. Although these estimates should be interpreted with caution (given that our small sample size may prevent us from detecting these differences), correlations with some other variables are nevertheless instructive. Of particular interest, given current policy discussions, is the relationship between a clinician's base pay (pay not tied to drug or service sales) and the quality of services provided and costs charged by providers. Specifically, separating clinician pay from drug sales may reduce over-prescription of drugs and inappropriate treatment, but may also weaken clinician incentives. While not statistically significant, clinicians with a base salary (of any amount) charge patients about 5 yuan (55%) less in total fees. Point estimates for having base pay are positive for process quality and the probability of correct treatment, but negative for the probability of a correct diagnosis.

Comparison to THCs

While based on a small sample of interactions (23 between the two diseases), we found the quality of primary care provided in THCs to be better than that provided in village clinics, but still poor generally. Possibly due to higher levels of education and qualifications among clinicians in THCs, process quality and the accuracy of diagnoses were better than that provided by village clinicians (Table 3). This improvement in measured process quality and diagnoses, however, was mainly due to improved scores on the angina case. For this case, THC clinicians asked 8% more of the recommended questions and performed 13% more of the recommended exams. In the dysentery case,

township clinicians asked approximately the same percentage of recommended questions as village clinicians. Of those questions asked, however, a higher proportion was considered essential. Still, although they performed better than village clinicians, township clinicians only addressed 38% of essential items for angina and 50% for dysentery.

In addition to better process quality, THC clinicians were also more likely to provide a correct diagnosis. In total, diagnoses were volunteered by the THC clinicians in 83% of the angina cases and 36% of the dysentery cases. These rates are slightly higher than that among village clinicians. Overall, 52% of diagnoses were correct, twice the rate of village clinicians. The largest difference was for the angina case; 58% of township clinicians correctly identified this disease compared with 25% of village clinicians.

Although surprising given better scores on other quality dimensions, the appropriateness of treatments given by THC clinicians was comparable to village clinicians. THC clinicians gave correct treatments for the angina case 6% more often than village clinicians, but 9% less often for the dysentery case. Perhaps most strikingly, THCs were significantly less likely to dispense medicine. Although village clinicians dispensed at least one medication in 75% of interactions, THC clinicians gave drugs to only 48% of the patients. The largest difference for the rate of prescription was in the dysentery case where village clinicians dispensed medicine 42% more often. Moreover, when THC clinicians did prescribe medicine, they were more likely than village clinicians to do so correctly. In the dysentery case, for example, the medicines prescribed by THC clinicians were correct nearly 30% more often. This suggests that THC clinicians were less likely to dispense unnecessary medications.

