Quantifying the Financial Impact of Overuse in Primary Care in China

Yafei Si (yafei.si@unsw.edu.au)
UNSW Sydney

Hazel Bateman
UNSW Sydney

Shu Chen
UNSW Sydney

Katja Hanewald
UNSW Sydney

Bingqin Li
University of New South Wales https://orcid.org/0000-0003-2240-0876

Min Su
Inner Mongolia University

Zhongliang Zhou
Xi'an Jiaotong University

Article

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

Posted Date: April 14th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1550259/v1

Abstract

Objective

Healthcare expenditures have increased very fast in many countries. Overuse of health care is a potential factor explaining the rapid increase. However, overuse, defined as 'the provision of health care service when its likely risk of harm exceeds its potential benefit' is difficult to measure [1]. This study employs a novel method using unannounced standardised patients (SPs) to identify overuse, document its patterns, and quantify its financial impact on patients in primary care in China.

Methods

We conducted an SP study in a capital city of western China in 2017 and 2018. We trained 18 SPs to present consistent cases of two common chronic diseases, unstable angina and asthma. The SPs recorded 492 physician-patient interactions in 63 public and private primary hospitals. Overuse, defined as the provision of unnecessary medical tests and drugs, was identified by a judging panel based on national clinical guidelines. We estimated linear regression models to quantify the financial impact of overuse after controlling for hospital, physician and patient characteristics and a series of fixed effects.

Findings

We found overuse in 72.15% (95% CI: 68.18–76.13%) of the SP visits, including overuse of medical tests (54.67%, 95% CI: 50.26–59.09%) and overuse of drugs (28.05%, 95% CI: 24.07–32.03%). 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 varied across patients presenting different diseases. Compared to the non-overuse group, the overuse had significantly increased the total cost by 118.8% (95% CI: 75.5–162.2%), the test cost by 60.0% (95% CI: 22.2–97.9%) and the drug cost by 100.2% (95% CI: 56.8–143.7%). The financial impact of overuse did not vary substantially across high-competence and low-competence physicians, male and female physicians, junior and senior physicians, and patients presenting different diseases. However, the financial impact of overuse was driven by physician-patient interactions in public hospitals rather than private hospitals, and it was driven by female patients rather than male patients.

Interpretation

Overuse of health care is pervasive in the primary care of China and leads to a significant increase in healthcare expenditure. The overuse in this study seems unlikely to be attributable to physician incompetence. We use a consultation-treatment transaction model to understand physician behaviour and show that overuse of medical tests and overuse of drugs are partial substitutes. These findings shed light on the cost escalation of primary care in China, which is a form of medical inefficiency and low quality of care that should be urgently addressed.

Introduction

Globally healthcare expenditure has risen steadily while it increased much faster in China, with an annual growth rate of over 10% in the past decade [2]. As a result, healthcare expenditure in China has 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 [3], accounting for over 29% of its government spending and 7% of its gross domestic product (GDP) in 2020. The reasons behind this rapid growth are multifold, including economic growth, rapid population ageing, an epidemiological transition to chronic diseases, high-risk health behaviours, and detrimental environmental factors [4]. As for healthcare financing, encouragingly, the increase was partly dominated by the Chinese government's injection of significant funding since 2009, reflecting its strong commitment of providing all citizens with equal access to basic primary care with reasonable quality and financial risk protection [5]. However, the out-of-pocket spending from individuals was positively associated with government spending and kept rising [6]. The striking increase in healthcare expenditure directly threatens health system sustainability and healthcare affordability.

Overuse could be an important factor contributing to the rapid increase in healthcare expenditures. Overuse, defined as 'the provision of health care service when its likely risk of harm exceeds its potential benefit' [1], has increasingly become a global concern for academics and medical practices [7–9]. Overuse is unlikely to make patients healthier, but harm patients physically, psychologically and financially and even damages an individual's intrinsic capacity in the late-life [10,11]. There is strong evidence for the widely acknowledged presence of overuse of a specific service. For example, overuse has been examined in infusion [12], screening tests [13–16], and antibiotics [17,18]. Overuse can be conceptualised as an agency problem [19–21], where the interest of physicians and patients conflict [22–24]. Since there is a widely acknowledged information asymmetry between medical professionals and customers [19,25], physicians can induce patients to use more health care services than the amount they need [25–27]. This is the case in both developed and developing countries, including China [18,28–30]. Failure to understand the nature of overuse and the extent of it can cause serious distortion and inefficiency, therefore leading to unintended consequences of programs designed to reduce healthcare expenditure.

However, measuring systematic overuse and developing robust evidence for its prevalence among patients is a significant challenge [9]. First, the identification of overuse in developed countries such as Australia[31], Israel [32], Spain [33] and USA [21], relies heavily on complete administrative data, which is hard to obtain for less developed countries. Second, defining 'the appropriateness of care' is difficult since harms are poorly documented in many clinical practices [8]. Also, even if clinical guidelines are available for determining 'the appropriateness of care', sufficient details about patients are needed [34]. Third, evidence-based diagnostic and treatment guidelines generally overlook comorbidity [35], which is common in real medical practice [36]. Fourth, patient preferences may affect physician decision making, which makes it more challenging to define 'the appropriateness of care' [37]. With the challenges documented above, researchers are beginning to conceptualise overuse as a general system problem and to develop system-level metrics [38]. Yet existing evidence on the scope of overuse is far more complete, and its potential financial impact on patients is poorly understood.

In this study, we use a standardised patient (SP) method to accurately identify overuse, document its patterns and quantify its financial impact on patients in the primary care of China. SPs are well trained "acting patients" who can present symptoms of a disease like any other regular patients but with a standardised background setting. The standardised patient setting enables researchers to set benchmarks for the likely benefits and harms of a specific medical test and drug [39]. Second, physician practices can be compared with predefined guidelines because the actual illness and the optimal care associated are known [40,41]. Third, well-trained SPs can consistently present diseases, thus controlling for unobservable factors that are related to patient preference, health-seeking behaviours and communication style [42]. Fourth, SP can record the detailed process of the physician-patient interactions and all

prices charged. Thus a SP study is less subject to recall bias than approaches relying on recall-based patient exit surveys and chart abstraction [43].

In this study, we aim to identify overuse, document its patterns, and quantify its financial impact on patients in primary care in China. 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 of China. We recruited and trained 18 SPs from local communities. For the direct evidence of overuse, we use the SP method to rule out confounding factors from the patient side and then compare physician practice with predefined guidelines to identify the presence of overuse [44]. To quantify the financial impact of overuse, our identification strategy is similar to Mullainathan and Obermeyer (2021). SPs were sent randomly 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 [45]. Conditioning on the day of week and location of an SP visit, this provides plausible exogenous variation in the overuse. Finally, we examined heterogeneity in the prevalence of overuse and its financial impact among hospitals, physicians, and patients.

Results

Of the 492 physician-patient interactions between 18 SPs and 264 physicians, 78 interactions (15.85%) were in private hospitals, 217 interactions (44.11%) were involved by high-competence physicians, 247 interactions (50.20%) were performed by patients presenting asthma, 224 interactions (45.53%) were performed by male physicians, and 81 interactions (16.46%) were performed by 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 72.15% of the SP visits (95% CI: 68.18–76.13%) involved overuse of health care (Table S2), including overuse of medical tests (54.67%, 95% CI: 50.26–59.09%), overuse of drugs (28.05%, 95% CI: 24.07–32.03%), or both (10.57%, 95% CI: 7.8–13.30%). The average number of unnecessary items was 1.88 (95% CI: 1.77 to 1.99), including 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).

The high prevalence of overuse held among public and private hospitals, low-competence and high-competence physicians, male and female physicians, junior and senior physicians, male and female patients, and patients presenting asthma and unstable angina (Fig. 1). However, the overuse rate among low-competence physicians (81.09%, 95% CI: 76.43–85.75%) was significantly higher than that among high-competence physicians (60.83%, 95% CI: 54.28–67.38%); the overuse rate among patients presenting asthma (86.23%, 95% CI: 81.91–90.56%) was significantly higher than that among patients presenting unstable angina (57.96%, 95% CI: 51.73–64.18%).

The pattern held for overuse of medical tests, except that the rate of overuse of medical tests among male patients (66.67%, 95% CI: 56.18–77.16%) was significantly higher than that among female patients (52.31%, 95% CI: 47.46–57.16%). Interestingly, we found that the rate of overuse of drugs was almost the same across low-competence physicians and high-competence physicians; that the rate of overuse of drugs among patients presenting asthma (23.48%, 95% CI: 18.16–28.80%) was significantly lower than that among patients presenting unstable angina (32.65%, 95% CI: 26.74–38.57%); that the rate of overuse of drugs among male patients (11.11%, 95% CI: 4.12–

18.10%) was significantly lower than that among female patients (31.39%, 95% CI: 26.88–35.89%). The distribution of unnecessary items was similar to the distribution of overuse (Table S2).

