

# Primary care quality and provider disparities in China: a standardized-patient-based study

Mian Huang,<sup>a</sup> Scott Rozelle,<sup>c</sup> Yiming Cao,<sup>d</sup> Jian Wang,<sup>b</sup> Zhang Zhang,<sup>e</sup> Zhijie Duan,<sup>e</sup> Shuyi Song,<sup>e</sup> and Sean Sylvia<sup>e,\*</sup>

<sup>a</sup>School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

<sup>b</sup>Dongfureng Institute of Economic and Social Development, Wuhan University, Wuhan, China

<sup>c</sup>Stanford Center on China's Economy and Institutions, Stanford University, Stanford, USA

<sup>d</sup>Faculty of Business and Economics, the University of Hong Kong, Hong Kong SAR, China

<sup>e</sup>Department of Health Policy and Management, University of North Carolina at Chapel Hill, Chapel Hill, USA



## Summary

**Background** Primary health care is the foundation of high-performing health systems. Achieving an improved primary care system requires a thorough understanding of the current quality of care among various providers within the system. As the world's largest developing country, China has made significant investments in primary care over the past decade. This study evaluates the quality of primary care across different provider types in China, offering in-sights for enhancing China's primary care system.

**Methods** We merged data from four standardized patient (SP) research projects to compare the quality of five major primary care providers in China: rural clinics, county hospitals, migrant clinics, urban community health centers (CHCs), and online platforms. We evaluated quality of care across process quality (e.g., checklist completion), diagnosis quality (e.g., diagnostic accuracy), and case management (e.g., correct medication), employing multiple regression analyses to explore quality differences by provider type, and their associations with physician characteristics.

**Findings** We document a poor quality of primary care in China, with no-table disparities across different providers. CHCs emerge as relatively reliable primary care providers in terms of process quality, diagnostic accuracy, and correct medication prescriptions. Online platforms outpace rural clinics, county hospitals, and migrant clinics in many areas, showcasing their potential to enhance access to quality healthcare resources in under-resourced rural regions. We observe a positive association between the qualifications of physicians and the quality of primary care, underscoring the necessity for a greater presence of more highly qualified practitioners.

**Interpretation** Primary care quality in China varies greatly among providers, reflecting inequalities in healthcare access. While online platforms indicate potential for improving care in under-resourced areas, their high referral rates suggest they cannot completely substitute traditional care. The findings emphasize the need for more qualified practitioners and stringent regulation to enhance care quality and reduce unnecessary treatments.

**Funding** No funders had a role in the study design, data collection, data analysis, data interpretation, or writing of the report. We have acknowledged this in the revised manuscript.

**Copyright** © 2024 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** Primary care quality; Standardized patient; China

## Introduction

Primary health care serves as the cornerstone of high-performing health systems, a fact universally acknowledged and particularly emphasized in many developing countries.<sup>1,2</sup> Attaining an enhanced primary care system necessitates a comprehensive understanding of the present quality of care among

various providers within the existing system. This information is pivotal for the effectiveness and efficiency of policy decisions as it unveils existing system weaknesses, highlights potential areas for improvement, and could more productively steer the allocation of resources and funding. Consequently, understanding the current state of primary care is a vital prerequisite

\*Corresponding author.

E-mail address: [sean.sylvia@unc.edu](mailto:sean.sylvia@unc.edu) (S. Sylvia).

### Research in context

#### Evidence before this study

Before this study, research on comparing primary care quality among different providers in China was scarce. In June 2022, we conducted a comprehensive search across PubMed, Google Scholar, Web of Science, and the China National Knowledge Infrastructure. Using keywords like “primary care”, “care quality”, “care disparities”, and “telemedicine”, we identified limited studies comparing primary care quality across diverse providers in China. Notably, no study had systematically examined primary care quality across rural, urban, and online platforms using a consistent and objective method.

#### Added value of this study

This study is the first to provide a broad comparison of primary care providers across rural, urban, and online settings in China using a consistent and objective method. It is also the first to utilize pooled data from unannounced standardized patient (SP) projects for assessing China’s primary care quality. This method involves the use of trained individuals to present specific disease cases in a standardized manner to unsuspecting care providers, which is more objective compared to recall-based surveys and enables comparisons across different modalities. Our findings reveal a concerning overall poor quality of primary care, with

significant quality disparities among providers. Urban community health centers (CHCs) were relatively more reliable in process quality, diagnostic accuracy, and correct prescriptions. Online platforms, although excelling over rural clinics, county hospitals, and migrant clinics, are limited in their capacity to fully supplant traditional care. A positive correlation between physician qualifications and primary care quality was also observed, highlighting the need for more skilled practitioners. Additionally, our study highlights the critical need for stricter regulations, particularly in response to the low rate of appropriate prescription practices observed across providers.

#### Implications of all the available evidence

The study’s findings call for urgent policy measures to improve primary care quality, particularly in rural and migrant clinics, to achieve equitable healthcare access across China. The crucial role of physician quality in these disparities in primary care quality underscores the need for more qualified practitioners. Telemedicine platforms, as a supplement or alternative, could improve care quality for both rural and urban populations. Moreover, a balanced compensation system and more rigorous regulations are essential to guide both on-line and offline providers towards reducing unnecessary care and prescriptions.

for driving meaningful and productive transformations in health care policy.

Obtaining accurate information on primary care quality in developing countries is a challenging task for two reasons. First, the quality of care may vary substantially across sectors for a country in the process of rapid urbanization and structural transformation. Second, evaluating the quality of care is inherently challenging due to the presence of recall bias, the incompleteness of medical records for chart abstraction, and the difficulty in ensuring comparability across different care settings and patient populations. Such complexities underscore the importance of employing robust and reliable methods to assess primary care quality, especially in contexts with diverse health care environments and rapidly evolving medical practices.

As the world’s largest developing country, China has invested substantially in primary care since its healthcare reform that was initiated in 2009, which has led to a tenfold increase in primary health care funding, jumping from 19 billion RMB in 2008 to a staggering 197 billion RMB in 2018.<sup>3,4</sup> The commitment to improve the healthcare system was significantly reinforced by the introduction of the Healthy China 2030 Plan in 2016, focusing on enhancing the primary healthcare sector.<sup>5,6</sup> Despite these policy initiatives and their significant potential, the current state of healthcare quality in China, as well as how it varies among different care providers,

is not well understood. Understanding this is crucial not only for policymaking within China but also offers valuable insights for other developing countries facing similar challenges.

In this research, we have merged data from four standardized patient (SP) projects that we and our collaborators carried out earlier. Our goal is to examine the quality of primary healthcare across different types of providers and identify what might have influenced the difference. These SP projects used trained people to act out pre-determined, standardized cases for doctors, thus avoiding problems with measurement bias and missing information. This approach offers an objective and unbiased assessment of primary care quality, and allows comparison across different modalities. The main advantage of the SP method is that it evaluates healthcare providers realistically without them knowing they’re being assessed. This way, they respond to the same standard cases in different settings.<sup>7–12</sup> Thus, the outcomes of the evaluation are generally regarded as reliable and consistent across the different healthcare settings.<sup>8,13</sup>

This study concentrates on five types of primary care providers: rural clinics, county hospitals, migrant clinics, urban community health centers (CHCs), and online platforms. These five categories cover the main types of primary care in both rural and urban settings. This includes services within the formal healthcare

system and also covers additional and nontraditional care options that many people seek out. Rural clinics, which consist of township health centers and village clinics, along with county hospitals, serve as the three tiers of China's rural health system. In urban settings, Community Health Centers (CHCs) act as the official primary care providers. Although these three official providers are designed to serve distinct population groups,<sup>14–17</sup> patients can, in practice, choose to access any of these providers freely due to the absence of a gatekeeping general practitioner system. Migrant clinics are private health care centers that operate outside the official system. They provide medical assistance to rural people who have relocated to urban areas but can't access Community Health Centers because they don't have medical insurance. This is often due to restrictions tied to their household registration (Hukou) status. Finally, we look at telemedicine, which has grown quickly as an option, especially after the COVID-19 pandemic. As a new way to provide care, telemedicine could supplement traditional care providers. It has the unproven possibility to transform how we access primary health care.<sup>5,18–20</sup>

This study is the first to offer a systematic comparison of the quality of care among various primary care providers in China, using a consistent and objective approach. Previous studies have either discussed the unequal access to healthcare services without systematically comparing their quality<sup>1,5,6</sup> or have used the objective standardized patient method to evaluate the quality of a subset of these providers.

