Use of standardised patients to assess gender differences in quality of tuberculosis care in urban India: a two-city, cross-sectional study

Benjamin Daniels*, Ada Kwan*, Srinath Satyanarayana*, Ramnath Subbaraman, Ranendra K Das, Veena Das, Jishnu Das†, Madhukar Pai†

Summary

Background In India, men are more likely than women to have active tuberculosis but are less likely to be diagnosed and notified to national tuberculosis programmes. We used data from standardised patient visits to assess whether these gender differences occur because of provider practice.

Methods We sent standardised patients (people recruited from local populations and trained to portray a scripted medical condition to health-care providers) to present four tuberculosis case scenarios to private health-care providers in the cities of Mumbai and Patna. Sampling and weighting allowed for city representative interpretation. Because standardised patients were assigned to providers by a field team blinded to this study, we did balance and placebo regression tests to confirm standardised patients were assigned by gender as good as randomly. Then, by use of linear and logistic regression, we assessed correct case management, our primary outcome, and other dimensions of care by standardised patient gender.

Findings Between Nov 21, 2014, and Aug 21, 2015, 2602 clinical interactions at 1203 private facilities were completed by 24 standardised patients (16 men, eight women). We found standardised patients were assigned to providers as good as randomly. We found no differences in correct management by patient gender (odds ratio 1.05; 95% CI 0.76–1.45; p=0.77) and no differences across gender within any case scenario, setting, provider gender, or provider qualification.

Interpretation Systematic differences in quality of care are unlikely to be a cause of the observed under-representation of men in tuberculosis notifications in the private sector in urban India.

Funding Grand Challenges Canada, Bill & Melinda Gates Foundation, World Bank Knowledge for Change Program.

Introduction

Multiple systematic reviews,1–3 including a review of 56 prevalence surveys from 24 countries, have found that men are more than twice as likely to have active tuberculosis but are considerably less likely than women to be diagnosed and notified to national tuberculosis programmes. In India, this represents a reversal of the usual pattern of disadvantage for women in use of health care.4 Although India’s national tuberculosis programme receives 1-9 notifications regarding men for every notification regarding women,5 tuberculosis population prevalence is even higher among men, suggesting that men access tuberculosis care at substantially lower rates than women.6 Understanding the sources of this gender imbalance is crucial to identifying and treating the missing millions of patients with tuberculosis globally.8

After an individual develops active tuberculosis, they must traverse a process of care-seeking, diagnosis, linkage to treatment, treatment initiation, and notification to national tuberculosis programmes.9–11 Men might face a disadvantage at any or all of these stages, and identifying the stage of the care cascade that contributes most to the relative undernotification of men could help to focus interventions on the most important gaps in care.

Of these stages, our study focuses on understanding gender imbalance in the diagnostic process. This stage is challenging to evaluate and represents a crucial point in care where changes in provider behaviour could improve case detection.11–14 Additionally, gender differences in the quality of care delivered by health-care providers, whether biased against women or men, have important implications for health equity and social justice.

Incomplete data (eg, patient charts) on clinical processes in these settings combined with the complexity of the diagnostic process and differences in presentations between men and women makes it difficult to answer a conceptually simple question: when men and women with the same tuberculosis symptoms visit the same health-care providers, do they receive the same quality of care during initial clinical evaluations? In the absence of (complete) medical records,15 we use standardised patients to answer this question.

Standardised patients are people recruited from local populations and extensively trained to portray a...
Research in context

Evidence before this study
In India, as in many other countries with a high tuberculosis burden, men substantially outnumber women among notified people in the national tuberculosis programme. However, men are usually underrepresented in these programmes relative to their share of the tuberculosis disease burden, meaning that men are disadvantaged at some point in the care seeking process relative to women. Whether health providers themselves contribute to these differences through delays in diagnosis (or misdiagnosis) of men or women is unknown. Reliance on administrative programme data and provider interviews is insufficient for discerning potential gender differentials, because of biases created from potentially different care-seeking patterns across gender. Therefore, how gender differences in notification rates reflect access to or provision of health services and differential quality of tuberculosis care for men and women is uncertain. The standardised, simulated patient method, which is considered the gold standard method to assess provider practice, allows us to rigorously address the quality of tuberculosis care by gender.

Added value of this study
The standardised patient method, by ensuring that the case presentations and provider selections of men and women are identical, is used here to determine whether a significantly different response occurs on the provider side when treating otherwise identical case presentations from men and women. Our study used a large, representative sample of private health-care providers in two urban Indian settings with standardised patients to compare the quality of tuberculosis care received by men and women. Because standardised patients were not assigned to providers with a random assignment function from a computer program, but instead by a field team blinded to this study, we did tests that confirmed that standardised patients were assigned by gender as good as randomly. We found that men and women do not receive different health-care experiences in terms of quality of care (ie, correct case management). Providers were as likely to correctly manage men as they were women, extending to several management decisions among all provider types.

Implications of all the available evidence
Systematic differences in quality of care are unlikely to be a cause of the observed under-representation of men in tuberculosis notifications in the private sector in urban India.

Methods

Study design
We analysed clinical interactions done by standardised patients portraying four tuberculosis case scenarios at health facilities in the two Indian cities of Mumbai (population 12 million) and urban Patna (2 million).23 Health facilities were representatively selected in each city using random sampling described previously,23 stratified by qualification and pilot tuberculosis programme engagement status. The health facilities included providers across the range of qualifications available in urban India. On one end of this range were chest specialists and providers with Bachelor of Medicine, Bachelor of Surgery (MBBS) degrees. At the other were providers without an MBBS, including practitioners trained in ayurveda, yoga and naturopathy, unani, siddha, and homeopathy, as well as registered medical practitioners and those without any formal medical training at all. In each city, stratified sampling was used to randomly oversample providers enrolled in pilot tuberculosis programmes in the private health sector.

