Research

# BMJ Open

# Efficiency and productivity measurement of rural township hospitals in China: a bootstrapping data envelopment analysis

Zhaohui Cheng, Miao Cai, Hongbing Tao, Zhifei He, Xiaojun Lin, Haifeng Lin, Yuling Zuo

CrossMark

Department of Health Management, School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei Province, People's Republic of China

**Correspondence to**
Professor Hongbing Tao;
hhbtao@hust.edu.cn

## ABSTRACT

**Objective:** Township hospitals (THs) are important components of the three-tier rural healthcare system of China. However, the efficiency and productivity of THs have been questioned since the healthcare reform was implemented in 2009. The objective of this study is to analyse the efficiency and productivity changes in THs before and after the reform process.

**Setting and participants:** A total of 48 sample THs were selected from the Xiaogan Prefecture in Hubei Province from 2008 to 2014.

**Outcome measures:** First, bootstrapping data envelopment analysis (DEA) was performed to estimate the technical efficiency (TE), pure technical efficiency (PTE) and scale efficiency (SE) of the sample THs during the period. Second, the bootstrapping Malmquist productivity index was used to calculate the productivity changes over time.

**Results:** The average TE, PTE and SE of the sample THs over the 7-year period were 0.5147, 0.6373 and 0.7080, respectively. The average TE and PTE increased from 2008 to 2012 but declined considerably after 2012. In general, the sample THs experienced a negative shift in productivity from 2008 to 2014. The negative change was 2.14%, which was attributed to a 23.89% decrease in technological changes (TC). The sample THs experienced a positive productivity shift from 2008 to 2012 but experienced deterioration from 2012 to 2014.

**Conclusions:** There was considerable space for TE improvement in the sample THs since the average TE was relatively low. From 2008 to 2014, the sample THs experienced a decrease in productivity, and the adverse alteration in TC should be emphasised. In the context of healthcare reform, the factors that influence TE and productivity of THs are complex. Results suggest that numerous quantitative and qualitative studies are necessary to explore the reasons for the changes in TE and productivity.

## INTRODUCTION

China is a populous country, and its rural residents comprise 45.23% of the country's population.[1] Health policymakers face a daunting

**Strengths and limitations of this study**

- This study introduced the bootstrapping data envelopment analysis approach to measure the efficiency and productivity of township hospitals (THs) in China. This process is advantageous to draw reliable and robust conclusions.
- This study provided empirical evidence of the efficiency and productivity changes of THs before and after the healthcare reform implementation in 2009, as well as identified the possible reasons for the changes.
- The information on the case-mix index and the patients' outcome quality for each TH was excluded because of the dated hospital information system.
- THs in the other regions of China were excluded in this study.

challenge as they seek to meet the healthcare demands of this large group. The healthcare system of China comprises community health centres, and secondary and tertiary hospitals in the urban areas, whereas village clinics, township hospitals (THs) and county hospitals (from low-levels to high-levels) cover the rural areas.[2] THs are important components of the rural healthcare system that link village clinics and county hospitals, as wells as serve as the main providers of primary healthcare services in the rural areas. THs offer an extensive range of healthcare services by delivering curative activities and public health services. These services include outpatient visits, inpatient care, health education, vaccinations, family electronic health records (EHR) management (ie, from record archives establishment to regular management) and chronic patient health management. THs also supervise healthcare delivery at the grassroots level and act as gatekeepers that direct patients to specialised health facilities.[3]

The rural healthcare system of China has changed significantly since the economic reforms in 1978. From 1975 to 1990, the efficient three-tier healthcare delivery system and the Cooperative Medical System (CMS) nearly collapsed.[4][5] Government support to THs became scarce and insufficient. Hence, THs had to support their operations with substantial medical service charges and drug sales, thereby causing the enduring problem of covering hospital expenses with pharmaceutical revenue.[3][4] Consequently, several negative externalities surfaced: medical technical examinations were overused, drugs were overprescribed, healthcare quality declined and preventive measures were seldom assumed by THs.[3][6–8] The collapse of CMS also affected the referral system; thus, patients bypassed THs and decided to receive treatment in high-level healthcare facilities.[7] This situation was intensified by the increase of rural incomes, thereby leading patients to seek quality services.[4] Therefore, the extent of the TH activities was reduced. In 2003, the government implemented the New Rural Cooperative Medical Scheme (NCMS) with two objectives: to offer an insurance system to rural residents, which will reduce the financial barriers to healthcare access,[9] and to redirect patients back towards THs.[5] In 2009, the Chinese government launched a new round of healthcare reform, which had the primary goal of rebuilding an effective primary healthcare system. Accordingly, several measures have been adopted to achieve this objective. First, the government invested an additional US$127 billion to enhance the infrastructure of primary healthcare facilities, particularly those in the rural areas.[10] Second, government subsidies, instead of pharmaceutical revenues, would be the main financial source to support for the operation of primary care facilities.[11] In 2009, the National Essential Medicines Policy (NEMP) was launched, and one of its priorities was the Essential Medicine List (EML), which aimed to select low-cost medicine for common treatable diseases. All public primary healthcare facilities (including all THs) were required to purchase and prescribe only drugs in EML and to sell them at cost price (zero profit).[12] By the end of 2013, China had 37 015 THs.[13] The efficiency and productivity of THs are crucial with regard to these hospitals' strategic position in the rural healthcare delivery chain and the changes they experienced over the years, as well as in the context of scarce resources, vertical and horizontal competition, and health insurance and healthcare reforms.[14] The efficiency and productivity are important considerations of THs' managers and policymakers under these conditions.

