Use of Standardized Patients to Assess Quality of Health Care in Nairobi, Kenya: A Pilot, Cross-Sectional Study with International Comparisons

SUPPLEMENTARY APPENDICES

Authors: Benjamin Daniels[1], Amy Dolinger[1], Guadalupe Bedoya[1], Khama Rogo[2], Ana Goicoechea[3], Jorge Coarasa[2], Francis Wafula[2,4], Njeri Mwaura[2], Redemptar Kimeu[5], Jishnu Das[1,6]

[1] Development Economics Research Group, The World Bank
[2] Health, Nutrition and Population Global Practice, The World Bank
[3] Trade and Competitiveness Global Practice, The World Bank
[4] Institute of Healthcare Management, Strathmore University, Nairobi, Kenya
[5] Talana Specialists Centre, Nairobi, Kenya
[6] Centre for Policy Research, New Delhi, India

Corresponding author:

Jishnu Das
Development Economics Research Group, The World Bank
1818 H Street NW, Washington, DC 20433, USA

Email: jdas1@worldbank.org
Phone: +1 (202) 472-2781
Fax: +1 (202) 477-6391

Contents:

Appendix A. Supplementary Methods

1. Description of tracer conditions

Four tracer conditions were developed to validate the SP methodology and provide a first account of quality of care among recruited health care facilities in the Nairobi, Kenya. These were:

- **Asthma:** Classic case of untreated asthma in a 25-year-old male or 24-year-old female, with attack episodes having specific triggers. The SP begins the interaction with the opening statement (in Swahili): *"Doctor, last night I had a lot of difficulty with breathing."* The case should be treated with an inhaler, oral corticosteroid, or bronchodilator.
- **Child Diarrhea:** Classic case of watery diarrhea in an 18-month-old child lasting two days, presented in the clinic by the mother *without* the child's presence. The SP begins the interaction with the opening statement: *"My child has been having diarrhea."* The case should be treated with ORS.
- **Tuberculosis:** Classic case of presumed TB in a 30-year old male or 35-year old female with 2-3 weeks of productive cough and fever. The SP presents to the providers and begins the interaction with the opening statement: *"Doctor, I have been having cough for a while."* The case should be treated with a recommendation for sputum testing (AFB smear).
- **Unstable Angina:** Classic case of unstable angina in a 40-year-old male with crushing chest pain recurring 2-3 times in the past year. The SP presents to the provider with the opening statement: *"Doctor, this morning I had pain in my chest."* The case should be treated with a referral, an ECG, or the immediate provision of aspirin.

2. Essential history question items for each case

Table A1. Essential history question items

Question	Proportion who asked	Question	Proportion who asked
Asthma		**Tuberculosis**	
Breathing Difficulty Details	100%	Duration of cough	90%
Cough	64%	Sputum produced	64%
Pattern of cough	5%	Blood in sputum	29%
Constant or Episodic	21%	Cough pattern	10%
Triggers	33%	Any fever	31%
Expectoration	10%	Chest pain	38%
Previous Breathing Problems	62%	Night sweats	52%
How Often	29%	Loss of appetite	26%
Effective treatments	26%	Weight loss	40%
Fever	17%	**Unstable Angina**	
Chest Pain	36%	Location of pain	62%
Weight Loss	5%	How long pain lasted	26%
Child Diarrhea		Activity when pain began	33%
Age of Child	98%	Quality of pain	40%
Illness duration	65%	Pain radiation	19%
Frequency	57%	History of similar pain	36%
Quality of stool	38%	Pain change with breathing	24%
Blood in stool	10%	Sweating	10%
Child fever	43%	Smoking	7%
Vomiting	47%	Taking Other Medications	33%
Taking fluids	17%		

3. Standardized patient recruitment and training

SPs were recruited from various areas of Nairobi using a recruiting database and network of candidates hosted by a local survey firm. We interviewed 50 potential SPs over the course of two days. Initial screening was based on the required age and gender characteristics for the selected conditions, flexibility in schedule for the duration of the pilot period, level of education and previous work experience. The initial screening of SPs was done by showing the candidates videos of an average health care experience in a health facility in a low-income area of Nairobi. Afterwards, each candidate responded to a writing prompt asking about their observations during the video and for a description of the health care experience. Candidates were interviewed individually to speak about the videos and to recount positive and negative interactions with health care providers in their own personal experiences. We used interview forms to evaluate

whether SPs held any strong opinions towards health care services, self-assessed their ease spending time in health care settings, and their observation and communication skills.