Procedures
Data collection for the study was done between Nov 21, 2014, and Aug 21, 2015, in both cities, as part of a larger study on tuberculosis care among private health-care providers in urban India.21 Details of the standardised patient method used in this study are discussed in our previous publications,20–23 including our validation study of the standardised patient method for tuberculosis in India.22 The validation study22 showed that standardised patients were able to recall clinical encounters accurately, that detection of standardised patients as fake patients was low, and that the method posed no major risks to either providers or the standardised patients themselves.
Standardised patients portrayed four different standardised cases (table 1) developed with the support of a technical advisory group comprised of clinicians, public health experts, economists, anthropologists, and experts in both the Standards for TB Care in India (STCI) and the International Standards of TB Care.26,27 Each standardised patient primarily portrayed only one of the four cases. Several standardised patients worked in both cities, and, in Patna, all standardised patients did some Case 1 interactions, even if this was not their primary presentation. All medicines prescribed or offered to the standardised patients were independently coded and classified. Further details regarding standardised patient recruitment, training, sampling, and data collection are provided in the appendix.

This study was granted clearance by the ethics committees at McGill University Health Centre in Montreal, Canada, and the Institute for Socio-Economic Research on Development and Democracy in New Delhi, India. All the standardised patients were hired as field staff and participated in training and refresher training to mitigate any potentially harmful events, such as injections, invasive tests, and consuming any medicines during encounters (appendix). As described in our previous publication that uses the same data,23 we sought a waiver of provider informed consent based on the research ethics provisions from the Government of Canada Panel on Research Ethics and a study commissioned by the United States Department of Health and Human Services to assess the ethics of standardised patient detection, and (2) no more than minimal risk is associated with participation for the standardised patients or the providers, as reported in our validation study.20 All questionnaires and case scripts are available from the authors upon request.

**Outcomes**

Table 2 shows the outcome measures that we used and how they were measured. Our primary outcome was correct case management (table 1), which is a pre-determined, case-specific outcome benchmarked against the STCI and approved by the technical advisory group convened before standardised patient data collection.26 Secondary outcome measures included the following dimensions of process quality indicators: history questions asked, time spent with the patient, diagnosis provided, medicines given, whether the standardised patient was counselled on treatment, their assessment of whether the environment was private or distracting, whether they liked the doctor, whether they would go to the provider again, and whether the provider seemed knowledgeable and addressed their concerns seriously. We assessed the number and types of medications prescribed to the standardised patients for each of these cases. In addition to assessing the number of different medications prescribed or dispensed in an interaction, we also report the use of broad-spectrum antibiotics, anti-tuberculosis prescriptions, fluoroquinolone antibiotics, and steroids, which can have negative patient-specific or public health consequences. Treatment with fluoroquinolone antibiotics or steroids can mask primary tuberculosis symptoms, leading to a delay in accurate diagnosis.29,30

### Table 1: Standardised patient case scenario descriptions

| Case description | Presentation of patient | Expected correct case management |
|-------------------|-------------------------|----------------------------------|
| Case 1            | Classic case of presumed tuberculosis with 2–3 weeks of cough and fever | Presents with presumptive tuberculosis, for the first time, to a private health-care provider, saying “Doctor, I have cough that is not getting better and some fever too” | Recommendation for sputum testing, chest radiograph, or referral to a public DOTS centre or qualified provider |
| Case 2            | Classic case of presumed tuberculosis in a patient who has had 2–3 weeks of cough and fever. The patient has taken a broad-spectrum antibiotic (amoxicillin) given by another health-care provider for 1 week with no improvement. He also carries an abnormal chest x-ray suggestive of tuberculosis | Presents after an initial, failed (empirical) treatment for symptoms with broad-spectrum antibiotics and a diagnostic chest x-ray, saying “I have cough and fever which is not getting better. I went to a doctor and took the medicines he gave me and have also had an x-ray done.” The chest x-ray and blister pack for the antibiotics are shown if the provider asks | Recommendation for sputum testing, chest radiograph, or referral to a public DOTS centre or qualified provider |
| Case 3            | Chronic cough with a positive sputum smear report for tuberculosis from a public health facility | Presents with evidence of microbiologically confirmed tuberculosis, saying “I am having cough for nearly a month now and also have fever. I visited [the local government hospital] and they gave me some medicines and did a sputum test.” The sputum report is shown if the provider asks | Either referral to a public DOTS centre, a qualified provider or specialist, or (in the case of a qualified private provider) initiation of treatment with standard, four-drug, first-line anti-tuberculosis therapy (isoniazid, rifampicin, pyrazinamide, and ethambutol [the HRZE regimen]) |
| Case 4            | Chronic cough, and, if asked, elaborates a history of previous, incomplete treatment for tuberculosis, which would raise the suspicion of multidrug-resistant tuberculosis | Presents as a previously treated patient with tuberculosis with recurrence of the disease (ie, suspicion of drug resistance), saying “Doctor, I am suffering from a bad cough. One year ago I had got treatment in [the local public hospital], and it had got better. But now I am having cough again” | Recommendation for any drug-susceptibility test (culture, line probe assay, or Xpert MTB/RIF) or referral to a public DOTS centre or qualified provider |

DOTS=Directly observed treatment, short-course.
Table 2: Standardised patient variable descriptions

| Balance variable                                      | Measurement method                      |
|-------------------------------------------------------|-----------------------------------------|
| MBBS provider                                         | Recorded in provider data               |
| Provider younger than 30 years of age                 | Assessed by standardised patient        |
| Provider 30–50 years of age                           | Assessed by standardised patient        |
| Provider older than 50 years of age                   | Assessed by standardised patient        |
| Provider male                                         | Observed by standardised patient        |
| Patients waiting on arrival                           | Observed by standardised patient        |
| Patients waiting on departure                         | Observed by standardised patient        |
| Provider has clinic assistant                         | Observed by standardised patient        |