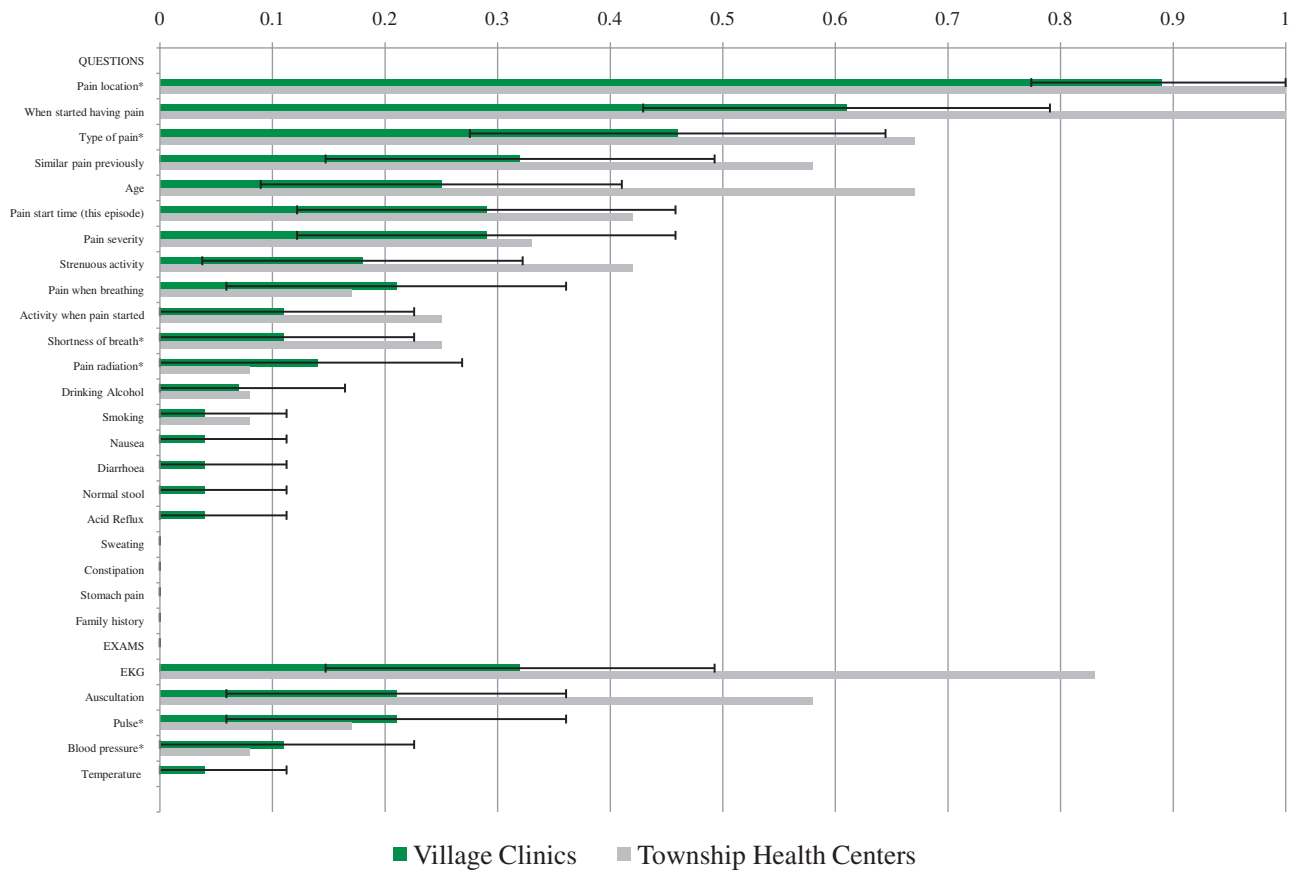


Figure 2 Adherence to clinical checklist, unstable angina case

Source: Authors' analysis.

Notes: Figure shows means and 95% confidence intervals for clinician adherence to dysentery case clinical checklist items. Items with asterisks (*) are considered essential. EKG refers to either performance of an electrocardiogram or referral for an electrocardiogram.

Using multivariate regression analysis (full specification and results shown in [Supplementary Table A6](#)), we find that THC clinicians address around 5% more of the recommended questions and physical exams ($P < 0.05$). They are also around 22 percentage points more likely to give a correct diagnosis ($P < 0.05$). Once clinician and medical qualifications are adjusted for, however, there is no significant difference between the process quality provided at THCs and village clinics.

Discussion

Interactions between incognito SPs and village clinicians showed the quality of care provided in China's rural clinics to be poor. On average, village clinicians completed only 18% of the items on a clinical checklist of recommended questions and physical exams, only 26% of clinician diagnoses were fully correct, and clinicians provided medicine that was unnecessary or harmful in 64% of the interactions.

The factors most strongly correlated with the quality of care were the educational attainment and medical qualification of clinicians. These factors also appear to explain a substantial portion of the difference in quality observed between village clinics and THCs at the next tier in the rural health system.

THC clinicians performed significantly better than village clinicians; however, the quality of care in THCs was also found to be low.