In terms of the unequal access to healthcare services, it has been noted that 80% of China's healthcare resources are concentrated in cities, primarily within megacities.<sup>21–23</sup> Similarly, the qualifications of physicians in Township Health Centers (THCs) are significantly lower than their counterparts in urban Community Health Centers (CHCs).<sup>5</sup> Although these findings suggest the potential quality disparities within the healthcare system, there is still an absence of evidence that directly compare primary care quality across these different providers.

Regarding the objective evaluation of primary care quality, several studies have employed the standardized

patient method to assess the quality of care in specific provider types, such as village clinics.<sup>24–26</sup> The study most closely related to our research is Sylvia et al. (2017), which found significantly lower correct management rates in Village Clinics (VCs, 28%) and Township Health Centers (THCs, 38%) compared to county hospitals (90%).<sup>7</sup> However, Sylvia et al. (2017) primarily focused on rural area providers and examined only a single disease case. In contrast, Su et al. (2021) assessed primary care quality between public and private providers in urban settings.<sup>17</sup> Our study builds on previous research by comparing a wider range of healthcare providers, including those in rural and urban areas, as well as migrant clinics and online services that operate outside the official system. This gives a fuller picture of the primary care options available to people in China. Additionally, by including a broader variety of diseases, the findings of our study are more representative and applicable to a wider range of scenarios. This comprehensive method allows for a more detailed and nuanced grasp of the differences in the quality of primary care within China's healthcare system.

## Methods

### Sampling and data collection

The data used for this study were aggregated from four standardized patient research projects conducted by the authors and collaborators in China between 2015 and 2019.<sup>7,17,27,28</sup> This pooled dataset covers five types of primary care providers: rural clinics, county hospitals, migrant clinics, community health centers (CHCs), and online platforms. Throughout this paper, we refer to these five types of institutions/platforms as “providers” and the individual doctors who offer the service as “physicians.” [Table 1](#) summarizes the time, location, facilities included, number of interactions, and disease cases of the four SP research projects from which we obtained our data. [Supplementary Materials B.1 and B.2](#) offer detailed information on the sampling and data collection methods for each project. The selection of sites included providers from migrant and urban areas

Provider	SP project	Facilities included	Year of data collection	Location	Number of interactions	Cases (number)
Rural	3-Provinces SP Project	209 VCs 209 THCs	2015	3 prefectures in Shaanxi, Sichuan, Anhui	909	diarrhea (285), unstable angina (315), TB (309)
County		21 county hospitals				
Migrant	Migrant SP Project	48 migrant clinics	2015	one prefecture in Shaanxi	155	diarrhea (51), unstable angina (51), TB (53)
Urban	Xi'an CHCs SP Project	63 CHCs	2017	urban Xi'an, Shaanxi	78	unstable angina (78)
Online	Online SP Project	36 identified DTC telemedicine platforms	2019	online	160	diarrhea (27), unstable angina (63), TB (70)

Source: Metadata of the four SP programs conducted by the authors and their collaborators between 2015 and 2019.

**Table 1: Sample of standardized patient studies.**

in Shaanxi, as well as those from rural and county areas in Shaanxi, Sichuan, and Anhui. The online project was not restricted to any specific region.

Since the data from various providers were gathered from four existing research projects, it's crucial to assess how comparable the data is across these projects. First, it is important to highlight that the same research team carried out all these studies using the standardized patient method. This approach's key benefit is that it standardizes patients and their case presentations, allowing us to compare visits across different modalities as accurately as possible. This eliminates confounding factors, such as variations in patients' or users' characteristics and their ability in accurately describing diseases across different providers. In subsection, we discuss the details of the recruitment, training, and implementation of the SPs that may vary across research projects and explore how these differences may affect the interpretations of our findings.

Second, we select three diseases that are most comparable across different projects: unstable angina, viral diarrhea, and tuberculosis. We chose these diseases for three main reasons. First, they are widespread across China, which makes them good indicators of quality among various healthcare providers. Second, they typically don't show visible symptoms and are unlikely to require invasive tests during consultations. This simplicity ensures that standardized patients (SPs) can consistently present these conditions, whether in person or during remote consultations. Third, in all four projects, we applied these diseases across various types of healthcare providers. This method allows us to assess care quality while controlling for the disease. The sole exception is in Community Health Centers (CHC), where only unstable angina was included during the visits.

Additionally, we have aligned the definitions and measurement methods for variables collected across different projects. This ensures that each variable included in our study is measured consistently and can be directly compared with one another. We have outlined the specifics of these definitions and the standardization process in subsection.

We recognize that the data for various providers come from different provinces and were collected in different years. Yet, it's important to mention that these provinces share similarities in several key socioeconomic factors. These include the level of economic development (measured by GDP per capita), urbanization rate, population age distribution, literacy levels, and per capita income. Moreover, there weren't significant changes in public health expenditure during the time these Standardized Patient (SP) projects were carried out. The similarities in these fundamental variables indicate that the data from these projects are comparable enough to offer insightful information about the differences in quality of care among various providers,

despite the differences in location and time of data collection.

### Standardized patients

As we mentioned before, we gathered our data from four SP projects. Each project includes aspects tailored to its specific context, which can differ from one project to another. This subsection explains the consistent elements as well as those that can vary across projects, and discusses their implications for interpreting our findings.

#### Recruitment

In every standardized patient (SP) project, we hired standardized patients using consistent criteria: a) they must be in excellent health and free from confounding symptoms; b) they cannot be physicians, to avoid biases from having medical knowledge; c) they need to have good intelligence, memory, and communication skills; and d) their age must match the requirements of the specified cases.<sup>9,29</sup> For visits to rural clinics, county hospitals, migrant clinics, and urban health centers in person, we hired local SPs. This was to ensure they matched the patient population in accent and local norms. For online consultations on telemedicine platforms, SPs were selected from Sun Yat-Sen University in Guangdong province. Sixty-three Standardized Patients (SPs) took part in visits to rural clinics, county hospitals, and migrant clinics, with this same group being shared among these providers. Twelve SPs were involved in visits to urban community health centers, while a different group of fifty-two SPs participated in consultations on online platforms.

#### Training

All SPs underwent rigorous centralized training led by a team of researchers and medical consultants. This training aimed to prepare the SPs to consistently and discreetly act out their assigned disease cases during interactions with healthcare providers. Medical experts provided in-depth explanations of the symptoms associated with the designated illnesses and discussed how real patients typically behave and present themselves. After the training, evaluations were conducted to confirm that the SPs: a) consistently and accurately depicted the diseases; b) correctly remembered their interactions with healthcare providers, which included recording conversations with a hidden device to review the exchange; and c) remained undetected by providers.

#### Procedure of SP visits

All SPs involved were randomly allocated to different healthcare facilities and online platforms. They followed the same procedures a real patient would adhere to during a walk-in consultation at offline facilities or when scheduling appointments on online platforms. Upon consultation, SPs presented the primary symptom(s)

associated with the disease case they were portraying (e.g., fever and cough for TB; chest pain for angina). For the case of childhood diarrhea, SPs introduced themselves as parents of an absent child suffering from the symptoms. SPs adhered to a predetermined script when answering provider inquiries.

Consultation fees were remitted at the end of in-person visits and at the appointment scheduling stage for online consultations. SPs visiting physical facilities also purchased all prescribed medications, which are commonly sold directly by China's healthcare providers.

Interactions between the SPs and the providers were discreetly recorded, using a concealed device for in-person visits and screen recording software for online consultations. Post-consultation, SPs were debriefed using a structured questionnaire, with their responses cross-verified against the recorded interactions to ensure accuracy and impartiality. Section 2.2 of the [Supplementary Material](#) offers further details, including a link to the Mandarin SP case script with an English translation.