The unadjusted results above could be confounded by unobservable hospital and physician factors not included in the model. Using logistic regressions by adjusting for day of week, month, year and county fixed effect, we found no or weak evidence that the rate of overuse varied across public and private hospitals, low-competence and high-competence physicians, male and female physicians, junior and senior physicians (Table S3). The probability of overuse among patients presenting asthma was significantly higher than that of patients presenting unstable angina. So was the probability of overuse of medical tests. However, the probability of overuse of drugs among male patients was significantly lower than that among female patients. The number of unnecessary items was almost the same across these hospital, physician and patient characteristics using negative binomial regressions (Table S3).

Financial Impact of Overuse

First, we compared the healthcare expenditure between the overuse and non-overuse group and found that the healthcare expenditure in the overuse group was much higher than that of the non-overuse group (Fig. 2), including the total cost, the test cost and the drug cost. The differences were statistically significant at 1% levels (Table S4). Surprisingly, 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 (Fig. 3 Panel A and Table S5), the overuse significantly led to a 27.45 CNY (95% CI: 17.42 to 37.48) increase in total cost, a 14.42 CNY (95% CI: 7.11 to 21.73) increase in test cost, and a 12.90 CNY (95% CI: 5.85 to 19.95) increase drug cost, respectively, corresponding to a 118.8% (95% CI: 75.5–162.2%) increase in total cost, a 60.0% (95% CI: 22.2–97.9%) increase in test cost, and a 100.2% (95% CI: 56.8–143.7%) increase in drug cost, compared with the non-overuse group. Second, overuse of medical tests significantly led to a 73% (95% CI: 41–105%) increase in test cost and a 67% (95% CI: 29–105%) decrease in drug cost, without a significant change in total cost. Third, overuse of drugs significantly led to a 307% (95% CI: 274–341%) increase in drug cost and a 163% (95% CI: 120–206%) increase in total cost, but no significant change in test cost.

At the intensive margin (Fig. 3 Panel B and Table S5), the cost increase incurred by an additional item of unnecessary service is comparably small. For example, a one-item increase in unnecessary service significantly led to a 20.8% (95% CI: 3.1–38.6%) increase in total cost and a 38.3% (95% CI: 17.9–58.7%) increase in drug cost, respectively, but no significant change in test cost. We did not find that a one-item increase in unnecessary tests (/drugs) would significantly decrease drug (/test) cost, although the coefficients were still negative. Also, a one-item increase in unnecessary tests led to no significant change in total cost and test cost. However, a one-item increase in unnecessary drugs significantly led to a 60.2% (95% CI: 27.8–92.5%) in total cost and a 139.9% (95% CI: 103.2–176.7%) increase in drug cost.

Heterogeneity of the Financial Impact

We examined the heterogeneity of the financial impact of overuse by estimating econometric models using hospital, physician and patient subgroups. Overall, the results did not vary substantially across low-competence and high-competence physicians, male and female physicians, junior and senior physicians, and patients presenting different diseases (Table 1). For example, we found that the overuse significantly led to an 82.7% (95% CI: 0.6–163.8%) increase in total cost among low-competence physicians. In contrast, overuse significantly led to a 129.4% (95% CI:

74.6–184.2%) increase in total cost among high-competence physicians. Yet a *Chow-Test* indicated that there was no statistically significant difference between the two coefficients (*F = 0.58, p-value = 0.450*) [46].

Table 1
Heterogeneity of financial impact of overuse

	Public hospitals			***Private hospitals***		
Panel A	(1)	(2)	(3)	(4)	(5)	(6)
	Total cost	Test cost	Drug cost	Total cost	Test cost	Drug cost
Overuse	1.171***	0.546**	1.071***	1.113*	0.864	0.579
	[0.678,1.664]	[0.131,0.961]	[0.565,1.577]	[-0.239,2.464]	[-0.715,2.443]	[-1.139,2.296]
N	414	414	414	78	78	78
Adjusted R^2	0.33	0.25	0.28	0.56	0.53	0.23

	Low competence			***High competence***		
Panel B	(1)	(2)	(3)	(4)	(5)	(6)
	Total cost	Test cost	Drug cost	Total cost	Test cost	Drug cost
Overuse	0.827**	0.347	0.867**	1.294***	0.685***	1.083***
	[0.006,1.648]	[-0.342,1.035]	[0.054,1.680]	[0.746,1.842]	[0.215,1.156]	[0.435,1.731]
N	275	275	275	217	217	217
Adjusted R^2	0.26	0.23	0.13	0.38	0.26	0.45

	Unstable angina			***Asthma***		
Panel C	(1)	(2)	(3)	(4)	(5)	(6)
	Total cost	Test cost	Drug cost	Total cost	Test cost	Drug cost
Overuse	1.019***	0.423	1.126***	1.665***	1.126***	0.687*
	[0.336,1.702]	[-0.172,1.017]	[0.563,1.688]	[0.947,2.382]	[0.363,1.888]	[-0.066,1.440]
N	245	245	245	247	247	247
Adjusted R^2	0.33	0.15	0.44	0.40	0.39	0.11

	Female physicians			*Male physicians*		
Panel D	(1)	(2)	(3)	(4)	(5)	(6)
	Total cost	Test cost	Drug cost	Total cost	Test cost	Drug cost
Overuse	1.403***	0.640**	1.217***	1.082**	0.517	0.853*
	[0.879,1.928]	[0.124,1.156]	[0.651,1.783]	[0.229,1.935]	[-0.241,1.276]	[-0.014,1.720]
N	268	268	268	224	224	224
Adjusted R^2	0.40	0.26	0.26	0.26	0.23	0.29

	Female patients			*Male patients*		
Panel E	(1)	(2)	(3)	(4)	(5)	(6)
	Total cost	Test cost	Drug cost	Total cost	Test cost	Drug cost
Overuse	1.280***	0.625***	1.110***	0.681	0.422	0.321
	[0.783,1.777]	[0.173,1.076]	[0.603,1.617]	[-2.757,4.119]	[-2.886,3.730]	[-3.524,4.165]
N	411	411	411	81	81	81
Adjusted R^2	0.33	0.24	0.32	0.32	0.37	-0.01

	Junior physicians			*Senior physicians*		
Panel F	(1)	(2)	(3)	(4)	(5)	(6)
	Total cost	Test cost	Drug cost	Total cost	Test cost	Drug cost
Overuse	1.335**	0.479	0.934	1.145***	0.516*	1.059***
	[0.202,2.469]	[-0.590,1.548]	[-0.185,2.052]	[0.540,1.750]	[-0.068,1.100]	[0.431,1.686]
N	152	152	152	340	340	340
Adjusted R^2	0.35	0.33	0.27	0.32	0.29	0.21

Note: Healthcare expenditure was used in its logarithmic form. Covariates include physician gender, physician age, patient gender, hospital ownership, and participating health alliance or not, except for the specific heterogeneity analysis. County, hospital, day of week, year, case, and patient fixed effect 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.

Interestingly, the financial impact of overuse in the study was driven by interactions in public hospitals instead of by interactions in private hospitals. For example, overuse significantly led to a 117.1% (95% CI: 67.8–166.4%) increase in total cost in public hospitals, while we only found weak evidence that the overuse was associated with a 111.3% (95% CI: -23.9–246.4%) increase in total cost in private hospitals. Although the comparably small number of private hospitals could make the estimation less precise, the sample size reflected a small market share occupied by private

hospitals in China. Moreover, we found the financial impact of overuse was driven by interactions with female patients (128.0%, 95% CI: 78.3–177.7%) rather than by interactions with male patients (68.1%, 95% CI: -275.7–411.9%).

Discussion

There is an increasing global interest in optimising health care delivery and improving the quality of care [47]. However, previous research has primarily focussed on reducing underuse and the failed delivery of needed services [48], while its counterpart, overuse, is comparatively less studied [49]. To our knowledge, this is the first study using a standardised patient 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 that the overuse led to a significant increase in healthcare expenditure. Compared to the non-overuse group, overuse increased the total cost by 118.8%, the exam cost by 60.0% and the drug cost by 100.2%. Also, we showed that in China public hospitals provided as many unnecessary services as private hospitals did in China, leading to a significant increase in health care expenditure. The financial impact of overuse was driven by interactions with female patients rather than male patients.

Using the SP method to measure overuse offers a unique opportunity to explore physician practices and accurately estimate overuse. Our results are comparable to other studies in China and other developing countries. For example, the prevalence of overuse in this study was overall higher than that in rural China [17,40], and similar to the level overuse found in India and Kenya [44,50]. One typical difference of this study was the focus on community health centres in China. Community health centres are generally well equipped with medical equipment. For example, more than half of community health centres in China provided inpatient care in 2017 and even had specialist departments [51]. Moreover, in our study setting community health centres can provide (almost) all the essential medical tests for the two tracer conditions [52]. Therefore, we identified pervasive overuse of unnecessary medical tests, which were generally absent from the earlier investigations in (comparably poor) rural areas [17,40,41] and less developed regions [50,53].