Process indicator

| Provider used cell phone                              | Observed by standardised patient        |
| Other people in room during interaction               | Observed by standardised patient        |
| Television on during interaction                      | Observed by standardised patient        |
| Essential checklist %                                 | Calculated from standardised patient data|
| Time with provider (min)                              | Measured by standardised patient        |
| Did the provider create a private environment?        | Assessed by standardised patient        |
| Did the provider explain about your illness?          | Assessed by standardised patient        |
| Did the provider explain your treatment plan?         | Assessed by standardised patient        |
| Did you like this doctor?                            | Assessed by standardised patient        |
| Would you go to this doctor again?                    | Assessed by standardised patient        |
| Did the provider seem knowledgeable about your illness| Assessed by standardised patient        |
| Did the provider address your worries seriously?      | Assessed by standardised patient        |
| How would you rate the provider? (1–10)               | Assessed by standardised patient        |

Quality outcome

| Correct management                                    | Calculated from standardised patient data|
| Referred case                                         | Reported by standardised patient        |
| Tuberculosis suspicion                                | Reported by standardised patient        |
| Chest x-ray                                           | Reported by standardised patient        |
| Sputum acid-fast bacillus                             | Reported by standardised patient        |
| Xpert MTB/RIF                                         | Reported by standardised patient        |
| Any medicine                                          | Reported by standardised patient        |
| Polypharmacy                                          | Reported by standardised patient        |
| Anti-tuberculosis treatment                           | Determined by analysis team             |
| Fluoroquinolone                                       | Determined by analysis team             |
| Other antibiotic                                      | Determined by analysis team             |
| Steroids                                              | Determined by analysis team             |

Statistical analysis

A primary concern for the validity of the standardised patient method is that the individuals presenting the cases were not actually ill and, therefore, do not automatically present the correct clinical appearance, despite their training. To ensure that standardised patient presentations were convincing, such that detailed questioning and examination of the standardised patient did not lead providers to conclude that the standardised patients were healthy, we examined associations between case management and checklist completion for each standardised patient.

We did two sets of comparisons to establish that providers visited by women were, on average, identical to providers visited by men presenting the same standardised patient case scenario. In our study, women and men were recruited as standardised patients in all case scenarios in each city. However, fieldwork conditions precluded the explicit random assignment of standardised patients to providers, and the assignment of standardised patients to providers was done in the field by supervisors who were blinded to the fact that we would do a gender analysis. This made it highly unlikely that standardised patients who were men (or women) were assigned to providers with perceived higher or lower quality. Nevertheless, to verify that the assignment was as good as random, we used balance tests and placebo regressions as randomisation tests.

Specifically, we did a randomisation test to establish that we could obtain unbiased estimates of the effect of patient gender on quality of care outcomes by using placebo regressions based on standardised patient gender for each case. The rationale for this randomisation test is as follows: if standardised patient gender was correlated with provider quality within each case, then the providers who saw women present Case 1 should differ from those who saw men present Case 1 only within Case 1. By contrast, all other comparisons between those two groups of providers should be uncorrelated with that assignment. Therefore, for each case, we split the sample into the group of providers who saw any man present the case, and the providers who saw any woman present that case (these groups could overlap). Using these two groups, we ran six placebo regressions comparing how those two groups treated all other potential standardised patient cases (a total of 24 regressions). We report additional information regarding power calculations for such an audit study using standardised patients in the appendix.

We used ordinary least squares regression and logistic regression to assess differences in clinical care processes and case management across standardised patients by gender. In these specifications, we controlled for differences that arose from the study design. These included the location, the case scenario, and whether the provider had an MBBS qualification. We have shown previously that these differences affect the care that is provided and were components of our study design.22 We complemented ordinary least squares regressions with logistic regressions where appropriate, reporting odds ratios by standardised patient gender for dichotomous outcomes and illustrating appropriate CIs for these estimates. We clustered standard errors at the individual standardised patient level when calculating gender differences, and we used inverse-probability-weighted estimates based on our sampling strategy to arrive at city-representative interpretations of our outcome measures. Therefore, reported estimates correspond to the expected average quality of care outcomes and gender differences if providers were chosen at random by a patient from each city with each city contributing equal weight (appendix).23,24 All data analyses were performed with Stata 15.
Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

2602 interactions were done by 24 unannounced standardised patients at 1203 different health-care facilities across two cities (table 3). 1900 (73%) interactions were done by men, who made up 16 of our 24 individual standardised patients (table 3). We describe the findings in three parts: (1) whether assignment of standardised patients in the field produced as good as random allocations of women and men for valid inference; (2) how the objective and subjective patient experience varied between women and men; and (3) how provider case management decisions and quality of care varied between women and men.

We found that increased clinical scrutiny was associated with higher propensity to treat the patient as though they had tuberculosis, which suggests that providers in general were convinced by the presentation of our standardised patients. A 100% completion rate for the essential history question checklist for each case was associated with a 4% (95% CI –7 to 16; p=0.443) change in the likelihood of correct treatment compared with no questions asked, a 14% (2 to 26; p=0.018) increase in the likelihood of giving any medication, an 18% (6 to 29; p=0.0032) increase in the likelihood of any verbal diagnosis, and a 16% (6 to 26; p=0.0011) increase in the likelihood of a verbal tuberculosis diagnosis. Although we cannot totally reject provider response to individually varying standardised patient characteristics on all outcomes, the measured height, weight, and age of the standardised patients jointly had no effect on correct management decisions (p=0.125) and their inclusion as controls does not systematically affect our main results, ruling out confounding due to gender-correlated physical attributes that might prompt clinical conclusions.