There has been extensive literature dealing with the efficiency and productivity of healthcare organisations, and data envelopment analysis (DEA) is extensively acknowledged and used. DEA was first applied by Sherman[15] to identify and measure hospital efficiency. Thereafter, many studies have adopted DEA to evaluate healthcare efficiency using micro-level[16–20] or macro-level data sets.[21][22] The reviews of Hollingsworth[21] and

O'Neill et al[22] indicated that the literature on the efficiencies of health facilities mainly involves North American and European case studies. Although an increasing number of studies on hospital efficiency have been conducted in developing countries, only a few studies have investigated this issue in China. Ng[23] evaluated the productivity of 463 hospitals in Guangdong Province and discovered a deterioration in productivity with advancements in technology. Audibert et al[14] measured the effects of NCMS on the technical efficiency (TE) of 24 THs in Weifang and determined that TE declined from 2000 to 2008. Li et al[24] explored the productivity of 12 tertiary hospitals in Beijing from 2006 to 2009 and found that the sample hospitals experienced productivity growth with technological advancements. Cheng et al[25][26] analysed the TE and productivity of county hospitals in Henan and Jiangsu Provinces during the reform process and found a considerable space for TE improvement.

The objectives of this study were to explore the efficiency and productivity changes in the sampled THs in rural China before and after the healthcare reform implementation, as well as to discuss the possible reasons for the changes.

## METHODOLOGY

TE was developed by Farrell in 1957[27] and refers to the capacity of a decision-making unit (DMU) to transform a quantity of inputs into an amount of outputs. Two main methods have been extensively applied to measure hospital efficiency: the parametric method of stochastic frontier analysis (SFA) and the non-parametric method of DEA. SFA needs to construct a function and is limited to a single output, whereas DEA is limited to its non-parametric nature that restricts the possibility of including covariates in the model.[28] The current study applied DEA because it is the most extensively employed method in production frontier analysis, does not require assumptions on the functional form and can handle multiple inputs and outputs.[29]

### Data Envelopment Analysis

The CCR (developed by Charnes, Cooper and Rhodes),[30] BCC (developed by Banker, Charnes and Cooper)[31] and the Malmquist productivity index (MPI)[32–34] are DEA models that have been extensively used to measure efficiency and productivity. The CCR model assumes that production is constant return to scale (CRS), where an increase in the input(s) will result in a proportional increase in the output(s).[30] The BCC model assumes that production is variable return to scale (VRS), that is, an increase in the input(s) will result in either an increase or decrease in the output(s) when units are not operating at optimum scale. TE measured using the CCR model may be altered by scale efficiency (SE).[31] VRS has two dimensions: increasing returns to scale (IRS), where a 1% increase in inputs is

followed by over 1% increase in outputs; and decreasing returns to scale (DRS), where a 1% increase in inputs will result in below 1% increase in output.[25] The BCC model calculates the pure technical efficiency (PTE) that incorporates the effect of SE:[35]

$$TE_{DEA-CRS} = PTE_{DEA-VRS} \times SE$$

## Malmquist productivity index

Productivity measures changes in a production unit's efficiency in transforming inputs into outputs from time t to time t+1.[23] MPI was named after Malmquist[33] and was introduced by Caves, Christensen and Diewert to evaluate productivity changes among different production units.[32] MPI (also called total factor productivity changes (TFPC)) can be decomposed into technical efficiency changes (TEC) and technological changes (TC). TEC can also be decomposed into pure technical efficiency changes (PTEC) and scale efficiency changes (SEC).[34]

$$TFPC = TEC \times TC = (PTEC \times SEC) \times TC$$

## Output orientation

TE can be evaluated via an input-oriented or output-oriented model that depends on the reasons for conducting DEA. Pelone et al[36] systematically reviewed 39 DEA studies on primary healthcare efficiency evaluation and found that 24 and 9 studies used an input-oriented and output-oriented models, respectively. The current study used an output-oriented model, which defines the frontier by maximising the outputs of THs while holding the amounts of inputs constant. This preference is justified by the Chinese government's implementation of a series of reform measures to improve the capability of THs and establish a well-organised hierarchical treatment and referral system.[11 37] The service quality and reputation of THs have gradually improved, and numerous patients have once again availed of the services of THs; hence, THs can augment their outputs at their given inputs to increase efficiency.[38] Moreover, TH managers have limited control over their inputs. Hospital decisions involving staff recruitment, appointment and investment are controlled by multiple government departments.[39] Thus, the current study prefers an output-oriented model than the alternative input-oriented model.

## Bootstrapping DEA and bootstrapping MPI

In the DEA models, only DMUs operating in the frontier are considered efficient. However, all DMUs are subject to environmental and random factors, that is, their efficiency scores fall into a fluctuating range.[40] Thus, efficiency scores generated using the traditional DEA models may be biased.[41] Simar and Wilson et al[28 42–45] introduced the bootstrapping DEA method to correct the bias in the efficiency and productivity scores. Efron[46]

introduced the bootstrapping technique by simulating the data-generating process to obtain a new estimate for each simulated sample. The simulated data set is approximately equivalent to the original one; therefore, the sampling distributions and SDs approximate the original ones. By applying bootstrapping DEA, the estimated efficiency can be drawn by conducting repeated sampling, thereby resulting in bias-corrected efficiency and confidential intervals at the α level; the estimates obtained would mimic the distribution of the real estimator, thereby making the efficiency scores considerably accurate.[42] Simar and Wilson[42 47] developed a smooth bootstrapping procedure to accurately estimate the efficiency and productivity scores. The bootstrapping DEA method has been applied internationally in the relative efficiency and productivity measurements.[48] However, to the best of our knowledge, this method has yet to be introduced in China to measure the efficiency of healthcare facilities in this country.[24 49] Online appendices 1 and 2 illustrate the entire procedure algorithm that calculates the efficiency and productivity using the bootstrapping DEA and bootstrapping MPI. The smooth bootstrapping procedures were implemented using the R software and command *boot. sw98* in the FEAR package.[50]

## DATA AND VARIABLES
## Data

The original data set comprises 48 selected THs in the Xiaogan Prefecture in Hubei Province; these hospitals were observed from 2008 to 2014. Data were collected from the National Health Statistical Information Report System (NHSIRS) of the Xiaogan Health Bureau Database. No patient information was involved in this study.