### Measurements

This study considers quality measures and physician characteristics consistently gathered in all four SP research projects. We have standardized the definitions and measurements of these variables. This ensures that every variable included in this study is coded uniformly and is comparable across different provider types.

#### Primary care quality measures

We assess care quality in three key dimensions: 1) process quality, 2) diagnostic accuracy, and 3) case management. All four SP research projects followed the same protocol created by the authors and collaborators to evaluate care quality. Where the original studies' definitions and standards slightly differed, we standardized the coding of those variables to guarantee they could be compared across various projects.

First, process quality was gauged by the physician's completion rate of the recommended questions and examinations that they should ideally perform during each consultation, as stipulated on a predefined checklist. This checklist, adapted for use in China, was based on Das et al. (2016) and revised in line with national and international guidelines by Sylvia et al. (2017).<sup>7,30</sup> See [Table A4](#). This measure is applied consistently across all providers in the four SP projects.

Second, the quality of diagnoses was evaluated in accordance with both China's national standards and international benchmarks. This aspect of quality was measured by a binary variable, assigned a value of one if physicians provided any of the correct diagnoses as per predetermined criteria. The detailed definitions are summarized in [Table A1](#) and are adopted consistently across projects.

Third, we evaluate case management in a few different ways. First, we look at referrals. For situations

like TB and angina, referring patients to an upper-level facility is the right action, but it's not appropriate for diarrhea, which requires immediate treatment. Our second criterion is medication prescriptions. We check if any medication was prescribed, whether the prescription was appropriate, and if antibiotics were unnecessarily prescribed (in all cases in our study, antibiotics were never appropriate). In addition, we construct a summary measure that indicates if the case management was handled correctly. We adopt a broad definition of correctness here, considering the management correct if the patient was appropriately referred, received the proper medication, or underwent the correct examinations (for example, through X-ray testing for TB).

#### Physician characteristics

Our analysis considers key characteristics of physicians, such as their gender, age, and level of certification. One issue we encounter is that certifications were defined differently across various research projects. To solve this, we have simplified the classification of physician titles into three main groups according to their certification levels: a) physicians with a practicing physician certificate or higher; b) physicians with an assistant practicing physician certificate; and c) physicians with a rural physician certificate. Out of these, the "Practicing Physician" certification ranks the highest, followed by "Assistant Practicing Physician" and "Rural Physician" certifications. It's important to note that the certifications for practicing and assistant practicing physicians are provided by the national Ministry of Health and are recognized across the country. On the other hand, rural physician certifications are given out after passing local exams at the provincial or municipal level, so they are only valid in specific regions. In addition, we have created a separate category for titles that do not fit into these groups or when this information is missing.

#### Statistical analysis

We employ both summary statistics and regression analyses to assess the quality of primary care and its variation across different providers. We consider the measures in terms of process quality, diagnostic accuracy, and various components of case management. For the summary statistics presented in [Table 2](#), we calculate the expected value and 95% confidence intervals for each outcome variable. This is done for the entire sample and also separately for each type of provider. For non-binary variables such as checklist completion rates, we also report its standard deviation, median, and interquartile range (IQR). We use *t*-tests to compare the mean values across groups.

We then use the regression analysis to estimate the differences in care quality across provider types. The advantage of this approach is that it allows us to control

	Full sample	Rural	County	Migrant	CHC	Online
	Mean (SD)/n (%)	95% CI				
<b>Process quality</b>						
Proportion of recommended questions and examinations	0.214 (0.121)	0.206 (0.114)	0.217 (0.137)	0.180 (0.093)	0.305 (0.160)	0.250 (0.122)
	[0.21–0.22]	[0.20–0.21]	[0.18–0.25]	[0.17–0.19]	[0.27–0.34]	[0.23–0.27]
Median	0.19	0.19	0.19	0.18	0.40	0.24
IQR	(0.12–0.28)	(0.12–0.28)	(0.12–0.29)	(0.12–0.24)	(0.20–0.40)	(0.14–0.35)
<b>Diagnosis quality</b>						
Correct diagnosis	232 (17.8%)	110 (13.0%)	12 (19.4%)	13 (8.4%)	33 (42.3%)	64 (40.0%)
	[16%–20%]	[11%–15%]	[9%–29%]	[4%–13%]	[31%–53%]	[32%–48%]
<b>Case management</b>						
Referred patients	488 (37.5%)	263 (31.1%)	1 (1.6%)	60 (38.7%)	20 (25.6%)	144 (90.0%)
	[35%–40%]	[28%–34%]	[–2% to 5%]	[31%–46%]	[16%–35%]	[85%–95%]
<b>Medication</b>						
Medications prescribed	678 (52.1%)	493 (58.2%)	28 (45.2%)	81 (52.3%)	17 (21.8%)	59 (36.9%)
	[49%–55%]	[55%–62%]	[33%–58%]	[44%–60%]	[13%–31%]	[29%–44%]
Correct medications, if any	63 (9.3%)	29 (5.9%)	2 (7.1%)	2 (2.5%)	13 (76.5%)	17 (28.8%)
	[7%–11%]	[4%–8%]	[–3% to 17%]	[–1% to 6%]	[56%–97%]	[17%–40%]
Correct medications	63 (4.8%)	29 (3.4%)	2 (3.2%)	2 (1.3%)	13 (16.7%)	17 (10.6%)
	[4%–6%]	[2%–5%]	[–1% to 8%]	[0%–3%]	[8%–25%]	[6%–15%]
Unnecessary antibiotics prescribed, if any	336 (50.8%)	262 (53.1%)	6 (21.4%)	50 (61.7%)		18 (30.5%)
	[47%–55%]	[49%–58%]	[6%–37%]	[51%–72%]		[19%–42%]
Unnecessary antibiotics prescribed	336 (25.8%)	262 (30.9%)	6 (9.7%)	50 (32.3%)		18 (11.3%)
	[23%–28%]	[28%–34%]	[2%–17%]	[25%–40%]		[6%–16%]
Correct case management	594 (45.6%)	341 (40.3%)	26 (41.9%)	61 (39.4%)	29 (37.2%)	137 (85.6%)
	[43%–48%]	[37%–44%]	[30%–54%]	[32%–47%]	[26%–48%]	[80%–91%]
Observations	1302	847	62	155	78	160

Note: Data are mean (SD) for continuous variables and n (%) for binary variables. Process quality for rural and urban providers are measured as the proportion of recommended questions and examinations. Online process quality is measured as the proportion of recommended questions.

**Table 2: Summary of quality outcomes.**

for diseases, which might be a confounding factor to the differences across provider types.

The regression takes the following form:

$$y_i = \alpha + \beta_1 I_i^{County} + \beta_2 I_i^{Migrant} + \beta_3 I_i^{CHC} + \beta_4 I_i^{Online} + \delta_1 I_i^D + \delta_2 I_i^{TB} + \varepsilon \quad (1)$$

where  $y_i$  is one of the set of quality measures,  $I_i^{County}$  is a binary indicator of county hospital visits,  $I_i^{Migrant}$  migrant clinic visits,  $I_i^{CHC}$  CHC visits, and  $I_i^{Online}$  online platform visits. Visits to rural clinics are designated as the base group.  $I_i^D$  and  $I_i^{TB}$  are the binary indicators that control for the differences across diseases (treating unstable angina as the base disease group). For numerical quality indicators like checklist completion rates, we apply the ordinary least squares (OLS) estimator. For binary quality metrics, such as correct management, we use logistic regression models. Table 3 reports the estimated coefficients from OLS and the average marginal effects in terms of proportion difference from logistic models.<sup>5</sup> Observations associated with categorical variables that perfectly predicted the outcome were excluded from this

analysis by the statistical package to prevent the perfect separation problem in logistic regressions. We use cluster-robust standard errors, which does not make any assumptions about independence within a cluster.<sup>31</sup> We define clusters at the relevant administrative level, which varies across different standardized patient (SP) projects. For rural clinics and county hospitals, both sourced from the “3-Province SP project,” the relevant administrative level is the county. For migrant clinics, it is a migrant community. For community health centers (CHCs), the relevant administrative level is an urban district of the city (equivalent to a county of a prefecture/city). For online providers, the relevant administrative level is the telemedicine platform. The normality assumption for OLS is met asymptotically due to our large sample size, which allows the sampling distribution of the estimates to approximate normality, as per the Central Limit Theorem. The linearity assumption is satisfied because all predictors in the model are binary variables, ensuring a correctly specified relationship between the predictors and the log-odds of the outcome.