We found no differences in the provider’s qualification, age, gender, caseload (measured by queue length at arrival and departure), or the presence of a clinic assistant (table 4). Across the 19 comparisons for which we had sufficient statistical power for logistic regression (five pairs were not computable), one was significant at p<0.1, one at p<0.05, and one at p<0.01 (table 5). The two at p<0.1 and p<0.05 are expected by chance with 19 simultaneous comparisons and are, therefore, rejected as statistically insignificant by Bonferroni multiple-hypothesis, critical value adjustments. The one at p<0.01 occurred in a sample of 35 interactions, and our results are robust to their exclusion. This result is consistent with the assessment that women and men were as good as
randomly assigned to providers during this study and reinforces our conclusion that women and men visited equivalent providers during the fieldwork. Consequently, observed differences in clinical interactions between women and men can be attributed to the gender of the standardised patient rather than variation in undetected provider differences.

We compared process indicators by standardised patient gender as well as the standardised patients’ own subjective experiences of the interactions (table 6). We observed one significant difference between the reported experiences of women and men: providers spent significantly less time with men (8.24 min vs 5.57 min; 95% CI –3.69 to –1.59; p < 0.0001). On all other observed dimensions we found no differences in interactions by standardised patient gender.

Comparisons of the standardised patients’ subjective assessment of quality, however, show that men were less likely to agree that the provider seemed knowledgeable about the illness or that the provider addressed their worries seriously (table 6). Men also rated the providers lower than did women on a subjective scale (table 6). Under the assumption that the men did not have different initial perceptions or attitudes towards the providers, the subjective assessment mirrors the shorter interactions with men.

Additionally, differences occurred in the types of history questions asked of women and men (appendix). Men were more likely to be asked about smoking (with estimates ranging from 12 percentage point difference to 24 percentage point difference, depending on case) and drinking (3–18% percentage point difference, depending

---

**Table 5: Randomisation test across providers who saw differently gendered case presentations**

| Case | Saw a woman presenting | Saw a man presenting | Adjusted odds ratio (regression estimate) | 95% CI | p value |
|------|------------------------|----------------------|------------------------------------------|--------|---------|
|      | Number of interactions | Correct management proportion | Number of interactions | Correct management proportion |          |         |
| Case 1 |                        |                      |                                      |        |         |
| Case 1 (Women) | 50 | 0.70 | 161 | 0.63 | 1.08 | 0.55–2.11 | 0.82 |
| Case 1 (Men) | 44 | 0.66 | 165 | 0.66 | 0.91 | 0.43–1.94 | 0.81 |
| Case 2 (Women) | 39 | 0.33 | 81 | 0.25 | 0.83 | 0.27–2.55 | 0.74 |
| Case 2 (Men) | 48 | 0.25 | 220 | 0.31 | 1.22 | 0.55–2.72 | 0.63 |
| Case 3 (Women) | 37 | 0.22 | 95 | 0.13 | 0.65 | 0.14–2.91 | 0.57 |
| Case 3 (Men) | 46 | 0.17 | 310 | 0.10 | 0.45 | 0.10–1.04 | 0.31 |
| Case 4 (Women) | 46 | 0.17 | 310 | 0.10 | 0.45 | 0.10–1.04 | 0.31 |
| Case 4 (Men) | 46 | 0.17 | 310 | 0.10 | 0.45 | 0.10–1.04 | 0.31 |

For each case scenario, this table shows a test of balance across the providers who saw a man present that case and the providers who saw a woman present that case. For each other gender-case presentation, it assesses whether any significant difference exists between those two groups of providers. The table presents the N, mean correct management proportion, odds ratio, 95% CI, and p value for differences in correct management between those two groups. Reported odds ratios are logistic regression coefficients on the gender of the standardised patient, controlling for city, case scenario, and provider qualification, and standard errors are clustered at the health care facility level.
### Table 6: Differences in interaction process indicators by patient gender

| Women | Mean (SD) | Men | Mean (SD) | Difference | 95% CI       | p value |
|-------|-----------|-----|-----------|------------|--------------|---------|
| Provider used cell phone | 0.13 (0.34) | 0.10 (0.30) | -0.02 | -0.07 to 0.03 | 0.42 |
| Other people in room during interaction | 0.17 (0.37) | 0.14 (0.35) | -0.05 | -0.12 to 0.02 | 0.18 |
| TV on during interaction | 0.03 (0.17) | 0.03 (0.17) | 0.00 | -0.02 to 0.03 | 0.78 |
| Essential checklist % | 0.50 (0.25) | 0.50 (0.25) | -0.01 | -0.12 to 0.11 | 0.93 |
| Time with provider (min) | 8.24 (6.19) | 5.57 (3.60) | -2.64 | -3.69 to -1.59 | <0.0001 |

This table reports estimated differences by gender across individual interactions as observed or assessed by the standardised patient during or after the interaction. It then reports linear differences and 95% CIs and p values for those differences. Linear differences are estimated controlling for city and case scenario and standard errors are clustered at the individual standardised patient level. All variables are binary, except time with provider (expressed in min), the provider rating (1–10), and essential checklist (%).