## Input and output variables

The selection of input and output variables in this study was guided by previous empirical studies[51 52] and depended on the availability of data routinely compiled by THs.

Labour and capital were considered important input variables in the delivery of health services. This study used three personnel (ie, labour) variables: total number of medical staff (I1), including the number of physicians and nurses; total number of other technicians (I2), namely, pharmacists, clinical laboratory technicians, medical imaging technicians and other medical technical staff members; and total number of non-medical staff members (I3), namely, administrative, ground skilled staff members and other technical workers. For capital, most studies considered 'the number of beds' as a proxy of capital inputs. The dynamic situation of hospitalisation in China is characterised by idle beds or the use of many temporary beds; thus, "the actual number of open beds" (I4) instead of "the number of authorised beds" was adopted in this study.

For the output indicator, THs in China mainly deliver curative and public health services. The hospital efficiency studies conducted by Ineveld et al[53] and Torabipour et al,[54] as well as several studies that estimated Chinese hospital efficiency,[24 25 55] assisted to characterise the product of curative activities. The outputs were represented by two indicators: "the number of outpatient and emergency visits (O1)", including the number of patients asking for outpatient and emergency diagnostic services, and "the number of inpatient (O2)", which was argued to be a better output than inpatient days. To characterise the outputs of public health services, Deidda et al[56] selected "the average number of notes per patient within the electronic health records (EHR)" as an output. They argued that EHR that covers patient demographics, medical history, reasons for appointment, clinical notes, diagnoses, prescriptions, laboratory test results, test and examination orders should be considered a result of patients' visits. The EHR management is also an important activity in Chinese public health services, including the establishment of EHR to its regular management. With the availability of data in NHSIRS and following several

studies on estimating the efficiency of THs in China,[57–59] "the number of family EHRs under management (O3)" and "the number of chronic diseases (including high blood pressure and diabetes) patients under management (O4)" were selected as outputs. The dated hospital information system has resulting in missing data and statistics on other public health services, including the number of vaccinations and elderly under health management. Thus, these data were excluded in this study. The case mix was not considered in this study because the hospital information systems in the sample THs are outdated and lack timely updates. Moreover, no complete data from 2008 to 2014 were available to accurately measure the case-mix index.

## RESULTS
### Descriptive statistics of the studied THs
Table 1 presents a descriptive overview of the set of input–output variables. The average I1 and I4 increased from 29 and 29 in 2008 to 42 and 50 in 2014, respectively. In contrast, the average I2 and I3 remained nearly unchanged. Furthermore, the averages of O1 and O2 experienced a remarkable increase from 2008 to 2012

**Table 1** Summary statistics of input–output variables (2008–2014)

| | Inputs | | | | Outputs | | | |
|---|---|---|---|---|---|---|---|---|
| | I1 | I2 | I3 | I4 | O1 | O2 | O3 | O4 |
| **2008** | | | | | | | | |
| Mean | 29 | 18 | 9 | 29 | 31 430 | 956 | 14 328 | 774 |
| Max | 78 | 47 | 51 | 96 | 111 920 | 4300 | 58 316 | 4689 |
| Min | 6 | 2 | 2 | 10 | 3514 | 114 | 1000 | 85 |
| **2009** | | | | | | | | |
| Mean | 29 | 18 | 10 | 34 | 35 389 | 1289 | 14 058 | 890 |
| Max | 75 | 55 | 51 | 100 | 102 633 | 3868 | 58 956 | 4504 |
| Min | 5 | 3 | 2 | 10 | 2678 | 255 | 1234 | 78 |
| **2010** | | | | | | | | |
| Mean | 28 | 17 | 10 | 35 | 37 131 | 1283 | 15 475 | 1082 |
| Max | 81 | 49 | 51 | 100 | 117 556 | 4217 | 58 560 | 3588 |
| Min | 5 | 3 | 2 | 10 | 5994 | 127 | 1280 | 80 |
| **2011** | | | | | | | | |
| Mean | 29 | 16 | 9 | 38 | 41 542 | 1358 | 22 972 | 1840 |
| Max | 106 | 60 | 42 | 106 | 126 712 | 4958 | 50 831 | 5347 |
| Min | 5 | 3 | 2 | 10 | 6111 | 120 | 3750 | 115 |
| **2012** | | | | | | | | |
| Mean | 31 | 17 | 10 | 42 | 51 070 | 1734 | 25 827 | 2285 |
| Max | 106 | 54 | 32 | 155 | 151 291 | 7968 | 65 324 | 6466 |
| Min | 5 | 3 | 2 | 10 | 5022 | 118 | 1345 | 118 |
| **2013** | | | | | | | | |
| Mean | 32 | 18 | 10 | 45 | 49 715 | 1612 | 28 073 | 2682 |
| Max | 79 | 70 | 31 | 155 | 156 927 | 8027 | 67 568 | 7277 |
| Min | 5 | 3 | 2 | 10 | 2600 | 116 | 5433 | 120 |
| **2014** | | | | | | | | |
| Mean | 42 | 17 | 10 | 50 | 51 860 | 1642 | 31 193 | 2864 |
| Max | 187 | 46 | 67 | 155 | 144 497 | 9899 | 68 665 | 7501 |
| Min | 5 | 2 | 2 | 10 | 12 739 | 112 | 4985 | 121 |

I1, total number of medical staff; I2, total number of other technicians; I3, total number of non-medical staff members; I4, actual number of open beds; O1, the number of outpatient and emergency visits; O2, the number of inpatient; O3, the number of family electronic health records under management; O4, the number of chronic diseases patients under management.

