Methodological Appendix

1.1 Description of SP Case

The SP case was designed to depict a classic case of pulmonary TB, childhood viral gastroenteritis, or unstable angina. For TB case, the SP begins the interaction with providers using the opening statement: "Doctor, I have a cough that is not improving and a fever." For gastroenteritis case, the opening statement is "Doctor, my child has diarrhea", and for angina case with "Doctor, I have chest pain recently". The SP then answers any questions asked by the providers and receives any (non- invasive) exams. Upon appropriate questioning by the provider, the SP reveals symptoms consistent with a classic case of three disease case. For TB, these symptoms include cough durations for 2-3 weeks, fever with night sweats, and loss of appetite and weight, which are consistent with a classic case of presumed TB and, according to national guidelines, should not be prescribed appropriate antibiotics until confirmation. For viral gastroenteritis, appropriate questioning would reveal symptoms of a two-year old child consistent with a viral infection (including watery diarrhea without blood or mucus, no fever or vomiting and little change in behavior). The chief complaint of the angina case, chest-pain, is less likely consistent with a disease requiring antibiotics but provides insight into antibiotic prescription for less ambiguous cases.

1.2 SP script development

The TB case used in this study was adapted from an earlier validation study in India[1] and more detail information is in another study.[2] The gastroenteritis and angina cases were adapted from an earlier pilot study in China.[3] That study demonstrated (1) participation in the study had minimal to no risk for the SPs or health care providers, (2) the likelihood of SP detection among visited providers was low, confirming that SPs were considered real by health providers who were visited, and (3) the abilities of the SPs to recall what occurred during the interaction was strongly correlated with what actually happened. Additionally, because the SPs pay the fees requested by the healthcare provider, there is no loss to provider income from participation in the study.

For the purposes of the current study, the SP case was adapted to the Chinese context with the help of an advisory panel consisting of medical experts in China. Adaptation of the scripts included: (1) ensuring that the clinical presentation of the SP would be interpreted clearly given local context; (2) ensuring that SP responses were prepared for any likely questioning by providers (3) developing SP interaction protocols given clinical settings in China, including protocols for avoiding invasive procedures; (4) developing detailed background histories for SPs to minimize the threat of SP detection as fake patients. Additionally, because local dialects varied across the three regions of the study, scripts with alternative phrasing to match the local dialect (where appropriate) were developed from a version in standard Mandarin. These were small alterations to phrasing or vocabulary and were chosen to convey the exact information as the standard script. SP script adaptations took place as an iterative process including field pretesting (in local but not project areas) with 6 pre-trained SPs.

The SP Scripts (Including background and dialog) are available for download at web for TB case (https://seansylvia.web.unc.edu/sp-case-script-tuberculosis/), gastroenteritis case (https://seansylvia.web.unc.edu/sp-case-script-gastroenteritis/) and angina case (https://seansylvia.web.unc.edu/sp-case-script-angina/).

1.3 SP recruitment and characteristics

To ensure low likelihood of detection among visited providers, SPs were recruited from local areas. This meant that SPs were similar to patients typically seen by the clinics in terms of language (dialect), mannerisms, and dress. Across the three provinces, a total of 21 individuals (14 males and 7 females) were hired from an initial group of 63 to act as SPs. The SPs, although recruited specifically to fit the three disease cases in terms of health and physical characteristics, differed in age, gender height and weight. Our recruitment standard for SPs was that they be average weight and height, and healthy with no obvious signs of illness or other conditions that could prejudice diagnoses. The average age of recruited SPs for TB was 37, the youngest was 28 and the oldest 43; the average age of recruited SPs for gastroenteritis was 29, the youngest was 25 and the oldest 38; the average age of recruited SPs for angina was 41, the youngest was 34 and the oldest 50.

1.4 SP training

All SPs underwent a centralized intensive two-week training. The aim of SP training was to ensure that they (a) correctly presented the cases in a consistent way, (b) correctly recalled the interaction with providers, (c) avoided detection and (d) SPs would be able to complete interactions safely without being exposed to invasive tests or procedures.

These aims were achieved through classroom training in case presentation and testing of recall, as well as mock interviews and dry runs that were supervised in the field. The training started with a focus on the cases and the development of scripts and proceeded to memorization and appropriate role-playing. SPs were taught to internalize completely the characters and the details of their background stories. Mock interviews were conducted with trainers as well as providers working with the research team. These mock interviews initially asked only potentially clinically relevant questions, but in later rounds added additional questions about family or neighborhood to ensure that SPs could answer appropriately and convincingly. Mock interviews also simulated "threats" of invasive procedures.