SPs were also given health history forms to assess their current and past health status, as well as a basic vital examination in which the SPs were normal. There were two candidates who were asthmatic, and one with a prior history of TB and HIV diagnosis. After evaluation, the initial decision was to retain them but to keep a close watch to ensure that the modeled SP case was accurately depicted without interference from their personal experience with the conditions. Eventually, one of the asthmatic SPs was retained for the final exercise.

At the end of the initial screening process, 26 SPs were invited to be part of the training. Training began with an introduction to the SP methodology and the selected cases, and an overview of what the researchers had learned about specific qualities that tend make a good SP. The training process opened with a general introduction phase involving additional script development, during which SPs were encouraged to build the narrative of the patient. This included deciding where the patients are from, what they do for a living, a typical day's concerns, and the development of the condition that prompted the patient to seek health care. SPs discussed what they knew of the medical cases from friends, family or personal experience, and the questions they would like to ask a doctor to help them better internalize and portray the condition. Emphasis was placed on the opening statement for each case, which is standardized and should lead the health care provider to subsequent history taking and examinations.

In the second step, SPs learned the basic medical details, symptoms and risk factors for each case to build their familiarity with the SP background stories and understand how to present the

physical aspects of the case. Each script was reviewed in detail in groups by the SPs who were to present that case, and SPs were given a chance to see visual materials and begin practicing the physical aspects as presented by patients with the symptoms. We also reviewed medical and technical terms, but made sure SPs understood their language should remain consistent with a real patient's understanding of the terminology.

SPs were taken through the layout of actual health facilities and the patient experiences, with detailed information on what to expect in terms of on-site patient registration, waiting times, pharmacy encounters, medical consultations, lab encounters and payment. This included how to deal with potentially risky situations that the SPs could encounter during visits and the development of risk avoidance strategies. SPs were instructed to undergo all basic vital measurements, triage and most physical examinations conducted by the health care provider. However, SPs were instructed to avoid blood tests (free or otherwise, including rapid diagnostic tests), injections, x-rays, and other minimally invasive or intrusive procedures, including taking medications, while at the health care facility. Strategies to avoid these risks were actively developed with the SPs through the training period and included the following:

1. Say they do not have enough money for the test or injection.
2. Say they have fainted after an injection/blood draw and would like a relative present.
3. Say they have not eaten that day and would like to return after eating.
4. Mimic taking the medicines and place the medicine in a pocket, if offered medicine.
5. Cough and ask that the thermometer be placed under the arm if attempted by mouth.

The SPs are to recall the names of all investigations and tests to record as part of the exit interview. In cases where these orders were recorded on forms for the patient, SPs retain and

hand over all forms with medicine information as part of the exit interview. SPs are also instructed to accept all prescriptions for medicines, and to purchase drugs at the pharmacy in the clinic or a nearby chemist after their visits before meeting their supervisors. For unlabeled medicines, SPs ask the health care worker for the name of the medicine to purchase more at a chemist when they run out, ask the health care worker to write the name of the medicine, or the SP will recall the name if given verbally for the exit interview following their encounter.

A set of mock interviews was carried out to demonstrate increasingly complex scenarios that SPs could face in real health care settings. The first scenarios focused on strategies for improving recall and mastering scripts. The second set introduced the possibility of unscripted provider behaviors and situations of risk, and teach the SPs how they should respond appropriately. For instance, scenarios may include the complete misdiagnosis of the SP case, probing for medical history not prepared in the SP script, or being asked to perform a blood test or take tablets. A final set of scenarios mimicked the full encounter an SP may experience, beginning with approaching a health care facility, registering as a patient, potential situations involving other patients, interactions with health care workers, laboratory tests, treatment, payment, and protocols after leaving the facility. These scenarios were reenacted for multiple contexts, from large public facilities to private single-provider health clinics. Details were altered to include long and short consultation times and different attitudes held by health care workers. Following all mock scenarios, SPs conducted structured exit interviews as they would in actual encounters.

