ARC Centre of Excellence in Population Ageing Research

Working Paper 2023/02

Quantifying the Financial Impact of Overuse in Primary Care in China: A Standardised Patient Study

Yafei Si, Hazel Bateman, Shu Chen, Katja Hanewal, Bingqin Li, Min Su and Zhongliang Zhou

---

Quantifying the Financial Impact of Overuse in Primary Care in China: A Standardised Patient Study

Yafei Si MA [1, 2], Hazel Bateman PhD [1, 2], Shu Chen MSc [1, 2], Katja Hanewald PhD [1, 2], Bingqin Li PhD [3], Min Su PhD [4], and Zhongliang Zhou PhD [5]

[1] *ARC Centre of Excellence in Population Ageing Research (CEPAR), University of New South Wales, Sydney, Australia.*

[2] *School of Risk & Actuarial Studies, University of New South Wales, Sydney, Australia.*

[3] *Social Policy Research Centre, University of New South Wales, Sydney, Australia.*

[4] *School of Public Administration, Inner Mongolia University, Hohhot, China.*

[5] *School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an, China.*

Address correspondence to:

Zhongliang Zhou, School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an, China. E-mail: zzliang1981@xjtu.edu.cn

Yafei Si, ARC Centre of Excellence in Population Ageing Research (CEPAR), The University of New South Wales, Sydney, Australia. E-mail: yafei.si@unsw.edu.au.

Abstract

Overuse of health care is a potential factor in explaining the rapid increase in health care expenditure in many countries; however, it is difficult to measure overuse. This study employed the novel method of using unannounced standardised patients (SPs) to identify overuse, document its patterns and quantify its financial impact on patients in primary care in China. We trained 18 SPs to present consistent cases of two common chronic diseases and recorded 492 physician-patient interactions in 63 public and private primary hospitals in a capital city in western China in 2017 and 2018. Overuse, defined as the provision of unnecessary medical tests and drugs, was identified by a panel of medical experts based on national clinical guidelines. We estimated linear regression models to investigate how hospital, physician and patient characteristics were associated with overuse and to quantify the financial impact of overuse after controlling for a series of fixed effects. We found overuse in 72.15% of the SP visits. The high prevalence of overuse was similar among public and private hospitals, low-competence and high-competence physicians, male and female physicians, junior and senior physicians and male and female patients, but it varied between patients presenting different diseases. Compared to the non-overuse group, overuse significantly increased the total cost by 117.8%, the test cost by 58.8% and the drug cost by 100.3%. The financial impact of overuse was consistent across the aforementioned hospital, physician and patient characteristics. We suggest that the overuse observed in this study is unlikely to be attributable to physician incompetence but rather to the financing framework for primary care in China. These findings illuminate the cost escalation of primary care in China, which is a form of medical inefficiency that should be urgently addressed.

Keywords: health care expenditure, overuse, primary care, standardised patient, China

Introduction

Health care expenditure has risen steadily worldwide but has increased much faster in China, with an annual growth rate of over 10% in the past decade (Chang et al., 2019). As a result, health care expenditure in China soared from under 500 billion yuan (≈78.5 billion US dollars) in 2000 to over 7.2 trillion yuan (≈1.1 trillion US dollars) in 2020 (Statista, 2022), accounting for over 29% of government spending and 7% of gross domestic product (GDP) in 2020. The reasons for this rapid growth are numerous and include economic growth, rapid population ageing, an epidemiological transition to chronic diseases, high-risk health behaviours and detrimental environmental factors (WHO, 2019a). As for health care financing, encouragingly, the increase was partly influenced by the Chinese government's significant injection of funding since 2009, reflecting the country's strong commitment to provide all citizens with equal access to basic primary care with reasonable quality and financial risk protection (Yip et al., 2019). However, individuals' out-of-pocket spending has been positively associated with government spending and has continued to rise (Zhang and Rahman, 2020). The striking increase in health care expenditure directly threatens health system sustainability and health care affordability.

Overuse could be an important factor contributing to the rapid increase in health care expenditures. Overuse, defined as 'the provision of health care service when its likely risk of harm exceeds its potential benefit' (Chassin and Galvin, 1998), has increasingly become a global concern for academics and medical practices (Brownlee et al., 2017; Morgan et al., 2015; Mulley, 2013). As services occur along a continuum of benefits and harms, overuse is unlikely to make patients healthier but can harm them physically, psychologically and financially. Overuse can even damage an individual's intrinsic capacity in later life (Cross et al., 2016; WHO, 2019b) and is often considered to be a low-value health care service (Colla, 2014). There is strong evidence for the widely acknowledged presence of overuse in Australia (Elshaug et al., 2012), Israel (Goldberg et al., 2008), Spain (Aguilar et al., 2001) and the United States (Berwick and Hackbarth, 2012), and overuse has been examined specifically with regard to infusions (Zhang et al., 2019), screening tests (Alber et al., 2018; Mathias et al., 2012; Schwartz et al., 2004; Tan et al., 2014) and antibiotics (Currie et al., 2014; Xue et al., 2019).

Overuse can be conceptualised as an agency problem (Arrow, 1963; Berwick and Hackbarth, 2012; Orszag, 2008), where the interest of physicians and patients conflict (Balafoutas et al., 2017; Dulleck and Kerschbamer, 2006; Gottschalk et al., 2020). However, the reasons for overuse are multifold, and a recent review has suggested that the drivers of overuse often fall into three domains (Saini et al., 2017b). The first consists of finance and organisation. For

example, physicians can induce patients to use more health care services than they need (Johnson and Rehavi, 2016; Rizzo and Blumenthal, 1996; Si et al., 2020) since there is a widely acknowledged information asymmetry between medical professionals and patients (Arrow, 1963; Johnson and Rehavi, 2016). Second, thinking frameworks (i.e., knowledge, beliefs, assumptions, biases and uncertainty) can influence physicians' decision-making. For example, high expectations from patients and incorrect beliefs about treatment effectiveness held by physicians can contribute to the overuse of health care (Weeks et al., 2012). Third, overuse can be attributed to power and human relationships, such as poor physician-patient relationships, which can drive both overuse and underuse (He, 2014). Failure to understand the nature of overuse and its extent can cause severe distortion and inefficiency, leading to unintended consequences for programs that are designed to reduce health care expenditure.

However, measuring systematic overuse and developing robust evidence for its prevalence among patients poses a significant challenge (Brownlee et al., 2017). First, the identification of overuse in developed countries relies heavily on complete administrative data, which is hard to obtain in developing countries. Second, defining 'the appropriateness of care' is difficult since harms are poorly documented in many clinical practices (Morgan et al., 2015). Moreover, even when clinical guidelines are available for determining 'the appropriateness of care', sufficient details about patients are unavailable (Huang and Rosenthal, 2015). Third, evidence-based diagnostic and treatment guidelines generally overlook comorbidities (Boyd et al., 2005), which are common in real medical practice (Van Weel and Schellevis, 2006). Fourth, patient preferences may affect physician decision- making, which increases the challenge of defining 'the appropriateness of care' (Blank et al., 2006). In the context of the challenges documented above, researchers are beginning to conceptualise overuse as a general system problem and develop system-level metrics (Segal et al., 2014). However, existing evidence on the scope of overuse is far from complete, and its potential financial impact on patients is poorly understood.

In this study, we used a standardised patient (SP) method to accurately identify overuse, document its patterns and quantify its financial impact on primary care patients in China. SPs are well-trained 'acting patients' who can present symptoms of a disease like any other patient but in a standardised background setting. The SP approach enables researchers to set benchmarks for the likely benefits and harms of a specific medical test or drug (King et al., 2019). Another benefit of the SP method is that it allows physician practices to be compared with predefined guidelines because the actual illness and the optimal care associated with it are known (Sylvia et al., 2017, 2015). Furthermore, well-trained SPs can consistently present the symptoms of diseases, thus controlling for unobservable factors that are related to

patient preference, health-seeking behaviours and communication styles (Kwan et al., 2019). Moreover, SPs can record the detailed process of the physician-patient interactions and the prices charged. Thus, an SP study is less likely to be subject to recall bias than approaches that rely on recall-based patient exit surveys and chart abstraction (Spivak et al., 2016).

In this study, we used two tracer conditions (i.e., SP case presentations), unstable angina and asthma, to collect physician-patient visit data in a primary care setting. We recruited and trained 18 SPs from the local community. To obtain direct evidence of overuse, we used the SP method to rule out confounding factors from the patient side and then compared physician practices with predefined guidelines to identify the presence of overuse (Sulis et al., 2020). Then we constructed statistical models to document how hospital, physician and patient characteristics were associated with overuse. To quantify the financial impact of overuse, we used an identification strategy similar to that of Mullainathan and Obermeyer (2021). SPs were randomly assigned to visit hospitals and physicians on a workday without a scheduled appointment. The arrival time of an SP determined which physician saw them, and physicians varied in their tendency to provide overuse (Mullainathan and Obermeyer, 2021). Conditioning on the day of week and location of an SP visit provided plausible exogenous variation in the overuse. Finally, we examined heterogeneity in the financial impact of overuse among hospitals, physicians and patients.

Methods

This section describes the institutional context, the SP method, the definitions of overuse and health care expenditure in this study and the econometric models that were used to document the pattern of overuse and identify its financial impact.