**Table 2** Bootstrapping estimated efficiency (annual average and 95% CIs)

| Year | TE | | | | | PTE | | | | | SE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Estimated eff. | Bias corrected | Bias | Lower bound | Upper bound | Estimated eff. | Bias corrected | Bias | Lower bound | Upper bound | Bias corrected |
| 2008 | 0.5911 | 0.4495 | 0.1416 | 0.4320 | 0.5799 | 0.6905 | 0.5659 | 0.1246 | 0.5430 | 0.6835 | 0.7036 |
| 2009 | 0.6397 | 0.4914 | 0.1483 | 0.4833 | 0.6278 | 0.7238 | 0.6048 | 0.1190 | 0.5779 | 0.7166 | 0.7266 |
| 2010 | 0.6261 | 0.4870 | 0.1391 | 0.4833 | 0.6144 | 0.7030 | 0.5941 | 0.1089 | 0.5682 | 0.6960 | 0.7548 |
| 2011 | 0.6706 | 0.5092 | 0.1614 | 0.4886 | 0.6583 | 0.7689 | 0.6322 | 0.1367 | 0.6088 | 0.7612 | 0.6906 |
| 2012 | 0.7671 | 0.5871 | 0.1800 | 0.5774 | 0.7527 | 0.8413 | 0.6915 | 0.1498 | 0.6668 | 0.8328 | 0.7149 |
| 2013 | 0.7176 | 0.5438 | 0.1738 | 0.5362 | 0.7036 | 0.8130 | 0.6876 | 0.1255 | 0.6534 | 0.8051 | 0.6970 |
| 2014 | 0.7137 | 0.5348 | 0.1789 | 0.5167 | 0.7005 | 0.8272 | 0.6851 | 0.1421 | 0.6573 | 0.8186 | 0.6687 |
| Mean | 0.6751 | 0.5147 | 0.1604 | 0.5025 | 0.6625 | 0.7668 | 0.6373 | 0.1295 | 0.6108 | 0.7591 | 0.7080 |

TE scores were calculated by authors using bootstrapping DEA$_{CRS}$ (CCR model); PTE and SE scores were calculated by authors using bootstrapping DEA$_{VRS}$ (BCC model).
PTE, pure technical efficiency; TE, technical efficiency.

and plateaued or declined after 2012. In contrast, the averages of O3 and O4 constantly increased during the study period.

### Bootstrapping bias estimated efficiency scores

Table 2 summarises the annual mean TE, PTE and SE of the sample THs from 2008 to 2014. The average bias-corrected TE, PTE and SE for the 7-year period were 0.5147, 0.6373 and 0.7080, respectively. This result implied that if THs were running efficiently, then these hospitals should have increased their outputs by 36.27% with the same volume of inputs. The TE and PTE improved during 2008 and 2012 but declined after 2012. Moreover, the mean of the estimated efficiency is located at the left part of the bootstrapping estimated CIs in all cases, thereby indicating that efficiency should be improved.

Table 3 presents the TE, PTE and SE according to two criteria: size and type of THs. TE of THs does not significantly differ in the group according to their sizes, as shown by the mean and the Kruskal–Wallis test (p=0.369 >0.05). In contrast, PTE and SE are significantly different (p=0.001 <0.05). TE, PTE and SE do not differ significantly between the central and general THs, as

shown by the mean and the Mann–Whitney test (p>0.05).

### Bootstrapping MPI and its components

MPI was measured to analyse the productivity changes from 2008 to 2014. Table 4 presents the bootstrapping MPI (TFPC) and the 95% bootstrapping confidence intervals summary of the annual geometric means for TFPC and its components. When the confidence interval includes the value of one, we do not reject the null hypothesis of the no productivity change. The same applies to the TFPC decompositions. Over the 7-year period, the sample THs experienced a negative shift in the production frontier, and the negative TFPC change was large at 2.14% (although statistically insignificant). For its components, TEC increased by 28.58% and TC decreased by 23.89%. Therefore, the TC decrease was the main contributor for the TFPC deterioration. From 2008 to 2014, 17 (35.4%) hospitals had TFPC scores of above 1, thereby indicating growth in productivity. In contrast, 31 (64.6%) hospitals had TFPC scores of below 1, thereby indicating deterioration in productivity. From 2008 to 2012, the sample THs experienced positive productivity changes, thereby indicating that these hospitals were able to provide additional care with the same

**Table 3** Summary of the efficiency scores according to size and type of THs (2008–2014)

| | Observations | TE | | PTE | | SE | |
|---|---|---|---|---|---|---|---|
| | | Mean | p Value | Mean | p Value | Mean | p Value |
| Efficiency and size of THs (Kruskal–Wallis test) | | | | | | | |
| Small | 182 | 0.5077 | 0.369 | 0.6129 | 0.001** | 0.7184 | 0.001** |
| Medium | 95 | 0.5339 | | 0.6465 | | 0.7362 | |
| Large | 59 | 0.5051 | | 0.6977 | | 0.6308 | |
| Efficiency and type of THs (Mann–Whitney test) | | | | | | | |
| Central | 102 | 0.5040 | 0.391 | 0.6424 | 0.518 | 0.7232 | 0.531 |
| General | 234 | 0.5193 | | 0.6351 | | 0.7014 | |

The size of THs is defined by the number of beds. Fewer than 31 beds, THs are considered as small. They are classified as medium from 31 to 59 beds and as large when the number of beds is equal or higher than 60. Single asterisk (*) denotes significant differences at 90%, double asterisk (**) denotes significant differences at 95% and triple asterisk (***) denotes significant differences at 99%.
PTE, pure technical efficiency; TE, technical efficiency; THs, township hospitals.