The high prevalence of overuse in the study was dominated by physician behaviours rather than patient behaviours. In theory, overuse can be influenced by hospital, physician and patient characteristics. For example, patients may prefer better technology and medical tests for reassurance [54]; may demand more services than those needed when covered by health insurance and not bearing the entire costs [55]; and may request costly medications due to the direct-to-consumer marketing strategy for new drugs [56]. However, in our setting, SPs were required not to request any service initially and therefore, we had a clean estimation of overuse that is only derived from provider behaviour. The reasons why physicians overprovide services are multifold. For example, physicians may be trained to identify (to exclude or confirm) all possible diagnoses using medical tests when they were medical students [57]. Furthermore, physicians may show so-called second-degree moral hazard when patients are covered by health insurance [23]. Also, physicians may prefer to be paid and do something in a fee-for-service system, even although these services did not add much value for patient health [58]. Also, concerns about uncertain malpractice lawsuits in the future may also lead physicians to do more to defend themselves [59–61], especially for physicians with experience of previous medical disputes [62].

We provide new evidence on the physician-induced demand theory and a clean quasi-experimental estimation of the economic significance of overuse. Physicians tend to prescribe more services in response to either an increase or a

decrease in the related reimbursement rate [25,27,63–65]. The economic significance is impressive, but similar to findings from other credence good markets [66]. In 2018, there were 4.41 billion outpatient visits in the primary care, accounting for 53.04% of total outpatient visits in China, and the outpatient healthcare expenditure per capita in primary care was 156.8 CNY [67], representing a market of 691.5 billion CNY (≈108.6 billion US dollars). The average cost of SP visit in the study was 35.0 CNY, with a standard deviation of 41.3. The relatively low cost is reasonable because we only used two non-complicated and common chronic diseases. We estimated the financial burden incurred by the overuse on the specific population that is affected by the two chronic diseases by considering the prevalence and the increased cost. In our sample, it was estimated that the overuse led to a 56.59% increase in total cost, a 54.50% increase in exam cost, and a 61.20% increase in drug cost in the population with the two chronic diseases. Our results strongly suggest that overuse can potentially place enormous costs on patients and whole economies. Eliminating the use of unnecessary medical tests and drugs would substantially reduce physical harm and the risk of poly-pharmacy [11], and simultaneously reduce the financial burden on the population.

In general, the unregulated provision of health care by the private sector is not socially desired [50], since providers may over-respond to demand, leading to socially inefficient provision. Alternatively, the default policy approach to delivering health care, especially primary care, is through public hospitals. In this study, we found that the overuse rate in public hospitals was as high as that in private hospitals. Moreover, the financial impact of overuse was driven by interactions in public hospitals rather than private hospitals. This may suggest that financial incentives in the public health sector in China are as strong as in the private health sector, or even stronger. Previous research indicated that financial incentives among public hospitals in China originated from market-oriented reforms since the 1980s [68]. From that time, 1) health care professionals did not receive a wage from the government anymore; 2) the prices charged for physician service time were strictly capped by the government; 3) hospitals were permitted to earn profits from prescribing new drugs and high-tech examinations [69] and physicians received substantial bonuses from hospital profits. This context led to the pervasive over-prescription of unnecessary and expensive drugs and high-tech examinations [18].

Many studies suggest that physicians are substantially underpaid in public hospitals of China [70] due to the regulation on prices of physician service time. Physicians often have no legally mandated social benefits in primary care [51], and are commonly burned out [71]. However, the last decade witnessed a significant increase in the health care workforce in China, including the physician numbers in primary care [72]. The expansion of primary care under an underpayment system might be explained by the pervasive overuse in China [73]. Overuse led to a significant increase in healthcare expenditure, which may proxy for an underlying market price to compensate for physician nominal wage. The financial incentives persisted after the latest health system reform in China, including the introduction of an essential medicine list and a *zero-markup* drug policy [74,75]. The consequence is likely that physicians used more unnecessary medical tests when drug prescription was heavily regulated, as we found in the study.

However, financial incentives may not be the main issue since overuse can also be influenced by physician knowledge [17,76,77], professional ethics [78], altruistic behaviours [79], and practice norms [80]. Our study is limited in that we did not measure physician knowledge and other characteristics at the data collection [17]. However, the financial impact of overuse was similar across high-competence and low-competence physicians. Since physicians may have given a correct diagnosis by luck but no other service, 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 (Fig. 4), and the continuum should serve as a more precise proxy for physician competence in the study. Also, we found similar

results using physician age as a proxy for experience in medical practice. These results suggest the pervasive overuse and its associated cost in our study is unlikely to be attributable to physician incompetence.

We did not find that the overuse and its financial impact varied across female and male physicians, although recent studies suggest female physicians did better in metrics of physician-patient relationships and health outcomes. Surprisingly, the financial impact of overuse was driven by interactions with female patients rather than male patients. Theoretically, the overuse-associated physician agency problem is related to information asymmetry between physicians and patients. It is a relative position because patients varied in health literacy for making informed decisions [25,27], and therefore physician behaviours changed correspondingly. Women and girls in China are more vulnerable, less educated and earned less compared to their male counterparts [81,82], especially among the older cohorts like our setting. These disadvantages were finally manifested as low health literacy among female adults [83] and therefore physicians could become more aggressive in the provision of unnecessary medical tests and treatment.

This study has several limitations. First, the main limitation of the SP method is that only a few types of diseases (two in the study) can be presented. However, the SP method has good external validity when the detection rate (zero in the study) is reasonably low. For example, physician behaviours are consistent between SPs visits and real patient visits [50,53]. Our results were robust across the two diseases, but more evidence is needed in the future. Second, the SP method focuses on actual medical practice rather than physician knowledge [40,84]. 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. For example, growing evidence indicates a gap between physician knowledge and practice [84–86]. Third, we present the evaluation for one-time new patient interactions with different physicians, and the situation perhaps will change when the continuity of care for the same physician-patient model is promoted. Further research could assess patient preference for specific physicians after his/her initial interaction. Finally, whether and how the overuse would influence physician decision making in improving other metrics of quality of care and patient wellbeing is beyond the scope of this study, and these remaining questions will be evaluated in follow-up projects.

Overall, the study has important implications for health system reforms in China. China has undergone tremendous demographic and epidemiological transitions during the past thirty years [5]. The increasing burden from chronic diseases and an ageing population have presented great health care challenges for the country [87]. Recent health system reforms encouraged the first contact with lower-tier health care providers and primary care providers [88], which may substantially affect patients with chronic diseases. Our study shows that overuse of health care is pervasive in the primary care of China and leads to a significant increase in health care expenditure. These findings shed light on the cost escalation of primary care in China, where overuse is a form of medical inefficiency and low quality of care that should be urgently addressed.

First, while we cannot exclude many factors incurring overuse from the physician side, overuse in this study seems unlikely to be attributable to physician incompetence. Thus additional clinical training does not seem to be an effective policy intervention because its impact on the quality of care depends on how much the additional training increases physician competence; and how much the (increased) competence, in turn, is reflected in practice. Previous studies indicated that both effects could be minimal in medical practice [53,84]. This may also help explain why a previous study did not find a positive impact of clinical training on the quality of care in the same setting [89]. Similar to the case in India, overuse among qualified physicians with sufficient knowledge is as common as that

among unqualified providers [50]. The clinical training significantly improved the quality of health care provided but failed to reduce the adverse effect of overuse [53].

Second, we show that overuse of medical tests and overuse of drugs are partial substitutes. We find that 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 use a consultation-treatment transaction model to understand physician behaviour (Fig. 5) [50,77]. The main insight from the model is that 1) physicians can induce patients to consume more services either at the consultation stage (medical tests) or the treatment stage (drugs), and 2) physicians can shift efforts and medical resources between the two stages to make the global budget seemingly balanced [60]. Therefore, the consultation stage could be as expensive as the treatment stage in practice, but the consultation stage has been comparably less studied in the previous literature [18,77]. The good news of this finding is that, in a sense, even without government regulation, market accountability (i.e., competition, reputation) can motivate physicians to keep the total health care expenditure to a particular upper bound. The cost indicator could be easily observed by patients and compared within their social network, especially for common diseases. It is well possible that patients would be unsatisfied if the health care expenditure is somewhat too high and absurd. In this case, the overuse leads to a significant but finite increase in health care expenditure.

Methods

This section describes the institutional context, the SP method, the definition of overuse and healthcare expenditure in this study, and the econometric model to identify the financial impact of overuse.

Institutional Context

In China, the primary care system provides general clinical care and basic public health services. Primary care hospitals provide outpatient and limited inpatient services, mainly for common clinical conditions. Primary care requires multidisciplinary professional teams, including physicians, nurses and pharmacists etc. Despite being the backbone of the primary care workforce, primary care physicians are unevenly distributed across China and are poorly trained in some less developed areas like our setting. In general, community health centres provide service in urban areas while township health centres and village medical centres operate in rural areas.