We next present the summary statistics of physician characteristics (age, gender, and title) for the entire



	Proportion of recommended questions and examinations	Correct diagnosis	Referred patients	Medications prescribed	Correct medications, if any	Correct medications	Unnecessary antibiotics prescribed, if any	Unnecessary antibiotics prescribed	Correct management
<b>County</b>	0.011	0.073	-0.337	-0.130	-0.015	-0.004	-0.223	-0.198	0.020
P-value	(0.5316)	(0.1345)	(<0.0001)	(0.0571)	(0.8246)	(0.9223)	(0.0049)	(<0.0001)	(0.5473)
95% CI	[-0.025 to 0.047]	[-0.023 to 0.168]	[-0.421 to -0.253]	[-0.263 to 0.004]	[-0.144 to 0.115]	[-0.094 to 0.085]	[-0.378 to -0.068]	[-0.260 to -0.135]	[-0.045 to 0.085]
<b>Migrant</b>	-0.025	-0.057	0.078	-0.066	-0.072	-0.047	0.076	0.008	-0.001
	(0.0233)	(0.0674)	(0.0244)	(0.1544)	(0.1607)	(0.0300)	(0.1132)	(0.7867)	(0.9781)
	[-0.046 to -0.003]	[-0.118 to 0.004]	[0.010-0.145]	[-0.156 to 0.025]	[-0.174 to 0.029]	[-0.089 to -0.005]	[-0.018 to 0.169]	[-0.048 to 0.063]	[-0.074 to 0.072]
<b>CHCs</b>	0.065	0.188	-0.200	-0.118	0.487	0.157			-0.193
	(0.0013)	(<0.0001)	(<0.0001)	(0.1187)	(0.0002)	(0.0010)			(<0.0001)
	[0.026-0.104]	[0.116-0.260]	[-0.257 to -0.144]	[-0.266 to 0.030]	[0.232-0.743]	[0.064-0.251]			[-0.266 to -0.119]
<b>Online</b>	0.046	0.252	0.552	-0.180	0.395	0.154	-0.239	-0.203	0.423
	(0.0001)	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)	(0.0004)	(<0.0001)	(<0.0001)	(<0.0001)
	[0.023-0.069]	[0.168-0.337]	[0.486-0.617]	[-0.266 to -0.094]	[0.259-0.531]	[0.069-0.239]	[-0.330 to -0.148]	[-0.247 to -0.159]	[0.356-0.490]
<b>Disease cases</b>									
Diarrhea	-0.034	-0.119	-0.375	0.381	-0.114	0.032	0.184	0.252	-0.402
	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)	(0.0057)	(0.2425)	(0.0004)	(<0.0001)	(<0.0001)
	[-0.050 to -0.018]	[-0.170 to -0.068]	[-0.417 to -0.333]	[0.322-0.440]	[-0.194 to -0.033]	[-0.022 to 0.086]	[0.082-0.286]	[0.190-0.314]	[-0.466 to -0.337]
TB	-0.071	-0.078	-0.268	0.297			0.551	0.506	-0.189
	(<0.0001)	(0.0015)	(<0.0001)	(<0.0001)			(<0.0001)	(<0.0001)	(<0.0001)
	[-0.084 to -0.058]	[-0.126 to -0.030]	[-0.308 to -0.228]	[0.251-0.342]			[0.484-0.617]	[0.464-0.549]	[-0.235 to -0.142]
Observations	1302	1302	1302	1302	407	870	661	1224	1302

Note: Data are average marginal effects, p-values, and 95% confidence intervals. Average marginal effects are calculated as the OLS coefficients for numerical variables and mean proportion difference from logistic regressions for binary variables. We compute the cluster-robust standard errors. We define clusters at the relevant administrative level, which corresponds to a county for rural clinics and county hospitals, a migrant community for migrant clinics, an urban district of the city (equivalent to a county of a prefecture/city) for CHCs, and a platform for telemedicine platforms. Observations with perfect separation were excluded from the logistic regression analysis to ensure valid parameter estimates.

**Table 3: Association between quality outcomes and provider types.**

sample and separately for each type of provider. The mean of each characteristic along with its 95% confidence interval is visually represented in Fig. 1.

To understand how the differences in physician characteristics are associated with quality outcomes, we run regressions of each quality outcome on the set of physician characteristics (controlling for diseases) and report the estimation results in Table 4.

Our final analysis explores how the differences in physician characteristics may be associated with quality outcomes. We do so by regressing each quality outcome on the set of physician characteristics, while controlling for disease-case fixed effects.

The regression takes the following form:

$$y_i = \alpha + \beta_1 I_i^{Male} + \beta_2 I_i^{Age \geq 50} + \beta_3 I_i^{AssCert} + \beta_4 I_i^{RuralCert} + \beta_5 I_i^{OtherCert} + \delta_1 I_i^D + \delta_2 I_i^{TB} + \varepsilon \quad (2)$$

where  $I_i^{Male}$  is a binary indicator for male, and  $I_i^{Age \geq 50}$  is a binary indicator for age over 50 years.  $I_i^{AssCert}$ ,  $I_i^{RuralCert}$ , and  $I_i^{OtherCert}$  are binary indicators for physicians with Assistant Practicing Physician Certificate, Rural Physician Certificate, and other or missing certificates. Physicians with Practicing Physician Certificate or higher serves as the base group. As before, we implement the OLS estimators for continuous outcomes and logistic regressions for binary outcomes. Table 4 reports the estimated coefficients from OLS and the average

marginal effects in terms of proportion difference from logistic models. As in previous analyses, we use cluster-robust standard errors, which does not make any assumptions about independence within a cluster. The normality assumption for OLS is met asymptotically due to our large sample size, which allows the sampling distribution of the estimates to approximate normality, as per the Central Limit Theorem. The linearity assumption is satisfied because all predictors in the model are binary variables, ensuring a correctly specified relationship between the predictors and the log-odds of the outcome.

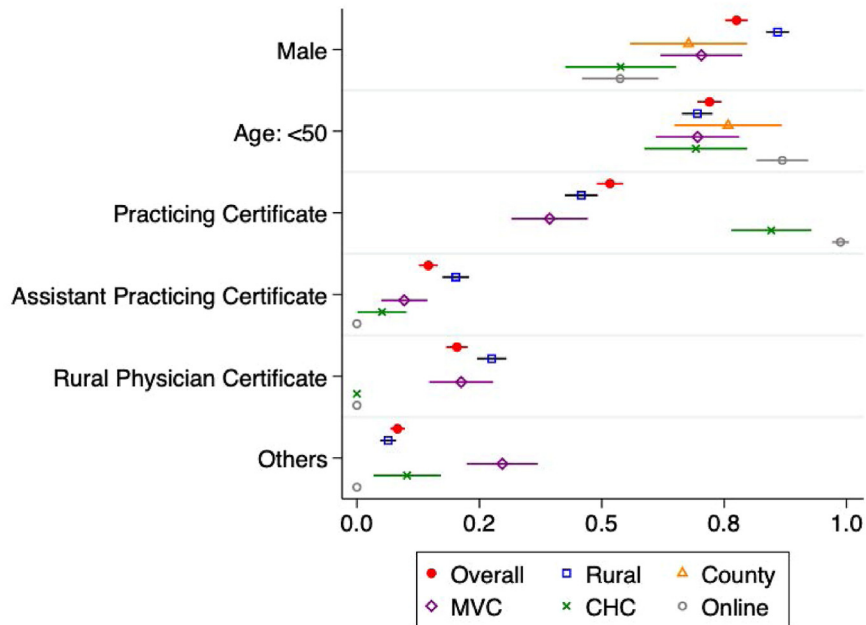
For all analyses, we utilize Stata 16.1 (Stata Corporation, College Station, TX).