**Table 4** Average change in THs productivity

| Item | 2008–2009 | | | 2009–2010 | | | 2010–2011 | | | 2011–2012 | | | 2012–2013 | | | 2013–2014 | | | 2008–2014 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Observed | Lower | Upper | Observed | Lower | Upper | Observed | Lower | Upper | Observed | Lower | Upper | Observed | Lower | Upper | Observed | Lower | Upper | Observed | Lower | Upper |
| TFPC | 1.0642 | 0.8988 | 1.1814 | 1.0174 | 0.8865 | 1.1130 | 1.0477 | 0.8581 | 1.1679 | 1.0001 | 0.8492 | 1.0901 | **0.8472***** | 0.7385 | 0.9293 | **0.9139**** | 0.8012 | 0.9952 | 0.9786 | 0.8369 | 1.0756 |
| TEC | **1.2894***** | 1.0557 | 1.6733 | 1.1912 | 0.9754 | 1.5527 | **1.4207***** | 1.1737 | 1.8358 | **1.2768***** | 1.0169 | 1.6457 | 1.2369 | 0.9787 | 1.5361 | **1.3116***** | 1.0228 | 1.6468 | **1.2858***** | 1.0352 | 1.6456 |
| TC | **0.8254***** | 0.4608 | 0.9101 | **0.8541***** | 0.4980 | 0.9308 | **0.7375***** | 0.4539 | 0.7955 | **0.7833***** | 0.4613 | 0.8530 | **0.6850***** | 0.4731 | 0.7685 | **0.6968***** | 0.4467 | 0.7843 | **0.7611***** | 0.4653 | 0.8381 |
| PTEC | 1.2305 | 0.8165 | 1.6211 | 1.0872 | 0.7550 | 1.4504 | 1.2124 | 0.8701 | 1.5631 | 1.0686 | 0.6946 | 1.3896 | 1.0644 | 0.8305 | 1.3207 | 1.0923 | 0.8350 | 1.3321 | 1.1239 | 0.7981 | 1.4419 |
| SEC | 1.0479 | 0.8100 | 1.2785 | 1.0956 | 0.7643 | 1.3178 | 1.1718 | 0.7754 | 1.4090 | 1.1948 | 0.8725 | 1.4142 | 1.1621 | 0.8650 | 1.3382 | 1.2007 | 0.9637 | 1.4201 | 1.1441 | 0.8391 | 1.3619 |
| Frequency distribution | >1 | =1 | <1 | >1 | =1 | <1 | >1 | =1 | <1 | >1 | =1 | <1 | >1 | =1 | <1 | >1 | =1 | <1 | >1 | =1 | <1 |
| TFPC observed | 26 (54.2%) | 0 (0.0%) | 22 (45.8%) | 24 (50.0%) | 0 (0.0%) | 24 (50.0%) | 27 (56.4%) | 0 (0.0%) | 21 (43.8%) | 24 (50.0%) | 0 (0.0%) | 24 (50.0%) | 11 (22.9%) | 0 (0.0%) | 37 (77.1%) | 18 (37.5%) | 0 (0.0%) | 30 (62.5%) | 17 (35.4%) | 0 (0.0%) | 31 (64.6%) |
| TEC observed | 34 (70.8%) | 4 (8.4%) | 10 (20.8%) | 28 (58.3%) | 7 (14.6%) | 13 (27.1%) | 36 (75.0%) | 2 (4.2%) | 10 (20.8%) | 33 (68.8%) | 6 (12.4%) | 9 (18.8%) | 27 (56.2%) | 7 (14.6%) | 14 (29.2%) | 33 (68.8%) | 3 (6.2%) | 12 (25.0%) | 42 (87.5%) | 0 (0.0%) | 6 (12.5%) |
| TC observed | 6 (12.5%) | 0 (0.0%) | 42 (87.5%) | 9 (18.8%) | 0 (0.0%) | 39 (81.2%) | 2 (4.2%) | 0 (0.0%) | 46 (95.8%) | 7 (14.6%) | 0 (0.0%) | 41 (85.4%) | 0 (0.0%) | 0 (0.0%) | 48 (100.0%) | 2 (4.2%) | 1 (2.1%) | 45 (93.7%) | 0 (0.0%) | 0 (0.0%) | 48 (100.0%) |
| PTEC observed | 29 (60.4%) | 7 (14.6%) | 12 (25.0%) | 19 (39.6%) | 17 (35.4%) | 12 (25.0%) | 31 (64.6%) | 6 (12.5%) | 11 (22.9%) | 21 (43.8%) | 18 (37.4%) | 9 (18.8%) | 25 (52.1%) | 9 (18.8%) | 14 (29.1%) | 27 (56.2%) | 9 (18.8%) | 12 (25.0%) | 35 (72.9%) | 0 (0.0%) | 13 (27.1%) |
| SEC observed | 27 (56.3%) | 5 (10.4%) | 16 (33.3%) | 29 (60.4%) | 7 (14.6%) | 12 (25.0%) | 34 (70.8%) | 2 (4.2%) | 12 (25.0%) | 35 (72.9%) | 6 (12.5%) | 7 (14.6%) | 28 (58.3%) | 8 (16.7%) | 12 (25.0%) | 34 (70.8%) | 3 (6.3%) | 11 (22.9%) | 37 (77.1%) | 0 (0.0%) | 11 (22.9%) |

Geometric mean and bootstrapping results (95% CIs).