In general, primary care hospitals provide walk-in services for patients without an appointment. Patients need to cover the cost for outpatient visits almost entirely out-of-pocket, both in public and private hospitals in our setting. This is because social health insurance policies mainly provide coverage for inpatient services but limited coverage for outpatient care through setting low annual caps for total reimbursement. The settings lead to patients using more inpatient services, even for minor health conditions, rather than primary outpatient care. This could be one driver of the overuse in primary care in China.

Standardised Patient

In this research, we used data collected from a standardised patient study in a capital city of 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 symptoms of an illness to a physician like any other regular patients. SPs are used routinely in the training and evaluation of medical students in high-income countries [90]. SPs were coached to present their initial symptoms and answer any questions that the physician may ask as part of history taking, in a manner consistent with the underlying condition [50,91]. The hospitals and physicians granted approval for the study three months

before. 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 presented by SPs), unstable angina and asthma, to identify overuse in this study. The incidence of the two tracer conditions was high in the study region. Asthma is a condition in which a person's airways become inflamed, narrow, swell and produce extra mucus, which makes it difficult to breathe. Asthma can usually be managed with rescue inhalers to treat symptoms and controller inhalers that prevent symptoms. Asthma affects 1–18% the populations in a range of countries [92]. The symptoms of asthma can be minor, can interfere with daily activities, but may also lead to a life-threatening attack. Asthma is prevalent but largely undiagnosed and undertreated in China [93]. Unstable angina is also referred to as acute coronary syndrome. The disease is a type of chest discomfort caused by poor blood flow through the blood vessels (coronary vessels) of the heart muscle (myocardium). It may lead to a heart attack and should be treated as an emergency as suggested by the *American Health Association*. In 2017, stroke and ischaemic heart disease were the leading causes of death in China [94], with 149 and 124 in every 100,000 deaths attributed to stroke and heart attack [95].

In the study, SPs were randomly assigned one of the two tracer conditions with standardised background information. For unstable angina, a 50-year-old patient had worsening chest pain recently. For asthma, a 40-year-old patient had a worsening breathing problem since last night. The hypothetical patient was covered by a moderate social health insurance in China, the Urban Resident Health Insurance [96]. The two hypothetical cases are not complicated but were specifically chosen so that the opening statement by the SPs would be consistent with multiple underlying illnesses. In addition, the diseases can be portrayed easily and have a low risk of receiving invasive examinations by design. The two tracer conditions had explicitly predefined guidelines for physician practice and were adapted from earlier studies in India and China [40,41,91]. A panel of doctors advised 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, different 18 SPs (person-years) were selected after interviewing, including 10 SPs in 2017 and 8 SPs in 2018. The SPs participated in rigorous training before visiting the physicians working in the primary care (i.e., community health centres in our settings). We included all community health centres (CHCs) in 7 districts of the capital city. To exclude a potential "sorting effect" where patients deliberately selected 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 hospitals and physicians. The SPs were randomly assigned to a CHC on a workday like a regular patient and then visited the first physician in the first office of that particular CHC.

Physicians may change their behaviours when being aware of being observed. None of the 18 SPs reported that their acting in visits had been revealed. Therefore, we were measuring physician practice in a real-world scenario and *Hawthorne effects* are not a concern because the physicians did *not know* that they were being observed [97]. However, the physicians had given their consent to be included in the study and did *know* that some SPs would randomly visit them in a given period. The SPs accepted all non-invasive medical tests and rejected all invasive medical tests according to 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. After the visit, SPs were required to report the physician-patient interaction immediately using a structured questionnaire. Importantly, the detailed interaction between the SPs and physicians was 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 12 SPs and 269 physicians.

Healthcare Expenditure

Healthcare expenditure (measured in Chinese Yuan, CNY; exchange rate, 6.37 CNY ≈ 1 US dollar) mainly included a fixed consultation fee, medical test fee and drug fee the primary medical setting of China. We include all costs of medical tests that physicians performed and planned to perform for one specific visit, although SPs rejected all invasive tests by giving reasonable excuses (Note, in China, patients usually do not reject medication prescribed by a physician). Consistent with this approach, we used three continuous variables to measure healthcare expenditure, the cost of medical tests (the test cost) and the cost of the prescribed medications (the drug cost) and the total cost (including the test cost, the drug cost and the consultation fee). The healthcare expenditure was also transformed using a natural logarithm to enhance interpretation.

Overuse of Health Care

The SP method allows us to directly measure the overuse in CHCs of China. Consistent 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 and over-treatment, mostly referring the inappropriate usage of types of services. Based on the case-specific interaction (Supplement 1), each medical test was classified as essential or unnecessary while each drug was classified as correct or unnecessary (and even harmful), based on the predefined guidelines and the inputs from a judging panel of doctors and pharmacists. We used one binary variable to measure the presence of overuse of any unnecessary service (0-no, 1-yes). Another two binary variables were also 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 unnecessary tests and drugs prescribed (the total unnecessary items).

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 government. The difference of funding sources is likely to affect physician decision of providing unnecessary services [50,52]. (2) Physician competence: we used one binary variable, whether physicians giving a correct diagnosis for this interaction, to proxy for physician competence. It was classified as high-competence if yes or low-competence if no. Additionally we used a predicted probability of physician giving a correct diagnosis for the interaction in a sensitivity analysis [76,77]. (3) Diseases: we used two different chronic diseases, asthma and unstable angina. The diversity of diseases is related to external validity of the study. (4) Physician gender: we examined whether physician behaviour in overuse varied across physician gender because growing evidence supports that female physicians perform better than their male counterparts in adherence to clinical protocols [98], psychosocial consultation [99,100], performance on standardised tests [101], and readmission rate [102]. (5) Patient gender: patient gender may play a role in influencing the overuse via physician’s taste-based discrimination [103], or internalised negative beliefs about one’s gender [104]. (6) Physician age: we used physician age as a proxy of working experience. All physicians were classified as the junior (aged under 40) and the senior (aged 40 and above).

Econometric Model

The estimation strategy to identify the financial impact of overuse is straightforward in the study. The econometric specification is:

$$y_{ijt}^{k} = \beta_0 + \beta_1 Overuse_{ijt}^{m} + \beta_2 X_{ijt} + \beta_3 W_{ijt} + \pi_i + \delta_j + \phi_t + \mu_{ijt} + \omega_{ijt} + \epsilon_{ijt}^{k}, \quad (1)$$

where y_{ijt}^{k} represents the health care expenditure indicator k (CNY or its log form) that is analysed in hospital i district j on day t. $Overuse_{ijt}^{m}$ is the variable of interest, indicating overuse (m = ext) and the number of unnecessary items (m = int). X_{ijt} is a set of the observable demographic correlates of the physicians. W_{ijt} is a set of the observable hospital characteristics. π_{ij} are hospital fixed effects, δ_j are district fixed effects and μ_{ijt} are disease fixed effect. ϕ_t are day of week, month, and year fixed effect. ω_{ijt} are SPs fixed effect. ϵ_{ijt}^{k} is the error term. Robust standard errors were clustered at the CHC level.

Using the econometric specification, we examined the financial impact of overuse using the original continuous variables of healthcare expenditure and its natural logarithm forms, at the extensive margin (m = ext) and the intensive margin (m = int). For the extensive margin, we used three binary variables, the total overuse, overuse of medical tests and overuse of drugs; for the intensive margin, we used three count variables, the total unnecessary items, the unnecessary tests and drugs.

Declarations

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.

Code availability: 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, USA; National Natural Science Foundation of China, China; The University of New South Wales, Australia; ARC Centre of Excellence in Population Ageing Research (CEPAR), 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; 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