**Ethical approval**

Informed consent was obtained from all physicians participating of in-person consultations about two months prior to the visits. The informed consent requirement was waived for online consultations on the basis of minimal risk and that no individually identifiable information on physicians was recorded (see Supplementary Materials B.2 Data Collection for further details).

**Role of the funding source**

There are no founders had a role in study design, data collection, data analysis, data interpretation, or writing of the report.



**Fig. 1:** Physician characteristics by provider type. Source: Author’s analysis of data from the four SP programs conducted by the authors and their collaborators between 2015 and 2019. Note: This figure displays the mean values of physician characteristics, complemented by their 95% confidence intervals. The regression analyses underlying the figure is reported in Table A3 of the Supplementary material.



	Proportion of recommended questions and examinations	Correct diagnosis	Referred patients	Medications prescribed	Correct medications to if any	Correct medications	Unnecessary antibiotics prescribed, if any	Unnecessary antibiotics prescribed	Correct management
<b>Physician gender (male physician)</b>	-0.02	-0.06	-0.13	0.06	-0.02	-0.00	0.02	0.06	-0.09
P-value	(0.0174)	(0.0348)	(0.0006)	(0.0223)	(0.5429)	(0.8222)	(0.6633)	(0.0263)	(0.0180)
95% CI	[-0.04 to -0.00]	[-0.11 to -0.00]	[-0.21 to -0.05]	[0.01-0.12]	[-0.13 to 0.07]	[-0.05 to 0.04]	[-0.07 to 0.11]	[0.01-0.11]	[-0.16 to -0.01]
<b>Physician age (years)</b>									
Physician age >50 years	-0.02	-0.10	-0.12	0.07	-0.03	-0.00	0.05	0.06	-0.12
	(0.0062)	(<0.0001)	(0.0001)	(0.0492)	(0.5148)	(0.9118)	(0.2609)	(0.0558)	(0.0004)
	[-0.04 to -0.01]	[-0.14 to -0.06]	[-0.18 to -0.06]	[0.00-0.15]	[-0.13 to 0.07]	[-0.05 to 0.05]	[-0.04 to 0.13]	[-0.00 to 0.12]	[-0.19 to -0.05]
<b>Physician Title</b>									
Physician with Assistant Practicing Physician Certificate	-0.01	-0.11	-0.11	0.13	-0.10	-0.03	0.05	0.10	-0.09
	(0.3385)	(<0.0001)	(0.0079)	(0.0001)	(0.0124)	(0.1237)	(0.1624)	(<0.0001)	(0.0538)
	[-0.04 to 0.01]	[-0.15 to -0.07]	[-0.20 to -0.03]	[0.07-0.20]	[-0.18 to -0.02]	[-0.07 to 0.01]	[-0.02 to 0.12]	[0.05-0.15]	[-0.17 to 0.00]
Physician with Rural Physician Certificate	-0.03	-0.11	0.01	0.09	-0.09	-0.03	-0.04	0.03	-0.01
	(0.0026)	(<0.0001)	(0.8318)	(0.0159)	(0.0115)	(0.1035)	(0.4707)	(0.3479)	(0.7337)
	[-0.06 to -0.01]	[-0.16 to -0.06]	[-0.07 to 0.08]	[0.02-0.16]	[-0.17 to -0.02]	[-0.07 to 0.01]	[-0.14 to 0.07]	[-0.03 to 0.08]	[-0.09 to 0.06]
Other title/certificate	-0.01	-0.04	-0.17	0.01	-0.06	-0.01	-0.06	-0.06	-0.01
	(0.6169)	(0.2196)	(<0.0001)	(0.7560)	(0.1043)	(0.6729)	(0.3303)	(0.1734)	(0.8181)
	[-0.03 to 0.02]	[-0.10 to 0.02]	[-0.24 to -0.10]	[-0.07 to 0.10]	[-0.14 to 0.01]	[-0.06 to 0.04]	[-0.18 to 0.06]	[-0.14 to 0.02]	[-0.08 to 0.06]
<b>Disease cases</b>									
Diarrhea	-0.05	-0.14	-0.37	0.40	-0.17	0.00	0.16	0.27	-0.39
	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)	(0.0005)	(0.9598)	(<0.0001)	(<0.0001)	(<0.0001)
	[-0.06 to -0.03]	[-0.19 to -0.10]	[-0.43 to -0.32]	[0.34 to 0.45]	[-0.26 to -0.08]	[-0.04 to 0.05]	[0.08-0.24]	[0.21-0.33]	[-0.46 to -0.32]
TB	-0.08	-0.09	-0.22	0.30			0.63	0.52	-0.14
	(<0.0001)	(0.0003)	(<0.0001)	(<0.0001)			(<0.0001)	(<0.0001)	(<0.0001)
	[-0.09 to -0.06]	[-0.13 to -0.04]	[-0.28 to -0.16]	[0.25-0.34]			[0.56-0.69]	[0.48-0.56]	[-0.20 to -0.08]
Observations	1261	1261	1261	1261	397	843	644	1261	1261

Note: Data are average marginal effects, p-values, and 95% confidence intervals. Average marginal effects are calculated as the OLS coefficients for numerical variables and mean proportion difference from logistic regressions for binary variables. We compute the cluster-robust standard errors. We define clusters at the relevant administrative level, which corresponds to a county for rural clinics and county hospitals, a migrant community for migrant clinics, an urban district of the city (equivalent to a county of a prefecture/city) for CHCs, and a platform for telemedicine platforms. Observations with perfect separation were excluded from the logistic regression analysis to ensure valid parameter estimates.

Table 4: Association between physician characteristics and quality outcomes.

## Results

### Care quality across providers

#### Summary statistics

Our analysis has revealed significant concerns with the overall quality of primary care in China, as shown in [Table 2](#), Column 1. We find that the required checklist items were completed just 21.4% of the time across all types of providers (median: 19%; IQR: 12%–28%). Correct diagnoses were made in only 17.8% of cases using standardized patients. Correct management, including appropriate prescriptions and referrals, was achieved in just 45.6% of these cases. It is even more worrying that the majority of this correct management was due to referrals, accounting for 37.5% of the interactions. Among the 52.1% of patient interactions that resulted in prescriptions, a mere 9.3% were appropriate. Moreover, over half of these prescriptions were for unnecessary antibiotics.

Our data also highlight considerable quality disparities among the five types of providers, as shown in [Table 2](#), Columns 2–6. In terms of process quality, which we measure by the proportion of completed checklist items, community health centers (CHCs) (mean: 31.5%; median: 40%; IQR: 20%–40%) and online platforms (mean: 25.0%; median: 24%; IQR: 14%–35%) outperform county hospitals (mean: 21.7%; median: 19%; IQR: 12%–29%), rural clinics (mean: 20.6%; median: 19%; IQR: 12%–28%), and migrant clinics (mean: 18.0%; median: 18%; IQR: 12%–24%). There is significant difference between CHC and county hospitals ( $p = 0.0007$ ). Diagnosis quality, indicated by the correct identification of the disease, is also notably higher in CHCs (42.3%) and online platforms (40.1%), compared to county hospitals (19.4%), rural clinics (13.0%), and migrant clinics (8.4%). The differences between CHCs and county hospitals ( $p = 0.0035$ ) and the difference between online platforms and county hospitals ( $p = 0.0037$ ) are both significant at the 1% level.

We next look into the various measures of case management. When it comes to referrals, 25.6% of visits to CHCs lead to patients being referred to higher-level health-care facilities. This rate is significantly lower compared to migrant clinics at 38.7% ( $p = 0.0477$ ). Since county hospitals represent the highest of three tiers of China's rural health system, it's rare for patients to be referred elsewhere, thus the notably low referral rate of only one in 62 visits. Conversely, online platforms display a remarkably high referral rate of 90.0%.