Single asterisk (*) denotes significant differences from unity at 90%, double asterisk (**) denotes significant differences from unity at 95% and triple asterisk (***) denotes significant differences from unity at 99%. Significant results in bold.

$$TFPC = TEC \times TC = (PTEC \times SEC) \times TC$$

A score >1 indicates growth, a score=1 signifies stagnation and a score <1 indicates decline or deterioration.

PTEC, pure technical efficiency changes; SEC, scale efficiency changes; TC, technological changes; TEC, technical efficiency changes; TFPC, total factor productivity changes.

level of resources during this period. From 2012 to 2014, the sample THs went through a negative productivity shift, thereby indicating that these hospitals provided limited care with the same level of inputs.

The decomposition of TFPC indicated that TEC improved by 28.58% from 2008 to 2014, that is, THs managed to catch up nearly the entire potential efficiency improvement; the result was statistically significant. The decomposition of TEC estimated that PTEC and SEC improved by 12.39% and 14.41%, respectively. From 2008 to 2014, all THs experienced negative TC (ie, the geometric mean for TC is 0.7611), thereby indicating a large technical decrease of 23.89%. However, this result was statistically insignificant. The combined results of these changes produced a net negative impact on TFPC.

## Return to scale

Table 5 presents the nature of the scale inefficiency. For each of the 7 years, we reported the number of THs that were operating at the most productive scale size (MPSS), IRS or DRS. A few of the sample THs (eg, 10% in 2008 and 20% in 2014) were operating at MPSS, thereby indicating that they operated at their most productive size. The majority of the sample THs (ie, 54% in 2008 and over 68% in 2014) were operating at DRS. Hence, these THs were scale inefficient. The binomial test rejected

| Table 5 | Return to scale | | | |
|---|---|---|---|---|
| Year | RTS | N | Percentage | Sig. (two tailed)† |
| 2008 | MPSS | 5 | 10.42 | 0.222‡ |
| | IRS | 17 | 35.42 | |
| | DRS | 26 | 54.17 | |
| 2009 | MPSS | 6 | 12.50 | 0.164 |
| | IRS | 16 | 33.33 | |
| | DRS | 26 | 54.17 | |
| 2010 | MPSS | 5 | 10.42 | 0.014** |
| | IRS | 13 | 27.08 | |
| | DRS | 30 | 62.50 | |
| 2011 | MPSS | 6 | 12.50 | 0.003** |
| | IRS | 11 | 22.92 | |
| | DRS | 31 | 64.58 | |
| 2012 | MPSS | 13 | 27.08 | 0.002** |
| | IRS | 8 | 16.67 | |
| | DRS | 27 | 56.25 | |
| 2013 | MPSS | 9 | 18.75 | 0.000*** |
| | IRS | 8 | 16.67 | |
| | DRS | 31 | 64.58 | |
| 2014 | MPSS | 10 | 20.83 | 0.000*** |
| | IRS | 5 | 10.42 | |
| | DRS | 33 | 68.75 | |

Double asterisk (**) denotes significant differences from unity at 95%, and triple asterisk (***) denotes significant differences from unity at 99%.
†Binomial test, null hypothesis: no difference in the proportion of hospitals under IRS or DRS.
‡Based on Z approximation.
DRS, decreasing returns to scale; IRS, increasing returns to scale; MPSS, most productive scale size; RTS, return to scale.

the null hypothesis that the percentage of the sample THs under IRS or DRS is the same for all years, except in 2008 and 2009.

## DISCUSSION

In this study, the results indicate that the average TE and PTE of the sample THs were relatively low (ie, 0.5147 and 0.6373, respectively), thereby indicating a huge potential for efficiency improvement. In contrast, the average SE was relatively high (ie, 0.7080). Thus, the pure technical inefficiency may mainly account for the overall technical inefficiency. The TE of the sample THs in Xiaogan was lower than those reported in other countries or other regions of China. For example, Hollingsworth[21] measured the efficiencies of different types of hospitals and reported a mean efficiency score of 0.835. The average TE of primary care facilities in the Nouna health district is 0.862,[60] 0.833 in the Spanish region of Extremadura[61] and 0.620 in the Weifang Prefecture of China.[14] The results of our study also indicated that medium-sized THs have a relatively high SE compared with the small and large THs. The most scale-inefficient THs were in DRS, thereby suggesting that managers should explore the optimum operation scale of THs and rationalise these hospitals' supply.

TFPC of the sample THs declined by 2.14% during the 7-year period. Although TEC has improved with an average of 28.58%, the average 23.89% decrease of TC leads to the decrease of TFPC. Therefore, improving TC to boost TFPC is essential. The results we obtained in the Xiaogan Prefecture are not distinct. Torabipour et al[54] studied the productivity changes of hospitals in Ahvaz county from 2007 to 2010 and found that TC played a major role in productivity changes. Li explored the changes in efficiency and productivity of 45 THs in Anhui Province and determined that the efficiency of these hospitals exhibited a downward trend from 2009 to 2011; TEC and TC also decreased by 3.8% and 14.1%, respectively.[62]