Institutional Context

In China, the primary care system provides general clinical care and basic public health services. Primary care facilities are *de-facto* 'ambulatory care hospitals', providing outpatient services, mainly for common clinical conditions, and very limited inpatient services. For example, the facilities in our sample had, on average, 13.78 (ranging from 5 to 41) physicians and 16.27 (ranging from 0 to 89) beds. No referral is required for patients to use either outpatient or inpatient services. Primary care requires multidisciplinary professional teams that include physicians, nurses and pharmacists. Despite serving as the backbone of medical professionals, physicians are unevenly distributed across China and are often poorly trained in less developed areas like our setting. Community health centres provide primary care facilities in urban areas, while these facilities are provided by township health centres and village medical centres in rural areas.

Both the public and private sectors operate primary care facilities.

In general, primary care facilities provide walk-in outpatient services. Patients were required to cover the cost for outpatient visits almost entirely out-of-pocket in both the public and private facilities in our study. This is because social health insurance policies mainly provide coverage for inpatient services but offer limited coverage for outpatient care due to low annual caps for total reimbursement. These settings cause patients to increase their use of inpatient services, even for minor health conditions, rather than using primary outpatient care. The disadvantage of attracting patients this way is that it may motivate physicians in primary care to provide more services per patient.

Primary care facilities commonly use a fee-for-service approach in China. First, patients generally pay a fixed fee to initiate the consultation with physicians, while the fee is heavily capped or sometimes free of charge (An, 2013). During and after the consultation, physicians decide on the tests and medications to be prescribed, which will be charged separately at the end of visits. It is important to note that the charge for a specific test or drug is almost the same across facilities, although hospitals have the autonomy of determining the price. Second, physicians' incomes in public and private hospitals in China have two elements: a fixed base salary based on qualifications and experience and performance-based bonuses from the hospital. The performance-based bonus typically depends on the number of patients served and the revenue the physician generates for the hospital. This payment structure incentivises the overuse of health care. Furthermore, the operation of public and private hospitals is very similar in China, apart from their ownership. Private hospitals provide about 10% of total health care services in China, with the remainder provided by public hospitals (Su et al., 2021). The percentage is similar to our sample distribution.

Standardised Patient

In this study, we used data collected from an SP audit study that involved all of the primary care facilities/hospitals in the urban area of a capital city in western China from August 17 to 28, 2017, and from July 30 to August 10, 2018 (*Supplement 1*). SPs are well-trained 'acting patients' who present the symptoms of an illness to a physician just like any other patient. SPs are used routinely to train and evaluate medical students in high-income countries (Fenton et al., 2016), and the method has been widely used to measure physician practice in developing countries (King et al., 2019). In this study, SPs were coached to present their initial symptoms and answer any questions that the physician asked during history taking, in a manner consistent with the underlying condition (Das et al., 2012, 2016b). The hospitals and physicians granted approval for the study three months prior to the visits but were unaware of the diseases that would

be tested. The study was approved by the Ethics Committee of Xi'an Jiaotong University Health Science Centre (Approval number: 2015-406) and the Ethics Committee of the University of New South Wales (HC210354).

We used two tracer conditions (i.e., common chronic diseases to be presented by the SPs) to identify overuse in this study: unstable angina and asthma. The incidence of the two tracer conditions was high in the study region. The first condition, asthma, causes a person's airways to become inflamed, narrow and swollen and produce extra mucus, making it difficult to breathe. Asthma is usually managed with rescue inhalers that treat symptoms and controller inhalers that prevent symptoms. Asthma affects 1–18% of the population in a range of countries (Global Initiative for Asthma, 2018). The symptoms of asthma can be minor or can interfere with daily activities; however, they may also lead to a life-threatening attack. Asthma is a prevalent but largely undiagnosed and undertreated condition in China (Huang et al., 2019). The second condition, unstable angina, is also known as acute coronary syndrome. The disease is a type of chest discomfort caused by poor blood flow through the vessels (coronary vessels) of the heart muscle (myocardium). Unstable angina may lead to a heart attack, and the American Health Association suggests that it should be treated as an emergency. In 2017, stroke and ischaemic heart disease were the leading causes of death in China (Luan et al., 2021), with 149 and 124 in every 100,000 deaths attributed to stroke and heart attack, respectively (Zhou et al., 2019).

In the study, SPs were randomly assigned to present one of the two tracer conditions using standardised background information. For unstable angina, each of the SPs 'acted' as a 50-year-old who had recently experienced worsening chest pain. For asthma, each of the SPs 'acted' as a 40-year-old patient who had experienced a breathing problem that had worsened since the previous night. The hypothetical patient carried a form of moderate social health insurance, the Urban Resident Health Insurance (Su et al., 2018). The two hypothetical cases were not complicated, but were specifically chosen so that the opening statement by the SPs would be consistent with multiple underlying illnesses. In addition, the tracer conditions could be portrayed easily and presented a low risk of prompting invasive examinations. Explicit, predefined guidelines for physician practice were available for the two tracer conditions and had been adapted from earlier studies in India and China (Das et al., 2012; Sylvia et al., 2017, 2015). A panel of doctors determined that appropriate history-taking and examinations should lead providers toward the correct diagnosis and treatment.

The SPs were recruited from local communities and trained by the research team. Overall, 10 SPs in 2017 and 8 SPs in 2018 were selected after being interviewed, and we have 18 SPs (person-years) in the study. The SPs participated in rigorous training before visiting the primary care hospital physicians. We included all public and private primary care

hospitals in seven districts of the capital city. In order to avoid a potential 'sorting effect' in which patients deliberately select clinics based on personal preference or illness characteristics, which is common in the observation of real patient-provider interactions, the SPs were not allowed to choose their hospitals and physicians. The SPs were randomly assigned to a hospital on a workday and then visited the physician in the first office they encountered in that particular hospital.

Physicians may alter their behaviours when they are aware of being observed. None of the 18 SPs reported that their acting during the visits was revealed. Therefore, we were able to measure physician practice in a real-world scenario in which Hawthorne effects were not a concern because the physicians did not know that they were being observed (Leonard and Masatu, 2010a). However, the physicians had previously given their consent to participate in the study and did *know* that a number of SPs would randomly visit them during a given period. The SPs accepted all non-invasive medical tests and rejected all invasive medical tests in accordance with a specific script. In addition, the SPs paid the visit fee and purchased all prescribed non-invasive medical tests and medications, which were reimbursed by the research team. The study did not involve any insurance claiming. After the visit, the SPs were required to report the physician-patient interaction immediately using a structured questionnaire. In order to ensure accuracy, the detailed interactions between the SPs and physicians were recorded, and thus the SPs' responses were double-checked by instructors using verbal recordings. It is worth noting that the verbal recordings were not used as data in this study. Overall, the sample included 492 interactions between 18 SPs and 269 physicians.

Health Care Expenditure

We measured health care expenditure in the primary medical setting in China by totalling the cost of the fixed consultation fee, fees for medical tests conducted and fees for drugs prescribed (measured in Chinese Yuan, CNY; exchange rate, 6.37 CNY ≈ 1 US dollar on May 18, 2018). We include all costs for the medical tests that physicians performed and planned to perform for one specific visit, noting that the SPs rejected all invasive tests by giving reasonable excuses (Note that in China, patients usually do not reject medication prescribed by a physician). Our approach used three continuous variables to measure health care expenditure: the cost of medical tests (test cost), the cost of the prescribed medications (drug cost) and the total cost (including the test cost, the drug cost and the consultation fee). The health care expenditure variable was transformed using a natural logarithm to enhance our interpretation.

Overuse of Health Care

The SP method allowed us to directly measure the overuse of health care in primary care hospitals. In accordance with Brownlee et al. (2017), we use the term 'overuse' to refer to any services that are unnecessary in any way. The related terms in medical settings and other studies include over-testing, over-diagnosis, over-treatment and low-value care, which mostly refers to the inappropriate usage of these services. Based on the case-specific interaction (*Supplement 1*), each medical test was classified as either essential or unnecessary, while each drug was classified as correct or unnecessary (and even harmful) in accordance with predefined guidelines and input from a judging panel of doctors and pharmacists. We used one binary variable to measure the presence of the overuse of any unnecessary service (0 = no, 1 = yes). Two binary variables were used to measure two specific kinds of overuse: overuse of medical tests (0 = no, 1 = yes) and overuse of drugs (0 = no, 1 = yes). Finally, three count variables were used to measure the number of overprovided unnecessary services: the unnecessary tests, the unnecessary drugs and the total number of unnecessary items prescribed (both unnecessary tests and drugs).

Hospital, Physician and Patient Characteristics

We included six variables of interest to examine how the overuse and its financial impact varied among hospital, physician and patient characteristics.

1. Public vs. private hospitals: private hospitals usually operate as financially self-sufficient institutions, while public hospitals receive substantial funding from the government. The difference in funding sources is likely to affect physician decisions with respect to the provision of unnecessary services (Das et al., 2016b; Su et al., 2021).
2. Physician competence: we used one binary variable, whether the physician gave a correct diagnosis for the interaction, as a proxy for physician competence. The physician was classified as high-competence if yes and low-competence if no. Additionally, we used a predicted probability of a physician giving a correct diagnosis for the interaction when conducting the sensitivity analysis (Currie and MacLeod, 2017, 2020).
3. Diseases: we used two chronic diseases, asthma and unstable angina. The use of two different diseases enhanced the external validity of the study.
4. Physician gender: we examined whether physician behaviour in overuse varied based on physician gender because a growing body of evidence finds that female physicians outperform their male counterparts in adherence to clinical protocols (Bertakis et al., 1995), effectiveness in psychosocial consultation (Bertakis et al., 2009; Roter et al., 2002), performance on standardised tests (Kim et al., 2005) and patient readmission rate (Tsugawa et al., 2017).