As for appropriate prescriptions, CHCs lead the way with a 76.5% accuracy rate. This significantly outperforms online platforms ( $p = 0.0003$ ), which have a much lower accuracy rate of 28.8%. On the other hand, nearly all prescriptions from county hospitals, rural clinics, and migrant clinics were inappropriate, with correct prescription rates at 7.1%, 5.9%, and 2.5%, respectively. Unnecessary prescribing of antibiotics was

more common in migrant clinics (61.7%) and rural clinics (53.1%) compared to online platforms (30.5%) and county hospitals (21.4%). The difference between rural clinics and online platforms is statistically significant ( $p = 0.0010$ ).

Finally, we examine the summary measure of correct management, which includes correct referral, correct medication, and any necessary examinations (e.g., X-Ray tests for TB). For this measure, CHCs with a 37.2% correct rate do not seem to perform significantly better than rural clinics at 40.3% or migrant clinics at 39.4% ( $p > 0.1$ ). This is mainly because rural and migrant clinics have higher referral rates. Online platforms lead with the best management accuracy among all providers at 85.6%.

[Table A2](#) investigates variations in primary care quality across three disease cases: angina, diarrhea, and TB. In line with overall findings, CHCs and online platforms generally provided superior primary care compared to rural clinics, county hospitals, and migrant clinics. However, disease-specific outcomes varied, and was especially notable in angina case management. Migrant clinics outperformed others with a 74.5% correct management rate, followed by rural clinics at 64.6%, although the difference is not statistically significant ( $p = 0.1697$ ). Conversely, CHCs and county hospitals lagged significantly at 37.2% and 9.5% ( $p < 0.0001$ ), respectively. This variance largely stems from differing referral rates for SPs with angina, being higher in rural (60.9%) and migrant clinics (74.5%), compared to lower rates in CHCs (25.6%) and none in county hospitals. The difference between CHCs and rural clinics is statistically significant ( $p < 0.0001$ ). Additionally, the small sample size, especially with only 21 county providers, might also have led to a bias.

In summary, CHCs and online platforms deliver comparatively superior primary care overall. County hospitals marginally outperform rural clinics, and migrant clinics deliver the lowest quality of care among the surveyed institutions.

#### Regression analysis

In [Table 3](#), we employ regression analysis to show the associations between provider types and quality outcomes. We run the regression outlined in [Equation \(1\)](#), where rural clinics are used as the base group for comparison.

The findings from the regression analysis, which controls for diseases, show results similar to what we observed through summary statistics. In particular, when looking at providers that mostly serve rural communities, including migrant workers, we found that county hospitals don't significantly differ from rural clinics in many aspects. The main differences are that county hospitals have a significantly lower referral rate ( $p < 0.0001$ ) and less frequent use of unnecessary antibiotics ( $p = 0.0049$ ). Migrant clinics completed 2.5%

fewer items on the checklist ( $p = 0.0233$ ) and were 4.7% less likely to prescribe the correct medications ( $p = 0.0300$ ). The probability of the overall correct management rate is, however, similar to rural clinics due to the higher referral rate at migrant clinics ( $p = 0.0244$ ).

CHCs outperformed rural clinics significantly. They completed 6.5% more items on the checklist ( $p = 0.0013$ ), made accurate diagnoses 18.8% more frequently ( $p < 0.0001$ ), and were 48.7% more likely to prescribe the correct medications ( $p = 0.0002$ ). Although the overall rate of correct management was 19.3% lower in CHCs than in rural clinics ( $p < 0.0001$ ), this gap was solely due to the lower rate of referrals at CHCs. Correct management includes referrals, which were 20% less frequent at CHCs ( $p < 0.0001$ ).

Online platforms also had clear advantages over rural clinics. They completed 4.6 percentage points more checklist items ( $p = 0.0001$ ) and made correct diagnoses 25.2 percentage points more frequently ( $p < 0.0001$ ) compared to rural clinic providers. Correct case management was 42.3 percentage points higher for online providers ( $p < 0.0001$ ), and online physicians referred SPs 55.2 percentage points more often ( $p < 0.0001$ ). Online physicians were 18.0 percentage points less likely to prescribe medications ( $p < 0.0001$ ) but prescribed correct medications 39.5 percentage points more often ( $p < 0.0001$ ) and prescribed 23.9 percentage points fewer unnecessary antibiotics ( $p < 0.0001$ ).

In summary, the regression analysis shows that, after accounting for diseases, CHCs and online platforms still demonstrate superior primary care quality compared to rural clinics.

### Physician characteristics across providers

To explore what may account for differences in healthcare quality among varied providers, we examine the characteristics of physicians, which are the direct service providers to patients. We specifically focus on their demographic attributes and qualifications related to the provision of healthcare services (Fig. 1).

We begin by examining the demographic profile of physicians. Overall, most physicians are males (77.5%) and are under 50 years old (72.0%). Physicians providing online services are younger, with 86.9% under 50 years old ( $p < 0.0001$  for online vs. rural,  $p = 0.0449$  for online vs. county,  $p = 0.0004$  for online vs. migrant,  $p = 0.0010$  for online vs. CHC).

Physician qualifications vary considerably across provider types. Overall, 51.7% of the physicians hold practicing physician certificates (the highest certificate level). However, 98.8% of online providers and 84.6% of CHC providers hold practicing physician certificates, whereas only 45.8% of rural and 39.4% of migrant providers hold this certificate ( $p < 0.0001$  for the difference between CHC and rural providers) (Table A3).

### Association between physician characteristics and quality outcomes

To understand how the differences in physician characteristics are associated with quality outcomes, we run regressions of each quality outcome on the set of physician characteristics (controlling for diseases) and report the estimation results in Table 4.

When considering the gender of physicians, our estimates indicate that male physicians performed poorer than female physicians, including checklist completion ( $p = 0.0174$ ), correct diagnoses ( $p = 0.0348$ ), and overall correct management ( $p = 0.0180$ ). Male physicians were also less willing to make referrals ( $p = 0.0006$ ) and more inclined to prescribe unnecessary antibiotics ( $p = 0.0263$ ).

In terms of physician age, our data indicates that physicians older than 50 years exhibited poorer performance compared to younger ones in checklist completion ( $p = 0.0062$ ), accurate diagnoses ( $p < 0.0001$ ), and overall correct management ( $p = 0.0004$ ). The lower correct management rate comes mainly from their reluctance to make referrals ( $p = 0.0001$ ).

Regarding physician titles, our analysis reveals a largely positive association between a physician's certificate level and the quality of primary care across all measures. Physicians with lower certificates demonstrated a lower likelihood to complete check-lists and correctly diagnose conditions. Moreover, these physicians were less prone to prescribing correct medications.

In conclusion, our findings indicate a strong association between physician characteristics and the quality of primary care. Specifically, care quality tends to be lower among male physicians, older physicians, and those with lower-ranking certificates. These groups of physicians are more prevalent in rural and migrant clinics and less common in CHCs and online platforms. Therefore, the disparities in healthcare quality among various providers can be largely attributed to the differences in the physicians serving at these locations.

## Discussion

### Interpretation of results

In this section, we will discuss the findings of this study and their implications for policy. First, our results show a serious issue with the overall quality of primary care in China. Across all providers, only about 21.4% of the checklist items were completed, and just 17.8% of the patients were accurately diagnosed. Further, correct case management, including making appropriate referrals, happened in less than half of the instances. Although earlier studies have highlighted that some healthcare providers deliver poor quality care,<sup>7,17,24,27</sup> this study reveals that the issue is not limited to specific providers. Instead, it is a common problem affecting all types of primary care providers that most people can access,

indicating that the overall quality of primary care in China is significantly low.

For the first time, this research uses a consistent and objective approach to show significant differences in the quality of primary care offered by various types of providers across China. In general, Community Health Centers (CHCs) and online platforms provide better primary care compared to rural clinics, migrant clinics, and county hospitals. Among the latter group, county hospitals outperform rural clinics in performance, whereas migrant clinics show the lowest quality. Our results further reveal that this quality disparity across providers is largely attributed to the qualifications of the physicians practicing in these locations or platforms.

The disparity in primary care quality across providers implies substantial inequalities in healthcare access across different groups of people. Rural residents are particularly vulnerable to substandard healthcare services given the poor performance of rural clinics. This disadvantage persists even after they migrate to urban regions, as migrant clinics, primarily serving urban dwellers who have migrated from rural areas, demonstrated the poorest quality of care among all types of providers. As a result, the quality of primary care that rural people can access does not improve but rather deteriorate after they relocate to urban regions.

These alarming findings necessitate immediate policy interventions to elevate the quality of primary care, with a specific emphasis on improving care in rural areas to ensure universal and equitable healthcare access for all citizens. Given the pivotal role of physician quality in driving these disparities, potential measures could include the establishment of training projects for the existing primary healthcare workforce, and the restructuring of financial compensation systems to attract more highly qualified physicians to these underserved areas.

Our findings also yield significant implications for the gatekeeping policy that the Chinese government intend to implement in the primary care system. The objective of this policy is to improve the efficiency of primary care by directing rural patients to village and township clinics (both are referred to as rural clinics in this study) as their initial point of contact before they could visit county hospitals. Without a significant improvement in rural healthcare quality, however, this the well-meant policy might unintentionally lead to more cases of wrong diagnosis and incorrect treatments for those living in rural regions,<sup>1</sup> especially given the strikingly low rate of correct medication use when no referral is made. As a result, the policy could worsen their healthcare conditions and increase current healthcare disparities instead of reducing them.

The findings indicate both the potential benefits and challenges of promoting telemedicine platforms as either an alternative or additional option to bridge the rural-urban gap in primary care quality. Our findings

indicate that rural residents are likely to get more accurate diagnoses through online consultations compared to visiting rural clinics or even county hospitals. Yet, online platforms do not show the capacity of replacing offline services because 90% of the online consultations ended up with a referral to offline facilities. That is, most patients still need to visit offline providers after the online consultation. This diminishes the marginal benefit of online platforms and its potential to complement the existing primary care system. Nonetheless, we believe that telemedicine platforms can effectively extend rural populations' access to better medical resources and elevate quality of primary care that they receive despite the limitations.<sup>23</sup>

Lastly, we draw attention to the grave issue of inappropriate prescription practices, which not only widen the quality gaps between different providers, but could potentially amplify broader problems such as antibiotic resistance resulting from overuse.<sup>32</sup> Merely 9.3% of standardized patients were prescribed the appropriate medications, while a troubling 50.8% were given unnecessary antibiotics. This scenario is partly fueled by China's current compensation system, where healthcare providers derive a significant proportion of their income from drug sales, thereby incentivizing overprescription, including the dispensation of unnecessary antibiotics. As such, the implementation of a more balanced compensation system and incentive scheme, along with stricter regulations, is crucial for both online and offline providers to curb unnecessary care and prescriptions.<sup>33-37</sup>

While this research specifically examines care quality in China, its results could be relevant for other developing countries who are undergoing rapid urbanization and structural transformation.<sup>3</sup> The differences in care quality among various provider types that we've uncovered highlight the need to take these discrepancies into account when evaluating healthcare quality in other developing countries.

### Limitations

We need to acknowledge several limitations of this study. First, the data for various providers come from different provinces. Yet, it's important to mention that these provinces share similarities in several key socioeconomic factors. These include the level of economic development (measured by GDP per capita), urbanization rate, population age distribution, literacy levels, and per capita income. The similarities in these fundamental variables indicate that the data from these projects are comparable enough to offer insightful information about the differences in quality of care among various providers, despite the differences in location of data collection.

Another limitation of our study is that we conducted standardized patient visits to various types of healthcare providers at different times. Specifically, visits to rural

clinics, county hospitals, and migrant clinics occurred in 2015, to urban Community Health Centers (CHCs) in 2017, and to online platforms in 2019. As a result, some of the differences we observed among these providers might be influenced by changes over time. However, our findings would only be fully explained by temporal changes if there were substantial improvements in the quality of healthcare from 2015 to 2017. For instance, we would need to see an increase of 4.6 percentage points (or 40.4 percent) in process quality, 29.3 percentage points (or 225.4 percent) in diagnosis accuracy, and 70.6 percentage points (or 1196.6 percent) in the prescription of correct medications. Given the relatively stable changes in public health investment in China during this period — healthcare spending as a percentage of GDP grew by just 0.72 percentage points, or 12 percent, over four years — such dramatic improvements seem unlikely.

Moreover, we recognize that, like any observational study, there may be additional unmeasured confounding factors that influence the causal interpretation of our findings. However, the goal of this study is not primarily to measure the effect of provider type on quality, but rather to compare the quality of services as fairly as possible. We find this comparison crucial in its own right, and see it as a starting point for future studies to further explore the causal relationship.

Our study is limited by the data and variables that are uniformly available in all standardized patient projects. This confines our scope to three diseases and overlapping variables across studies. Although this may limit the broad applicability of our conclusions, we believe that it ensures a more focused examination of the most comparable variables across different studies.

Lastly, our data only covers three provinces, which might not completely reflect the entire country. However, since these three provinces are considered average on a national scale and are alike in terms of socioeconomic factors, we believe that the findings from these provinces can reasonably indicate the healthcare quality in China's mid-income regions.

## Conclusion

This paper examines the quality of primary healthcare and its associations with physician characteristics in China, the world's largest developing country. By merging data from four standardized patient (SP) research projects, we evaluate the performance of five primary care providers that cover most population: rural clinics, county hospitals, migrant clinics, urban community health centers (CHCs), and online platforms. Our findings reveal a poor overall quality of primary care in China, with notable disparities across different providers. CHCs emerge as relatively reliable primary care providers in terms of process quality, diagnostic accuracy, and correct medication prescriptions. Online platforms outpace rural clinics, county hospitals, and

migrant clinics in many areas, showcasing their potential to enhance access to quality healthcare resources in under-resourced rural regions. We observe a positive association between the qualifications of physicians and the quality of primary care, underscoring the necessity for a greater presence of highly qualified practitioners. Furthermore, we highlight the critical need for regulations to curb unnecessary care and prescription practices. Over time, these actions can enable the government to allocate healthcare resources more efficiently, yielding societal benefits through improved, more equitable access to universal primary care.

## Contributors

Mian Huang: data curation, formal analysis, methodology, software, visualization, writing—original draft, writing—review & editing. Scott Rozelle: conceptualization, project administration, supervision, writing—review & editing. Yiming Cao: validation, writing—review & editing. Jian Wang: supervision. Zhang Zhang: validation. Zhijie Duan: validation. Shuyi Song: validation. Sean Sylvia: conceptualization, methodology, project administration, resources. All authors read and approved the final manuscript.

## Data sharing statement

All data and code are available on the Harvard Dataverse: <https://doi.org/10.7910/DVN/5WYWL>.

## Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT 4 in order to improve language and readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

## Declaration of interests

The authors declare no competing interests.

## Acknowledgements

Sean Sylvia gratefully acknowledges support from the National Institute of Allergies and Infectious Diseases (Award Number 5K01AI159233).

## Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.lanwpc.2024.101161>.

## References

- Xu J, Powell-Jackson T, Mills A. Effectiveness of primary care gate-keeping: difference-in-differences evaluation of a pilot scheme in China. *BMJ Glob Health*. 2020;5(8):e002792. <https://doi.org/10.1136/bmjgh-2020-002792>.
- World Health Organization. *Ethics guidance for the implementation of the end TB strategy*. Number WHO/HTM/TB/2017.07. World Health Organization; 2017. ISBN 978-92-4-151211-4. Section: 64 p.
- Chen Z. Launch of the health-care reform plan in China. *Lancet*. 2009;373(9672):1322–1324. [https://doi.org/10.1016/S0140-6736\(09\)60753-4](https://doi.org/10.1016/S0140-6736(09)60753-4).
- National Health Commission of the People's Republic of China. *China health statistical yearbook 2019*. Peking union medical college publishing house; 2019.
- Li X, Krumholz HM, Yip W, et al. Quality of primary health care in China: challenges and recommendations. *Lancet*. 2020;395(10239):1802–1812. [https://doi.org/10.1016/S0140-6736\(20\)30122-7](https://doi.org/10.1016/S0140-6736(20)30122-7).
- You X, Gu J, Xu DR, et al. Impact of the gate-keeping policies of China's primary healthcare model on the future burden of tuberculosis in China: a protocol for a mathematical modelling study. *BMJ Open*. 2021;11(8):e048449. <https://doi.org/10.1136/bmjopen-2020-048449>.