1. Chassin, M. R. & Galvin, R. W. The urgent need to improve health care quality: Institute of Medicine National Roundtable on Health Care Quality. JAMA **280**, 1000–1005 (1998).
2. Chang, A. Y. *et al.* 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 (2019).
3. Statista. China: health expenditure. *Statista* https://www.statista.com/statistics/279400/health-expenditures-in-china/.
4. WHO. *Healthy China: Deepening Health Reform in China: Building High-Quality and Value-Based Service Delivery*. (World Bank Publications, 2019).
5. Yip, W. *et al.* 10 years of health-care reform in China: progress and gaps in universal health coverage. The Lancet **394**, 1192–1204 (2019).
6. Zhang, D. & Rahman, K. A. 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 (2020).
7. Mulley, A. G. 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 (2013).
8. Morgan, D. J. *et al.* Setting a research agenda for medical overuse. *BMJ* h4534 (2015) doi:10.1136/bmj.h4534.
9. Brownlee, S. *et al.* Evidence for overuse of medical services around the world. The Lancet **390**, 156–168 (2017).
10. Cross, A. J. *et al.* Potentially inappropriate medications and anticholinergic burden in older people attending memory clinics in Australia. Drugs Aging **33**, 37–44 (2016).
11. WHO. *Integrated care for older people (ICOPE): guidance for person-centred assessment and pathways in primary care*. (2019).
12. Zhang, Y., Zhou, Z. & Si, Y. When more is less: What explains the overuse of health care services in China? Soc. Sci. Med. **232**, 17–24 (2019).
13. Schwartz, L. M., Woloshin, S., Fowler Jr, F. J. & Welch, H. G. Enthusiasm for cancer screening in the United States. Jama **291**, 71–78 (2004).
14. Mathias, J. S., Gossett, D. & Baker, D. W. Use of electronic health record data to evaluate overuse of cervical cancer screening. J. Am. Med. Inform. Assoc. **19**, e96–e101 (2012).
15. Tan, A., Kuo, Y.-F. & Goodwin, J. S. Potential overuse of screening mammography and its association with access to primary care. Med. Care **52**, 490 (2014).
16. Alber, J. M. *et al.* Reducing overuse of cervical cancer screening: A systematic review. Prev. Med. **116**, 51–59 (2018).
17. Xue, H. *et al.* Diagnostic ability and inappropriate antibiotic prescriptions: a quasi-experimental study of primary care providers in rural China. J. Antimicrob. Chemother. **74**, 256–263 (2019).
18. Currie, J., Lin, W. & Meng, J. Addressing antibiotic abuse in China: an experimental audit study. J. Dev. Econ. **110**, 39–51 (2014).
19. Arrow, K. J. Uncertainty and the welfare economics of medical care. Am. Econ. Rev. **53**, 941–973 (1963).
20. Orszag, P. R. The overuse, underuse, and misuse of health care. Wash. DC Congr. Budg. Off. (2008).
21. Berwick, D. M. & Hackbarth, A. D. Eliminating waste in US health care. JAMA **307**, 1513–1516 (2012).

22. Dulleck, U. & Kerschbamer, R. On doctors, mechanics, and computer specialists: The economics of credence goods. J. Econ. Lit. **44**, 5–42 (2006).
23. Balafoutas, L., Kerschbamer, R. & Sutter, M. Second-degree moral hazard in a real-world credence goods market. Econ. J. **127**, 1–18 (2017).
24. Gottschalk, F., Mimra, W. & Waibel, C. Health services as credence goods: A field experiment. Econ. J. **130**, 1346–1383 (2020).
25. Johnson, E. M. & Rehavi, M. M. Physicians treating physicians: Information and incentives in childbirth. Am. Econ. J. Econ. Policy **8**, 115–41 (2016).
26. Rizzo, J. A. & Blumenthal, D. Is the target income hypothesis an economic heresy? Med. Care Res. Rev. **53**, 243–266 (1996).
27. Si, Y. *et al.* Comparison of health care utilization among patients affiliated and not affiliated with healthcare professionals in China. BMC Health Serv. Res. **20**, 1118 (2020).
28. Currie, J., Lin, W. & Zhang, W. Patient knowledge and antibiotic abuse: Evidence from an audit study in China. J. Health Econ. **30**, 933–949 (2011).
29. Lu, F. Insurance coverage and agency problems in doctor prescriptions: evidence from a field experiment in China. J. Dev. Econ. **106**, 156–167 (2014).
30. Wu, B. Physician agency in China: evidence from a drug-percentage incentive scheme. J. Dev. Econ. **140**, 72–89 (2019).
31. Elshaug, A. G., Watt, A. M., Mundy, L. & Willis, C. D. Over 150 potentially low-value health care practices: an Australian study. Med. J. Aust. **197**, 556–560 (2012).
32. Goldberg, A., Yalonetsky, S., Kopeliovich, M., Azzam, Z. & Markiewicz, W. Appropriateness of diagnosis of unstable angina pectoris in patients referred for coronary arteriography. Exp. Clin. Cardiol. **13**, 133 (2008).
33. Aguilar, M. D., Fitch, K., Lázaro, P. & Bernstein, S. J. The appropriateness of use of percutaneous transluminal coronary angioplasty in Spain. Int. J. Cardiol. **78**, 213–221 (2001).
34. Huang, X. & Rosenthal, M. B. Overuse of Cardiovascular Services: Evidence, Causes, and Opportunities for Reform. Circulation **132**, 205–214 (2015).
35. Boyd, C. M. *et al.* Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA **294**, 716–724 (2005).
36. Van Weel, C. & Schellevis, F. G. Comorbidity and guidelines: conflicting interests. Lancet **367**, 550–550 (2006).
37. Blank, T., Graves, K., Sepucha, K. & Llewellyn-Thomas, H. Understanding treatment decision making: contexts, commonalities, complexities, and challenges. Ann. Behav. Med. **32**, 211–217 (2006).
38. Segal, J. B. *et al.* Identifying possible indicators of systematic overuse of health care procedures with claims data. Med. Care 157–163 (2014).
39. *et al.* 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 (2019).
40. Sylvia, S. *et al.* Survey using incognito standardized patients shows poor quality care in China's rural clinics. Health Policy Plan. **30**, 322–333 (2015).
41. Sylvia, S. *et al.* Tuberculosis detection and the challenges of integrated care in rural China: A cross-sectional standardized patient study. PLoS Med. **14**, (2017).
42. Kwan, A. *et al.* Use of standardised patients for healthcare quality research in low-and middle-income countries. BMJ Glob. Health **4**, e001669 (2019).

43. Spivak, E. S., Cosgrove, S. E. & Srinivasan, A. Measuring appropriate antimicrobial use: attempts at opening the black box. Clin. Infect. Dis. **63**, 1–6 (2016).
44. Sulis, G. *et al.* 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 (2020).
45. Mullainathan, S. & Obermeyer, Z. Diagnosing Physician Error: A Machine Learning Approach to Low-Value Health Care. Q. J. Econ. qjab046 (2021) doi:10.1093/qje/qjab046.
46. Wooldridge, J. M. *Introductory econometrics: A modern approach*. (Cengage learning, 2015).
47. Macleod, M. R. *et al.* Biomedical research: increasing value, reducing waste. The Lancet **383**, 101–104 (2014).
48. Glasziou, P. *et al.* Evidence for underuse of effective medical services around the world. The Lancet **390**, 169–177 (2017).
49. Saini, V., Brownlee, S., Elshaug, A. G., Glasziou, P. & Heath, I. Addressing overuse and underuse around the world. The Lancet **390**, 105–107 (2017).
50. Das, J., Holla, A., Mohpal, A. & Muralidharan, K. Quality and accountability in health care delivery: audit-study evidence from primary care in India. Am. Econ. Rev. **106**, 3765–99 (2016).
51. Li, X. *et al.* The primary health-care system in China. The Lancet **390**, 2584–2594 (2017).
52. Su, M. *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 **18**, 5060 (2021).
53. Das, J., Chowdhury, A., Hussam, R. & Banerjee, A. V. The impact of training informal health care providers in India: A randomized controlled trial. Science **354**, (2016).
54. van Bokhoven, M. A. *et al.* Why do patients want to have their blood tested? A qualitative study of patient expectations in general practice. BMC Fam. Pract. **7**, 1–8 (2006).
55. Folland, S., Goodman, A. C. & Stano, M. *The Economics of Health and Health Care: Pearson New International Edition*. (Routledge, 2016).
56. Daubresse, M. *et al.* Effect of direct-to-consumer advertising on asthma medication sales and healthcare use. Am. J. Respir. Crit. Care Med. **192**, 40–46 (2015).
57. Danielson, A. R., Venugopal, S., Mefford, J. M. & Clarke, S. O. How do novices learn physical examination skills? A systematic review of the literature. Med. Educ. Online **24**, 1608142 (2019).
58. Zuvekas, S. H. & Cohen, J. W. Fee-for-service, while much maligned, remains the dominant payment method for physician visits. Health Aff. (Millwood) **35**, 411–414 (2016).
59. Green, A. R., Tung, M. & Segal, J. B. Older adults' perceptions of the causes and consequences of healthcare overuse: a qualitative study. J. Gen. Intern. Med. **33**, 892–897 (2018).
60. Si, Y., Zhou, Z., Su, M. & Chen, X. Revisiting gender gap in quality of health care in urban China: a standardised patient audit study. The Lancet **394**, S25 (2019).
61. Keane, M. P., McCormick, B. & Poplawska, G. Health care spending in the US vs UK: The roles of medical education costs, malpractice risk and defensive medicine. Eur. Econ. Rev. **124**, 103401 (2020).
62. He, A. J. 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 (2014).
63. Cromwell, J. & Mitchell, J. B. Physician-induced demand for surgery. J. Health Econ. **5**, 293–313 (1986).
64. Dranove, D. & Wehner, P. Physician-induced demand for childbirths. J. Health Econ. **13**, 61–73 (1994).