In the final week of training, SPs conducted supervised dry runs in clinics nearby the training site. Since it is common in China that more than one patient is present in the examination room at the same time, dry runs were conducted in which a supervisor was present and thus could watch the interaction and offer corrections later. Dry runs were also conducted during a two-day practice round in local areas to ensure that the SPs were accustomed to local conditions before starting data collection.

Enumerators accompanying SPs also attended the full training in order to familiarize themselves with the survey process and the SP visit.

1.5 Assignment of SPs to providers

All SPs were randomly assigned to clinics. Each survey team (comprising of three SPs for different cases and two enumerators) were randomly assigned to two counties within each province. Within each county, teams were assigned a random half of sample townships. SPs were never assigned to their home township where they would risk being recognized.

Each survey team first visited their assigned townships (both township health centers and village clinics in each township) in their first assigned county and then traveled to their second assigned county. County hospitals were visited by the second team assigned to each county.

Within each facility, SPs visited the doctor following the normal procedures for any walk-in patient. Given a choice of which doctor to visit, SPs randomly chose a doctor following a pre-determined randomization protocol. All three cases SPs Our results therefore approach the care a walk-in patient would receive at each of the sampled facilities.

1.6 Consent

Informed consent was obtained verbally from all providers participating in the study. To prevent influence on the study, a procedure was approved whereby providers consented to SP visits "at some point in the next six months." Consent from village and township providers was obtained as part of the facility survey approximately 5 weeks before SP visits using the script below. All individuals who participated as SPs were trained to protect themselves from any invasive tests or procedures.

Following the conclusion of the baseline survey, consent was obtained verbally from providers using the following script: *"At some point during the next six months, we may send a fake patient may visit your facility to seek care. If you believe that a patient is a fake patient, please record the patients name, symptoms, the timing of the visit, but do not directly ask patients if they are fake patients. Do you agree for a fake patient to visit your facility?"*

1.7 Detection of SPs

To assess the rate at which SPs were detected (as fake patients), a detection survey was launched 2-3 weeks following SP visits. Providers were told at the time of giving consent to participate in the study to record information on any patient that suspected as a fake patient. During the detection survey following SP visits, providers were asked whether they suspected anyone as a standardized patient and, if so, to report the characteristics of detected SPs and the specific symptoms provided by the suspected patient.

All sampled township providers completed the detection survey. All providers completed the detection survey. Of these, 9 (2%) reported that they suspected someone as an SP. Of the 9 total possible SP detections, 6 physician descriptions matched the standardized patient. In no cases did the provider voice suspicion during the interaction.

Appendix Table 1: Variable descriptions

Variable name	Description
THC Characteristics (n=207)	
Population in catchment area	The general population in the area served by the health center
Patient visits in 2014	Total number of patient visits in calendar year 2014
Full-time providers	Number of providers working full time at the facility
Central THC	In large townships with >1 township health center, the designated primary township health center
Record storage system	An indicator for whether the THC has established protocols for medical record storage
Prescription review	An indicator for whether the prescriptions are reviewed by hospital staff other than the provider as a quality control
Prescription quality as part of provider work evaluations	An indicator for whether providers' evaluations include an assessment of prescription quality
Provider Characteristics (n=383)	
Age (years)	The provider's age in years
Male	Whether the provider is male
Monthly income (1,000 RMB)	The provider's monthly income in 1,000 RMB
Medical work experience (years)	The number of years the provider has worked as a health care provider
Bachelor's degree or higher	An indicator for providers having attained a bachelor's degree or higher
Three levels of provider certificates	
Rural doctor certificate	Lowest level of certification, necessary to practice in rural areas
Assistant practicing doctor certificate	Mid-level certification in rural facilities
Practicing provider certificate	The highest level of certification in rural facilities
Village doctor experience	Indicator for whether the provider has previous experience working in a village clinic
Disease type (n=620)	
Angina	SP presented with symptoms of unstable angina
Diarrhea	SP presented child with symptoms of gastroenteritis
TB	SP presented with symptoms of pulmonary tuberculosis

Appendix Table 2: Characteristics of sample township health centers and providers