The results we found for village clinicians in China are comparable to those found in a recent study that used the same SP methodology to measure the quality of care provided by rural providers in Madhya Pradesh, one of the poorest states in India ([Das et al. 2012](#)). In that study, it was found that the average percentage of 'essential' clinical checklist items completed by rural providers was 34% (for dysentery and unstable angina cases). In our study—which closely replicated the Indian study, down to the same list of 'essential' checklist items used to measure process quality—we found that village clinicians completed 36% of these on average. In terms of diagnoses (for which assessment criteria were comparable), the study in India found that 56% of providers gave correct or partially correct diagnoses, compared with 59% in our sample. It is important to note, however, that the methodology for obtaining a diagnosis from providers varied slightly across the two studies: in the India study, diagnoses were obtained only from providers who volunteered one without prompting from the SP, whereas in the China study SPs were instructed to prompt providers for a diagnosis if none was volunteered. The study in India found a much lower rate of 'correct' treatments (22% compared with 53% in our study); however, these figures

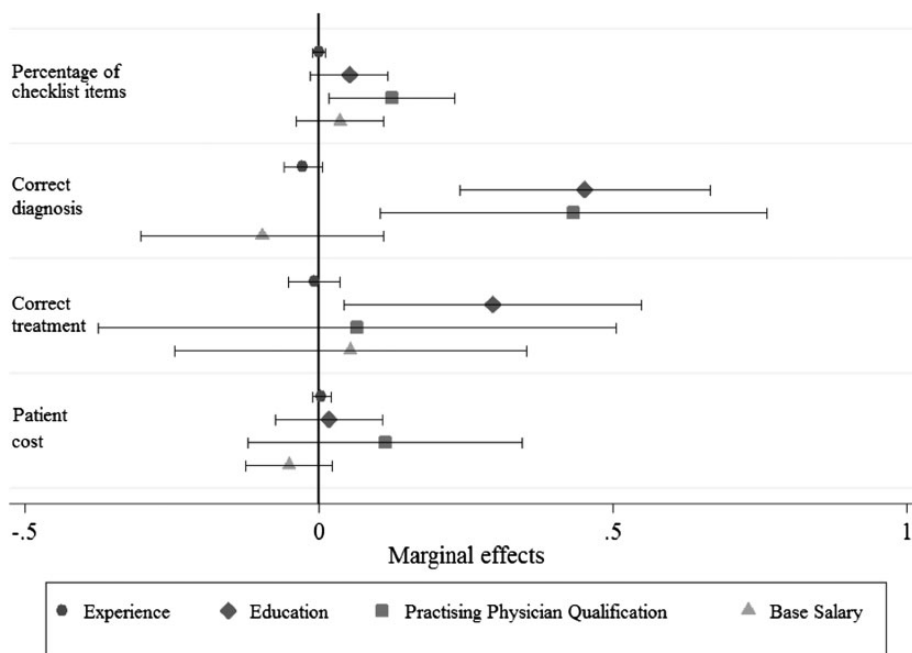


Figure 3 Marginal effects on process quality, diagnosis, treatment and patient costs in village clinics

Source: Authors' analysis.

Notes: Figure graphs the marginal effects and 95% confidence intervals of village clinician experience (in years), education (whether the clinician has an upper secondary degree or higher), qualification (whether the clinician has a 'Practising Physician' certificate), and compensation (whether the clinician has a base salary) on dimensions of quality and fees charged during standardized patient visits. For clinics with more than one practicing clinician, the average of clinician characteristics was used. Marginal effects on each dimension of quality and cost were estimated in separate regressions. Outcome variables for each regression were as follows: percentage of checklist items addressed by clinicians were rescaled to be between 0 and 1 (so 1 is 100%); correct diagnosis and treatment are 1 if clinicians gave a correct diagnosis and 0 otherwise; and patient costs are expressed in 100 yuan. Regressions for the percentage of checklist items and patient costs were estimated using ordinary least squares and correct diagnosis and treatment were estimated using probit regressions. Each regression also controls for clinician age and gender, the total value of the provider's medical instruments, patient load (measured as the number of patients in the clinic at the time of the SP visit) the SP case (dysentery or unstable angina) and country fixed effects.

are also less comparable due to differences in standards used to assess 'correct' and 'incorrect' treatments.

That the quality of care we find in our sample in China is similar to that found among rural providers in Madhya Pradesh is surprising considering that our study area is much more developed economically. In 2011, per capita gross domestic product (GDP) was \$583 in Madhya Pradesh and \$3179 in Shaanxi, the province in China where our study was conducted. Moreover, 67% of the clinicians in the Indian study had no medical qualifications at all whereas those in our sample all had at least a 'Rural Physician' qualification.

Establishing the exact source of the low quality of health care in rural China is beyond the scope of this study. However, one possible reason might be related to the state of the rural labour market today. Because of the low incomes of villagers in our study areas, the fees that village clinicians can reasonably charge patients are low. Moreover, the government only pays a small salary (and this has only started in recent years). Given these realities, only low-skilled workers will select themselves into the rural health sector. The low human capital of village clinicians (less than half have completed any form of secondary education—Babiarz *et al.*, 2013) supports this notion. Without strong financial incentives, individuals have little reason to be attracted to the field and, once they are in it, may have little reason to work diligently.

Another possible explanation for the surprising combination of low quality of care and high per capita GDP in rural China may stem from the perverse financial incentives facing providers. In rural China, as has already been described, providers depend on large profits from drug sales to supplement their nominal base salary (Eggleston *et al.* 2008; Yip and Hsiao 2008; Yip *et al.* 2010). Currie *et al.* (2011) have shown that physician induced demand is in part responsible for the high rates of unnecessary antibiotic prescriptions in China. A more efficient system would regulate profits from drug sales so as to discourage over-prescription of unnecessary medications; indeed, recent policy pronouncements indicate that this is the direction in which China seems to be moving (Chen 2009; Yip *et al.* 2012).

As with other studies using SPs, our study faces a number of limitations. First, the cases that SPs can present to physicians are necessarily limited to diseases with no obvious physiological symptoms and for which there is no risk of invasive procedures to the SP. Thus, quality measured based on these diseases may not be representative of the broader spectrum of diseases treated by clinicians.

Second, the criteria on which the quality of care is judged in SP studies are based on a western model of patient care. This may be of particular concern given that around 15% of clinicians in western China specialize in traditional Chinese

medicine (7.7% nationally) (Li *et al.* 2013). Note, however, that—regardless of specialty—national guidelines for diagnosing and treating are consistent across providers and adhere to international standards of western medical treatment (Fang *et al.* 1998; Chinese Society of Cardiovascular Diseases of Chinese Medical Association 2007).

Third, our study focuses on actual clinical practice, not on clinicians' knowledge of best practices. There is therefore no way to determine whether the poor quality of care observed in our sample clinics is due to lack of clinician knowledge, or to a gap between clinician knowledge and practice, or to a combination of the two. Other researchers (Das and Hammer, 2007; Das *et al.* 2008) have found that the 'know-do gap' is a serious contributor to the low quality of health care in developing settings, but further research is required to determine whether this is the case in rural China specifically.

An additional limitation of our study is size and representativeness of our sample. Our sample was large enough to give informative estimates of the quality of care provided among village clinicians and to detect differences between providers; however, a larger sample would make these estimates more precise. Future studies should also draw their samples in such a way as to represent clinicians over a wider geographical area.

Conclusion

Our results raise doubts that village clinicians are currently prepared to shoulder the increased role intended by reforms. Although our results do not speak to the ability of village clinicians to effectively deliver more general public health services (vaccinations, health education, etc.), they do suggest that village clinicians may not be able to serve as effective primary first points of contact for ill patients. If village clinicians are unable to effectively diagnose and treat common diseases, reforms that aim to reduce patient loads at upper tiers by encouraging or mandating patients to first visit village clinicians are premature. This is particularly true given the emergence of non-communicable diseases as prominent public health problems in rural China. That these diseases have become prevalent at a time when the burden of many infectious diseases remains significant means that rural clinicians need to be prepared to deal with a wide array of health problems.

What is clear from this study is that rural clinicians are not yet prepared to act as the front line fighters in China's primary care system. If recent reforms that expand access to rural health care are to lead to actual gains in population health, more attention should be paid to improving the quality of care delivered by the rural health system. Poor quality care may dampen the effect of expanded access on health outcomes not only due to worse outcomes for those patients that seek care but also because it may dissuade individuals from utilizing the health system in the first place. Recent government initiatives that focus on delivery are a step in the right direction, but more needs to be known about current quality of care delivered and policy approaches to improve that care. Little is known, for example, about to what degree lack of clinician training or misaligned incentives affect the care that patients receive. China has recently made enormous strides in providing access

to health care for rural residents; however, efforts focused on improving the quality of care delivered by grassroots providers will be essential for reforms to lead to actual gains in health. Further research is needed in order to better measure the scope of this problem, and to identify specific policy interventions that can improve quality of care in these areas.

Supplementary Data

Supplementary data are available at *HEAPOL* online.

Acknowledgements

The authors are grateful to Grant Miller and Manoj Mohanan for their invaluable advice in the early planning stages of this study. They would also like to offer a special thank you to John Kennedy, who kindly let us borrow his sampling frame for this survey, and without whom this study would have been impossible. Lastly, Sean Sylvia would like to thank the Stanford Center for International Development (SCID) for their support.

Conflict of interest statement: None declared.

References

- Babiarz KS, Miller G, Yi H, Zhang L, Rozelle S. 2010. New evidence on the impact of China's New Rural Cooperative Medical Scheme and its implications for rural primary healthcare: multivariate difference-in-difference analysis. *British Medical Journal* **341**: c5617.
- Babiarz KS, Miller G, Yi H, Zhang L, Rozelle S. 2012. China's new Cooperative Medical Scheme improved finances of township health centers but not the number of patients served. *Health Affairs* **31**: 1065–74.
- Babiarz KS, Yi H, Luo R. 2013. Meeting the health-care needs of the rural elderly: the unique role of village doctors. *China & World Economy* **20**: 44–60.
- Bloom G. 2011. Building institutions for an effective health system: lessons from China's experience with rural health reform. *Social Science & Medicine* **72**: 1302–9.
- Chaudhury N, Hammer J, Kremer M, Muralidharan K, Rogers FH. 2006. Missing in action: teacher and health worker absence in developing countries. *Journal of Economic Perspectives* **20**: 91–116.
- Chen Z. 2009. Launch of the health-care reform plan in China. *The Lancet* **373**: 1322–4.
- Chinese Society of Cardiovascular Diseases of Chinese Medical Association, Editorial Board of Chinese Journal of Cardiology. 2007. Unstable angina and non ST segment elevation myocardial infarction diagnosis and treatment guidelines [in Chinese]. *Chinese Journal of Cardiology* **35**: 295–304.
- CNBS (China National Bureau of Statistics). 2011. *China National Statistical Yearbook, 2011*. Beijing, China: China State Statistical Press.
- Currie J, Lin W, Zhang W. 2011. Patient knowledge and antibiotic abuse: evidence from an audit study in China. *Journal of Health Economics* **30**: 933–49.
- Das J, Hammer J. 2007. Money for nothing: the dire straits of medical practice in Delhi, India. *Journal of Development Economics* **83**: 1–36.