- 7 Sylvia S, Xue H, Zhou C, et al. Tuberculosis detection and the challenges of integrated care in rural China: a cross-sectional standardized patient study. *PLoS Med*. 2017;14(10):e1002405. <https://doi.org/10.1371/journal.pmed.1002405>.
- 8 Das J, Holla A, Das V, Mohanan M, Tabak D, Chan B. In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. *Health Aff*. 2012;31(12):2774–2784. <https://doi.org/10.1377/hlthaff.2011.1356>. Publisher: Health Affairs.
- 9 Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart AbstractionA prospective validation study of 3 methods for measuring quality. *JAMA*. 2000;283(13):1715–1722. <https://doi.org/10.1001/jama.283.13.1715>.
- 10 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(11):644–653. [https://doi.org/10.1016/S1070-3241\(00\)26055-0](https://doi.org/10.1016/S1070-3241(00)26055-0).
- 11 Leonard KL, Masatu MC. Using the Hawthorne effect to examine the gap between a doctor's best possible practice and actual performance. *J Dev Econ*. 2010;93(2):226–234. <https://doi.org/10.1016/j.jdevco.2009.11.001>.
- 12 Das J, Kwan A, Daniels B, et al. Use of standardised patients to assess quality of tuberculosis care: a pilot, cross-sectional study. *Lancet Infect Dis*. 2015;15(11):1305–1313. [https://doi.org/10.1016/S1473-3099\(15\)00077-8](https://doi.org/10.1016/S1473-3099(15)00077-8).
- 13 Abe K, Roter D, Erby L, Ban N. A nationwide survey of standardized patients: who they are, what they do, and how they experience their work. *Patient Educ Counsel*. 2011;84:261–264. <https://doi.org/10.1016/j.pec.2010.07.017>.
- 14 Wu D, Lam TP. Underuse of primary care in China: the scale, causes, and solutions. *J Am Board Fam Med*. 2016;29(2):240–247. <https://doi.org/10.3122/jabfm.2016.02.150159>. Publisher: American Board of Family Medicine Section: Family Medicine World Perspective.
- 15 Hort K, Gilbert K, Basnayaka P, Leslie Annear P. *Strategies to strengthen referral from primary care to secondary care in low- and middle- income countries*. World Health Organization. Regional Office for South-East Asia; 2019. ISBN 978-92-9022-709-0.
- 16 Peng X. China's demographic history and future challenges. *Science*. 2011;333(6042):581–587. <https://doi.org/10.1126/science.1209396>. Publisher: American Association for the Advancement of Science.
- 17 Su M, Zhou Z, Si Y, et al. Comparing the quality of primary care between public and private providers in urban China: a standardized patient study. *Int J Environ Res Public Health*. 2021;18(10):5060. <https://doi.org/10.3390/ijerph18105060>. Number: 10 Publisher: Multidisciplinary Digital Publishing Institute.
- 18 LeRouge C, Garfield MJ. Crossing the telemedicine chasm: have the u.s. barriers to widespread adoption of telemedicine been significantly reduced? *Int J Environ Res Public Health*. 2013;10(12):6472–6484. <https://doi.org/10.3390/ijerph10126472>.
- 19 Zobair KM, Sanzogni L, Sandhu K. Expectations of telemedicine health service adoption in rural Bangladesh. *Soc Sci Med*. 2019;238:112485. <https://doi.org/10.1016/j.socscimed.2019.112485>.
- 20 Huang M, Wang J, Nicholas S, Maitland E, Guo Z. Development, status quo, and challenges to China's health informatization during covid-19: evaluation and recommendations. *J Med Internet Res*. 2021;23(6):e27345.
- 21 Li X, Bai Z, Yao Q, Li H, Yao Y, Ding G. Status quo of telemedicine in china. *Chin J Evid Based Med*. 2013;13:1194–1199 [in Chinese].
- 22 Hua Y, Hua T. First exploration on the standards and academic organization of developed countries on telemedicine. *China Digital Med*. 2014;9:23–27 [in Chinese].
- 23 Zhai Y, Gao J, Chen B, et al. Design and application of a telemedicine system jointly driven by videoconferencing and data exchange: practical experience from henan province, China. *Telemed J E Health*. 2020;26(1):87–98. <https://doi.org/10.1089/tmj.2018.0240>. Publisher: Mary Ann Liebert, Inc., publishers.
- 24 Guo W, Sylvia S, Umble K, Chen Y, Zhang X, Yi H. The competence of village clinicians in the diagnosis and treatment of heart disease in rural China: a nationally representative assessment. *Lancet Reg Health Western Pacific*. 2020;2:100026.
- 25 Yi H, Liu H, Wang Z, et al. The competence of village clinicians in the diagnosis and management of childhood epilepsy in south-western China and its determinants: a cross-sectional study. *Lancet Reg Health Western Pacific*. 2020;3.
- 26 Wang Q, Poudel Adhikari S, Wu Y, et al. Consultation length, process quality and diagnosis quality of primary care in rural China: a cross-sectional standardized patient study. *Patient Educ Counsel*. 2022;105(4):902–908.
- 27 Xue H, Hager J, An Q, et al. The quality of tuberculosis care in urban migrant clinics in China. *Int J Environ Res Public Health*. 2018;15(9):2037. <https://doi.org/10.3390/ijerph15092037>.
- 28 Xue H, D'Souza K, Fang Y, et al. Direct-to-consumer telemedicine platforms in China: a national market survey and quality evaluation. *SSRN Electron J*. 2021. <https://doi.org/10.2139/ssrn.3944587>.
- 29 Wong S, Kung K, Griffiths S, et al. Comparison of primary care experiences among adults in general outpatient clinics and private general practice clinics in Hong Kong. *BMC Public Health*. 2010;10:397. <https://doi.org/10.1186/1471-2458-10-397>.
- 30 Das J, Mohpal A. Socioeconomic status and quality of care in rural India: new evidence from provider and household surveys. *Health Aff*. 2016;35(10):1764–1773. <https://doi.org/10.1377/hlthaff.2016.0558>. Publisher: Health Affairs.
- 31 Ali Mansournia M, Nazemipour M, Naimi AI, Collins GS, Campbell MJ. Reflection on modern methods: demystifying robust standard errors for epidemiologists. *Int J Epidemiol*. 2021;50(1):346–351. <https://doi.org/10.1093/ije/dyaa260>.
- 32 WHO. *Antimicrobial resistance*. 2018.
- 33 Li Y, Xu J, Wang F, et al. Overprescribing in China, driven by financial incentives, results in very high use of antibiotics, injections, and corticosteroids. *Health Aff*. 2012;31(5):1075–1082. <https://doi.org/10.1377/hlthaff.2010.0965>. Publisher: Health Affairs.
- 34 Eggleston K, Ling L, Qingyue M, Lindelow M, Wagstaff A. Health service delivery in China: a literature review. *Health Econ*. 2008;17(2):149–165. <https://doi.org/10.1002/hec.1306>.
- 35 Yip W, Hsiao WC. The Chinese health system at a crossroads. *Health Aff*. 2008;27(2):460–468. <https://doi.org/10.1377/hlthaff.27.2.460>. Publisher: Health Affairs.
- 36 Chi-Man Yip W, Hsiao WC, Chen W, Hu S, Ma J, Maynard A. Chapter 4: early appraisal of China's huge and complex health care reforms. In: *Health care policy in East Asia: a world scientific reference*. world scientific series in global health economics and public policy. World Scientific; 2019:51–83. ISBN 978-981-323-317-1. [https://doi.org/10.1142/9789813236134\\_0004](https://doi.org/10.1142/9789813236134_0004).
- 37 Currie J, Lin W, Meng J. Addressing antibiotic abuse in China: an experimental audit study. *J Dev Econ*. 2014;110:39–51. <https://doi.org/10.1016/j.jdevco.2014.05.006>.