### Table 7: Differences in quality of care by patient gender

| Women | Mean (SD) | Men | Mean (SD) | Difference | 95% CI       | p value |
|-------|-----------|-----|-----------|------------|--------------|---------|
| Correct management | 0.40 (0.49) | 0.36 (0.48) | 0.01 | -0.02 to 0.05 | 0.47 |
| Case 1 | 0.40 (0.49) | 0.39 (0.49) | 0.05 | -0.06 to 0.16 | 0.32 |
| Case 2 | 0.64 (0.48) | 0.68 (0.47) | 0.04 | -0.04 to 0.12 | 0.15 |
| Case 3 | 0.28 (0.45) | 0.31 (0.46) | -0.01 | -0.06 to 0.04 | 0.63 |
| Case 4 | 0.15 (0.36) | 0.10 (0.30) | -0.03 | -0.07 to 0.01 | 0.14 |
| MBBS provider | 0.47 (0.50) | 0.57 (0.50) | 0.04 | -0.04 to 0.11 | 0.34 |
| Non-MBBS provider | 0.33 (0.47) | 0.21 (0.41) | -0.02 | -0.13 to 0.08 | 0.69 |
| Male provider | 0.39 (0.49) | 0.36 (0.48) | 0.02 | -0.03 to 0.06 | 0.52 |
| Female provider | 0.49 (0.50) | 0.33 (0.47) | -0.05 | -0.18 to 0.08 | 0.41 |
| Patna | 0.31 (0.47) | 0.39 (0.49) | 0.01 | -0.03 to 0.05 | 0.67 |
| Mumbai | 0.50 (0.50) | 0.34 (0.47) | 0.02 | -0.04 to 0.08 | 0.56 |
| Referred case | 0.09 (0.28) | 0.07 (0.26) | 0.01 | -0.04 to 0.06 | 0.64 |
| Tuberculosis suspicion | 0.41 (0.49) | 0.37 (0.48) | 0.04 | -0.07 to 0.14 | 0.47 |
| Chest x-ray | 0.42 (0.49) | 0.40 (0.49) | 0.04 | -0.06 to 0.13 | 0.45 |
| Sputum acid-fast bacillus | 0.21 (0.41) | 0.13 (0.32) | -0.02 | -0.09 to 0.03 | 0.26 |
| Xpert Mtb/RIF | 0.05 (0.22) | 0.04 (0.19) | -0.01 | -0.03 to 0.01 | 0.18 |
| Any medicine | 0.83 (0.38) | 0.87 (0.33) | -0.02 | -0.07 to 0.03 | 0.38 |
| Polypharmacy | 3.29 (2.03) | 3.76 (2.04) | -0.48 | -0.48 to 0.04 | 0.089 |
| Anti-tuberculosis treatment | 0.07 (0.25) | 0.04 (0.19) | -0.01 | -0.06 to 0.04 | 0.68 |
| Fluoroquinolone | 0.12 (0.37) | 0.11 (0.31) | -0.01 | -0.05 to 0.04 | 0.71 |
| Other antibiotic | 0.45 (0.50) | 0.48 (0.50) | -0.06 | -0.14 to 0.02 | 0.16 |
| Steroids | 0.12 (0.32) | 0.17 (0.38) | 0.01 | -0.02 to 0.04 | 0.35 |

This table reports estimated differences by gender across management decisions in individual interactions as determined by the analysis team. It then reports linear differences and 95% CIs and p values for those differences. Linear differences are estimated controlling for city, case scenario, and provider qualification, and standard errors are clustered at the individual standardised patient level. The eight subsets reported under correct management report the results of the correct management difference regression among the individual case scenarios and among subsets of providers by qualification and provider gender. All variables are binary, except polypharmacy (the whole number of medications given).
Articles

on case) habits, whereas women were more likely to be asked about children (16–37 percentage point difference; appendix). Questions related to the tuberculosis diagnosis, however, did not vary systematically across the cases. Only small differences (directionally inconsistent across cases) occurred in the frequencies of essential questions like the duration of cough or whether the cough produces sputum (appendix). For physical examinations, there were small differences, with women more likely to have had their blood pressure taken (appendix).

Overall proportions of correct management were 40% for women and 36% for men (table 7). We detected no differences in this measure of STCI-compliant management or in any other key quality dimensions of care, such as medication use and laboratory testing, between women and men. All estimated differences in correct case management were statistically insignificant, of small absolute magnitude, and in varying directions by case (table 7). Estimates ranged from 3% less correct case management for men relative to women in Case 4, to 5% more in Case 1, highlighting the absence of any systematic difference in correct case management by gender (table 7). These differences remained insignificant in all subsamples—MBBS providers, non-MBBS providers, whether the provider was a man or a woman, or within either study city.

We also found no differences in any of the individual treatment behaviours composing the correct management index, including the decision to refer the case, whether the provider mentioned a suspicion of tuberculosis to the patient, or the choice among various types of tuberculosis testing. Similarly, we find no qualitatively large or statistically significant (at the 5% level) differences in the use of unnecessary medications. The overall absence of any large difference in case management outcomes is not an artifact of sample size or estimation methods—the null effects are precisely estimated with narrow confidence intervals and are robust to hypothesis testing using logistic regression (figure).

Discussion

Our study assessed differences by gender in Indian health providers’ management of tuberculosis patients with identical clinical profiles, using the gold-standard randomised patient method. We sought to understand whether gender-related differences in quality of care could be contributing to the relative under-diagnosis of men (and, therefore, their under-representation in notifications) with active tuberculosis in the general population, as has been found in previously published literature. We demonstrate a general absence of differences in provider behaviour between men and women presenting symptoms of tuberculosis in a high-burden setting with high levels of background gender inequality and large gender differences in both tuberculosis prevalence and notification rates.

Three characteristics of the study provide a unique opportunity to assess gender differences in tuberculosis diagnosis and treatment among Indian health-care providers. First, the study is at-scale across two cities with 2602 patient-provider interactions across 1203 health care providers. Second, the study is representative: in each city, we used a comprehensive list of all private health-care providers to randomly sample providers stratified by city, case scenario, and provider qualification, and standard errors are clustered at the individual standardised patient level. All variables are binary.

| Odds ratio | 95% CI | p value |
|------------|--------|---------|
| Correct management | 1.05 | 0.76–1.45 | 0.7658 |
| Referred case | 1.14 | 0.64–2.02 | 0.6521 |
| Tuberculosis suspicion | 1.08 | 0.56–2.09 | 0.8333 |
| Chest x-ray | 1.02 | 0.61–1.73 | 0.9293 |
| Sputum acid-fast bacilli | 0.84 | 0.59–1.21 | 0.3435 |
| Xpert MTB/RIF | 1.19 | 0.77–1.82 | 0.4335 |
| Amy medicine | 1.02 | 0.73–1.44 | 0.6978 |
| Anti-tuberculosis treatment | 1.38 | 0.64–2.97 | 0.4084 |
| Fluorquinolone | 0.85 | 0.69–1.04 | 0.1199 |
| Other antibiotic | 0.89 | 0.63–1.26 | 0.5114 |
| Steroids | 0.97 | 0.79–1.19 | 0.7514 |

Figure: Differences in quality of care by standardised patient gender

This figure illustrates estimated differences by gender across management decisions in individual interactions as determined by the analysis team. Odds ratios are estimated controlling for city, case scenario, and provider qualification, and standard errors are clustered at the individual standardised patient level. All variables are binary.
provider practice on any quality measures, for any case presentation, or for any level of provider qualification. Additionally, we find no evidence that providers of either gender behave differently when matched with patients of the same gender as themselves. Poor quality of care during the initial clinical evaluation for individuals with suspected tuberculosis, which we have discussed previously, affects women and men equally.32,33,34,35 However, men appear to have received care that involves less provider time and less detailed explanation, and reported significantly worse perceptions of the providers’ knowledge, seriousness, and overall satisfaction.

We cannot measure in our study whether any of these differences would lead men or women with tuberculosis symptoms to have different experiences of patient satisfaction or stigma on average, potentially shaping future care-seeking behaviour and engagement in tuberculosis care, although this is a distinct possibility. Systematic reviews and multisite studies show that in many settings, including India, men are less likely to complete tuberculosis therapy, more likely to die during treatment, and more likely to experience disease recurrence after completing treatment.1,3 As such, understanding whether gender-related differences in the time spent by providers, in the explanations given to patients, and in patient satisfaction remain consistent during subsequent stages of care, and whether these differences contribute to men’s poorer tuberculosis outcomes, is an important area for future research.

The strength of the study design and scale of implementation in the two large cities located in different regions of India, lead us to believe that the lack of systematic gender differences in the management of tuberculosis patients in urban India, including appropriate management, referral rates, choice of diagnostic, or use of unnecessary medication is a key and robust finding. Our null findings covered a broad range of process indicators and quality outcomes and have narrow confidence intervals around zero. Our strong balance and randomisation tests suggest that these estimated null results are unlikely to be confounded by observed or unobserved factors. Although we observed gender differences in history taking and examination, as well as differences in time spent with the patient and subjectively reported satisfaction, these differences appeared to have little consequence for case management.

Nevertheless, our study had several limitations. First, although the study was a population-weighted assessment of average behaviours for these provider types and cities, it was not necessarily statistically representative of the provider mix that patients face if women and men choose to visit different types of providers on average and might not replicate in other settings. Second, in this study, observed practice only reflected what health-care providers did when they came across a completely unknown or new patient seeking medical care in their first visit to the health care provider. Third, this study only covered private practitioners in two urban areas in India.

Additionally, inherent limitations exist in our approach of using standardised patients to assess gender differences. Although each standardised patient visited over 100 facilities on average, the decision to hire 24 individual standardised patients was based on fieldwork logistics and supervision constraints, which has implications for the power and bias in this study (discussed further in the appendix). Specifically, if gender and other disadvantages intersect (eg, care might be demonstrably worse for low-caste men), our study is externally valid only to the extent that these other disadvantages were also represented in our standardised patient selection.

We have compared the characteristics of our standardised patients with profiles of tuberculosis patients who visit private clinics in urban India using NFHS-4, which is a nationally representative sample of 601,509 households.36 Our standardised patients had similar age and education profiles, but no standardised patients were from the lowest wealth quintiles (17·6%), less than primary education (38·7%) and are children or elderly patients (34·9%). Therefore, our study only relates to the experience of 50–60% of patients who are in the middle and above wealth quintile, have secondary or higher education, and are between 18 and 59 years of age.

Standardised patient assignment was not randomised, so the credibility of the differences being due to gender is based on our evidence that the standardised patients were assigned as good as randomly across providers. Additionally, the standardised patient method is designed to provide objective estimates of actual provider behaviour; however, it provides little insight into why we observe gender-related differences in time spent with men. Detailed qualitative and ethnographic studies involving interviews with, and observation of, patients and providers might provide further insights into the social context that shapes these gender disparities in tuberculosis care.37,38

Despite these limitations, the major findings from this study have implications for public health. Our main results suggest that concerns about health-care providers being responsible for gender differences in diagnostic delays are unlikely to be well-founded, though less time and explanation given by providers to men on average could adversely affect outcomes for men in subsequent stages of the care cascade. Our findings should not be taken to imply that neither men nor women experience disease-related stigma or unique challenges in seeking or accessing tuberculosis care, but they do show that they do not face a systematic gender-related difference in care quality from health providers during the initial diagnostic evaluation for tuberculosis.39,40

As we discussed previously in multiple contexts regarding the supply-side of health care, the main cause
for concern is the low overall level of correct management for all patients, and its lack of strong correlation with provider characteristics like qualifications.\textsuperscript{22} For women and men, the average health-care provider did not ask the essential history questions that would lead to a tuberculosis diagnosis, did not mention tuberculosis suspicion to the patient (although this might be generally acceptable due to concerns of stress and stigma or caution ahead of receiving more convincing evidence), and did not order appropriate microbiological tests to diagnose tuberculosis as per the STCI and international recommendations. Instead, more than 80\% of interactions resulted in medicine prescriptions, half of which contained unnecessary antibiotics that do not have a role in tuberculosis care and have negative public health consequences. Prescriptions of fluoroquinolone antibiotics and steroids are particularly worrisome, as they can mask tuberculosis symptoms, leading to delays in diagnosis.\textsuperscript{20}

Given the lack of gender-related differences in quality of care delivered by providers, how can the relative under-notification of men be explained? As described previously, an individual with tuberculosis must traverse three stages to start treatment: care-seeking, diagnostic evaluation, and linkage to treatment. The absence of gender differences in diagnostic evaluation suggests potential barriers for men at the other two stages. Men might be less likely than women to seek care in the first place for their tuberculosis symptoms or to link to tuberculosis treatment after diagnosis in the Indian context. Although systematic reviews have summarised the literature in India on care-seeking, delays in reaching care, and linkage to treatment, they and many of the included studies did not specifically evaluate gender differences for each.\textsuperscript{11,12} Additionally, few of the studies included in those reviews evaluated patients seeking tuberculosis care in the private sector.

As such, research is needed to better understand gender differences across these other stages of the tuberculosis care cascade and to inform policy and programmes. The standardised patient method might be extended to assess gender differences in quality of tuberculosis care in other countries and in India’s public sector. In addition to evaluating the diagnostic workup for tuberculosis, the standardised patient method might also have utility for understanding gender differences in quality of care during tuberculosis treatment initiation. The standardised patient method is less useful for understanding gender differences during later stages of the care cascade that involve longitudinal follow-up, given the risk that standardised patients might be detected by providers. For these later stages, quality of care can potentially be measured using cohort studies to understand gender differences in patient outcomes and provider behaviour patterns, especially with regard to linkage to treatment, treatment completion, and recurrence-free survival. Additionally, future research to understand the issue of gender differentials in tuberculosis case notifications might focus on infection risks, the availability and accessibility of high-quality providers, and the decision making of symptomatic individuals.

Contributors
JD and MP conceptualised the study and acquired funding. BD, AK, SS, VD, JD, and MP designed the methods. AK, RKD, and VD supervised the fieldwork. SS and RS coded medicines and tests. BD, AK, and SS wrote the original draft. All authors reviewed and edited the manuscript.

Declaration of interests
We declare no competing interests.

Data sharing
Individual de-identified interaction data, including data dictionaries, will be available. All variables needed to recreate the results reported in this article will be included, as will the code required to reproduce these results. Data will be available indefinitely upon publication to anyone who wishes to access the data for any purpose. The data and code can be accessed at https://github.com/qutubproject/lancetgh2019.

Acknowledgments
This study was funded by Grand Challenges Canada, the Bill & Melinda Gates Foundation (OPP10919843), and the Knowledge for Change Program at the World Bank. SS is supported by the Center for Operational Research, The Union (Paris, France). RS acknowledges support from a grant from the Bill & Melinda Gates Foundation through the Arcady Group (OPP1134663) and a Doris Duke Clinical Scientist Development Award. MP is a recipient of a Tier I Canada Research Chair from Canadian Institutes of Health Research. We thank Puneet Dewan, Sameer Kunta, Shibul Vijayan, Sirisha Papineni, Nita Jha, Vaibhav Saran, and Anamta Dafty for useful input on implementation, analysis, and interpretation; Rajan Singh, Purshottam, Chinara Singh, Geeta, Devender, Varun Kumar, Anand Kumar, Balbulo, and Charu Nanda, who supervised and implemented the ISERDD field work; all the standardised patients for their dedication and hard work; and Caroline Vadnais for excellent administrative support. We thank Sumit Manandhar of the University of York Centre for Health Economics for constructing and sharing NFHS-4 results. The findings, interpretations, and conclusions expressed here are those of the authors and do not necessarily represent the views of the World Bank, its executive directors, or the governments they represent.

References
1 Horton KC, MacPherson P, Houben RM, White RG, Corbett EL. Sex differences in tuberculosis burden and notifications in low- and middle-income countries: a systematic review and meta-analysis. PLoS Med 2016; 13: e1002199.
2 Borgdorff M, Nagelkerke N, Dye C, Nunn P. Gender and tuberculosis: a comparison of prevalence surveys with notification data to explore sex differences in case detection. Int J Tuberc Lung Dis 2000; 4: 123–32.
3 Horton KC, White RG, Houben RMG. Systematic neglect of men as a key population in tuberculosis. Tuberculosis (Edinb) 2018; 113: 249–53.
4 Mukherjee AN, Karmakar K. Untreated morbidity and demand for healthcare in India: an analysis of national sample survey data. Econ Political Wkly 2008; 43: 71–77.
5 Sundar R, Sharma A. Morbidity and utilisation of healthcare services: a survey of urban poor in Delhi and Chennai. Econ Political Wkly 2002; 37: 4729–40.
6 Nandrap S, Madhiwala N, Sinha R, Jesani A. Women and health care in Mumbai. Mumbai: Centre for Enquiry into Health and Allied Themes, 2001.
7 Saikia N, Moradhav, Bora J.K. Gender difference in health-care expenditure: evidence from India human development survey. PLoS One 2016; 11: e0158532.
8 Central TB Division. TB India report 2018: revised national TB control programme annual status report. New Delhi: Central TB Division, 2018.
9 Balasubramanian R, Garg R, Santha T, et al. Gender disparities in tuberculosis: report from a rural DOTS programme in south India. Int J Tuberc Lung Dis 2004; 8: 323–32.
10 WHO. Global tuberculosis report 2016. Geneva: World Health Organization, 2016.
Subbaraman R, Nathavitharana RR, Satyanarayana S, et al. The tuberculosis cascade of care in India’s public sector: a systematic review and meta-analysis. PloS Med 2016; 13:e1002149.

Sreeramareddy CT, Qin ZZ, Satyanarayana S, Subbaraman R, Pai M. Delays in diagnosis and treatment of pulmonary tuberculosis in India: a systematic review. Int J Tuberc Lung Dis 2014; 18: 255–66.

Lozano R, Naghavi M, Foreman K, et al. Global and national mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2013; 380: 2095–128.

Murray CJL, Ortblad KF, Guinovart C, et al. Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014; 384: 1005–70.

Luck J, Peabody JW, Dresselhaus TR, Lee M, Glassman P. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. Am J Med 2000; 108: 642–49.

Das J, Holla A, Das V, Mohanan M, Talab D, Chan B. In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. Health Aff (Millwood) 2012; 31: 274–84.

Das J, Holla A, Mohpal A, Muralidharan K. Quality and accountability in health care delivery: audit-study evidence from primary care in India. Am Econ Rev 2016; 106: 3765–99.

Rethans J, Drop R, Sturmans F, van der Vleuten C. A method for introducing standardized (simulated) patients into general practice consultations. Br J Gen Pract 1991; 41: 94–96.

Glassman PA, Luck J, O’Gara EM, Peabody JW. Using standardized patients to measure quality: Evidence from the literature and a prospective study. Jt Comm J Qual Improv 2000; 26: 646–53.

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 2013; 13: 1105–13.

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.

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 2014; 30: 122–13.

Kwan Ada, Daniels B, Saria V, et al. Variations in the quality of tuberculosis care in urban India: a cross-sectional, standardized patient study in two cities. PLoS Med 2018; 15: e0206263.

Christian CS, Gerdhuma UG, Hompashe D, Smith A, Burger R. Measuring quality gaps in TB screening in South Africa using standardised patient analysis. Int J Environ Res Public Health 2018; 15: E729.

Office of the Registrar General Census Commissioner, India. Census of India 2011: Provisional population totals. Indian Census Bureau, 2011. http://censusindia.gov.in/2011-prov-results/index.html (accessed Jan 31, 2013).

WHO. Standards for TB care in India. Geneva: World Health Organization, 2014.

Hopewell PC, Pai M, Malher D, Uplekar M, Raviglione MC. International standards for tuberculosis care. Lancet Infect Dis 2006; 6: 719–25.

Rhodes KV, Miller FG. Simulated patient studies: an ethical analysis. Milbank Q 2012; 90: 706–24.

Hogan CA, Puri L, Gore G, Pai M. Impact of fluoroquinolone treatment on delay of tuberculosis diagnosis: a systematic review and meta-analysis. J Clin Tuberc Other Mycobact Dis 2017; 1: 1–7.

Chen T-C, Lu P-L, Lin C-Y, Lin W-R, Chen Y-H. Fluoroquinolones are associated with delayed treatment and resistance in tuberculosis: a systematic review and meta-analysis. Int J Infect Dis 2015; 15: e211–16.

Hosek J. An introduction to estimation with choice-based sample data. Santa Monica, CA: RAND Corporation, 1979.

Kapoor SK, Raman AV, Sachdeva KS, Satyanarayana S. How did the TB patients reach DOTS services in Delhi? A study of patient treatment seeking behavior. PLoS One 2012; 7: e24258.

Tamhane A, Ambe G, Vermund SH, Kohler CL, Karande A, Sathiakumar N. Pulmonary tuberculosis in Mumbai, India: factors responsible for patient and treatment delays. Int J Prev Med 2012; 3: 569–80.

Satyanarayana S, Subbaraman R, Shete P, et al. Quality of tuberculosis care in India: a systematic review. Int J Tuberc Lung Dis 2013; 19: 751–63.

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: 1261–68.

Watt C, Squire S. A systematic review of risk factors for death in adults during and after tuberculosis treatment. Int J Tuberc Lung Dis 2011; 15: 871–85.

Velayutham B, Chadha VK, Singla N, et al. Recurrence of tuberculosis among newly diagnosed sputum positive pulmonary tuberculosis patients treated under The Revised National Tuberculosis Control Programme, India: a multi-centre prospective study. PLoS One 2018; 13: e0200150.

International Institute for Population Sciences, ICF. National Family Health Survey (NFHS-4), 2015–16. India. Mumbai: International Institute for Population Sciences, 2017.

McDowell A, Pai M. Alternative medicine: an ethnographic study of how practitioners of Indian medical systems manage TB in Mumbai. Trans R Soc Trop Med Hyg 2016; 110: 192–98.

McDowell A, Pai M. Treatment as diagnosis and diagnosis as treatment: empirical management of presumptive prospective tuberculosis in India. Int J Tuberc Lung Dis 2016; 20: 536–43.

Thorson A, Hua N, Long N, Allebeck P, Diwan V. Do women with tuberculosis have a lower likelihood of getting diagnosed? Prevalence and case detection of sputum smear positive pulmonary TB, a population-based study from Vietnam. J Clin Epidemiol 2004; 57: 398–402.

Atre S, Kudale A, Morankar S, Goswinit D, Weiss M. Gender and community views of stigma and tuberculosis in rural Maharashtra, India. Glob Public Health 2011; 6: 56–71.