65. Fang, H., Lei, X., Shi, J. & Yi, X. *Physician-Induced Demand: Evidence from China's Drug Price Zero-Markup Policy*. (2021).
66. Kerschbamer, R., Neururer, D. & Sutter, M. Insurance coverage of customers induces dishonesty of sellers in markets for credence goods. *Proc. Natl. Acad. Sci.* **113**, 7454–7458 (2016).
67. China Statistics. China Statistical Yearbook for Health 2019. https://data.cnki.net/area/Yearbook/Single/N2020020200?z=D09.
68. Blumenthal, D. & Hsiao, W. Lessons from the East—China's rapidly evolving health care system. N. Engl. J. Med. **372**, 1281–1285 (2015).
69. Hesketh, T. & Zhu, W. X. Health in China: the healthcare market. Bmj **314**, 1616 (1997).
70. Jingang, A. Which future for doctors in China? The Lancet **382**, 936–937 (2013).
71. Ye, J. *et al.* Burnout among obstetricians and paediatricians: a cross-sectional study from China. BMJ Open **9**, e024205 (2019).
72. Tang, C. & Tang, D. The trend and features of physician workforce supply in China: after national medical licensing system reform. Hum. Resour. Health **16**, 1–8 (2018).
73. Li, Y. *et al.* Overprescribing in China, driven by financial incentives, results in very high use of antibiotics, injections, and corticosteroids. Health Aff. (Millwood) **31**, 1075–1082 (2012).
74. Xiao, Y., Zhao, K., Bishai, D. M. & Peters, D. H. Essential drugs policy in three rural counties in China: what does a complexity lens add? Soc. Sci. Med. **93**, 220–228 (2013).
75. Yi, H., Miller, G., Zhang, L., Li, S. & Rozelle, S. Intended and unintended consequences of China's zero markup drug policy. Health Aff. (Millwood) **34**, 1391–1398 (2015).
76. Currie, J. & MacLeod, W. B. Diagnosing expertise: Human capital, decision making, and performance among physicians. J. Labor Econ. **35**, 1–43 (2017).
77. Currie, J. M. & MacLeod, W. B. Understanding Doctor Decision Making: The Case of Depression Treatment. Econometrica **88**, 847–878 (2020).
78. Pellegrino, E. D. The metamorphosis of medical ethics: a 30-year retrospective. JAMA **269**, 1158–1162 (1993).
79. Olsen, K. R., Gyrd-Hansen, D., Boegh, A. & Hansen, S. H. GPs as citizens' agents: prescription behavior and altruism. Eur. J. Health Econ. **10**, 399–407 (2009).
80. Havighurst, C. C. The professional paradigm of medical care: obstacle to decentralization. Jurimetrics 415–429 (1990).
81. Ren, W., Rammohan, A. & Wu, Y. Is there a gender gap in child nutritional outcomes in rural China? China Econ. Rev. **31**, 145–155 (2014).
82. Wang, Y. Closing the gender gap in college attendance: Variation by family background in China over time. Soc. Sci. Res. **98**, 102578 (2021).
83. Yongbing, L. I. U. *et al.* The health literacy status and influencing factors of older population in Xinjiang. Iran. J. Public Health **44**, 913 (2015).
84. Das, J. & Hammer, J. Money for nothing: the dire straits of medical practice in Delhi, India. J. Dev. Econ. **83**, 1–36 (2007).
85. Leonard, K. L. & Masatu, M. C. Professionalism and the know-do gap: Exploring intrinsic motivation among health workers in Tanzania. Health Econ. **19**, 1461–1477 (2010).
86. Mohanan, M. *et al.* The know-do gap in quality of health care for childhood diarrhea and pneumonia in rural India. JAMA Pediatr. **169**, 349–357 (2015).

87. Glinskaya, E. & Feng, Z. *Options for aged care in China: Building an efficient and sustainable aged care system*. (The World Bank, 2018).
88. Zhou, Z. *et al.* Evaluating the effect of hierarchical medical system on health seeking behavior: A difference-in-differences analysis in China. Soc. Sci. Med. **268**, 113372 (2021).
89. Su, M., Zhou, Z., Si, Y. & Wei, X. 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 (2019).
90. Fenton, J. J. *et al.* Promoting patient-centered counseling to reduce use of low-value diagnostic tests: a randomized clinical trial. JAMA Intern. Med. **176**, 191–197 (2016).
91. Das, J. *et al.* 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 (2012).
92. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention 2018. (2018).
93. Huang, K. *et al.* Prevalence, risk factors, and management of asthma in China: a national cross-sectional study. The Lancet **394**, 407–418 (2019).
94. Luan, S., Yang, Y., Huang, Y. & McDowell, M. Public knowledge of stroke and heart attack symptoms in China: a cross-sectional survey. BMJ Open **11**, e043220 (2021).
95. Zhou, M. *et al.* 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 (2019).
96. Su, M. *et al.* 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 (2018).
97. Leonard, K. L. & Masatu, M. C. Using the Hawthorne effect to examine the gap between a doctor's best possible practice and actual performance. J. Dev. Econ. **93**, 226–234 (2010).
98. Bertakis, K. D., Helms, L. J., Callahan, E. J., Azari, R. & Robbins, J. A. The influence of gender on physician practice style. Med. Care (1995).
99. Roter, D. L., Hall, J. A. & Aoki, Y. Physician gender effects in medical communication: a meta-analytic review. JAMA **288**, 756–764 (2002).
100. Bertakis, K. D., Franks, P. & Epstein, R. M. Patient-centered communication in primary care: physician and patient gender and gender concordance. J. Womens Health **18**, 539–545 (2009).
101. Kim, C. *et al.* Is physician gender associated with the quality of diabetes care? Diabetes Care **28**, 1594–1598 (2005).
102. Tsugawa, Y. *et al.* Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians. JAMA Intern. Med. **177**, 206–213 (2017).
103. Becker, G. S. *The economics of discrimination*. (University of Chicago press, 2010).
104. Greenwood, B. N., Carnahan, S. & Huang, L. Patient–physician gender concordance and increased mortality among female heart attack patients. *Proc. Natl. Acad. Sci.* **115**, 8569–8574 (2018).

Figures

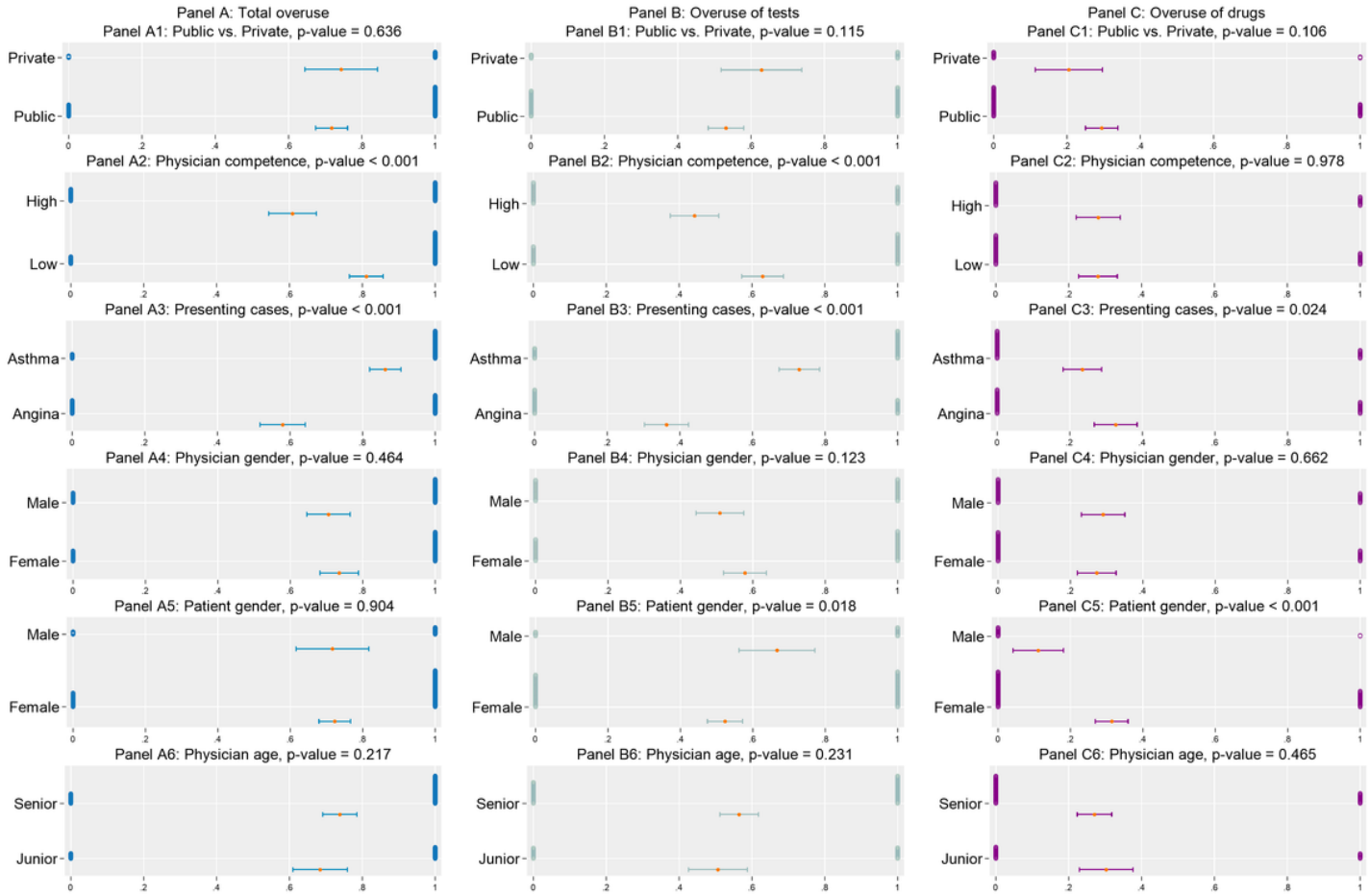

Figure 1

The prevalance of overuse among hospitals, physicians, and patients

Note: The graph reports mean and 95% confidence intervals for the prevalence of total overuse, overuse of tests and overuse of drugs. The bar chart suggets frequency of observations. The statistical differences are performed using chi-square test for binary variables.