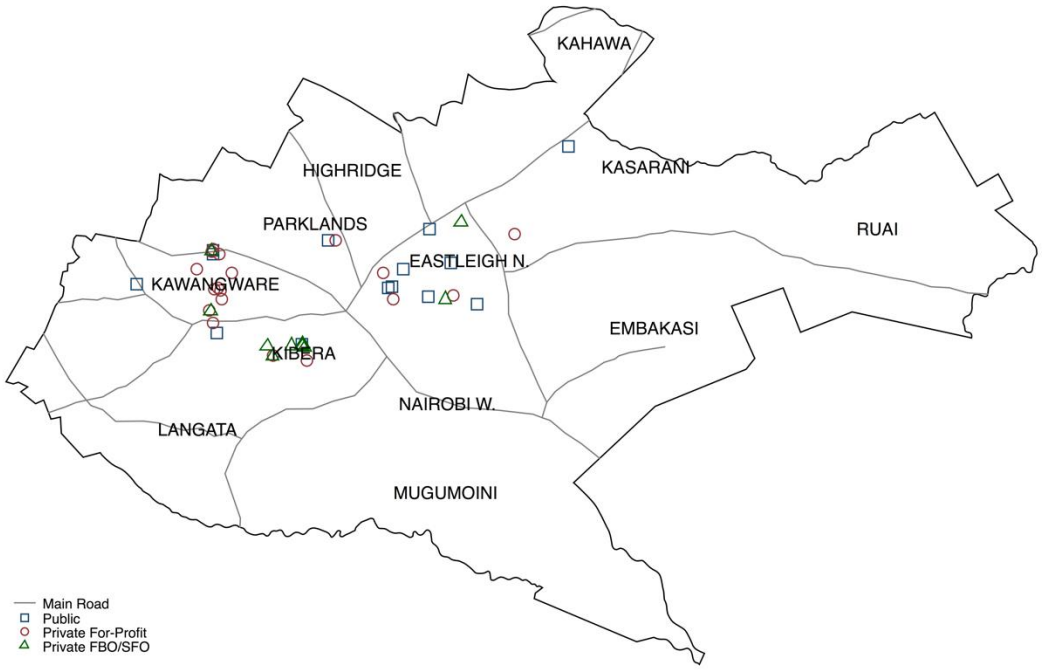
Finally, dry run case presentations to real doctors and nurses, were completed in clinics where SPs were pre-announced to the health care providers, who later gave feedback on the SP presentations. These health care workers assessed the standardization of presentations across SPs

for each case, while highlighting anything SPs said or did that could lead the provider dramatically away from the correct diagnosis, as well as anything in the presentations that led them to believe the SPs were not genuine patients. A second round of dry runs sent SPs unannounced to health care facilities in non-sampled areas of Nairobi to present their cases. The SPs successfully completed interactions, avoided risks and retained medicines from their encounters prior to conducting exit interviews in the field with supervisors. Full debriefing sessions where each SP recounted their experience took place at the training center every evening. After these interactions, 14 candidates were invited to participate in the final implementation based on their performance.

4. Sectoral comparison model

Our sample covers three types of facility ownership categories: 14 public sector facilities, 18 private for-profit facilities, and 10 private not-for-profit (FBO/SFO) facilities – which included 1 community clinic, 5 clinics operated by faith-based organizations and 4 clinics operated by social franchise operations – as shown in the sampling map (**Figure A1**).

Figure A1. Sampling Map



In our main results, we combine all privately-owned facilities into a single group for comparison against the public sector using linear and logistic regression models. Our rationale for combining for-profit and FBO/SFO providers, in addition to the small sample size that would prevent conclusive comparisons between the public sector and each of the subgroups separately, is that linear regression shows no significant differences between for-profit and not-for-profit private providers on our primary dimensions of analysis. **Figure A2** illustrates the comparison of all three sectors separately, and **Table A2** details private sector comparisons. Note that, although imprecise, the point estimates are qualitatively small, suggesting that the lack of significance for differences between the for-profit and not-for-profit private provider samples is not an artifact of imprecision.

Figure A2. Three-Sector Comparison of Quality Outcomes

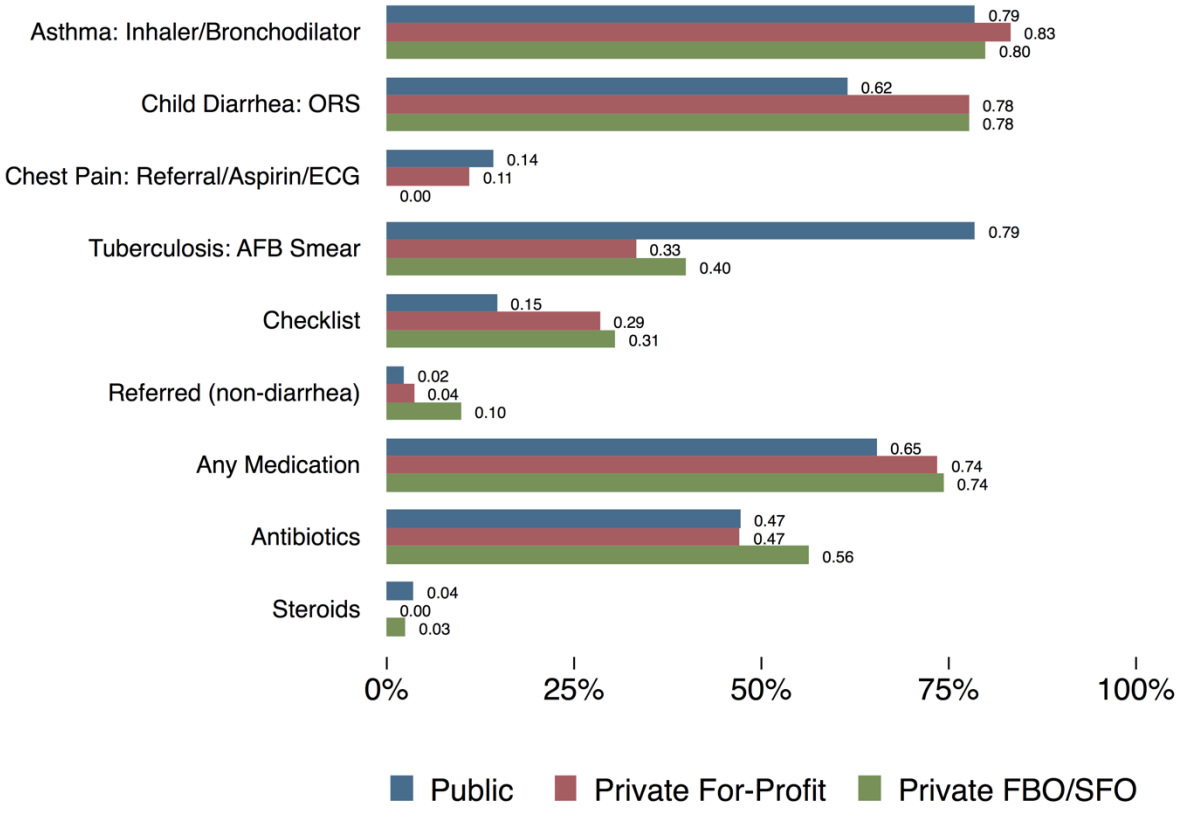


Table A2. Primary Outcomes for Standardized Patients Cases Among Private Facilities

	For-Profit Mean	FBO/SFO Mean	Difference
Preferred Management			
Asthma: Inhaler or Bronchodilator	0.83	0.80	-0.03
	(0.09)	(0.13)	(0.16)
Child Diarrhea: ORS	0.78	0.78	-0.00
	(0.10)	(0.15)	(0.18)
Unstable Angina: Referral, ECG, or Aspirin	0.11	0.00	-0.11
	(0.08)	(0.00)	(0.10)
Tuberculosis: Sputum Test	0.33	0.40	0.07
	(0.11)	(0.16)	(0.20)
Basic Statistics			
Time Waiting in Clinic (mins)	23.48	32.16	8.74
	(3.52)	(4.79)	(5.96)
Time with Provider (mins)	9.32	7.38	-2.04
	(0.92)	(0.68)	(1.26)
Checklist	0.43	0.47	0.04
	(0.03)	(0.04)	(0.05)
Price (Ksh)	574.76	541.41	-39.28
	(67.78)	(63.41)	(101.25)
Referred (non-diarrhea)	0.04	0.10	0.06
	(0.03)	(0.06)	(0.05)
Medications			
Any Medication	0.74	0.74	0.00
	(0.05)	(0.07)	(0.08)
Number of Medicines	1.78	2.00	0.22
	(0.17)	(0.24)	(0.29)
Antibiotics	0.47	0.56	0.09
	(0.06)	(0.08)	(0.10)
Steroids	0.00	0.03	0.03
	(0.00)	(0.03)	(0.02)
Number of Observations	*72*	*39*	