5. Patient gender: patient gender may play a role in influencing overuse as the result of a physician's taste-based discrimination (Becker, 2010) or the physician's internalised negative beliefs about the patient's gender (Greenwood et al., 2018).
6. Physician age: we used physician age as a proxy for professional experience. All physicians were classified as junior (aged under 40) or senior (aged 40 and above).

Econometric Model

Our primary interest was to document patterns of overuse and quantify its financial impact. We used a straightforward estimation strategy. The econometric specification for the patterns of overuse is as follows:

$$Overuse_{ijkt}^{m} = \beta_0 + \beta_1 X_{ijkt} + \beta_2 W_{ijkt} + \beta_3 Z_{ijkt} + \delta_k + \varphi_t + \epsilon_{ijkt}^{m} \quad , \qquad (1)$$

where $Overuse_{ijkt}^{m}$ indicates the overuse (m = ext.) or the number of unnecessary items (m = int.) for visit i at hospital j in district k on day t. X_{ijkt}, W_{ijkt} and Z_{ijkt} separately represent a set of the observable physician, hospital and patient characteristics. δ_k are district fixed effects, and φ_t are day of the week, month and year fixed effects. ϵ_{ijkt}^{m} is the error term. Robust standard errors were clustered at the hospital level.

The econometric specification for the financial impact of overuse was calculated as follows:

$$Cost_{ijkt}^{n} = \beta_0 + \beta_1 Overuse_{ijkt}^{m} + \beta_2 X_{ijkt} + \pi_j + \delta_k + \varphi_t + \mu + \omega + \epsilon_{ijkt}^{n} \quad , \qquad (2)$$

where $Cost_{ijkt}^{n}$ represents cost indicator n (i.e., total cost, test cost and drug cost). In the model, we further controlled for π_j hospital fixed effects, μ disease fixed effects and ω SP fixed effects. All fixed effects were included in the model as dummy variables. Based on the econometric specification (2), we examined the financial impact of overuse using the original continuous variables of health care expenditure (CNY) and its natural logarithm forms. For the extensive margin (m = ext.), we used three binary variables: total overuse, overuse of medical tests and overuse of drugs. For the intensive margin (m = int.), we used three count variables: the total of unnecessary items, unnecessary tests and unnecessary drugs.

Results

Of the 492 physician-patient interactions between 18 SPs and 269 physicians, 78 interactions (15.85%) occurred in private hospitals, 217 interactions (44.11%) involved high-competence physicians, 247 interactions (50.20%) involved patients presenting with asthma, 224 interactions (45.53%) involved male physicians, and 81 interactions (16.46%)

involved male patients. Junior physicians and senior physicians accounted for 152 interactions (30.89%) and 340 interactions (69.11%), respectively (*Table S1*).

Prevalence and Patterns of Overuse

We found that 72.15% of the SP visits (95% CI: 68.18% to 76.13%) involved overuse of health care (Table S2), including overuse of medical tests (54.67%, 95% CI: 50.26% to 59.09%), overuse of drugs (28.05%, 95% CI: 24.07% to 32.03%) or both (10.57%, 95% CI: 7.8% to 13.30%). The average number of unnecessary items was 1.88 (95% CI: 1.77 to 1.99), which included 1.26 items of unnecessary medical tests (95% CI: 1.56 to 1.77) and 0.62 items of unnecessary drugs (95% CI: 1.47 to 1.72).

We began our exploratory analyses by examining the bivariate association of the overuse with hospital, physician and patient characteristics. Our unadjusted analysis suggests a high prevalence of overuse among public and private hospitals, male and female physicians, junior and senior physicians and male and female patients. The overuse rate varied between physicians of low competence and high competence, between patients presenting asthma and unstable angina and between male and female patients. The distribution for unnecessary items was similar to the distribution for overuse (*Table S2*).

The unadjusted results above could be confounded by unobservable hospital and physician factors that were not included in the model. Through the use of logistic regressions based on specification (1), we found no evidence or weak evidence to indicate that the rate of overuse varied between public and private hospitals, low-competence and high-competence physicians, male and female physicians or junior and senior physicians (*Figure 1*). The probability of overuse among patients presenting with asthma was significantly higher than for patients presenting with unstable angina, as was the probability of the overuse of medical tests. However, the probability of the overuse of drugs among male patients was significantly lower than for female patients. Through the use of negative binomial regressions, we found that the number of unnecessary items was almost identical across these hospital, physician and patient characteristics. The marginal effects of the coefficients are reported in *Table S3*.

Financial Impact of Overuse

First, we compared the health care expenditure between the overuse and non-overuse group and found that the health care expenditure in the overuse group was much higher than that of the non-overuse group, including the total cost, the

test cost and the drug cost. The differences were statistically significant at 1% levels (*Table S4*). We found that the overuse of medical tests was significantly associated with a higher test cost, but a lower drug cost and overuse of drugs was significantly associated with a higher drug cost but a lower test cost.

We then used econometric models to estimate the financial impact of overuse at its extensive and intensive margins. At the extensive margin (*Figure 2 Panel A and Table S5*), overuse led to an increase of 27.39 CNY (95% CI: 17.35 to 37.43) in the total cost, of 14.32 CNY (95% CI: 6.97 to 21.67) in test costs and of 12.93 CNY (95% CI: 5.89 to 19.98) in drug costs. This corresponds to a 117.8% (95% CI: 73.9% to 161.7%) increase in the total cost, a 58.8% (95% CI: 20.1% to 97.6%) increase in test costs and a 100.3% (95% CI: 57.0% to 143.7%) increase in drug costs compared with the non-overuse group. Second, the overuse of medical tests led to a significant increase of 72.0% (95% CI: 39.0% to 105.0%) in the test cost and a significant decrease of 66.7% (95% CI: 29.1% to 104.3%) in the drug cost, without causing a significant change in the total cost. Third, the overuse of drugs led to a significant increase of 307.1% (95% CI: 273.6% to 340.6%) in drug cost and a significant increase of 163.1% (95% CI: 120.2% to 206.0%) in the total cost but resulted in no significant change in the test cost.

At the intensive margin (*Figure 2 Panel B and Table S5*), the cost increase incurred by an additional item of unnecessary service was relatively small. For example, a one-item increase in unnecessary service led to a significant increase of 20.9% (95% CI: 3.1% to 38.6%) in the total cost and a significant increase of 38.3% (95% CI: 18.0% to 58.6%) in the drug cost but resulted in no significant change in the test cost. We did not find that a one-item increase in unnecessary tests (/drugs) significantly changed either the total cost, the test cost, or the drug cost.

Heterogeneity of the Financial Impact

We examined the heterogeneity of the financial impact by interacting the overuse variables with hospital, physician and patient characteristics. Overall, the results suggest that the financial impact of total overuse did not vary substantially between male and female patients, male and female physicians, public and private hospitals, low-competence and high-competence physicians, junior and senior physicians or patients presenting the two different diseases (*Table 1*). However, the effect size of total overuse on the test cost among patients presenting asthma was significantly larger (79.2%, 95% CI: 4.6% to 153.9%), while the effect size of total overuse on the drug cost was lower (96.7%, 95% CI: 25.2% to 168.1%) than for patients presenting unstable angina.

We found very similar results when examining the heterogeneity of financial impact by examining the overuse of medical tests and the overuse of drugs. However, compared with the effect size of low-competence physicians, the effect size of the overuse of medical tests on drug cost was significantly larger (92.1%, 95% CI: 23.7% to 160.5%) among high-competence physicians. Compared with that of female SP visits, the effect size of overuse of drugs on drug cost was significantly larger (56.1%, 95% CI: 0.7% to 111.6%) among male SP visits. Compared with that of junior physicians, the effect size of overuse of drugs on test cost was significantly larger (62.1%, 95% CI: 2.0% to 122.2%) among senior physicians.

Discussion

There is an increasing global interest in optimising health care delivery and improving the quality of care (Macleod et al., 2014). However, previous research has primarily focused on reducing underuse and on the failure to deliver essential services (Glasziou et al., 2017), while its counterpart, overuse, has been studied less by comparison (Saini et al., 2017a). To our knowledge, this is the first study to use an SP audit-study design to examine the patterns of overuse and quantify its financial impact in primary care in China. We found a high incidence of overuse among physician-patient interactions in a real-world scenario and determined that the overuse led to a significant increase in health care expenditure. Compared to the non-overuse group, overuse increased the total cost by 117.8%, the test cost by 58.8% and the drug cost by 100.3%. We also demonstrated that the public hospitals in our study provided as many unnecessary services as the private hospitals, leading to a significant increase in health care expenditure.