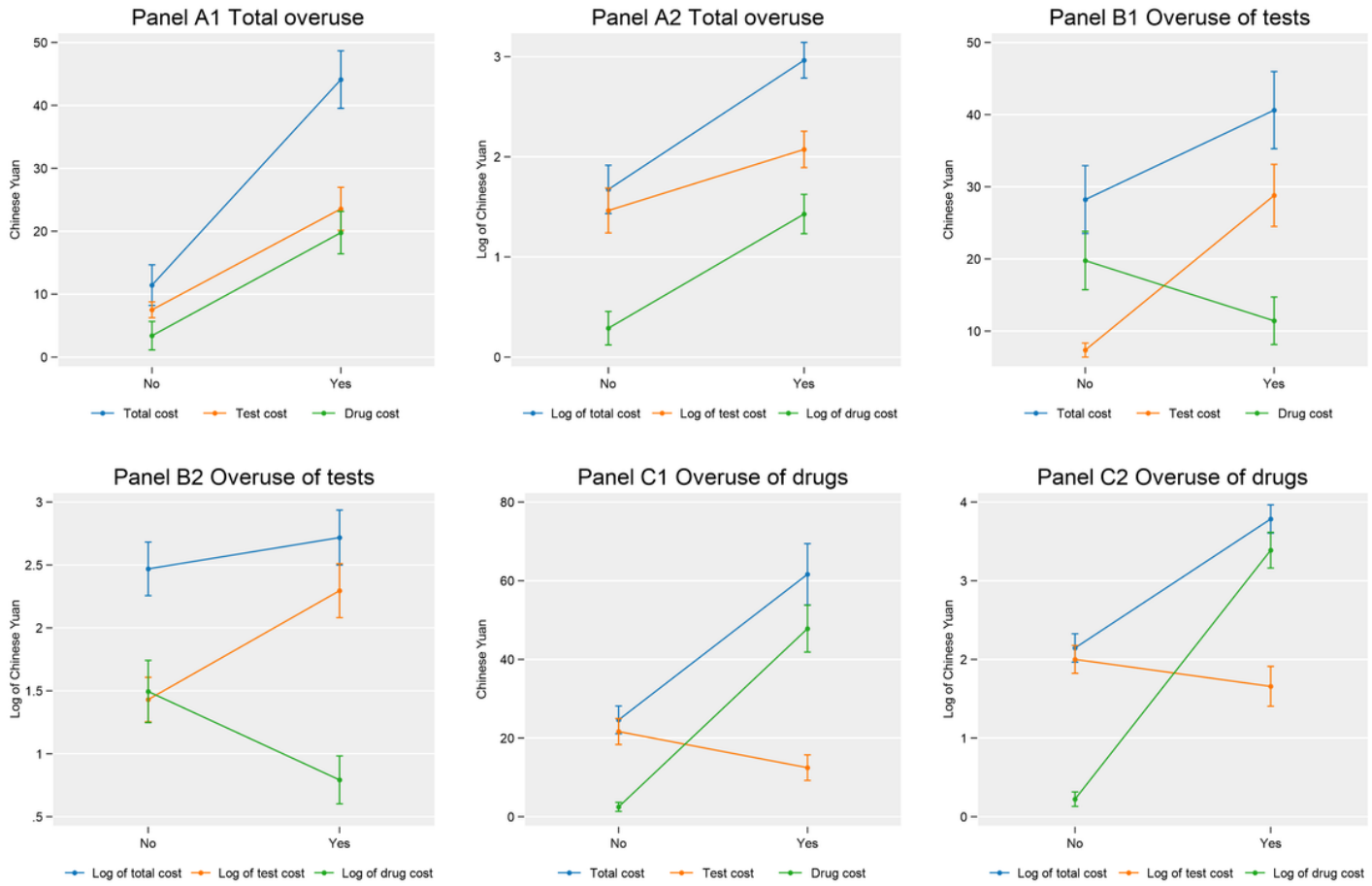

Figure 2

Healthcare expenditure and overuse of health care

Note: The distribution of health care expenditure was examined in Chinese yuan and its logarithmic form among three indicators, overuse of health care, overuse of tests, and overuse of drugs.