- Das J, Hammer J, Leonard K. 2008. The quality of medical advice in low-income countries. *The Journal of Economic Perspectives* **22**: 93–114.
- Das J, Holla A, Das V *et al.* 2012. In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. *Health Affairs* **31**: 2774–84.
- Eggleston K. 2012. *Healthcare for 1.3 Billion: An Overview of China's Health System*. Stanford, CA: Asia Health Policy Program. Working Paper No. 28.
- Eggleston K, Ling L, Qingyue M, Lindelow M, Wagstaff A. 2008. Health service delivery in China: a literature review. *Health Economics* **17**: 149–65.
- Fang H, Wei C, Duan S *et al.* 1998. Diarrhoeal disease diagnosis and treatment plan in China [in Chinese]. *Chinese Journal of Practical Pediatrics* **6**: 381–4.
- Glassman PA, Luck J, O'Gara EM, Peabody JW. 2000. Using standardized patients to measure quality: evidence from the literature and a prospective study. *Joint Commission Journal on Quality Improvement* **26**: 644–53.
- He J, Gu D, Wu X *et al.* 2005. Major causes of death among men and women in china. *New England Journal of Medicine* **353**: 1124–34.
- Leonard KL, Masatu MC. 2010. Using the Hawthorne effect to examine the gap between a doctor's best possible practice and actual performance. *Journal of Development Economics* **93**: 226–34.
- Li X, Liu J, Huang J, Qian Y, Che L. 2013. An analysis of the current educational status and future training needs of China's rural doctors in 2011. *Postgraduate Medical Journal* **89**: 202–8.
- Liang X, Guo H, Jin C, Peng X, Zhang X. 2012. The effect of new Cooperative Medical Scheme on health outcomes and alleviating catastrophic health expenditure in China: a systematic review. *PLoS One* **7**: e40850.
- Mohanam M, Das V, Tabak D *et al.* 2011. *Standardized Patients and the Measurement of Healthcare Quality*. Online Field Guide. http://spp.staging.utoronto.ca/sites/default/files/spmanualfieldguide_012012.pdf, accessed 21 February 2014.
- Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. 2000. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *Journal of American Medical Association* **283**: 1715–22.
- Rethans J-J, Gorter S, Bokken L, Morrison L. 2007. Unannounced standardised patients in real practice: a systematic literature review. *Medical Education* **41**: 537–49.
- Reynolds L, McKee M. 2009. Factors influencing antibiotic prescribing in China: an exploratory analysis. *Health Policy* **90**: 32–6.
- Wagstaff A, Lindelow M, Jun G, Ling X, Juncheng Q. 2009. Extending health insurance to the rural population: an impact evaluation of China's New Cooperative Medical Scheme. *Journal of Health Economics* **28**: 1–19.
- Wang H, Gusmano MK, Cao Q. 2011a. An evaluation of the policy on community health organizations in China: will the priority of new healthcare reform in China be a success? *Health Policy* **99**: 37–43.
- Wang H, Zhang L, Yip W, Hsiao W. 2011b. An experiment in payment reform for doctors in rural China reduced some unnecessary care but did not lower total costs. *Health Affairs* **30**: 2427–36.
- Wang L, Wang Y, Jin S *et al.* 2008. Emergence and control of infectious diseases in China. *The Lancet* **372**: 1598–605.
- Wang X, Tao F, Xiao D *et al.* 2006. Trend and disease burden of bacillary dysentery in China (1991–2000). *Bulletin of the World Health Organization* **84**: 561–8.
- Woodward CA, McConvey GA, Neufeld V, Norman GR, Walsh A. 1985. Measurement of physician performance by standardized patients: refining techniques for undetected entry in physicians' offices. *Medical Care* **23**: 1019–27.
- Yang G, Kong L, Zhao W *et al.* 2008. Emergence of chronic non-communicable diseases in China. *The Lancet* **372**: 1697–705.
- Yang W, Lu J, Weng J *et al.* 2010. Prevalence of diabetes among men and women in China. *New England Journal of Medicine* **362**: 1090–101.
- Yip W, Hsiao WC. 2008. The Chinese health system at a crossroads. *Health Affairs* **27**: 460–8.
- Yip W, Hsiao W. 2009. China's health care reform: a tentative assessment. *China Economic Review* **20**: 613–9.
- Yip WC-M, Hsiao WC, Chen W *et al.* 2012. Early appraisal of China's huge and complex health-care reforms. *The Lancet* **379**: 833–42.
- Yip WC-M, Hsiao W, Meng Q, Chen W, Sun X. 2010. Realignment of incentives for health-care providers in China. *The Lancet* **375**: 1120–30.