In the current study, TE, PTE and productivity experienced positive changes from 2008 to 2012. The possible explanations for such changes may be interpreted as follows. First, a primary objective of the new round of healthcare reform in 2009 is to strengthen the capability of primary healthcare facilities. From 2008 to 2012, the government of Xiaogan implemented a standardised construction programme in THs, in which the healthcare providers' abilities, medical techniques and other relevant equipment have been improved and standardised.[5] In addition, programmes that provided physicians from THs with on-the-job training in county and tertiary hospitals were launched. These measures have improved the service capability of THs, thereby increasing their TE, PTE and productivity.[11] Second, the expansion of the public health service increased the outputs of THs. From 2009 to 2012, the Chinese government funded THs to deliver a specific package for basic public health services (Xiaogan is no exception), which

includes health prevention; vaccinations; management of patients' chronic diseases; health examinations for pregnant women, children and the elderly; and the compilation of health records. These services increased the income of healthcare providers in THs, thereby encouraging them to further provide the necessary public health services.[63] Public health service as an important component of the daily responsibilities of THs has expanded their outputs, thereby improving their overall efficiency and productivity.[11] Third, the implementation of NCMS offered a comprehensive insurance system to the rural population, and over 98% of the rural residents in Xiaogan have been insured by NCMS. NCMS reduced the financial barriers to access healthcare and increased the healthcare demand. Babiarz et al[64] analysed the effect of NCMS and its implications on primary healthcare and determined that NCMS provides a financial risk protection for individuals. Hu et al[9] adopted DEA and regression analysis methods to explore the effects of NCMS on hospital efficiency and found that the NCMS exerts a considerable efficiency-enhancing effect of 11.1%, with numerous rural residents gaining access to low-cost medical service. Cheng et al[25] measured TE and productivity of county hospitals during the healthcare reform, and the results indicated that TE and productivity experienced a substantial increase. In the current study, the number of outpatient and emergency visits increased from 31 430 in 2008 to 51 070 in 2012, and the number of inpatients increased from 956 in 2008 to 1734 in 2012. These results suggested that the increase in demand, which is induced by the series of primary healthcare and NCMS reforms, may have enhanced the efficiency and productivity of the sample THs.

Although positive changes in the efficiency and productivity of THs from 2008 to 2012 have been observed, TE, PTE and SE were relatively low. In addition, the sample THs experienced a downward tendency in TE, PTE and productivity after 2012. Several possible reasons may justify the inefficiency and productivity deterioration. First, the "separating revenue from expenditure" policy stipulated that the total revenue of THs should be submitted to the financial sector, while the salaries of healthcare workers are restricted by their performance, thereby cutting off the interest chains among THs, medical personnel and pharmaceutical companies.[65] The decrease in personnel income aggravated the loss of quality medical professionals in THs, thereby possibly causing the decrease in PTE and TC.[66] Second, the implementation of the NEMP policy weakened the motivation of THs to provide healthcare. After the implementation of NEMP, only essential medicines (zero profit for THs) included in the EML could be prescribed.[67] Revenues from medicines, which was the main source of THs' income and bonus of medical personnel, have been eliminated. Li et al[68] and Fang et al[69] assessed the effects of NEMP on primary healthcare and determined that NEMP affected the doctor's diagnosis and treatment because of the limited medicines available, thereby resulting in the decrease in the number of patients. Hu[70] and Yang et al[71] found that medicines in EML cannot meet the requirements of clinical practice in THs, which caused patients to transfer directly to secondary or tertiary hospitals to seek appropriate treatment. Hu[70] explained that the flaws in NEMP and EML will skew demand, affect clinical performance, reduce income and contribute to disincentives for health workers to remain in primary healthcare facilities. In addition, Xiaogan lacks a defined regulation or legislation regarding the responsibility and accountability of the government to compensate the financial loss after implementing a zero-profit policy in THs. Limited public finance from local governments jeopardised the capacity of THs to meet the needs of local residents.[72] Considering that the inputs of the sample THs remained nearly unchanged from 2012 to 2014, the stagnation or decline in the outputs of these hospitals may have led to the deterioration of their efficiency and productivity.

The changes in TE and productivity of THs in the Chinese healthcare reform are complex; hence, we need to hold a cautious attitude towards the interpretation of relevant factors that may have led to these changes. Other possible reasons could explain the changes in TE, and productivity of THs should be explored in further studies. The managers of THs and policymakers will benefit from being aware of the efficiency deterioration of these hospitals and should be encouraged to address this issue. Additional research is necessary to support the discussions drawn from this study.

### Limitations and suggestions for further research

This study introduced the bootstrapping DEA method to evaluate the performance of hospital operations in China. However, we believe that this study can still be improved and extended in a number of directions. The first limitation is the limited number of THs available for the research. The sample THs in this study came from Xiaogan in central China, while the eastern and western regions were excluded, because of data limitations. Future research should be conducted and include THs in other areas, with their respective efficiency and productivity analysis for comparison. Second, similar to previous studies, we were unable to obtain the information on the case mix of each hospital or patient outcome quality; thus, we did not determine whether a few variations in the severity of cases treated in each TH are possible. Future studies should consider the in-depth effect of the quality of care if such data become available. Third, the DEA method has its natural limitations, such as the unreasonable weights assigned for the input and output variables. Thus, future studies may estimate efficiency using alternative approaches (eg, SFA) to compare with the results of this research and to evaluate robustness. Finally, a two-stage bootstrapping DEA method can be applied to explore the effects of environmental factors (eg, NCMS reform and NEMP policy) on

the efficiency scores to improve the reliability of the results. Despite these limitations, this study can be considered a beneficial attempt to measure the efficiency and productivity of THs in China.

## CONCLUSIONS

This study provides an empirical representation of the efficiency and productivity changes of THs before and after the healthcare reform implementation. The results indicated a considerable space for TE improvement in THs since the average TE was only 0.5147. Considering the nature of scale inefficiency, most scales of inefficient THs were in DRS, and considerable improvement room for output exists if this case is efficiently managed. The MPI results showed that the sample THs experienced a productivity decrease of 2.14%, which mainly resulted from the substantial deterioration in TC. The results indicated that the sample THs experienced positive TE, PTE and productivity growth from 2008 to 2012, but deteriorated from 2012 to 2014. The pure technical inefficiency and decrease of TC had the highest negative effect on the average TE and productivity changes.