	%/Mean (SD)
THC Characteristics (n=207)	
Population in the catchment area	26,125 (20,430)
Patients visits in 2014	18,706 (16,868)
Full-time providers	7.7 (7.2)
Central township health centers	48.6%
Record storage system	84.6%
Prescription review	55.3%
Work evaluation includes assessment of prescription quality	70.2%
Provider Characteristics (n=383)	
Age (years)	43.3 (10.0)
Male	85.9%
Monthly income (1,000 yuan)	3.4 (1.1)
Medical work experience (years)	21.0 (10.7)
Bachelor's degree or higher	18.8%
Rural doctor certificate	9.7%
Assistant practicing doctor certificate	23.2%
Practicing provider certificate	61.4%
Village doctor experience	26.9%

Appendix Table 3: Univariate analysis of differences in THC and provider characteristics by medical record existence

	With records (n=210)	Without records (n=420)	*P* value
	%/Mean (SD)	%/Mean (SD)	
THC Characteristics			
Patients in catchment area	26,456 (19,443)	26,041 (20,945)	0.81
Patients seen in 2014	21,256 (17,834)	17,462 (16,215)	0.01
Full-time providers	8.3 (7.8)	7.4 (6.8)	0.17
Central township health centers	48.67%	48.8%	0.96
Record storage system	87.5%	82.9%	0.14
Prescription review	56.3%	54.9%	0.75
Work evaluation includes assessment of prescription quality	74.6%	68.3%	0.11
Provider Characteristics			
Age (years)	44.7 (10.5)	42.1 (9.8)	0.00
Male	87.6%	84.4%	0.30
Monthly income (1,000 yuan)	3.5 (1.1)	3.4 (1.1)	0.16
Medical work experience (years)	22.4 (11.3)	19.8 (10.9)	0.01
Bachelor's degree or higher	18.2%	22.6%	0.20
Rural doctor certificate	8.6%	10.1%	0.56
Assistant practicing doctor certificate	24.4%	22.4%	0.57
Practicing provider certificate	61.7%	61.3%	0.92
Village doctor experience	26.3%	26.8%	0.91

Appendix Table 4: Demographic characteristics and quality of medical records by prefecture

Variables	Prefecture 1	Prefecture 2	Prefecture 3
Province	Anhui	Shaanxi	Sichuan
Area (square kilometers)	13,516	42,921	13,266
Rural population	2,826,260	2,855,300	2,702,600
Total number of townships	94	156	172
Existence (% with any record)	32.2%	28.1%	41.6%
Completeness (# of completed items)	4.3	3.5	5.4
Accuracy (# of accurate items)	4.3	3.4	5.3

Notes: The data on regional characteristics are in 2015 from the Anhui Statistical Yearbook 2016 (http://tongji.cnki.net/kns55/Navi/YearBook.aspx?id=N2017020263&floor=1), Shaanxi Statistical Yearbook 2017 (http://navi.cnki.net/knavi/YearbookDetail?pcode=CYFD&pykm=YSDIE), and Sichuan Statistical Yearbook 2016 (http://tongji.cnki.net/kns55/Navi/YearBook.aspx?id=N2017020274&floor=1), respectively.

Appendix Table 5: Multivariate regressions estimating the correlates medical record quality