The use of the SP method to measure overuse provided a unique opportunity to explore physician practices and accurately estimate overuse. Our results are comparable to studies in China and other developing countries. For example, the prevalence of overuse in this study was higher overall than in studies of rural China (Sylvia et al., 2015; Xue et al., 2019) and was similar to the level of overuse found in India and Kenya (Das et al., 2016b; Sulis et al., 2020). One key difference in this study was the focus on primary care hospitals in urban areas of China, which are generally well-equipped with medical equipment. For example, more than half of primary care hospitals in urban China provided inpatient care in 2017 and even had specialist departments (Li et al., 2017), and in our study setting, these hospitals were able to provide (almost) all the essential medical tests for the two tracer conditions (Su et al., 2021). Further, we note that no referral is needed for patients to use outpatient or inpatient services in these hospitals. We observed the pervasive overuse of unnecessary medical tests, which were generally absent from the earlier investigations in (comparably poor) rural areas (Sylvia et al., 2017, 2015; Xue et al., 2019) and less developed regions (Das et al., 2016b, 2016a). In addition,

our systematic identification of medical tests and drugs suggests a much higher prevalence of overuse and low-value care in China than in the United States or Australia (Badgery-Parker et al., 2019; Scott et al., 2022).

The high prevalence of overuse in the study was largely driven by physician behaviours rather than patient behaviours, and this study provided a clean quasi-experimental estimation of the economic significance of overuse. In 2018, there were 4.41 billion outpatient visits in primary care, which accounted for 53.04% of total outpatient visits in China. The outpatient health care expenditure per capita in primary care was 156.8 CNY (n.d.), representing a market of 691.5 billion CNY (≈ 108.6 billion US dollars). The average cost of an SP visit in the study was 35.0 CNY, with a standard deviation of 41.3. The relatively low cost is reasonable because we used two non-complicated and common chronic diseases. We estimated the financial burden of overuse that was incurred by the specific population affected by each of the two chronic diseases by considering the prevalence and the increased cost. For our sample, we estimated that the overuse led to a 56.59% increase in the total cost, a 54.50% increase in the test cost and a 61.20% increase in the drug cost for the population with the two chronic diseases. Our results strongly suggest that overuse has the potential to impose significant costs on patients and the economy as a whole. Eliminating the use of unnecessary medical tests and drugs would substantially reduce physical harm and the risk of polypharmacy (WHO, 2019b) while simultaneously reducing the financial burden on the population.

In general, the unregulated provision of health care by the private sector is not socially desirable (Das et al., 2016b) since providers may over-respond to demand, leading to socially inefficient provision of care. Alternatively, the default policy approach to delivering health care, especially primary care, occurs through public hospitals. In this study, we found that the overuse rate and its financial impact were equally high in both public and private hospitals, suggesting the presence of strong financial incentives in both the public and private health sectors in China. Previous research indicated that financial incentives among public hospitals in China originated from the market-oriented reforms that have been in place since the 1980s (Blumenthal and Hsiao, 2015). Since that time, health care professionals have not received a wage from the government, the prices charged for physician service time have been strictly capped by the government, hospitals have been permitted to earn profits by prescribing new drugs and high-tech medical tests (Hesketh and Zhu, 1997) and physicians have received substantial bonuses from hospital profits. These factors have led to the pervasive over-prescription of unnecessary and expensive drugs and high-tech tests (Currie et al., 2014).

Many studies suggest that physicians are substantially underpaid in public hospitals in China (An, 2013) due to the

regulation of fees for physician service time. Moreover, physicians working in primary care often lack coverage from welfare programs that are designed for government employees (Li et al., 2017), and they are commonly burned out (Ye et al., 2019). However, the last decade has witnessed a significant increase in the health care workforce in China, including the number of physicians in primary care (Tang and Tang, 2018). The expansion of primary care under an underpayment system might be explained by the pervasive overuse of health care (Li et al., 2012). Overuse has led to a significant increase in health care expenditure, which may proxy for an underlying market price to compensate for the physician's nominal wage. The financial incentives persisted after the latest wave of health system reforms in China since 2009, which included the introduction of an essential medicine list and a zero-markup drug policy (Xiao et al., 2013; Yi et al., 2015). A likely consequence is that physicians use a greater number of unnecessary medical tests when drug prescription is heavily regulated, as we found in this study.

However, financial incentives may not be the main issue since physician knowledge (Currie and MacLeod, 2017, 2020; Xue et al., 2019), professional ethics (Pellegrino, 1993), altruistic behaviours (Olsen et al., 2009), and practice norms (Havighurst, 1990) can also influence overuse. Our study is limited in that we did not measure physician knowledge and other characteristics in our data collection (Xue et al., 2019). However, the overall financial impact of overuse was similar across high-competence and low-competence physicians. Since physicians may have given a correct diagnosis through luck rather than the use of other services, we estimated the predicted probability of giving a correct diagnosis for each physician-patient interaction (*Table S6*). We found that the overuse was proportionately distributed among the continuum of predicted probability of providing a correct diagnosis (*Figure S1*), and the continuum should therefore serve as a precise proxy for physician competence in the study. We also found similar results when using physician age as a proxy for experience in medical practice. These results suggest that pervasive overuse and its associated cost are unlikely to be attributable to physician incompetence in our study.

We did not find that overuse and its financial impact varied between female and male physicians, although recent studies suggest that female physicians demonstrate superior performance in physician-patient relationships and health outcomes (Bertakis et al., 2009; Greenwood et al., 2018). However, physicians behaved differently in response to some patient characteristics. Male SP visits resulted in a significantly lower prevalence of drug overuse than female SP visits. Furthermore, the effect size of total overuse on test cost and drug cost varied significantly between patients presenting asthma and unstable angina, and the effect size of overuse of drugs on drug cost was larger for male SP visits than for female SP visits. Theoretically, the overuse-associated physician agency problem is related to information asymmetry

between physicians and patients. It is a relative position because patients vary in their health literacy as it pertains to informed decision-making (Johnson and Rehavi, 2016; Si et al., 2020), and physician behaviours can change accordingly. Moreover, as we found in the heterogeneity analysis, the effect size for the overuse of medical tests on drug cost was significantly larger among high-competence physicians than for low-competence physicians, indicating that physicians exhibit more aggressive prescription behaviours when their decisional uncertainty is low.

This study has several additional limitations. First, only a few types of diseases could be presented in the SP study. However, the SP method has good external validity when the detection rate is reasonably low (zero in the study) because physicians behave consistently in both SPs visits and real patient visits (Das et al., 2016b, 2016a). Second, we only used data from primary health care centres in the urban area of a western capital city in China. Future studies conducted nationwide at primary, secondary and tertiary hospitals are necessary to produce more generalisable evidence. Although the pervasive overuse in our study seems unlikely to be attributable to physician incompetence, more research is needed to understand how physician knowledge affects overuse, as growing evidence indicates a gap between physician knowledge and practice (Das and Hammer, 2007; Leonard and Masatu, 2010b; Mohanan et al., 2015). Third, our study involved the one-time interactions of new patients with different physicians, and the situation may differ when continuity of care within the physician-patient model is promoted. Further research could assess patient preferences for specific physicians after the initial interaction. Finally, examining whether and how overuse influences physician decision making in improving other metrics of health care quality and patient wellbeing was beyond the scope of this study, and these remaining questions will be evaluated in follow-up projects.

Overall, this study has important implications for health system reforms in China. The overuse observed in this study is unlikely to be attributable to physician incompetence but is instead associated with the financing framework for primary care, which is largely consistent nationwide. China has undergone tremendous demographic and epidemiological change during the past 30 years (Yip et al., 2019). The increasing burden from chronic diseases and an ageing population have presented significant health care challenges for the country (Glinskaya and Feng, 2018). Recent health system reforms tend to encourage initial contact with lower-tier health care providers and primary care providers (Zhou et al., 2021), which may substantially affect patients with chronic diseases. Our study shows that overuse is pervasive in primary care in China and results in higher-than-necessary health care expenditure. These findings illuminate the problem of cost escalation in Chinese primary care, where overuse constitutes a form of medical inefficiency that should be urgently addressed.

First, while we cannot exclude a variety of factors that may result in overuse from the physician's side, overuse in this study seems unlikely to be attributable to physician incompetence. Thus, further clinical training is unlikely to serve as an effective policy intervention because its impact on the quality of care depends on the extent to which the additional training increases physician competence and how much the (increased) competence, in turn, is reflected in practice. Previous studies have indicated that both of these effects could be minimal in medical practice (Das et al., 2016a; Das and Hammer, 2007). For instance, one previous study failed to find a positive impact of clinical training on the quality of care in primary care (Su et al., 2019). A study in India found that overuse among qualified physicians with sufficient knowledge was as common as overuse among unqualified providers (Das et al., 2016b). Clinical training significantly improved the quality of the health care provided but failed to reduce the adverse effect of overuse (Das et al., 2016a).

Second, this study demonstrated that the overuse of medical tests and the overuse of drugs are partial substitutes. Physicians provided fewer drugs (/tests) if they offered too many tests (/drugs), and overuse of drugs (/tests) was associated with a decrease in test (/drug) cost. We used a consultation–treatment transaction model to understand physician behaviour (*Figure 3*). The main insights from the model are that physicians induce patients to consume more services either at the consultation stage (medical tests) or the treatment stage (drugs), and physicians shift effort and medical resources between the two stages to make the global budget appear balanced (Si et al., 2019). Therefore, the consultation stage could be as expensive as the treatment stage in practice, but the former has been studied less frequently in previous literature (Currie et al., 2014; Currie and MacLeod, 2020). This finding has positive implications because it suggests that even without government regulation, market accountability (i.e., competition and reputation) may motivate physicians to maintain the total health care expenditure within a particular upper bound. The cost indicator can be easily observed by patients and compared within their social networks, especially as it pertains to common diseases. It is very possible for patients to be unsatisfied with health care expenditure that is too high or perceived as unreasonable. In this case, the overuse may lead to a significant but finite increase in health care expenditure.