Note: Data are means (standard errors) or estimated coefficients (standard errors). Preferred case management is defined as the use of an inhaler or bronchodilator for the asthma case, ORS for the child diarrhea 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 medication recommended by the provider. Stars indicate significance levels as follows: *** = $p<0.01$, ** = $p<0.05$, * = $p<0.1$.

Appendix B. Supplementary Results

1. Validation: Impact of SP Characteristics on Interaction Outcomes

A multivariate regression model was used to determine whether individual SP characteristics had a significant impact on provider behavior (**Table A3**). We find that SPs who were older and had a higher systolic blood pressure (although none was hypertensive) appear to have 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. As we note in the main results, these results document naturally-occurring variation in provider treatment in the treatment environment, and underscore the importance of (1) random assignment of SPs to providers to avoid unintentional bias in estimates of differences by ensuring that SP characteristics are not correlated with provider characteristics and (2) ensuring that SPs are reasonably representative of the patient population of interest for the study (in this case, the general population), to avoid estimates of quality that are not externally valid.

Table A3. Validation Regressions with Standardized Patient Characteristics

	Time with Provider (mins)	Essential Checklist Proportion	Preferred Case Management	Time Waiting in Clinic (mins)	Price (Ksh)
SP Age	0.604***	0.016***	0.014	-1.278	-1.268
	(0.203)	(0.005)	(0.010)	(1.546)	(12.249)
SP BMI	0.332*	0.001	0.003	-0.170	9.013
	(0.172)	(0.006)	(0.013)	(0.964)	(10.483)
SP Systolic BP	0.266**	0.011***	0.007	-0.026	2.247
	(0.099)	(0.003)	(0.005)	(0.671)	(5.222)
SP Male	-3.643	-0.215*	-0.214	-5.962	64.767
	(2.856)	(0.112)	(0.204)	(25.390)	(236.089)
Mean Value	7.171	0.379	0.530	49.118	425.097
Number of Observations	*166*	*166*	*166*	*166*	*165*

	Referred (non-diarrhea)	Any Medication	Number of Medications	Antibiotics	Steroids
SP Age	0.005	-0.004	-0.010	-0.018	-0.004
	(0.006)	(0.012)	(0.038)	(0.013)	(0.004)
SP BMI	0.008	-0.004	0.036	0.006	0.000
	(0.005)	(0.011)	(0.032)	(0.014)	(0.001)
SP Systolic BP	0.002	-0.009	-0.004	-0.009	0.001
	(0.002)	(0.006)	(0.021)	(0.007)	(0.002)
SP Male	0.105	0.164	0.527	0.319	0.029
	(0.134)	(0.209)	(0.643)	(0.210)	(0.043)
Mean Value	0.048	0.711	1.771	0.494	0.018
Number of Observations	*126*	*166*	*166*	*166*	*166*

Note: Data are estimated coefficients (standard errors). Preferred case management is defined as the use of an inhaler or bronchodilator for the asthma case, ORS for the child diarrhea 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 medication recommended by the provider. Stars indicate significance levels as follows: *** = $p<0.01$, ** = $p<0.05$, * = $p<0.1$.

2. Diagnoses Given by Providers

SPs were never instructed to prompt the provider for a verbal diagnosis; and providers in most cases (113/166) did not offer such a diagnosis to the SP. However, even where a diagnosis was offered, these varied widely. Across the four cases, 25 different diagnoses were given to patients in 53 instances, with the most common diagnosis being pneumonia, pronounced in 20 different interactions (and the most common diagnosis in the asthma, chest pain, and tuberculosis cases). **Table A4** provides the complete details.

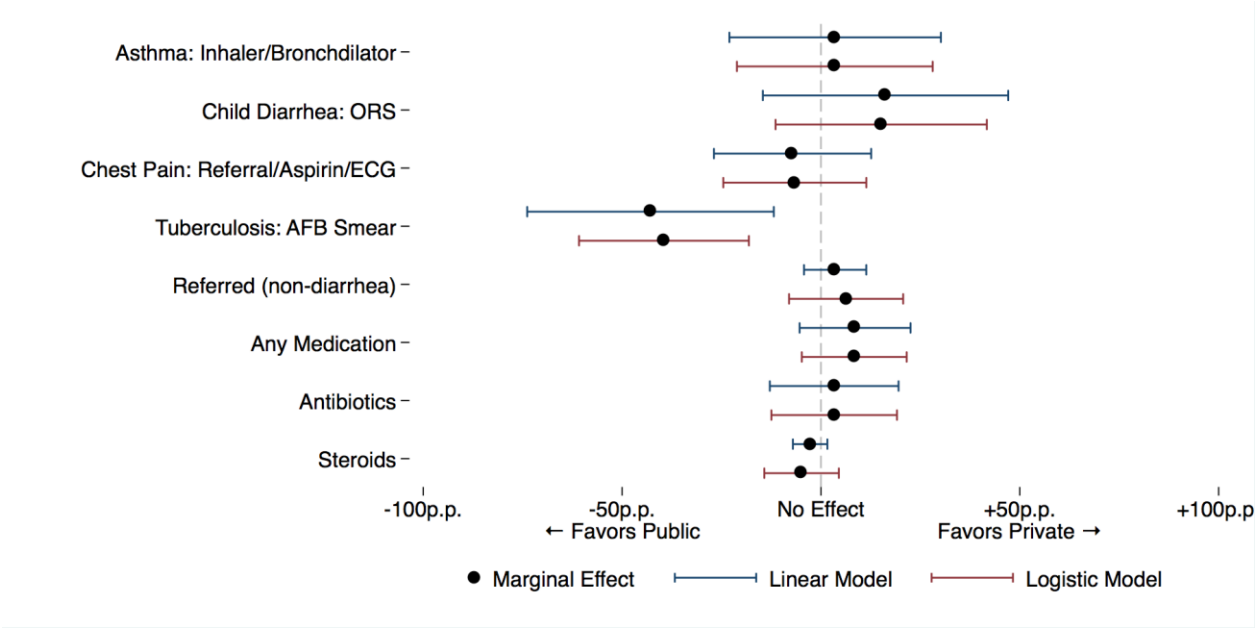
Table A4. Diagnoses Given, by Standardized Patient Case

	Asthma	Chest Pain	Child Diarrhea	TB Suspect
No Diagnosis Given	26	22	36	29
Acidity		2		
Allergy	3			
Amoeba			1	
Angina		2		
Asthma	2			
Bacterial Infection				1
Blood Infection		1		
Bronchitis				1
Brucellosis or TB				1
Chest Congestion				2
Chest Infection	1	2		1
Common Cough				1
High Blood Pressure	1			
Malaria			1	
No Sickness	1			
Pleuritis		1		
Pneumonia or TB		1		
Pneumonia	7	9		4
Pneumonia or Chest Infection		1		
Suspected Malaria and Chest Infection				1
The Heart Is Okay; It Could Be Muscle Pull		1		
Ulcers	1			
Upper Respiratory Infection				1
Viral Infection			1	
Watery Diarrhoea			1	
Number of Observations	42	42	40	42

3. Validation of linear and logistic models

In **Table 4**, we use both adjusted t-tests and logistic regressions to assess differences in outcomes between the public and private sector. As a robustness check for the functional form of the difference estimate in the t-tests, **Figure A3** reports marginal effects and standard errors computed from the logistic regressions and compares them to the linear estimates reported in the main text. We find that the two point estimates are never statistically significantly different, and do not vary systematically in direction or precision across the set of outcomes.

Figure A3. Comparison of marginal effects from linear and logistic specifications



4. Cross-country comparisons

Table A5 gives the complete results for cross-country comparisons.

Table A5. Primary Outcomes for Standardized Patient Cases by Country

	Preferred Case Management	Time Waiting in Clinic (mins)	Time with Provider (mins)	Common Checklist %	Total Price (USD Equivalent PPP)	Referred (non-diarrhea)	Any Medication	Number of Medications	Antibiotics	Steroids	*N*
Asthma											
India	0.57	9.20	4.32	0.22	2.06	0.07	0.92	3.19	0.47	0.26	*397*
	[0.52–0.62]	[7.83–10.56]	[3.95–4.68]	[0.21–0.23]	[1.84–2.27]	[0.04–0.09]	[0.89–0.95]	[3.05–3.33]	[0.42–0.52]	[0.21–0.3]	
Kenya	0.81	46.26	9.26	0.26	13.66	0.00	0.86	2.07	0.50	0.02	*42*
	[0.69–0.93]	[31.99–60.54]	[6.8–11.72]	[0.21–0.31]	[9.47–17.86]	–	[0.75–0.96]	[1.68–2.46]	[0.35–0.65]	[-0.02–0.07]	
Child Diarrhea											
China	0.00	1.03	1.13	0.17	2.73		0.76	1.57	0.43	0.00	*42*
	–	[0.37–1.69]	[0.88–1.39]	[0.14–0.21]	[1.82–3.63]		[0.63–0.89]	[1.21–1.93]	[0.28–0.58]	–	
India	0.18	9.97	1.57	0.14	1.22		0.77	1.59	0.63	0.01	*389*
	[0.14–0.22]	[5.34–14.6]	[1.42–1.71]	[0.13–0.16]	[1.07–1.37]		[0.73–0.81]	[1.47–1.72]	[0.59–0.68]	[0–0.01]	
Kenya	0.73	51.08	4.45	0.20	6.63		0.52	1.58	0.32	0.00	*40*
	[0.59–0.86]	[33.48–68.67]	[3.53–5.37]	[0.17–0.24]	[3.85–9.41]		[0.37–0.68]	[1.02–2.13]	[0.18–0.47]	–	
Tuberculosis Suspect											
India	0.04	6.44	5.96	0.24	6.92	0.03	0.96	4.84	0.61	0.00	*75*
	[0–0.08]	[4.84–8.05]	[5.2–6.71]	[0.21–0.26]	[5.85–7.98]	[-0.01–0.06]	[0.92–1]	[4.52–5.16]	[0.5–0.72]	–	
Kenya	0.50	45.64	6.72	0.26	12.65	0.12	0.57	1.48	0.55	0.00	*42*
	[0.35–0.65]	[28.76–62.52]	[5.11–8.33]	[0.21–0.31]	[8.1–17.2]	[0.02–0.22]	[0.42–0.72]	[1.05–1.91]	[0.4–0.7]	–	
Unstable Angina											
China	0.63	2.13	4.09	0.18	4.92	0.63	0.57	1.33	0.08	0.00	*40*
	[0.47–0.78]	[0.3–3.96]	[2.53–5.66]	[0.15–0.21]	[2.93–6.92]	[0.47–0.78]	[0.42–0.73]	[0.89–1.76]	[-0.01–0.16]	–	
India	0.41	9.94	3.56	0.25	1.67	0.25	0.88	2.67	0.20	0.05	*323*
	[0.36–0.47]	[5.29–14.59]	[3.23–3.89]	[0.23–0.27]	[1.45–1.9]	[0.21–0.3]	[0.84–0.91]	[2.51–2.84]	[0.16–0.25]	[0.03–0.08]	
Kenya	0.10	53.59	8.12	0.25	12.51	0.02	0.88	1.95	0.60	0.05	*42*
	[0.01–0.18]	[34.63–72.56]	[5.59–10.66]	[0.2–0.31]	[8.81–16.21]	[-0.02–0.07]	[0.78–0.98]	[1.63–2.28]	[0.45–0.74]	[-0.02–0.11]	

Note: Data are mean, [95%CI]. Preferred case management is defined as the use of an inhaler or bronchodilator for the asthma case, ORS for the child diarrhea case, a sputum test for the tuberculosis suspect, and a referral, ECG, or aspirin for the unstable angina case. Price includes all testing and medication recommended by the provider. Prices are standardized to 2014 purchasing-power adjusted U.S. dollars using World Bank conversion rates from http://data.worldbank.org/indicator/PA.NUS.PPP.