Dependent Variables	Existence	Completeness	Accuracy
	(1)	(2)	(3)
Number of patients in the catchment area	-0.01	0.03	0.00
	(-0.03, 0.12)	(-0.16, 0.23)	(-0.03, 0.03)
Number of patients seen in the calendar year 2014	0.01	0.01	0.00
	(-0.01, 0.04)	(-0.24, 0.26)	(-0.03, 0.04)
Number of providers working full time at the facility	0.00	-0.02	0.00
	(-0.01, 0.00)	(-0.04, 0.00)	(-0.00, 0.00)
Type of THCs (1=central THCs)	-0.01	0.08	-0.04
	(-0.10, 0.09)	(-0.74, 0.89)	(-0.16, 0.07)
Prescription review system (1=yes)	-0.05	-0.87*	-0.04
	(-0.17, 0.06)	(-1.68, -0.06)	(-0.16, 0.08)
Work evaluation includes prescription quality (1=yes)	0.05	-0.34	-0.05
	(-0.03, 0.14)	(-1.10, 0.42)	(-0.16, 0.06)
Records storage protocol (1=yes)	0.05	0.21	-0.07
	(-0.08, 0.17)	(-0.82, 1.24)	(-0.22, 0.08)
Medical education (1=bachelor's degree or higher)	-0.03	0.01	-0.03
	(-0.11, 0.05)	(-0.99, 1.01)	(-0.17, 0.11)
Sex (1=male)	0.07	0.19	-0.01
	(-0.04, 0.18)	(-0.21, 0.59)	(-0.06, 0.05)
Monthly income (1,000 yuan)	0.05*	-0.02	0.00
	(0.01, 0.09)	(-0.06, 0.01)	(-0.00, 0.01)
Medical work experience (years)	0.01	0.45	-0.01
	(-0.01, 0.01)	(-0.25, 1.15)	(-0.11, 0.09)
Practicing provider certificate (1=yes)	-0.04	0.37	-0.03
	(-0.12, 0.03)	(-0.38, 1.11)	(-0.14, 0.07)
Village doctor experience (1=yes)	0.02	0.11	-0.07
	(-0.07, 0.10)	(-0.71, 0.93)	(-0.19, 0.04)
Diarrhea case (1=yes)	-0.10*	1.77*	0.03
	(-0.19, -0.01)	(0.78, 2.76)	(-0.09, 0.16)
TB case (1=yes)	0.12*	1.33*	-0.06
	(0.05, 0.19)	(0.58, 2.07)	(-0.17, 0.04)
Mentioned items (number)	-	0.33*	-
	-	(0.20, 0.46)	-
Completed items (number)	-	-	0.97*
	-	-	(0.95, 0.99)
Mean of the outcome variable	33.80%	4.5	4.1
Sample size	620	210	210

Notes: Averaged marginal effects from multivariate regressions reported. Existence (a binary indicator of whether any medical records were kept) was estimated by logistic regression. *Completeness (the number*

9

of completed items in each medical record) and accuracy (the number of accurate items in each medical record) were estimated by OLS. Regressions also control for county fixed effects. * indicates P value less than 0.05

Appendix Table 6: The completeness and accuracy of medical records

Items	Mentioned in visits		Completeness		Accuracy	
	%	95% CI	%	95% CI	%	95% CI
Patient information						
Patient's name	85.7	80.2, 90.1	97.8	94.4, 99.4	100.0	-
Gender	88.6	83.5, 92.5	83.9	77.8, 88.8	100.0	-
Date of birth	80.0	73.9, 85.2	88.1	82.2, 92.6	99.3	96.3, 100.0
Nationality	0.5	0.01, 2.6	0.0	-	0.0	-
Marital status	0.5	0.01, 2.6	0.0	-	0.0	-
Occupation	33.3	27.0, 40.1	35.7	24.6, 48.1	100.0	-
Work unit	24.8	19.1, 31.2	0.0	-	0.0	-
Address	63.3	56.4, 69.9	61.6	52.8, 69.9	100.0	-
History of drug allergy	14.3	9.9, 19.8	0.0	-	0.0	-
Chief complaint						
Main symptoms	100.0	-	25.7	19.9, 32.2	81.7	68.6, 90.7
Duration	67.6	60.8, 73.9	9.9	5.5, 16.0	77.6	49.2, 95.3
History of present illness						
Onset time	67.6	60.8, 73.9	12.0	7.1, 18.5	70.4	44.0, 89.7
Main and accompanying symptoms	33.8	27.4, 40.6	36.7	25.5, 48.9	88.7	69.8, 97.6
Negative signs for differential diagnosis	22.9	17.4, 29.1	12.7	4.7, 25.2	100.0	-
Treatment situation during prior visits to other hospitals	31.0	24.8, 37.7	0.0	-	0.0	-
Past medical history						
Personal, past, or family history	18.6	13.6, 24.5	10.2	2.9, 24.2	100.0	-
Diagnosis						
Normative diagnosis name	75.2	68.8, 80.9	57.6	49.5, 65.4	83.6	74.3, 90.5
Treatment						
Non-drug treatment suggestions	75.2	68.8, 80.9	13.3	8.4, 20.0	100.0	58.4, 71.7
Drug treatment	30.0	23.9, 36.7	100.0	-	100.0	-
Referral or follow-up	29.5	23.4, 36.2	3.34	0.4, 11.2	100.0	-
Cautions	41.0	34.2, 47.9	0.0	-	0.0	-

Notes: Completeness percentages are conditional on items having been mentioned in visits ("Mentioned in visits") and Accuracy percentages are conditional on these items having been recorded ("Completeness").

11