Role of the funding source

The funders of the study, the China Medical Board, National Natural Science Foundation of China, the University of New South Wales, and the ARC Centre of Excellence in Population Ageing Research (CEPAR project CE170100005), had no role in study design, data collection, data analysis, data interpretation, or manuscript writing. The corresponding author had full access to all study data and decided to submit this manuscript for publication.

Data availability: The data of SP study that support the findings of this study are available on request from the corresponding authors. The data are not publicly available due to restrictions of ethic approval requirement for this study. The code scripts used in this analysis are available from the corresponding authors upon reasonable request.

Declaration of Funding: This study was funded by China Medical Board (15-227), USA; National Natural Science Foundation of China (71874137), China; The University of New South Wales, Australia; ARC Centre of Excellence in Population Ageing Research (CEPAR project CE170100005), Australia.

Acknowledgements: The authors are grateful to all standardised patients and student instructors for collecting the data. YS framed the research question, contributed to data management and analysis and drafted the first version of the paper; SC and MS helped frame the research question and analytical approach and contributed to drafting of the paper; MS verified the data. KH, BL, HB and ZZ oversaw the project, contributed to the framing of the research question and drafted significant portions of the paper. All authors approved the final version of the paper.

Declaration of Conflicts of Interests: We declare no competing interests.

References

Aguilar, M.D., Fitch, K., Lázaro, P., Bernstein, S.J., 2001. The appropriateness of use of percutaneous transluminal coronary angioplasty in Spain. Int. J. Cardiol. 78, 213–221.

Alber, J.M., Brewer, N.T., Melvin, C., Yackle, A., Smith, J.S., Ko, L.K., Crawford, A., Glanz, K., 2018. Reducing overuse of cervical cancer screening: A systematic review. Prev. Med. 116, 51–59. https://doi.org/10.1016/j.ypmed.2018.08.027

An, J., 2013. Which future for doctors in China? The Lancet 382, 936–937.

Arrow, K.J., 1963. Uncertainty and the welfare economics of medical care. Am. Econ. Rev. 53, 941–973.

Badgery-Parker, T., Pearson, S.-A., Chalmers, K., Brett, J., Scott, I.A., Dunn, S., Onley, N., Elshaug, A.G., 2019. Low-value care in Australian public hospitals: prevalence and trends over time. BMJ Qual. Saf. 28, 205–214.

Balafoutas, L., Kerschbamer, R., Sutter, M., 2017. Second-degree moral hazard in a real-world credence goods market. Econ. J. 127, 1–18.

Becker, G.S., 2010. The economics of discrimination. University of Chicago press.

Bertakis, K.D., Franks, P., Epstein, R.M., 2009. Patient-centered communication in primary care: physician and patient gender and gender concordance. J. Womens Health 18, 539–545.

Bertakis, K.D., Helms, L.J., Callahan, E.J., Azari, R., Robbins, J.A., 1995. The influence of gender on physician practice style. Med. Care.

Berwick, D.M., Hackbarth, A.D., 2012. Eliminating waste in US health care. JAMA 307, 1513–1516.

Blank, T., Graves, K., Sepucha, K., Llewellyn-Thomas, H., 2006. Understanding treatment decision making: contexts, commonalities, complexities, and challenges. Ann. Behav. Med. 32, 211–217.

Blumenthal, D., Hsiao, W., 2015. Lessons from the East—China's rapidly evolving health care system. N. Engl. J. Med. 372, 1281–1285.

Boyd, C.M., Darer, J., Boult, C., Fried, L.P., Boult, L., Wu, A.W., 2005. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA 294, 716–724.

Brownlee, S., Chalkidou, K., Doust, J., Elshaug, A.G., Glasziou, P., Heath, I., Nagpal, S., Saini, V., Srivastava, D., Chalmers, K., Korenstein, D., 2017. Evidence for overuse of medical services around the world. The Lancet 390, 156–168. https://doi.org/10.1016/S0140-6736(16)32585-5

Chang, A.Y., Cowling, K., Micah, A.E., Chapin, A., Chen, C.S., Ikilezi, G., Sadat, N., Tsakalos, G., Wu, J., Younker, T., 2019. Past, present, and future of global health financing: a review of development assistance, government, out-of-pocket, and other private spending on health for 195 countries, 1995–2050. The Lancet 393, 2233–2260.

Chassin, M.R., Galvin, R.W., 1998. The urgent need to improve health care quality: Institute of Medicine National Roundtable on Health Care Quality. JAMA 280, 1000–1005.

China Statistics, n.d. China Statistical Yearbook for Health 2019 [WWW Document]. URL https://data.cnki.net/area/Yearbook/Single/N2020020200?z=D09 (accessed 2.11.22).

Colla, C.H., 2014. Swimming against the current—what might work to reduce low-value care? N. Engl. J. Med. 371, 1280.

Cross, A.J., George, J., Woodward, M.C., Ames, D., Brodaty, H., Ilomäki, J., Elliott, R.A., 2016. Potentially inappropriate medications and anticholinergic burden in older people attending memory clinics in Australia. Drugs Aging 33, 37–44.

Currie, J., Lin, W., Meng, J., 2014. Addressing antibiotic abuse in China: an experimental audit study. J. Dev. Econ. 110, 39–51.

Currie, J., MacLeod, W.B., 2017. Diagnosing expertise: Human capital, decision making, and performance among physicians. J. Labor Econ. 35, 1–43.

Currie, J.M., MacLeod, W.B., 2020. Understanding Doctor Decision Making: The Case of Depression Treatment. Econometrica 88, 847–878.

Das, J., Chowdhury, A., Hussam, R., Banerjee, A.V., 2016a. The impact of training informal health care providers in India: A randomized controlled trial. Science 354.

Das, J., Hammer, J., 2007. Money for nothing: the dire straits of medical practice in Delhi, India. J. Dev. Econ. 83, 1–36.

Das, J., Holla, A., Das, V., Mohanan, M., Tabak, D., Chan, B., 2012. In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. Health Aff. (Millwood) 31, 2774–2784.

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

Dulleck, U., Kerschbamer, R., 2006. On doctors, mechanics, and computer specialists: The economics of credence goods. J. Econ. Lit. 44, 5–42.

Elshaug, A.G., Watt, A.M., Mundy, L., Willis, C.D., 2012. Over 150 potentially low-value health care practices: an Australian study. Med. J. Aust. 197, 556–560.

Fenton, J.J., Kravitz, R.L., Jerant, A., Paterniti, D.A., Bang, H., Williams, D., Epstein, R.M., Franks, P., 2016. Promoting patient-centered counseling to reduce use of low-value diagnostic tests: a randomized clinical trial. JAMA Intern. Med. 176, 191–197.

Glasziou, P., Straus, S., Brownlee, S., Trevena, L., Dans, L., Guyatt, G., Elshaug, A.G., Janett, R., Saini, V., 2017. Evidence for underuse of effective medical services around the world. The Lancet 390, 169–177. https://doi.org/10.1016/S0140-6736(16)30946-1

Glinskaya, E., Feng, Z., 2018. Options for aged care in China: Building an efficient and sustainable aged care system. The World Bank.