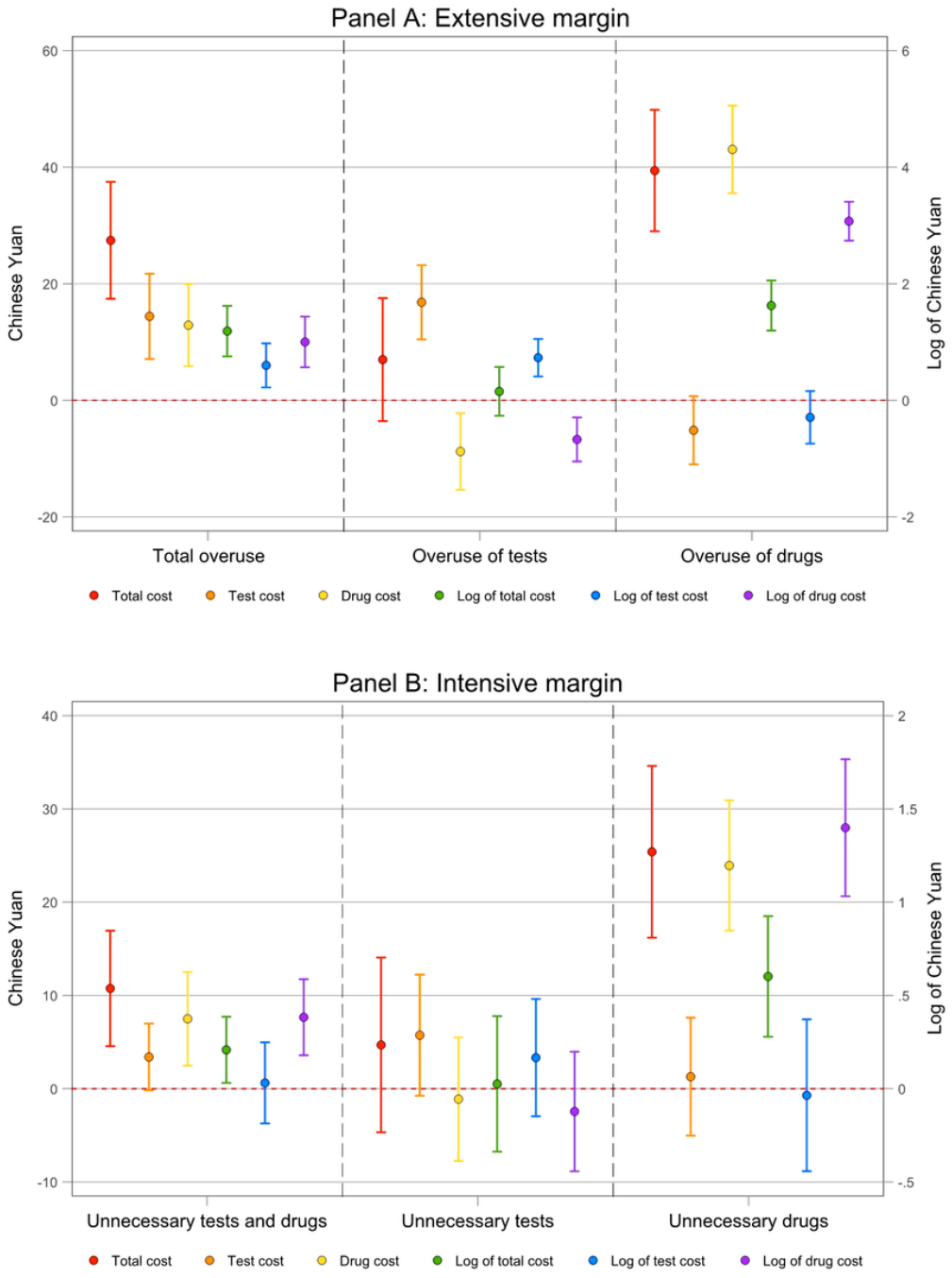

Figure 3

Financial impact of overuse of health care

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

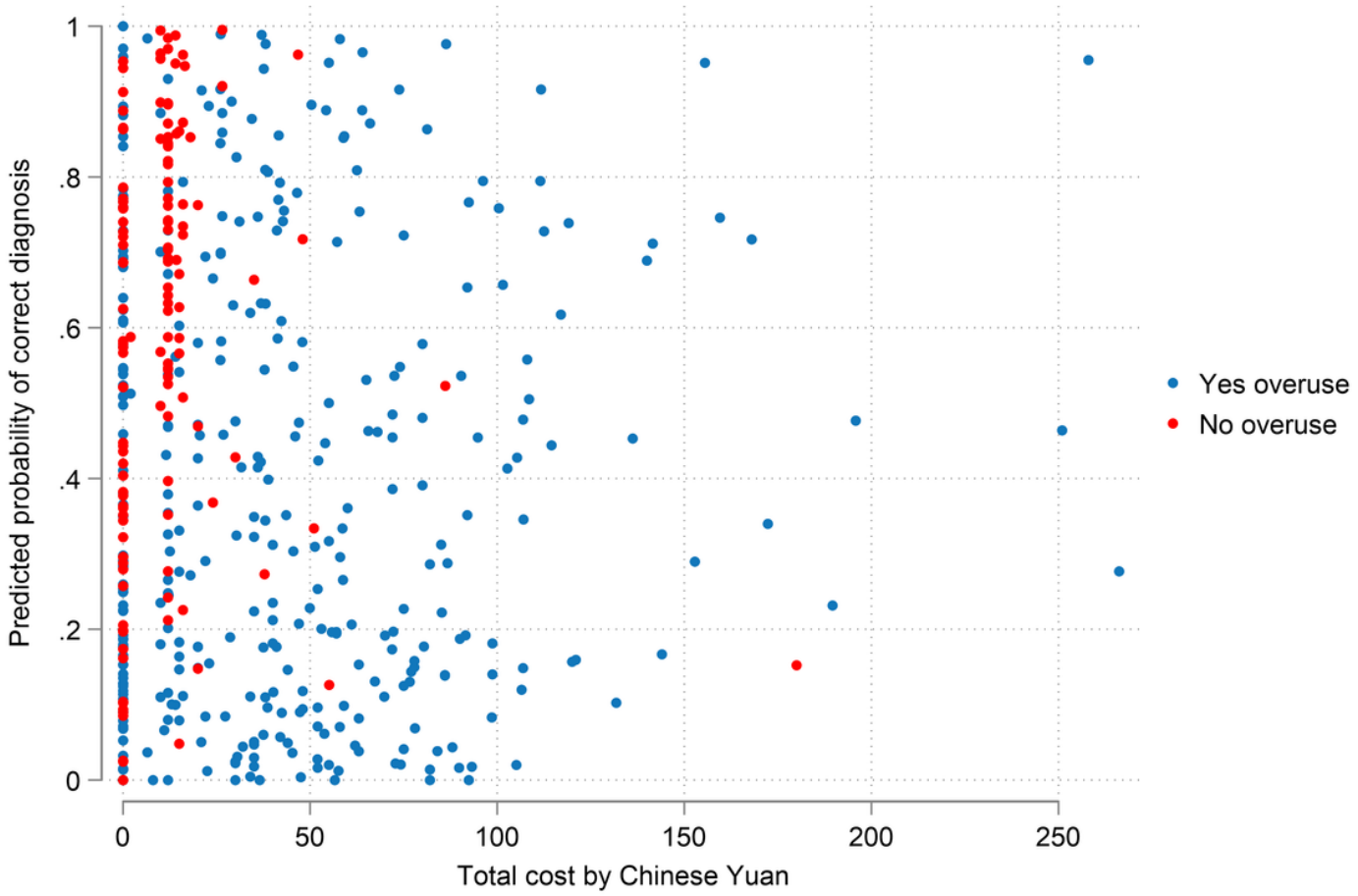

Figure 4

Physician competence and health care expenditure. Note: The predicted probability of correct diagnosis was estimated using a logistic regression.

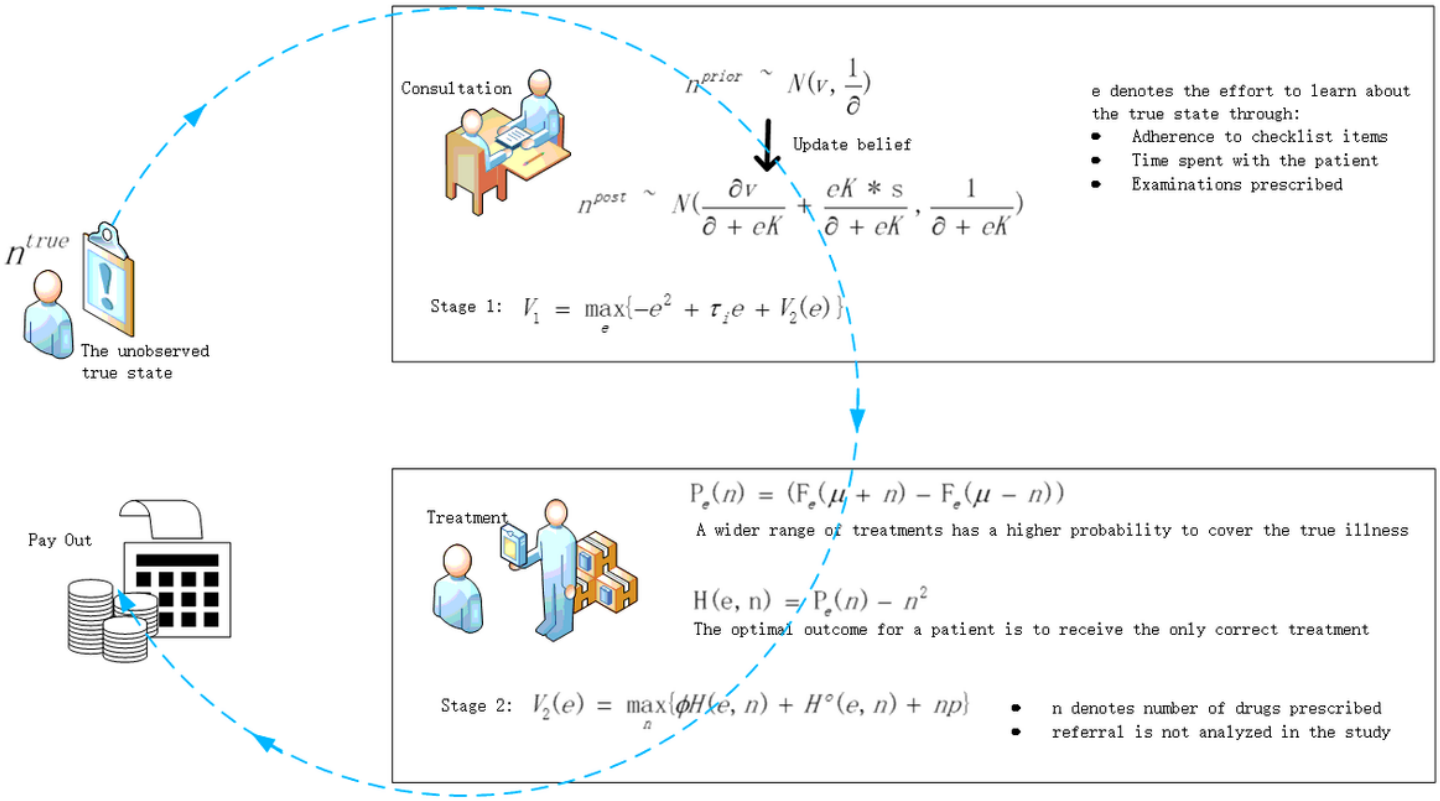

Figure 5

Theoretical model of physician behavior. Note: We assume the job of a physician is to identify the patient's true illness and perform effective and adequate treatments. At the beginning, the physician forms a prior belief about the diagnosis after the patient's opening statement. In the consultation stage, this is a *Bayesian Learning* process [50] where the physician continuously updates his belief about the true illness since there could be multiple underlying conditions. The physician is considered to exert costly effort (e) to learn about the true illness base on his knowledge level (k) through a number of checklist items, time spent with the patient, and examinations prescribed. It is noted that the greater the effort exerted, the more precise the diagnosis could be. At the end of the consultation stage, the physician would form a posterior belief about the true illness of the patient. In the treatment stage, the provider decides the number and types of treatment he plans to perform. This is modelled as a *Trials* process [77]. Given a wider range of treatments naturally leads to a higher probability of covering the correct treatment ( ) and curing the patient (

), it also increases health care expenditure for patients in a longer term. Therefore, the optimal outcome for a patient is to receive only the correct treatment but not any additional unnecessary treatments. In practice, providers will choose efforts and treatments to maximize their own utility too (

), which may not be aligned with those of patients. Here we (1) identify all overuse of health care at the consultation stage and the treatment stage, and (2) disentangle decision-making process of physicians across the consultation stage and the treatment stage.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementarymaterialsSPOveruse20220412.docx