

# Use of standardised patients to assess quality of healthcare in Nairobi, Kenya: a pilot, cross-sectional study with international comparisons

Benjamin Daniels,<sup>1</sup> Amy Dolinger,<sup>1</sup> Guadalupe Bedoya,<sup>1</sup> Khama Rogo,<sup>2</sup> Ana Goicoechea,<sup>3</sup> Jorge Coarasa,<sup>2</sup> Francis Wafula,<sup>2,4</sup> Njeri Mwaura,<sup>2</sup> Redemptar Kimeu,<sup>5</sup> Jishnu Das<sup>1,6</sup>

**To cite:** Daniels B, Dolinger A, Bedoya G, *et al.* Use of standardised patients to assess quality of healthcare in Nairobi, Kenya: a pilot, cross-sectional study with international comparisons. *BMJ Glob Health* 2017;**2**:e000333. doi:10.1136/bmjgh-2017-000333

Received 28 February 2017

Revised 25 April 2017

Accepted 30 April 2017



CrossMark

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjgh-2017-000333>).

<sup>1</sup>Development Economics Research Group, The World Bank, Washington, DC, USA

<sup>2</sup>Health, Nutrition and Population Global Practice, The World Bank, Washington, DC, USA

<sup>3</sup>Trade and Competitiveness Global Practice, The World Bank, Washington, DC, USA

<sup>4</sup>Institute of Healthcare Management, Strathmore University, Nairobi, Kenya

<sup>5</sup>Talana Specialists Centre, Nairobi, Kenya

<sup>6</sup>Centre for Policy Research, New Delhi, India

## Correspondence to

Dr Jishnu Das;  
jdas1@worldbank.org

## ABSTRACT

**Introduction** The quality of clinical care can be reliably measured in multiple settings using standardised patients (SPs), but this methodology has not been extensively used in Sub-Saharan Africa. This study validates the use of SPs for a variety of tracer conditions in Nairobi, Kenya, and provides new results on the quality of care in sampled primary care clinics.

**Methods** We deployed 14 SPs in private and public clinics presenting either asthma, child diarrhoea, tuberculosis or unstable angina. Case management guidelines and checklists were jointly developed with the Ministry of Health. We validated the SP method based on the ability of SPs to avoid detection or dangerous situations, without imposing a substantial time burden on providers. We also evaluated the sensitivity of quality measures to SP characteristics. We assessed quality of practice through adherence to guidelines and checklists for the entire sample, stratified by case and stratified by sector, and in comparison with previously published results from urban India, rural India and rural China.

**Results** Across 166 interactions in 42 facilities, detection rates and exposure to unsafe conditions were both zero. There were no detected outcome correlations with SP characteristics that would bias the results. Across all four conditions, 53% of SPs were correctly managed with wide variation across tracer conditions. SPs paid 76% less in public clinics, but proportions of correct management were similar to private clinics for three conditions and higher for the fourth. Kenyan outcomes compared favourably with India and China in all but the angina case.

**Conclusions** The SP method is safe and effective in the urban Kenyan setting for the assessment of clinical practice. The pilot results suggest that public providers in this setting provide similar rates of correct management to private providers at significantly lower out-of-pocket costs for patients. However, comparisons across countries are sensitive to the tracer condition considered.

## INTRODUCTION

Throughout Africa, healthcare markets are widely believed to provide low-quality care for patients, resulting in poor health outcomes

## Key questions

### What is already known about this topic?

- Standardised patients (SPs) have been safely and effectively used in urban and rural field settings in South Asia and East Asia, among other contexts, to measure quality of clinical practice for childhood and adult illnesses.
- These SP studies can reveal significant deficits in the quality of clinical practice, while avoiding the biases of other methods of clinical observation.
- Although SPs have been previously used in Sub-Saharan Africa, they have almost exclusively presented with family planning and sexually transmitted infections.

### What are the new findings?

- It is safe and effective to measure the quality of clinical primary care practice in urban Kenya using SPs presenting a range of presumptively ill adult cases with a variety of demographic backgrounds and physical characteristics.
- Correct management proportions ranged from 10% to 81% across the four tracer conditions, but public-private differences were significant in only one of the four conditions assessed.
- Quality of care, as measured by preferred case management and adherence to recommended history-taking checklists, was significantly higher in Nairobi relative to urban and rural India and rural China in three of the four cases assessed and significantly lower in the fourth.

### Recommendations for policy

- SPs can be broadly used in urban Sub-Saharan African contexts to produce accurate measurements of the quality of clinical care for a range of conditions or in conjunction with programmatic interventions.
- Quality-of-care comparisons across countries and sectors are sensitive to the specific tracer conditions under evaluation: Caution is thus required in extrapolating results from a particular study or condition to generic statements about care quality.

such as high child mortality.<sup>1</sup> Although countries like Kenya have made significant progress as far as affordable access is concerned,<sup>2</sup> there are few studies of actual quality.<sup>1</sup> Quality is difficult to define and measure, and traditional measures such as the availability of equipment and medicines are only weakly correlated with clinical performance.<sup>3–5</sup> To address these difficulties, in recent years alternative survey measures of quality have been developed to directly measure clinical performance and have been validated in field-based settings with large sample sizes and a variety of tracer conditions.<sup>6–8</sup>

One such method that is now gaining acceptance as close to a gold standard for the measurement of clinical practice is the use of ‘standardised patients’ (SPs) — people recruited from the local community and extensively trained to present exactly the same clinical condition to multiple healthcare providers in a study sample. Since case presentations are fully standardised and predetermined, the SP methodology allows for accurate quality comparisons across different types of providers and contexts and allows researchers to assess the accuracy of the condition-specific diagnosis and treatment, including the extent of unnecessary or inappropriate procedures and medications.<sup>6–7</sup> Further, as healthcare providers are blinded from SP conditions and assignments, their behaviour approximates the treatment of ‘real’ patients and is less prone to Hawthorne effects, whereby providers can alter their behaviour when they know they are under observation.<sup>9</sup> Finally, the SP methodology is less susceptible to recall errors among patients in exit interviews<sup>10</sup> and incomplete medical records or missing patient charts in resource-poor contexts.<sup>11–12</sup>

This contrasts with studies based on actual patients, which can confound true measurements of quality with differences in patient characteristics or large Hawthorne effects. It also contrasts with measures of provider knowledge, which have been shown to provide an upper bound of actual performance in the clinic.<sup>13</sup> In a number of recent studies, SPs have been used to estimate quality of care, explore practice quality variation and evaluate health interventions.<sup>14–20</sup>

In Sub-Saharan Africa, SP studies have been conducted among pharmacies and drug sellers, with a particular focus on family planning and sexually transmitted infections.<sup>21–25</sup> However, there has been no validation of this approach for a wider set of tracer conditions in public and private primary care settings, where a large fraction of care is provided.<sup>2</sup> This paper presents (1) validation results from the SP methodology for primary outpatient care, (2) quality-of-care results for four tracer conditions presented by SPs in Kenya’s capital city of Nairobi and (3) comparisons with similar cases presented in India and China.

In the Kenyan context, these results are of particular interest as user fees have been abolished in the public sector, and the private sector has grown to now account for 50% of all primary care facilities in the country and 50% of paediatric outpatient care.<sup>26–27</sup> Although recent

studies using vignettes show high levels of medical knowledge among Kenyan healthcare providers, little is known about the actual quality of care and its cost in the public and private sectors.<sup>28–29</sup> As SPs follow precisely the same route as ‘normal’ patients, the study is uniquely positioned to shed light both on the quality of care that patients receive in different sectors as well as ancillary information on wait times and out-of-pocket payments to providers.

In the remainder of this paper, we detail the study design and methodology and describe the validation results. We then compare Nairobi’s public and private sectors, examining differences both in terms of patient experience, waiting times and expense, and in terms of the appropriateness and accuracy of case management. Finally, although sampling strategies were different, we also present quality-of-care comparisons with similar studies in rural China and rural and urban India.

## METHODS

### Study design and participants

#### Case development and presentation

We worked with the Advisory Committee of Kenyan doctors and nurses to develop tracer conditions, construct medically relevant checklists and assign correct case management protocols, with review from medical professionals. Generally healthy individuals were selected for case presentations so that physical examination would not lead providers to alternative diagnoses, and SP responses to relevant history questions were scripted such that appropriate questioning and examinations by a healthcare provider should have led to correct diagnoses for all patients.

The case presentations and correct treatments are detailed in [table 1](#). The four cases were: (1) unstable angina, in which a 45-year-old man complains of being awoken suddenly by severe chest pain in the morning; (2) diarrhoea, in which the mother of an 18-month-old complains that her child has diarrhoea and requests medicines; (3) asthma, in which the patient complains of severe shortness of breath with specific triggers and, if asked, a history of similar attacks; and (4) tuberculosis (TB), in which the patient first complains of persistent cough, later revealing the duration and admitting to night sweats and weight loss.

The choice of these tracer conditions reflects their relevance in the Kenyan setting along with the availability of established medical protocols for triage, management and treatment. For example, among children aged 5 years or under, diarrhoea is the third most common cause of mortality and morbidity in Kenya.<sup>30</sup> In Nairobi, rates of specific symptoms associated with asthma in children were as high as 25% — twice as frequent as in surrounding rural areas.<sup>31</sup> TB is the fourth most lethal infectious disease in Kenya, with an incidence of 233 cases per 100 000 population and a high mortality rate of 20 deaths per 100 000 population; in addition, one

**Table 1** Standardised patient case descriptions

	Case description	Presentation	Preferred management
Asthma	A 25-year-old man or 24-year-old woman has had breathing episodes over the last year with increasing severity.	The patient enters the clinic saying that he or she had an especially difficult time with breathing the prior night.	Should be treated with an inhaler or bronchodilator such as salbutamol, cetirizine or prednisolone.
Child diarrhoea	A 1.5-year-old child has developed watery diarrhoea.	The 28-year-old mother enters the clinic alone saying she is on her way home from work and her child is at home suffering from diarrhoea.	Should be given or advised to take oral rehydration salts.
Tuberculosis suspect	A 35-year-old man/30-year-old woman has had a productive cough and fever for 2–3 weeks.	The patient enters the clinic and says that he/she has been coughing for a while.	Should be asked to undergo a sputum test.
Unstable angina	A 40-year-old man has experienced a crushing pain in his chest in the morning as well as several times previously.	The patient enters the clinic saying he had pain in his chest in the morning.	Should be referred to a higher level facility, ordered an ECG or given aspirin.

in three patients with TB whose HIV status is known is HIV-positive, posing an especially challenging comorbidity profile.<sup>32</sup> Finally, ischaemic heart disease is now the tenth leading cause of death in Kenya, just after malaria and TB, and may become an increasingly prevalent condition as the population ages.<sup>33</sup>

For asthma, diarrhoea and TB, the Kenyan Ministry of Health had already developed and disseminated its own diagnostic and treatment guidelines.<sup>34</sup> For unstable angina, guidelines had not yet been established in Kenya, and the Advisory Committee consulted the European Society of Cardiology and the American College of Cardiology and American Heart Association Joint Guidelines. These SP cases had previously been developed and deployed in various studies in India and China, therefore allowing for comparability with this study. In all instances, we use a lenient definition of ‘correct management’, where a condition is judged to be correctly managed as long as the necessary management was undertaken — *regardless* of whether other unnecessary medicines such as antibiotics were given as well. As in samples from other settings, adopting a more strict definition of correct management that accepts *only* correct management without any unnecessary medicines produces a proportion of ‘fully correct management’ near zero.

Together with the Advisory Committee, SP scripts were redeveloped and contextualised for the Kenyan setting, including adjustments to both the patient’s personal backgrounds and the vernacular manner of speaking. Standardised Swahili translations of scripts described symptoms and responses as patients themselves would express them. For example, when asked to describe the chest pain, an angina SP would say, *“It felt a little heavy, like something was sitting on my chest.”* The SPs themselves recommended revisions as broad as the occupation of the patient and as focused as the patient’s physical presentation, such as that tuberculosis SPs should wear loose

clothing to suggest weight loss. These recommendations were agreed on as part of the final phase of development, in which SPs presented their cases to Advisory Committee members and doctors for review.

#### SP cohort selection and training

From an initial list of 50 potential candidates recruited with the support of a local survey firm, 26 candidates were invited to attend 150 hours of training over 4 weeks after an extensive screening based on age, gender, education, work experience and health history. During training, SPs memorised medical details, symptoms and risk factors for each case with a qualified doctor and nurse, building their understanding of the background story and presentation. SPs were also extensively coached during training to avoid potentially dangerous situations that could arise during fieldwork. While they were allowed to undergo all basic vitals, triage and physical examinations offered by providers, SPs were coached to avoid taking medication or undergoing blood tests, injections, X-rays and other invasive or intrusive procedures during their encounters. Following the training, SPs presented the cases to doctors and nurses who were part of the study, and based on their feedback 14 SP candidates were selected for field implementation (see the online supplementary technical appendix for further details; the complete methodology and materials are available from the authors by request).

#### Study location and sample description

Health facilities were approached in a convenience sample designed to include low-income, middle-income and high-income neighbourhoods in various Nairobi neighbourhoods. Of 49 health facilities approached, 46 agreed to participate and SP interactions were completed in 42 facilities, with 4 randomly selected facilities held as

**Table 2** Facility summary statistics

	Public	Private	Faith-Based or Social Franchise Organisations (FBO/ SFO)	All facilities
All sampled facilities	14	18	10	42
Average staff size	4.0	2.6	2.3	3.0
Average daily patients	113	22	33	56
Level 2 facilities	4 (21%)	10 (53%)	5 (26%)	19
Level 3 facilities	10 (43%)	8 (35%)	5 (22%)	23
Facilities with pharmacy	14 (41%)	11 (32%)	9 (26%)	34
Facilities with laboratory	13 (38%)	12 (35%)	9 (26%)	34
Primary provider clinical officer	11 (39%)	12 (43%)	5 (18%)	28
Primary provider nurse	0 (0%)	1 (25%)	3 (75%)	4
Primary provider medical officer	3 (43%)	3 (43%)	1 (14%)	7
<b>All Nairobi facilities</b>	<b>528</b>	<b>138</b>	<b>192</b>	<b>858</b>
Nairobi level 2 facilities	431 (63%)	95 (14%)	161 (23%)	687
Nairobi level 3 facilities	80 (57%)	32 (23%)	28 (20%)	140
Nairobi level 4 facilities	14 (58%)	7 (29%)	3 (13%)	24
Nairobi level 5 facilities	3 (75%)	1 (25%)	0 (0%)	4
Nairobi level 6 facilities	0 (0%)	3 (100%)	0 (0%)	3

Percentages may not sum to 100% in each row, due to rounding.

reserves in case a sampled facility was closed or otherwise inaccessible. Of 168 potential interactions in these 42 facilities, we completed 166 for a completion rate of 98.5%. Of those facilities, 14 were public and 28 were privately owned and operated, with the private sample including 1 community clinic, 5 clinics (of 28) operated by faith-based organisations and 4 clinics operated by social franchise operations (table 2). In analysis, facilities are classified only as 'public' or 'private' due to the small sample size.

Kenyan health facilities are also classified by 'level', reflecting their size and technical capacity.<sup>35</sup> Levels 2 and 3 include smaller facilities offering basic primary care and preventive services and are typically staffed by clinical officers or trained nurse practitioners. Levels 4 and 5 cover facilities offering integrated care and inpatient services with increased specialisation, and will typically also be staffed by medical officers (the equivalent of an MBBS in the UK). We excluded level 4 and higher facilities from sampling, as the integrated service offered there could introduce complications to the SP interactions, such as the unstable angina case being immediately triaged into inpatient care.

In comparison with the universe of healthcare facilities present in Nairobi (table 2) according to the government roster, our final facility mix undersampled publicly owned facilities relative to privately owned facilities and oversampled level 3 facilities relative to level 2 facilities. We do not apply any weighting to our outcomes, since our sample is non-random and not intended to be representative of

the city's provider mix, but we report outcomes for public facilities separately in our main results and for private not-for-profit facilities separately in online supplementary appendix table A2.

Facilities were provided a consent form with details on the study, which included both a description of the process ('In the following 6 months, you will be visited by someone who has been trained by us to act as a patient. These patients are called 'standardized patients' and this approach has been used to assess quality of medical care. You will not know exactly when this standardized patient will visit you, but please note the date and time if/when you think you saw this standardized patient. No later than one month after this visit, our research team will contact you to find out if/when you saw our standardized patient'.) and informed consent ('The standardized patient who visits you for a consultation will pay your usual consultation fees. So, you will not suffer any economic loss due to participation in this study. While you will not directly benefit from the research, we hope that the information from this study will help us understand how the standardized patient approach can be used to evaluate quality of primary care in Kenya.'). There were no direct refusals to participate; however, 3 of the 49 facilities initially approached requested additional time before signing the consent form to obtain further authorisation from a main office that oversees the facility, and they were dropped from the sample.

Between 3 April 2014 and 9 June 2014, SPs were randomly allocated to different facilities with a visit

schedule that was optimised over travel and wait times. For instance, multiple SPs could visit a busy facility the same day of the week, but low-load facilities were capped at two SP interactions per day.

### Ethical clearance, consent and role of the funding source

Ethical clearance was granted by the review board at African Medical and Research Foundation (AMREF), Reference AMREF-ESRC P94/2013, with additional clearances from the Ministry of Health, Government of Kenya and each county in which the facilities were located. Every sampled facility was individually approached and written consent was obtained for the study. This paper used data from the Kenya Patient Safety Impact Evaluation project — a joint undertaking between The World Bank Group and the Kenyan Ministry of Health. Funding was provided by The World Bank Group through the Strategic Impact Evaluation Fund, the Impact Evaluation to Development Impact fund, the Trade and Competitiveness Impact programme, the Knowledge for Change Program and the Development Impact Evaluation health programme. The funders of the study had no role in study design, data collection, data analysis, data interpretation or writing of the report. The first author and corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

### Statistical analysis

We used *t*-tests and multiple logistic regression to compare the quality of care between government and privately operated facilities. Additional specifications using the SPs' age, sex, body mass index and blood pressure are calculated to provide validation results. We use the Wilson interval without continuity correction for all binary variable CIs. We analysed all data using Stata V.13.

## RESULTS

We divide the results into two sections: validation and pilot results. The validation exercise assessed whether SPs were able to (1) avoid potentially dangerous situations, (2) avoid suspicion that they are not true patients and (3) successfully complete interactions with providers without imposing a substantial time burden. It also assessed the sensitivity of provider behaviours to SP characteristics. The pilot results first describe the quality of clinical practice using outcome metrics for each SP case that capture both necessary and unnecessary care. Specifically, we assess desirable outcomes such as condition-specific adherence to a diagnostic checklist, preferred case management and undesirable outcomes such as the inappropriate use of medicines and antibiotics. We then assess variation across public and private sector clinics and compare the overall findings with those obtained in prior studies conducted in India and China.

## Validation

### Adverse events and detection of SPs

There were no adverse events reported by SPs during the fieldwork. The most common invasive treatment that SPs had to avoid was being given an injection, which was offered in 38 of 126 non-diarrhoea interactions (30%; 95% CI 23 to 39). To assess detection rates, each provider was administered a structured questionnaire within 2 weeks of the completion of fieldwork asking whether an SP had visited the clinic, and, if so, the characteristics of the SP. Providers claimed to have detected an SP in nine instances, but on further elicitation of the characteristics and presentations of the suspected cases, there was no match between the SPs we had sent and the presentation of the patient that the provider suspected to be an SP. We therefore conclude that detection rate for the SPs was zero in this study.

### Time burden for providers

The burden on providers was the time spent with SPs, which averaged 7.2 min per interaction (95% CI 6.1 to 8.2) for a total of 28.8 min per facility across the four cases. In all cases, SPs recorded the duration by checking watches or cellphones at the beginning and end of the interaction, for which they were specifically trained. In private facilities, the provider was fully compensated for this time, as the SP paid whatever fee was charged, according to regular clinic procedures. In public facilities, consultations were free and therefore the time spent is a true additional burden on the provider. In these facilities, average consultation time was 4.2 min per interaction (95% CI 2.6 to 5.8) for a total time burden of 16.8 min per facility. This time burden could be substantial if the delay affected seriously ill patients; however, in those circumstances, SPs were trained to allow others to bypass the queue and see the doctor immediately, as is typically customary; this situation did not arise during fieldwork.

### Impact of SP characteristics on interactions

A primary concern for the SP validation is that individuals presenting the cases are not actually ill. It is therefore important to ensure that case presentations are convincing, in the sense that continued questioning and examination of the SP will not lead the provider to conclude that the patient is healthy. We therefore examined associations between the likelihood of correct case management and checklist adherence in a multiple regression model with dummies for each case and SP. A 100% completion rate for the essential history question checklist (detailed in online supplementary appendix table A1) for each case versus no questions asked is associated with a 27 percentage point (pp) increase in the likelihood of correct treatment ( $p=0.123$ ; 95% CI -8 to 62), an 8 pp increase in the likelihood of giving any medication ( $p=0.662$ ; 95% CI -28 to 43) and a 32 pp increase in the likelihood of any verbal diagnosis ( $p=0.067$ ; 95% CI -2 to 66). The likelihood of completing at least one of these three behaviours increases by 24 pp

over the interval ( $p=0.041$ ; 95% CI 1 to 46). Although many of the verbal diagnoses given are incorrect (in particular, the angina case and the TB case were often verbally diagnosed with pneumonia or chest infections; see online supplementary appendix table A4 for the full list), the fact that the provider is not led to believe that the patient is presenting false symptoms when asking case-appropriate questions is a strong validation of the methodology in this context.

Since SPs were randomised across facilities, we can also assess whether SP characteristics affected clinical practice. To do so, we enter the body mass index, the systolic blood pressure, gender and age of the SP directly into regressions for various outcomes, controlling for case-specific dummies (online supplementary appendix table A3). SPs who were older and had a higher systolic blood pressure (although none was hypertensive) had higher consultation times and better checklist completion rates, but these characteristics did not affect the likelihood of correct treatment or the type and number of medications received. These differences are of interest in their own right, but do not present challenges to the use of the SP methodology or the accuracy of case depiction beyond documenting variation in care due to naturally occurring variation in the patient population. They do imply, however, that as SP characteristics need to be adjusted for when comparing across providers and ideally, studies should ensure that different providers are visited by similar SPs.

## Pilot results on clinical practice

### Aggregate results

The typical SP interaction lasted 7.2 min (95% CI 6.1 to 8.2), with providers completing 38% of the condition-specific checklist of history questions and examinations (95% CI 34% to 42%) and correctly managing 88 of 166 cases (53%; 95% CI 45 to 60) (see table 3 for full details). SPs were not instructed to request a diagnosis from the provider, and a verbal or written diagnosis was given to the patient in 53 of 166 cases (32%; 95% CI 25 to 39). These varied widely and are reported in online supplementary appendix table A4. Medication was offered in 118 of 166 cases (71%; 95% CI 64 to 77), with antibiotics, unnecessary in every case, offered in 82 of 166 interactions (49%; 95% CI 42 to 57). Unnecessary steroids, however, were used in only 3 of 166 interactions (2%; 95% CI 1 to 5). While we later present stratified results on differences between public and private providers, we note that the *average* wait time was 49 min (95% CI 40.6 to 57.7) and that the *average* interaction costs the SP KSh425 (equivalent to US\$5.10 at the exchange rate during the study) in out-of-pocket expenses.

Clinical practice varied substantially across conditions in terms of correct case management. For instance, the asthma case was treated with appropriate medication in 34 of 42 interactions (81%; 95% CI 67 to 90) compared with only 4 of 42 instances for angina (10%; 95% CI 4 to

22), with child diarrhoea and TB falling in between these two extremes. Notably, the use of medicines and (unnecessary) antibiotics was high across all cases, ranging from a low of 32% in child diarrhoea to 60% for the angina SP.

### Public-private variation

We observed few statistically significant differences in either preferred case management or the use of medicines across public and private sector clinics (table 4; figure 1), although public clinics were significantly more likely to correctly order a sputum test for the TB case. It is noteworthy that public clinics were just as likely to give unnecessary antibiotics as private clinics across all cases. In contrast, there were large and significant differences in patient experience, with private providers spending 4.5 more minutes with SPs when compared with public clinics ( $p<0.05$ ) and completing 20pp more of the required checklist (95% CI 16 to 24). Public-sector SPs waited an additional 68.2 min ( $p<0.001$ , 95% CI 53.1 to 83.3) and paid KSh426 less than in private clinics. These estimates are robust to alternate functional form assumptions, such as the logistic, as shown in online supplementary appendix figure A3.

### Cross-country comparison

We compare results from the Nairobi pilot against similar SP presentations from three prior studies: rural India (the state of Madhya Pradesh), urban India (Delhi) and rural China (Shaanxi province), shown in figure 2 and online supplementary appendix table A5.<sup>7 17 18</sup> For three conditions — asthma, TB and diarrhoea — the performance of the providers in the Nairobi sample is strikingly better. For example, in the child diarrhoea case, no provider in the sample from China gave oral rehydration salts (ORS) compared with 18% of Indian and 73% of Nairobi providers, and Nairobi providers gave fewer antibiotics than observed in the other settings. Providers in the Nairobi sample were also significantly more likely than those in the China sample to provide the correct treatment, but dispensed antibiotics at about the same rate. Similarly, Nairobi providers were significantly more likely to use microbiological testing for TB than those observed in Delhi. These trends in appropriate management are consistent with the significantly greater time that Nairobi providers spent with patients, which is typically twice as high as the other studies.

This advantage does not extend to all cases: In the Chinese and Indian samples, providers managed the angina case correctly for 63% and 41% of SPs, respectively, compared with only 10% of Nairobi providers. Further, despite wide variation in the proportion of correctly managed cases, there is little to no difference in the use of unnecessary antibiotics. In asthma, for instance, Nairobi providers were 23% more likely to correctly manage the case, but were equally likely to use unnecessary antibiotics. Finally, purchasing power-adjusted comparisons show that out-of-pocket expenses were twice as high in Nairobi relative to China and two to six times higher than in India for

**Table 3** Primary outcomes for standardised patient (SP) cases

	All SP Cases	Asthma	Child diarrhoea	Tuberculosis	Unstable angina
Preferred case management	0.53 (0.45 to 0.61)	0.81 (0.69 to 0.93)	0.73 (0.59 to 0.86)	0.50 (0.35 to 0.65)	0.10 (0.01 to 0.18)
Time waiting in clinic (min)	49.12 (40.57 to 57.67)	46.26 (31.93 to 60.59)	51.08 (33.41 to 68.74)	45.64 (28.69 to 62.58)	53.59 (34.55 to 72.63)
Time with provider (min)	7.17 (6.12 to 8.22)	9.26 (6.79 to 11.73)	4.45 (3.53 to 5.37)	6.72 (5.1 to 8.33)	8.12 (5.58 to 10.67)
Proportion history questions	0.38 (0.34 to 0.42)	0.34 (0.28 to 0.4)	0.47 (0.39 to 0.55)	0.42 (0.34 to 0.5)	0.29 (0.22 to 0.36)
Total price (KSh)	425.10 (350.77 to 499.43)	507.26 (350.81 to 663.72)	246.15 (142.49 to 349.82)	469.76 (300.14 to 639.39)	464.43 (326.38 to 602.47)
Referred (non-diarrhoea)	0.05 (0.01 to 0.09)	0.00 –		0.12 (0.02 to 0.22)	0.02 (–0.02 to 0.07)
Any medication	0.71 (0.64 to 0.78)	0.86 (0.75 to 0.96)	0.52 (0.37 to 0.68)	0.57 (0.42 to 0.72)	0.88 (0.78 to 0.98)
Number of medications	1.77 (1.55 to 1.99)	2.07 (1.68 to 2.47)	1.58 (1.01 to 2.14)	1.48 (1.05 to 1.91)	1.95 (1.63 to 2.28)
Antibiotics	0.49 (0.42 to 0.57)	0.50 (0.35 to 0.65)	0.32 (0.18 to 0.47)	0.55 (0.4 to 0.7)	0.60 (0.45 to 0.74)
Steroids	0.02 (0 to 0.04)	0.02 (–0.02 to 0.07)	0.00 –	0.00 –	0.05 (–0.02 to 0.11)
Number of observations	166	42	40	42	42

Data are mean (95% CI). Preferred case management is defined as the use of an inhaler or bronchodilator for the asthma case, oral rehydration salts for the child diarrhoea case, a sputum test for the tuberculosis suspect, and a referral, ECG or aspirin for the unstable angina case. Time waiting in clinic is the total time in the clinic not with the provider. Time with the provider is the duration of the consultation. Proportion history questions is the mean proportion of essential history questions asked during consultation (see online supplementary appendix). Price includes all consultation fees paid on the spot, as well as the prices of all testing and medications recommended by the provider. Referral is any case where the provider recommended the patient to see another provider.

the same cases, even though the Nairobi sample includes both public and private providers, with free consultations in the public sector. Waiting times of 95 min in public and 27 min in private clinics were also much higher in Nairobi, compared with 9.5 min for the average wait time in an Indian private clinic. Finally, similar to the 32% in Nairobi, only 32% of cases in rural Indian and 13% in urban India received any verbal diagnosis, which is much lower than the 80% of cases that received a diagnosis in the sample from China.

## DISCUSSION

This study assessed the feasibility of the SP methodology in Nairobi, Kenya, in diverse urban clinics that spanned multiple ownership structures and different levels of specialisation. We are able to report that the SP methodology was successfully validated without any adverse events or survey

detections, and was able to elicit context-appropriate reactions from clinical staff under field implementation conditions.

SPs in our study received a broad range of care outcomes: 53% of SPs were correctly managed, but this ranged from 10% of angina SPs to 81% of asthma SPs. Surprisingly, the use of unnecessary medicines and antibiotics was similarly high across all cases. International comparisons suggested little consistency across conditions; for instance, Indian and Chinese providers managed angina better but did not provide ORS for children with diarrhoea. Finally, in contrast to the substantial differences in correct case management between the public and private sector reported from India,<sup>18</sup> there was little difference in Nairobi.

These results are the first to look inside the ‘black box’ of patient experience in primary care settings in Kenya. They are of particular interest given the progressive elimination

**Table 4** Primary outcomes for standardised patient cases by sector

	Overall mean	Public mean	Private mean	Private–public difference (t-test)	Private–public difference (OR)
<b>Preferred management</b>					
Asthma: inhaler or bronchodilator	0.81	0.79	0.82	0.04	1.25
	(0.06)	(0.11)	(0.07)	(0.13)	(1.03)
Child diarrhoea: ORS	0.73	0.62	0.78	0.16	2.19
	(0.07)	(0.14)	(0.08)	(0.15)	(1.61)
Unstable angina: referral, ECG or aspirin	0.10	0.14	0.07	−0.07	0.46
	(0.05)	(0.10)	(0.05)	(0.10)	(0.49)
Tuberculosis: sputum test	0.50	0.79	0.36	−0.43***	0.15**
	(0.08)	(0.11)	(0.09)	(0.15)	(0.12)
<b>Basic statistics</b>					
Time waiting in clinic (min)	49.12	94.70	26.53	−68.19***	–
	(4.35)	(9.16)	(2.85)	(7.65)	
Time with provider (min)	7.17	4.21	8.64	4.45***	–
	(0.53)	(0.82)	(0.64)	(1.05)	
Checklist	0.38	0.25	0.44	0.20***	–
	(0.02)	(0.03)	(0.02)	(0.04)	
Total price (KSh)	425.10	141.54	563.05	426.66***	–
	(37.84)	(31.42)	(49.13)	(72.41)	
Referred (non-diarrhoea)	0.05	0.02	0.06	0.04	2.72
	(0.02)	(0.02)	(0.03)	(0.04)	(3.09)
<b>Medications</b>					
Any medication	0.71	0.65	0.74	0.09	1.60
	(0.04)	(0.06)	(0.04)	(0.07)	(0.62)
Number of medications	1.77	1.60	1.86	0.26	–
	(0.11)	(0.19)	(0.14)	(0.24)	
Antibiotics	0.49	0.47	0.50	0.03	1.15
	(0.04)	(0.07)	(0.05)	(0.08)	(0.39)
Steroids	0.02	0.04	0.01	−0.03	0.23
	(0.01)	(0.03)	(0.01)	(0.02)	(0.29)
Number of observations	166	55	111		

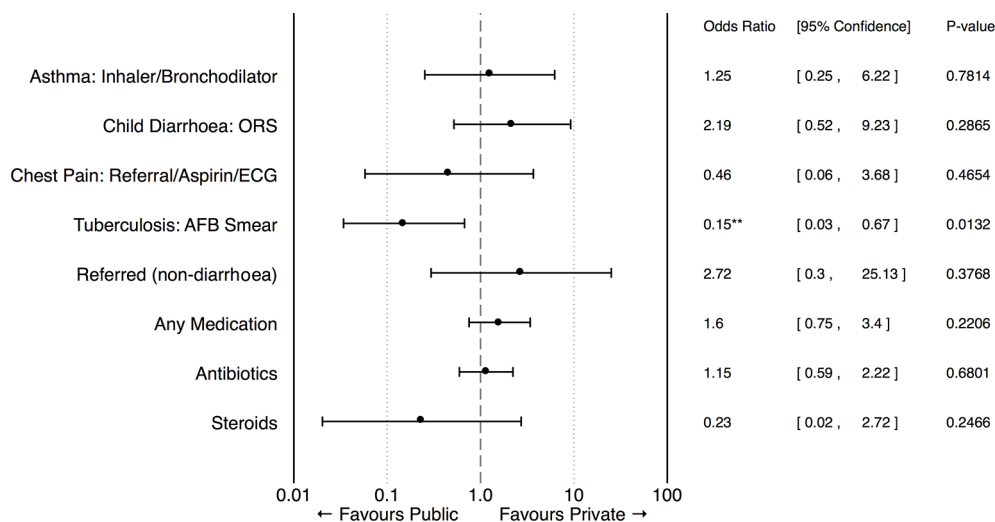
Data are means (SEs) or estimated coefficients (SEs). Preferred case management is defined as the use of an inhaler or bronchodilator for the asthma case, ORS for the child diarrhoea case, a sputum test for the tuberculosis suspect, and a referral, ECG or aspirin for the unstable angina case. Control indicator variables are included for each case. Price includes all testing and medications recommended by the provider. Asterisks indicate significance levels as follows: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ . ORS, oral rehydration salts.

of user fees and the widespread use of the private sector in the country.<sup>36–39</sup> The SPs in our study paid 76% less in the public sector, but appropriate case management was similar for three of four tracer conditions and was significantly *better* for TB, which carries a high public health risk. Although limited by sample size, an exploratory comparison between private for-profit providers and private not-for-profit providers found no significant differences

between the two in terms of case management indicators, waiting time, consultation time or cost to the patient.

Our data are consistent with the hypothesis that when patients visit the private sector, they are paying a premium for reduced waiting time, greater consultation time and higher adherence to checklists. If these results hold more widely in representative samples, they would confirm that the public sector in Kenya offers a viable alternative to the private sector with lower out-of-pocket





**Figure 1** Effect of sector on primary standardised patient outcomes. Adjusted ORs are illustrated for 55 interactions with public-sector providers versus 111 private-sector providers. Regressions are controlled for case. ORS, oral rehydration salts. AFB, acid-fast bacilli.

financial outlays for patients. Nevertheless, one issue of grave concern in both public and private clinics is the high use of unnecessary medicines, particularly antibiotics. This increases out-of-pocket payments and in the case of antibiotics carries significant public health risks due to rising antibiotic resistance.<sup>40 41</sup>

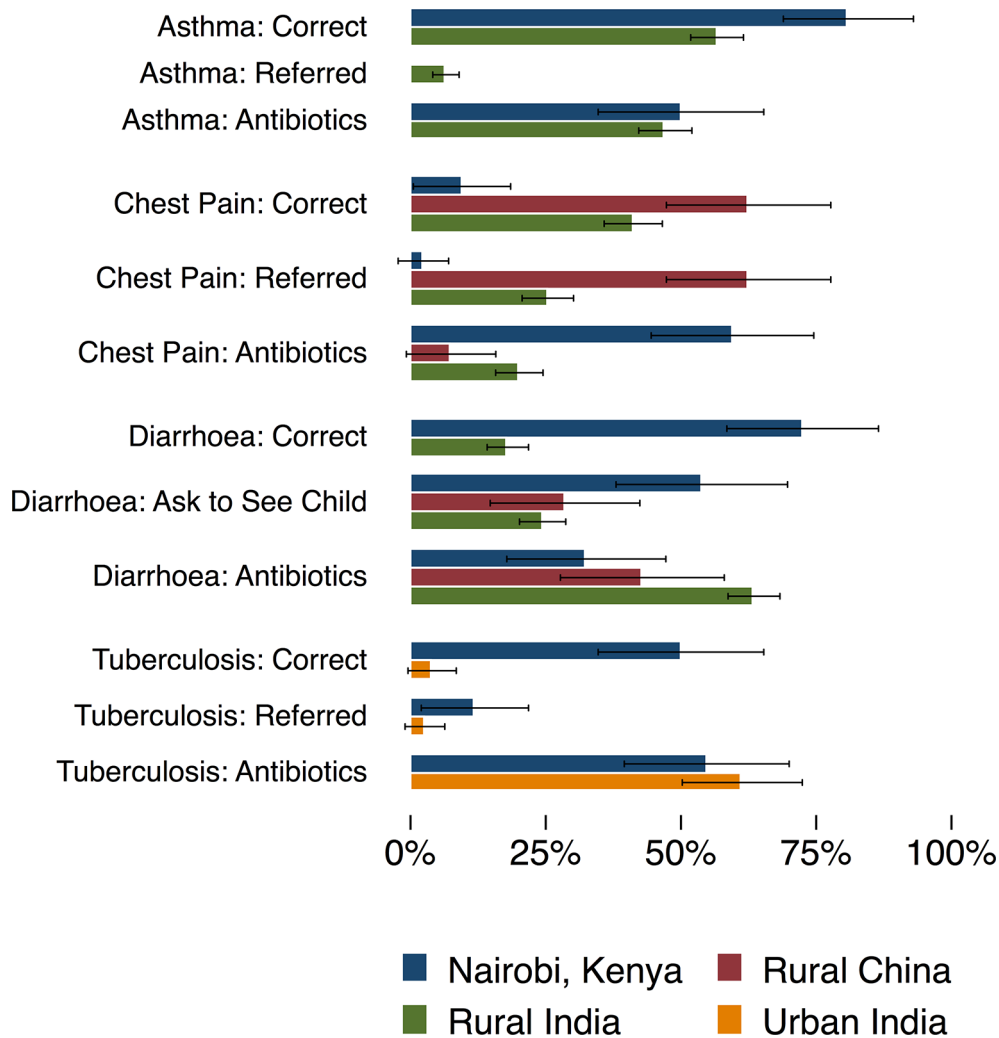
Like other SP studies, there are several practical limitations in analysis and interpretation. The study accurately describes clinical practice for anonymous initial interactions with SPs; however, overall quality of care also depends on whether patients correctly act on the providers' advice and medical prescriptions, which were not studied. It is also possible that clinical practice may vary when individuals or families frequently visit a particular provider and are well-known to them. Finally, avoidance behaviour on the part of the SPs yields another source of difference with real patients. One such example is the use of injections, which the SPs in our study would refuse, but which was offered in 38 of 126 non-diarrhoea interactions. The behaviour captured here thus represents care for a 'one-off' interaction with a new patient, which may not generalise to the patient population.

Additionally, as the sampling frames for the studies in Kenya, China and India were different, the cross-country comparison is indicative rather than definitive. Each sample was designed to cover a specific sampling frame in the target setting, and the results cannot be extrapolated to countrywide differences, which would require nationally representative samples. Even with representative samples, differences in provider caseload, disease prevalence, market structure and regulation, and patient choice within and across contexts can confound international comparisons. Despite this lack of general comparability, specific individual findings such as similar rates of inappropriate treatment across settings remain of policy interest.

Looking forward, our results also highlight the difficulties that arise when comparing quality across different settings. First, performance rankings are sensitive to the tracer conditions that are chosen: Ranked by correct management on the angina case, providers from Kenya underperformed; ranked by correct management on other conditions, they outperformed their counterparts in India and China. Second, performance rankings will depend on how much the use of unnecessary medicines and antibiotics is penalised. Despite substantial variation in the proportion of correctly managed cases, providers in all settings and all cases were equally likely to prescribe unnecessary medicines and antibiotics. Whether providers (wrongly) believe that antibiotics are efficacious for the case presented or they believe that patients are suffering from comorbidities that require antibiotics cannot be determined from these data but are urgent for further investigation. Third, the 'price' of care varies widely across countries. Comparisons will then depend on whether we compare 'average' providers or equally expensive providers. Societies with different preferences over affordability and quality may lead to very different rankings depending on these choices.

Combining these multiple dimensions of care into single indices will invariably require researchers to assign 'weights' to different conditions and to the use of unnecessary medicines. For instance, one could weigh correct management higher for infectious diseases to account for the further risk of contagion. In this regard, it is of interest that the performance gap of Kenyan providers was highest in TB and diarrhoea, whereas Indian and Chinese providers outperformed the Kenyan providers in the angina case.

Alternatively, these data can be used to better understand the institutional factors that mediate differences across settings. For instance, one hypothesis advanced



**Figure 2** Primary outcomes for standardised patient cases by setting. Overall proportions of preferred management (defined in table 1), referral to another location and use of unnecessary antibiotics are presented here from the four studies included in the cross-country analysis, with 95% CIs. The proportion of referrals for asthma in Kenya was zero and the proportion of correct oral rehydration salts use for the diarrhoea case in China was zero; all other missing bars indicate the condition was not studied in that context. The numbers of observations are as follows: asthma (rural India): 397; asthma (Kenya): 42; chest pain (China): 40; chest pain (rural India): 323; chest pain (Kenya): 42; child diarrhoea (China): 42; child diarrhoea (rural India): 389; child diarrhoea (Kenya): 40; tuberculosis (TB) (urban India): 75; tuberculosis (Kenya): 42.

by the Advisory Committee for the poor performance in the angina case was that the Kenyan government had not disseminated diagnostic and care guidelines for angina; the treatments the angina SPs received were consistent with a frequent misdiagnosis of pneumonia. However, this would require quality of care to respond to such guidelines, a correlation that has been difficult to show for patient safety in Kenya or quality of care in other settings.<sup>26</sup>

In either case, the validation of the SP methodology in Kenya opens up the possibility of its wider use in Sub-Saharan African countries that can lead to a more nuanced understanding of clinical care within and across a wide variety of international settings.

**Acknowledgements** This work is part of the Kenya Patient Safety Impact Evaluation Project, a joint project of the Development Economics Research Group and The World Bank Health in Africa initiative, in collaboration with the Kenyan

Ministry of Health. The KePSIE team includes Guadalupe Bedoya, Jorge Coarasa, Jishnu Das, Amy Dolinger, Ana Goicoechea, Njeri Mwaura and Khama Rogo, supported by Benjamin Daniels, Chex Yu, Seungmin Lee and Frank Wafula from The World Bank Group. The team works together with the Kenya Ministry of Health, the regulatory boards and councils. We thank the members of the Advisory Committee for standardised patients: Dr Redemptar Kimeu, Nairobi Hospital; Professor Ruth Nduati, University of Nairobi; and Drs Fatmah Abdallah, Sarah Chuchu, Rahab Maina, Izaq Odongo, Rachel Nyamai, Charles Kandie and John Kabanya from Kenya's Ministry of Health. We also thank Ada Kwan and Bernard Olayo from The World Bank for their research support. We are also grateful to The World Bank supervisors Pamela Kuya, Purity Kimuru, Leah Adero and Salome Omondi, and IPSOS staff, especially Samuel Muthoka, Oscar Mutinda and Eugene Wafula, for their support in the pilot study.

**Contributors** BD cleaned and analysed the data, and drafted and revised the paper. AD served as field coordinator throughout the project and analysed data. GB, KR, AG, JC and NM contributed to study design and field coordination. FW contributed to field coordination. RK contributed to study design and served as the lead SP trainer. JD contributed to study design, analysis, writing and revision, and is a guarantor.

**Competing interests** RK declares service as a paid consultant to the survey firm during the project.

**Ethics approval** African Medical and Research Foundation (AMREF), Reference AMREF-ESRC P94/2013.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** All documentation and data will be made available on publication by direct request to the corresponding author.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

- Berendes S, Heywood P, Oliver S, *et al*. Quality of private and public ambulatory health care in low and middle income countries: systematic review of comparative studies. *PLoS Med* 2011;8:e1000433.
- Republic of Kenya Ministry of Health. Household Health Expenditure and Utilisation Survey Report. 2003.
- Rethans JJ, Sturmans F, Drop R, *et al*. Does competence of general practitioners predict their performance? Comparison between examination setting and actual practice. *BMJ* 1991;303:1377–80.
- Das J, Gertler PJ. Variations in practice quality in five low-income countries: a conceptual overview. *Health Aff* 2007;26:w296–309.
- Leonard KL, Masatu MC, Vialou A. Getting doctors to do their best: the roles of ability and motivation in health care quality. *Journal of Human Resources* 2007;42:682–700.
- Das J, Holla A, Das V, *et al*. In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. *Health Aff* 2012;31:2774–84.
- Sylvia S, Shi Y, Xue H, *et al*. Survey using incognito standardized patients shows poor quality care in China's rural clinics. *Health Policy Plan* 2015;30:322–33.
- Rethans JJ, Gorter S, Bokken L, *et al*. Unannounced standardised patients in real practice: a systematic literature review. *Med Educ* 2007;41:537–49.
- Leonard KL, Masatu MC. Using the Hawthorne effect to examine the gap between a doctor's best possible practice and actual performance. *J Dev Econ* 2010;93:226–34.
- Onishi J, Gupta S, Peters DH. Comparative analysis of exit interviews and direct clinical observations in pediatric ambulatory care services in Afghanistan. *Int J Qual Health Care* 2011;23:mzq074.
- Luck J, Peabody JW. Using standardised patients to measure physicians' practice: validation study using audio recordings. *BMJ* 2002;325:679.
- Peabody JW, Luck J, Glassman P, *et al*. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* 2000;283:1715–22.
- Das J, Hammer J. Quality of primary care in Low-Income Countries: facts and Economics. *Annu Rev Econom* 2014;6:525–53.
- Friedman W, Woodman B, Chatterji M. Can mobile phone messages to drug sellers improve treatment of childhood diarrhoea?—A randomized controlled trial in Ghana. *Health Policy Plan* 2015;30(suppl 1):i82–92.
- Madden JM, Quick JD, Ross-Degnan D, *et al*. Undercover careseekers: simulated clients in the study of health provider behavior in developing countries. *Soc Sci Med* 1997;45:1465–82.
- Satyanarayana S, Kwan A, Daniels B, *et al*. Use of standardised patients to assess antibiotic dispensing for tuberculosis by pharmacies in urban India: a cross-sectional study. *Lancet Infect Dis* 2016;16:1261–8.
- Das J, Kwan A, Daniels B, *et al*. Use of standardised patients to assess quality of tuberculosis care: a pilot, cross-sectional study. *Lancet Infect Dis* 2015;15:1305–13.
- Das J, Holla A, Mohpal A, *et al*. Quality and Accountability in Health Care delivery: audit-study evidence from primary care in India. *Am Econ Rev* 2016;106:3765–99.
- Rethans JJ, Sturmans F, Drop R, *et al*. Assessment of the performance of general practitioners by the use of standardized (simulated) patients. *Br J Gen Pract* 1991;41:97–9.
- Planas ME, Garcia PJ, Bustelo M, *et al*. Effects of ethnic attributes on the quality of family planning services in Lima, Peru: a randomized crossover trial. *PLoS One* 2015;10:e0115274.
- Adu-Sarkodie Y, Steiner MJ, Attafuah J, *et al*. Syndromic management of urethral discharge in Ghanaian pharmacies. *Sex Transm Infect* 2000;76:439–42.
- Obare F, Liambila W. The provision of emergency contraceptives in private sector pharmacies in urban Kenya: experiences of mystery clients. *African Population Studies* 2013;24:1–2.
- Ross-Degnan D, Soumerai SB, Goel PK, *et al*. The impact of face-to-face educational outreach on diarrhoea treatment in pharmacies. *Health Policy Plan* 1996;11:308–18.
- Kwena ZA, Sharma A, Muga C, *et al*. Management of simulated patients with sexually transmitted infections by staff of retail pharmacies in Kibera slums of Nairobi. *East Afr Med J* 2008;85:419–24.
- Watson MC, Norris P, Granas AG. A systematic review of the use of simulated patients and pharmacy practice research. *Int J Pharm Pract* 2006;14:83–93.
- Bedoya G, Dolinger A, Rogo K, *et al*. *Observations of infection prevention and control practices in primary health care, Kenya*. Bulletin of the World Health Organization. Forthcoming.
- Kenya National Bureau of Statistics, Ministry of Health/Kenya, National AIDS Control Council/Kenya, Kenya Medical Research Institute, Population NCF, Development/Kenya. *Kenya Demographic and Health Survey 2014*. Rockville, MD, USA, 2015.
- Martin GP O. *Service delivery Indicators*. Kenya, 2013.
- Spreng CP, Ojo IP, Burger NE, *et al*. Does stewardship make a difference in the quality of care? Evidence from clinics and pharmacies in Kenya and Ghana. *Int J Qual Health Care* 2014;26:388–96.
- Republic of Kenya Ministry of Public Health and Sanitation. *Policy Guidelines on Control and Management of Diarrhoeal Diseases in Children Below Five Years in Kenya*, 2010.
- Ng'ang'a LW, Odhiambo JA, Mungai MW, *et al*. Prevalence of exercise induced bronchospasm in Kenyan school children: an urban-rural comparison. *Thorax* 1998;53:919–26.
- WHO. *Tuberculosis Profile*. Kenya, 2015.
- WHO. Kenya: WHO Statistical Profile, 2015.
- Republic of Kenya Ministry of Medical Services and Ministry of Public Health and Sanitation. Clinical management and Referral guidelines – Volume III: Clinical Guidelines for Management and Referral of Common Conditions at Levels 4–6. *Hospitals* 2009.
- WHO Regional Office for Africa. *Kenya Health System description*.
- Mwabu G. Health care reform in Kenya: a review of the process. *Health Policy* 1995;32(1-3):245–55.
- Oyaya CO, Rifkin SB. Health sector reforms in Kenya: an examination of district level planning. *Health Policy* 2003;64:113–27.
- Ridde V, Morestin F. A scoping review of the literature on the abolition of user fees in health care services in Africa. *Health Policy Plan* 2011;26:1–11.
- Hearn J. The 'NGO-isation' of Kenyan society: USAID & the restructuring of health care. *Rev Afr Polit Econ* 1998;25:89–100.
- Shallcross LJ, Davies SC. The World Health Assembly resolution on antimicrobial resistance. *J Antimicrob Chemother* 2014;69:2883–5.
- Levy SB, Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med* 2004;10:S122–9.