Global Initiative for Asthma, 2018. Global Strategy for Asthma Management and Prevention 2018.
Goldberg, A., Yalonetsky, S., Kopeliovich, M., Azzam, Z., Markiewicz, W., 2008. Appropriateness of diagnosis of unstable angina pectoris in patients referred for coronary arteriography. Exp. Clin. Cardiol. 13, 133.
Gottschalk, F., Mimra, W., Waibel, C., 2020. Health services as credence goods: A field experiment. Econ. J. 130, 1346–1383.
Greenwood, B.N., Carnahan, S., Huang, L., 2018. Patient–physician gender concordance and increased mortality among female heart attack patients. Proc. Natl. Acad. Sci. 115, 8569–8574.
Havighurst, C.C., 1990. The professional paradigm of medical care: obstacle to decentralization. Jurimetrics 415–429.
He, A.J., 2014. The doctor–patient relationship, defensive medicine and overprescription in Chinese public hospitals: Evidence from a cross-sectional survey in Shenzhen city. Soc. Sci. Med. 123, 64–71.
Hesketh, T., Zhu, W.X., 1997. Health in China: the healthcare market. Bmj 314, 1616.
Huang, K., Yang, T., Xu, J., Yang, L., Zhao, J., Zhang, X., Bai, C., Kang, J., Ran, P., Shen, H., 2019. Prevalence, risk factors, and management of asthma in China: a national cross-sectional study. The Lancet 394, 407–418.
Huang, X., Rosenthal, M.B., 2015. Overuse of Cardiovascular Services: Evidence, Causes, and Opportunities for Reform. Circulation 132, 205–214. https://doi.org/10.1161/CIRCULATIONAHA.114.012668
Johnson, E.M., Rehavi, M.M., 2016. Physicians treating physicians: Information and incentives in childbirth. Am. Econ. J. Econ. Policy 8, 115–41.
Kim, C., McEwen, L.N., Gerzoff, R.B., Marrero, D.G., Mangione, C.M., Selby, J.V., Herman, W.H., 2005. Is physician gender associated with the quality of diabetes care? Diabetes Care 28, 1594–1598.
King, J.J., Das, J., Kwan, A., Daniels, B., Powell-Jackson, T., Makungu, C., Goodman, C., 2019. How to do (or not to do)… using the standardized patient method to measure clinical quality of care in LMIC health facilities. Health Policy Plan. 34, 625–634.
Kwan, A., Daniels, B., Bergkvist, S., Das, V., Pai, M., Das, J., 2019. Use of standardised patients for healthcare quality research in low-and middle-income countries. BMJ Glob. Health 4, e001669.
Leonard, K.L., Masatu, M.C., 2010a. Using the Hawthorne effect to examine the gap between a doctor's best possible practice and actual performance. J. Dev. Econ. 93, 226–234.
Leonard, K.L., Masatu, M.C., 2010b. Professionalism and the know-do gap: Exploring intrinsic motivation among health workers in Tanzania. Health Econ. 19, 1461–1477.
Li, X., Lu, J., Hu, S., Cheng, K.K., De Maeseneer, J., Meng, Q., Mossialos, E., Xu, D.R., Yip, W., Zhang, H., 2017. The primary healthcare system in China. The Lancet 390, 2584–2594.
Li, Y., Xu, J., Wang, F., Wang, B., Liu, L., Hou, W., Fan, H., Tong, Y., Zhang, J., Lu, Z., 2012. Overprescribing in China, driven by financial incentives, results in very high use of antibiotics, injections, and corticosteroids. Health Aff. (Millwood) 31, 1075–1082.
Luan, S., Yang, Y., Huang, Y., McDowell, M., 2021. Public knowledge of stroke and heart attack symptoms in China: a cross-sectional survey. BMJ Open 11, e043220.
Macleod, M.R., Michie, S., Roberts, I., Dirnagl, U., Chalmers, I., Ioannidis, J.P., Salman, R.A.-S., Chan, A.-W., Glasziou, P., 2014. Biomedical research: increasing value, reducing waste. The Lancet 383, 101–104.
Mathias, J.S., Gossett, D., Baker, D.W., 2012. Use of electronic health record data to evaluate overuse of cervical cancer screening. J. Am. Med. Inform. Assoc. 19, e96–e101.
Mohanan, M., Vera-Hernández, M., Das, V., Giardili, S., Goldhaber-Fiebert, J.D., Rabin, T.L., Raj, S.S., Schwartz, J.I., Seth, A., 2015. The know-do gap in quality of health care for childhood diarrhea and pneumonia in rural India. JAMA Pediatr. 169, 349–357.
Morgan, D.J., Brownlee, S., Leppin, A.L., Kressin, N., Dhruva, S.S., Levin, L., Landon, B.E., Zezza, M.A., Schmidt, H., Saini, V., Elshaug, A.G., 2015. Setting a research agenda for medical overuse. BMJ h4534. https://doi.org/10.1136/bmj.h4534
Mullainathan, S., Obermeyer, Z., 2021. Diagnosing Physician Error: A Machine Learning Approach to Low-Value Health Care. Q. J. Econ. qjab046. https://doi.org/10.1093/qje/qjab046
Mulley, A.G., 2013. The global role of health care delivery science: learning from variation to build health systems that avoid waste and harm. J. Gen. Intern. Med. 28, 646–653.
Olsen, K.R., Gyrd-Hansen, D., Boegh, A., Hansen, S.H., 2009. GPs as citizens' agents: prescription behavior and altruism. Eur. J. Health Econ. 10, 399–407.
Orszag, P.R., 2008. The overuse, underuse, and misuse of health care. Wash. DC Congr. Budg. Off.
Pellegrino, E.D., 1993. The metamorphosis of medical ethics: a 30-year retrospective. JAMA 269, 1158–1162.
Rizzo, J.A., Blumenthal, D., 1996. Is the target income hypothesis an economic heresy? Med. Care Res. Rev. 53, 243–266.
Roter, D.L., Hall, J.A., Aoki, Y., 2002. Physician gender effects in medical communication: a meta-analytic review. JAMA 288, 756–764.
Saini, V., Brownlee, S., Elshaug, A.G., Glasziou, P., Heath, I., 2017a. Addressing overuse and underuse around the world. The Lancet 390, 105–107. https://doi.org/10.1016/S0140-6736(16)32573-9
Saini, V., Garcia-Armesto, S., Klemperer, D., Paris, V., Elshaug, A.G., Brownlee, S., Ioannidis, J.P.A., Fisher, E.S., 2017b. Drivers of poor medical care. The Lancet 390, 178–190. https://doi.org/10.1016/S0140-6736(16)30947-3
Schwartz, L.M., Woloshin, S., Fowler Jr, F.J., Welch, H.G., 2004. Enthusiasm for cancer screening in the United States. Jama 291, 71–78.
Scott, A., Li, J., Gravelle, H., McGrail, M., 2022. Physician Competition And Low-Value Health Care. Am. J. Health Econ. 8, 252–274.

Segal, J.B., Bridges, J.F., Chang, H.-Y., Chang, E., Nassery, N., Weiner, J., Chan, K.S., 2014. Identifying possible indicators of systematic overuse of health care procedures with claims data. Med. Care 157–163.
Si, Y., Zhou, Z., Su, M., Chen, X., 2019. Revisiting gender gap in quality of health care in urban China: a standardised patient audit study. The Lancet 394, S25.
Si, Y., Zhou, Z., Su, M., Hu, H., Yang, Z., Chen, X., 2020. Comparison of health care utilization among patients affiliated and not affiliated with healthcare professionals in China. BMC Health Serv. Res. 20, 1118. https://doi.org/10.1186/s12913-020-05895-y
Spivak, E.S., Cosgrove, S.E., Srinivasan, A., 2016. Measuring appropriate antimicrobial use: attempts at opening the black box. Clin. Infect. Dis. 63, 1–6.
Statista, 2022. China: health expenditure [WWW Document]. Statista. URL https://www.statista.com/statistics/279400/health-expenditures-in-china/ (accessed 2.10.22).
Su, M., Zhou, Z., Si, Y., Sylvia, S., Chen, G., Su, Y., Rozelle, S., Wei, X., 2021. Comparing the Quality of Primary Care between Public and Private Providers in Urban China: A Standardized Patient Study. Int. J. Environ. Res. Public. Health 18, 5060.
Su, M., Zhou, Z., Si, Y., Wei, X., 2019. Effect of health alliances on the quality of primary care in urban China: a coarsened exact matching difference-in-differences analysis. The Lancet 394, S86. https://doi.org/10.1016/S0140-6736(19)32422-5
Su, M., Zhou, Z., Si, Y., Wei, X., Xu, Y., Fan, X., Chen, G., 2018. Comparing the effects of China's three basic health insurance schemes on the equity of health-related quality of life: using the method of coarsened exact matching. Health Qual. Life Outcomes 16, 41. https://doi.org/10.1186/s12955-018-0868-0
Sulis, G., Daniels, B., Kwan, A., Gandra, S., Daftary, A., Das, J., Pai, M., 2020. Antibiotic overuse in the primary health care setting: a secondary data analysis of standardised patient studies from India, China and Kenya. BMJ Glob. Health 5, e003393.
Sylvia, S., Shi, Y., Xue, H., Tian, X., Wang, H., Liu, Q., Medina, A., Rozelle, S., 2015. Survey using incognito standardized patients shows poor quality care in China's rural clinics. Health Policy Plan. 30, 322–333.
Sylvia, S., Xue, H., Zhou, C., Shi, Y., Yi, H., Zhou, H., Rozelle, S., Pai, M., Das, J., 2017. Tuberculosis detection and the challenges of integrated care in rural China: A cross-sectional standardized patient study. PLoS Med. 14.
Tan, A., Kuo, Y.-F., Goodwin, J.S., 2014. Potential overuse of screening mammography and its association with access to primary care. Med. Care 52, 490.
Tang, C., Tang, D., 2018. The trend and features of physician workforce supply in China: after national medical licensing system reform. Hum. Resour. Health 16, 1–8.
Tsugawa, Y., Jena, A.B., Figueroa, J.F., Orav, E.J., Blumenthal, D.M., Jha, A.K., 2017. Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians. JAMA Intern. Med. 177, 206–213.
Van Weel, C., Schellevis, F.G., 2006. Comorbidity and guidelines: conflicting interests. Lancet 367, 550–550.
Weeks, J.C., Catalano, P.J., Cronin, A., Finkelman, M.D., Mack, J.W., Keating, N.L., Schrag, D., 2012. Patients' expectations about effects of chemotherapy for advanced cancer. N. Engl. J. Med. 367, 1616–1625.
WHO, 2019a. Healthy China: Deepening Health Reform in China: Building High-Quality and Value-Based Service Delivery. World Bank Publications.
WHO, 2019b. Integrated care for older people (ICOPE): guidance for person-centred assessment and pathways in primary care. World Health Organization.
Xiao, Y., Zhao, K., Bishai, D.M., Peters, D.H., 2013. Essential drugs policy in three rural counties in China: what does a complexity lens add? Soc. Sci. Med. 93, 220–228.
Xue, H., Shi, Y., Huang, L., Yi, H., Zhou, H., Zhou, C., Kotb, S., Tucker, J.D., Sylvia, S.Y., 2019. Diagnostic ability and inappropriate antibiotic prescriptions: a quasi-experimental study of primary care providers in rural China. J. Antimicrob. Chemother. 74, 256–263.
Ye, J., Wang, H., Wu, H., Ye, L., Li, Q., Ma, X., Yu, X., Zhang, H., Luo, X., 2019. Burnout among obstetricians and paediatricians: a cross-sectional study from China. BMJ Open 9, e024205.
Yi, H., Miller, G., Zhang, L., Li, S., Rozelle, S., 2015. Intended and unintended consequences of China's zero markup drug policy. Health Aff. (Millwood) 34, 1391–1398.
Yip, W., Fu, H., Chen, A.T., Zhai, T., Jian, W., Xu, R., Pan, J., Hu, M., Zhou, Z., Chen, Q., 2019. 10 years of healthcare reform in China: progress and gaps in universal health coverage. The Lancet 394, 1192–1204.
Zhang, D., Rahman, K.A., 2020. Government health expenditure, out-of-pocket payment and social inequality: A cross-national analysis of China and OECD countries. Int. J. Health Plann. Manage. 35, 1111–1126.
Zhang, Y., Zhou, Z., Si, Y., 2019. When more is less: What explains the overuse of health care services in China? Soc. Sci. Med. 232, 17–24. https://doi.org/10.1016/j.socscimed.2019.04.018
Zhou, M., Wang, H., Zeng, X., Yin, P., Zhu, J., Chen, W., Li, X., Wang, Lijun, Wang, Limin, Liu, Y., 2019. Mortality, morbidity, and risk factors in China and its provinces, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 394, 1145–1158.
Zhou, Z., Zhao, Y., Shen, C., Lai, S., Nawaz, R., Gao, J., 2021. Evaluating the effect of hierarchical medical system on health seeking behavior: A difference-in-differences analysis in China. Soc. Sci. Med. 268, 113372.

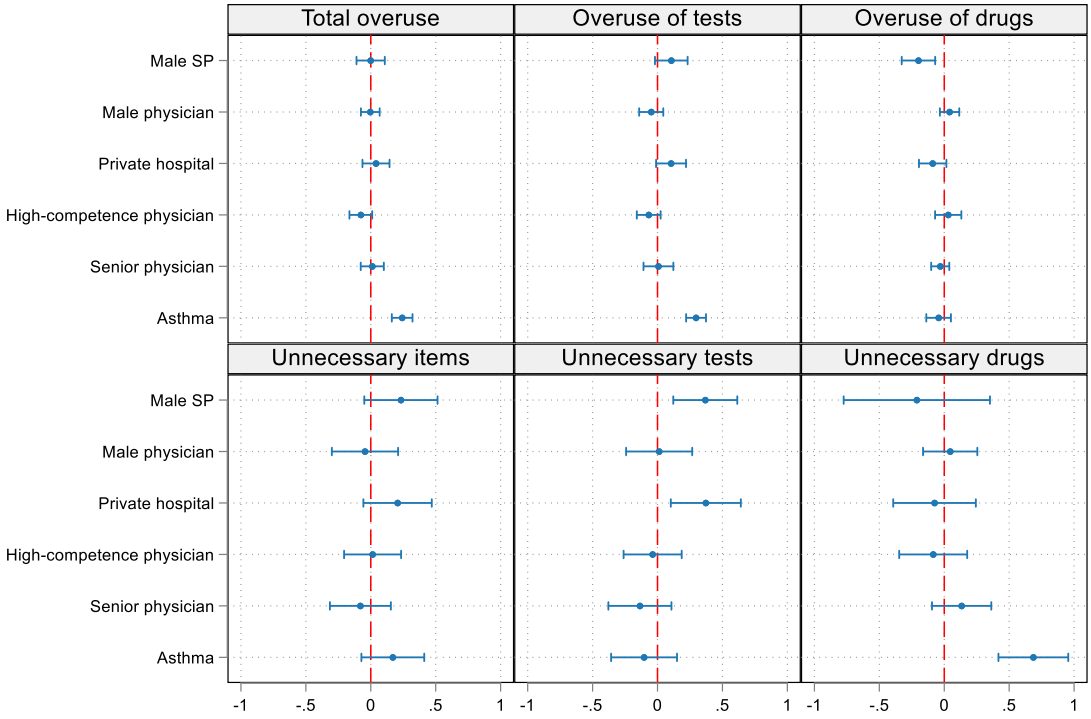
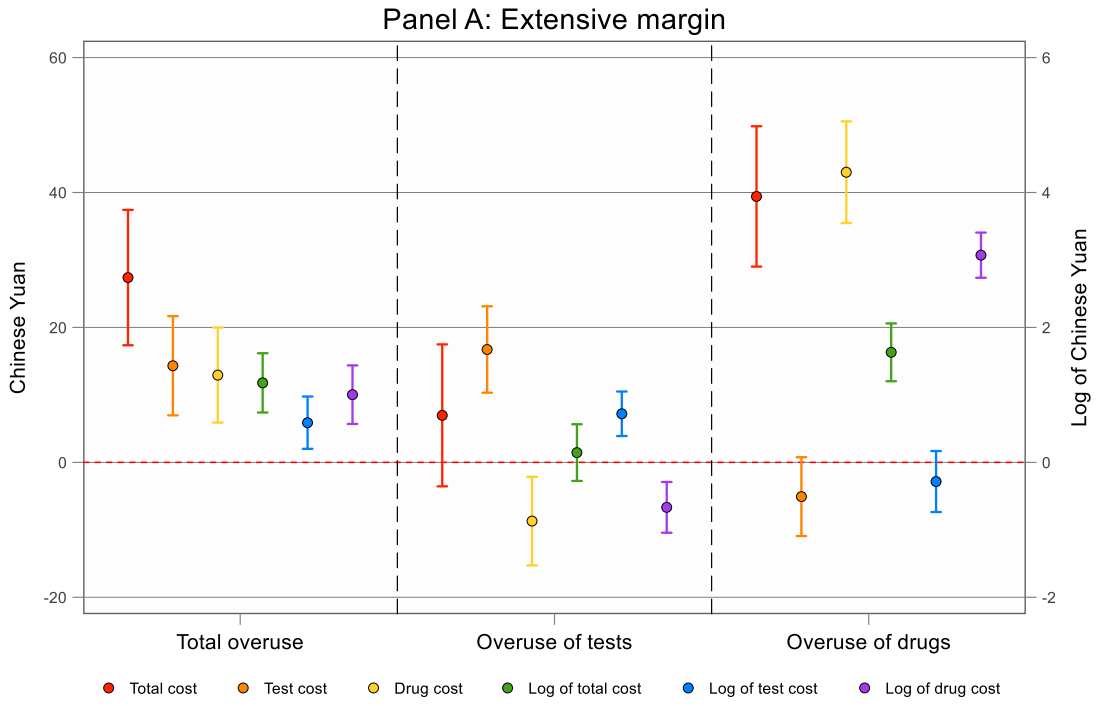


Figure 1 Patterns of overuse among hospitals, physicians and patients

Note: The graphs report the marginal effects and 95% confidence intervals for the presence of overuse and the number of unnecessary items.

Logistic and negative binomial regressions were performed to control for day of week, month, year and county fixed effects.



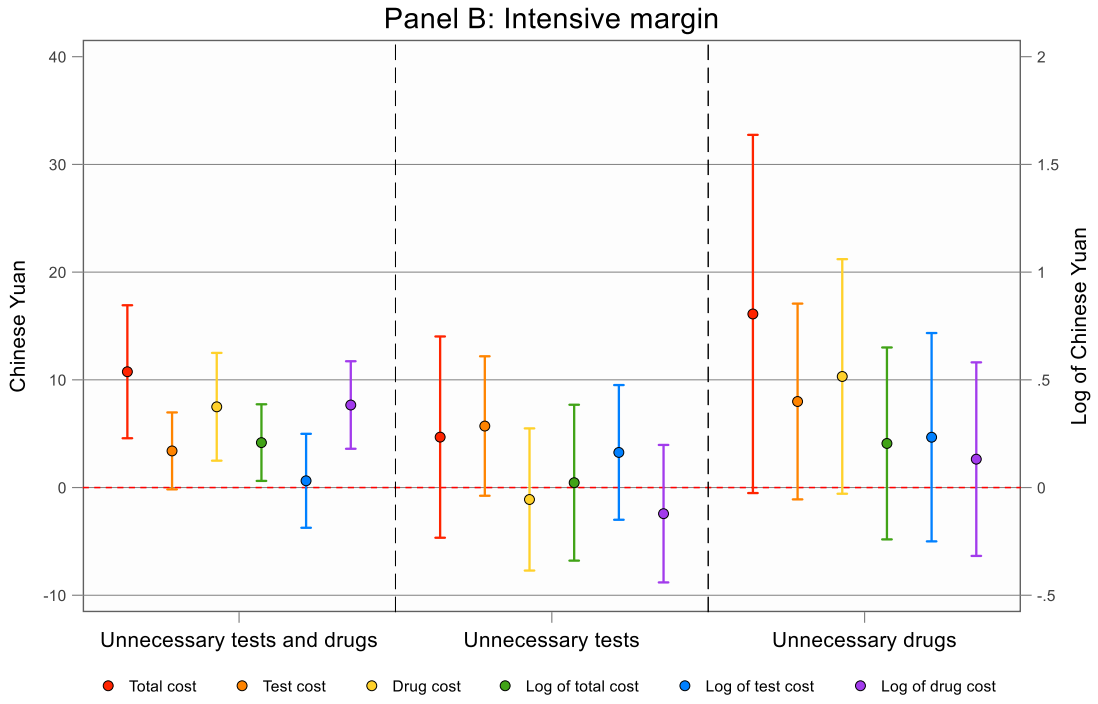


Figure 2 Financial impact of overuse of health care

Note: The financial impact was estimated using econometric models. The extensive margin denotes whether or not the physician-patient interactions involved overuse; the intensive margin denotes the number of items if overuse happened in the specific interaction.

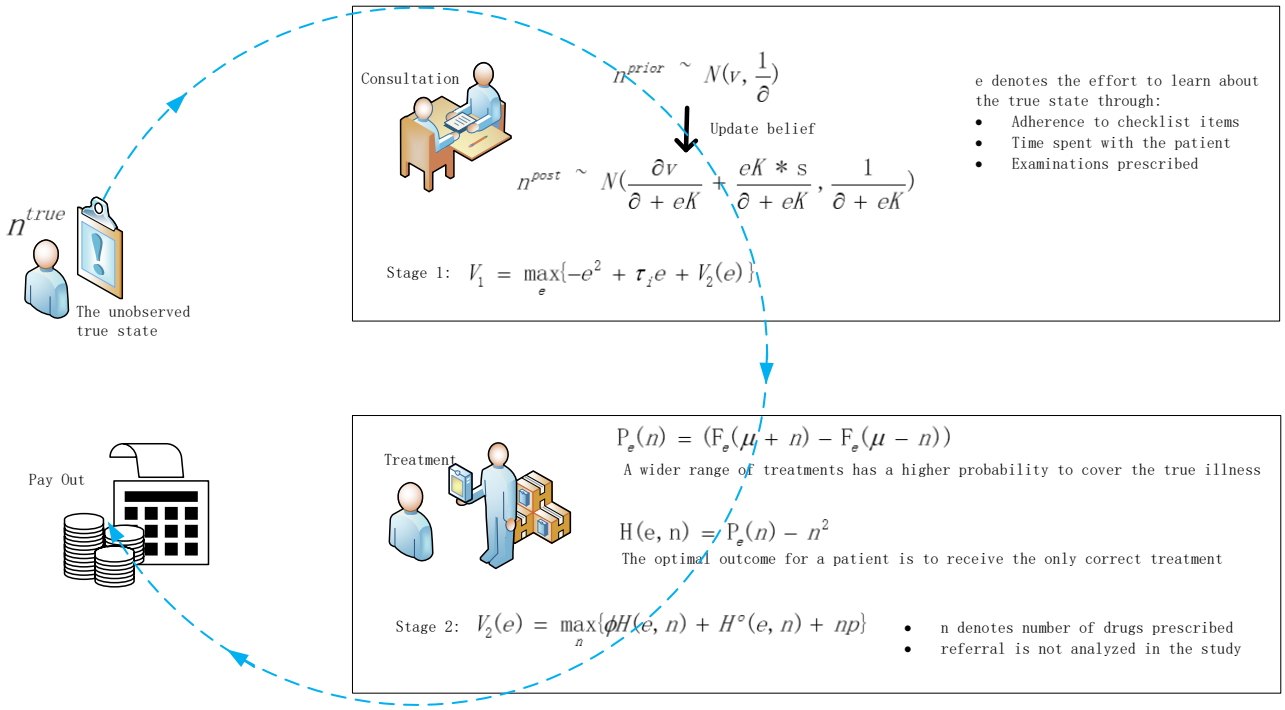
Table 1 Heterogeneity of financial impact of overuse

	(1) Total cost	(2) Test cost	(3) Drug cost	(4) Total cost	(5) Test cost	(6) Drug cost	(7) Total cost	(8) Test cost	(9) Drug cost
Overuse	1.081**	0.348	1.560***						
	[0.128, 2.034]	[-0.422, 1.117]	[0.633, 2.486]						
Overuse of tests				-0.407	0.098	-0.676			
				[-1.172, 0.359]	[-0.593, 0.788]	[-1.571, 0.218]			
Overuse of drugs							2.240***	1.077**	3.072***
							[1.433, 3.046]	[0.246, 1.909]	[2.469, 3.676]
Male SP	0.260	0.109	0.357	-0.073	0.078	-0.254	0.412*	0.355	0.285*
	[-0.419, 0.938]	[-0.474, 0.693]	[-0.402, 1.117]	[-0.794, 0.648]	[-0.366, 0.522]	[-1.061, 0.553]	[-0.043, 0.867]	[-0.108, 0.818]	[-0.044, 0.614]
Male physician	0.064	0.098	0.036	0.057	-0.019	0.032	-0.170	-0.129	-0.129
	[-0.622, 0.750]	[-0.503, 0.698]	[-0.392, 0.465]	[-0.546, 0.660]	[-0.535, 0.497]	[-0.516, 0.579]	[-0.660, 0.319]	[-0.618, 0.360]	[-0.313, 0.054]
Private hospital	0.108	0.227	-0.252	-0.336	0.223	-0.706**	0.422	0.539	-0.240**
	[-0.797, 1.014]	[-0.536, 0.991]	[-0.841, 0.336]	[-1.056, 0.385]	[-0.477, 0.922]	[-1.304, -0.108]	[-0.195, 1.039]	[-0.136, 1.214]	[-0.476, -0.004]
High-competence physician	0.070	-0.034	0.110	-0.294	-0.167	-0.296	0.089	-0.079	0.276*
	[-0.607, 0.746]	[-0.585, 0.516]	[-0.522, 0.743]	[-0.803, 0.215]	[-0.513, 0.179]	[-0.885, 0.293]	[-0.446, 0.624]	[-0.604, 0.446]	[-0.003, 0.555]
Senior physician	-0.013	-0.156	0.052	-0.046	-0.248	-0.009	0.462*	0.402*	0.075
	[-0.664, 0.638]	[-0.721, 0.408]	[-0.439, 0.543]	[-0.545, 0.452]	[-0.732, 0.235]	[-0.564, 0.547]	[-0.020, 0.943]	[-0.056, 0.859]	[-0.120, 0.271]
Asthma	-0.678*	-0.989***	0.272	-0.144	-1.146***	0.768**	0.125	0.121	-0.060
	[-1.382, 0.027]	[-1.485, -0.493]	[-0.336, 0.879]	[-0.674, 0.385]	[-1.490, -0.802]	[0.159, 1.377]	[-0.303, 0.552]	[-0.280, 0.522]	[-0.294, 0.175]
#Male SP	-0.066	0.407	-0.610	0.393	0.317	0.382	0.312	-0.017	0.561**
	[-0.969, 0.838]	[-0.421, 1.234]	[-1.518, 0.298]	[-0.570, 1.357]	[-0.409, 1.043]	[-0.584, 1.348]	[-0.383, 1.007]	[-1.266, 1.231]	[0.007, 1.116]
#Male physician	-0.227	-0.443	-0.033	-0.247	-0.285	-0.077	0.070	-0.383	0.197
	[-0.970, 0.515]	[-1.127, 0.241]	[-0.602, 0.536]	[-1.005, 0.511]	[-0.943, 0.373]	[-0.720, 0.566]	[-0.546, 0.687]	[-1.007, 0.241]	[-0.325, 0.720]
#Private hospital	-0.048	0.285	-0.460	0.670	0.248	0.267	-0.532	-0.266	0.001
	[-1.199, 1.104]	[-0.633, 1.204]	[-1.222, 0.303]	[-0.469, 1.810]	[-0.770, 1.265]	[-0.407, 0.942]	[-1.238, 0.174]	[-1.356, 0.825]	[-0.763, 0.765]
#High-competence physician	0.116	-0.194	0.308	0.612*	-0.041	0.921***	-0.310	-0.532	-0.450*
	[-0.649, 0.882]	[-0.876, 0.489]	[-0.446, 1.061]	[-0.104, 1.329]	[-0.643, 0.562]	[0.237, 1.605]	[-1.099, 0.479]	[-1.325, 0.261]	[-0.984, 0.084]
#Senior physician	0.288	0.335	-0.018	0.419	0.621**	-0.005	-0.600**	-0.761**	-0.023
	[-0.406, 0.982]	[-0.300, 0.970]	[-0.609, 0.573]	[-0.208, 1.046]	[0.020, 1.222]	[-0.674, 0.664]	[-1.170, -0.030]	[-1.391, -0.130]	[-0.522, 0.475]
#Asthma	0.363	0.792**	-0.967***	0.021	1.090***	-1.219***	-0.266	-1.171***	0.147
	[-0.506, 1.232]	[0.046, 1.539]	[-1.681, -0.252]	[-0.829, 0.872]	[0.378, 1.803]	[-2.048, -0.390]	[-0.998, 0.465]	[-1.822, -0.520]	[-0.475, 0.770]

N	492	492	492	492	492	492	492	492	492
Adjusted R^2	0.15	0.04	0.20	0.07	0.11	0.19	0.20	0.05	0.66

Note: # indicates an interaction term of the overuse variable with the hospital, physician or patient characteristics. Health care expenditure was used in its logarithmic form. County, day of week, month and year fixed effects were controlled for in the regression. Robust standard errors, clustered at the community health centre level, are presented in parentheses. *10% significance level. **5% significance level. ***1% significance level.

Figure 3 Theoretical model of physician behaviour



Note: The figure is based on Das et al. (2016b) and provides a conceptual explanation of physician behaviour regarding medical tests and drugs.