This study introduced the bootstrapping DEA approach, which overcomes several limitations of the DEA approach, to measure the efficiency and productivity of THs in China. Our results are based on a small THs sample; however, this study used all THs in Xiaogan during a 7-year period. Thus, this sample is representative of the status quo in Xiaogan. Although we focused on one prefecture, this Xiaogan case study is representative of the majority of THs in central China and can improve the understanding of the complex efficiency issues in the Chinese healthcare system.

## REFERENCES

1. National Bureau of Statistics, China. *China Statistical Yearbook 2014*. Beijing: China Statistics Press, 2014.
2. Yip W, Hsiao W, Meng QY, *et al*. Realignment of incentives for health-care providers in China. *Lancet* 2010;375:1120–30.
3. Hillier S, Shen J. Health care systems in transition: People's Republic of China. Part I: an overview of China's health care system. *J Public Health Med* 1996;18:258–65.
4. Liu X, Xu L, Wang S. Reforming China's 50,000 township hospitals—effectiveness, challenges and opportunities. *Health Policy* 1996;38:13–29.
5. Wagstaff A, Lindelow M, Wang S, *et al*. *Reforming China's Rural Health System*. Washington (DC): World Bank, 2009.
6. Hsiao WC. The Chinese health care system: lessons for other nations. *Soc Sci Med* 1995;41:1047–55.
7. World Bank. *Financing Health Care: Issues and Options for China*. Washington (DC): World Bank, 1997.
8. Eggleston K, Ling L, Meng QY, *et al*. Health service delivery in China: a literature review. *Health Econ* 2008;17:149–65.
9. Hu HH, Qi QH, Yang CH. Analysis of hospital technical efficiency in China: effect of health insurance reform. *China Econ Rev* 2012;23:865–77.
10. Liu Q, Wang B, Kong YY, *et al*. China's primary health-care reform. *Lancet* 2011;377:2064–6.
11. Yip W, Hsiao W. Harnessing the privatisation of China's fragmented health-care delivery. *Lancet* 2014;384:805–18.
12. State Council of China. *Implementation Plan for the Recent Priorities of the Health Care System Reform (2009–2011)*. Beijing: State Council, 2009.
13. National Health and Family Planning Commission. *China Statistical Yearbook of Health and Family Planning*. Beijing: China Union Medical University Press, 2014.
14. Audibert M, Mathonnat J, Pelissier A, *et al*. Health insurance reform and efficiency of township hospitals in rural China: an analysis from survey data. *China Econ Rev* 2013;27:326–38.
15. Sherman HD. Hospital efficiency measurement and evaluation—empirical test of a new technique. *Med Care* 1984;22:922–38.
16. Ozcan YA, Luke RD, Haksever C. Ownership and organizational performance. A comparison of technical efficiency across hospital types. *Med Care* 1992;30:781–94.
17. Grosskopf S, Valdmanis V. Evaluating hospital performance with case-mix-adjusted outputs. *Med Care* 1993;31:525–32.
18. Burgess JF, Wilson PW. Hospital ownership and technical inefficiency. *Manage Sci* 1996;42:110–23.
19. Steinmann L, Dittrich G, Karmann A, *et al*. Measuring and comparing the (in)efficiency of German and Swiss hospitals. *Eur J Health Econ* 2004;5:216–26.
20. Nayar P, Ozcan YA. Data envelopment analysis comparison of hospital efficiency and quality. *J Med Syst* 2008;32:193–9.
21. Hollingsworth B. The measurement of efficiency and productivity of health care delivery. *Health Econ* 2008;17:1107–28.
22. O'Neill L, Rauner M, Heidenberger K, *et al*. A cross-national comparison and taxonomy of DEA-based hospital efficiency studies. *Socio-Econ Plan Sci* 2008;42:158–89.
23. Ng YC. The productive efficiency of Chinese hospitals. *China Econ Rev* 2011;22:428–39.
24. Li H, Dong SP, Liu TF. Relative efficiency and productivity: a preliminary exploration of public hospitals in Beijing, China. *BMC Health Serv Res* 2014;14:158.
25. Cheng ZH, Tao HB, Cai M, *et al*. Technical efficiency and productivity of Chinese county hospitals: an exploratory study in Henan Province, China. *BMJ Open* 2015;5:e7267.
26. Cheng Z, Tao H, Cai M, *et al*. Using a two-stage data envelopment analysis to estimate the efficiency of county hospitals in China: a panel data study. *Lancet* 2015;386:S64.
27. Farrell MJ. The measurement of productive efficiency. *J R Stat Soc* 1957;120:253–81.
28. Daraio C, Simar L. *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*. New York: Springer, 2007.
29. Charnes A, Cooper WW, Lewin AY, *et al*. *Data Envelopment Analysis: Theory, Methodology and Applications*. New York: Kluwer Academic Publishers, 1997.
30. Charnes A, Cooper WW, Rhodes E. Measuring efficiency of decision making units. *Eur J Oper Res* 1978;3:429–44.
31. Banker RD, Charnes A, Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag Sci* 1984;30:1078–92.
32. Caves DW, Christensen LR, Diewert WE. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 1982;50:1393–414.

33. Malmquist S. Index numbers and indifference surfaces. *Trabajos de Estadística* 1953;4:209–42.
34. Färe R, Grosskopf S, Lindgren B, *et al*. Productivity developments in Swedish hospitals: a Malmquist output index approach. In: Charnes A, Cooper W, Lewin A, *et al*, eds. *Data Envelopment Analysis: Theory, Methodology and Applications*. Boston (MA): Kluwer, 1994:253–72.
35. Cooper WW, Seiford L, Tone K. *Data Envelopment Analysis a Comprehensive Text with Models, Applications, References and DEA-Solver Software*. New York: Springer, 2007.
36. Pelone F, Kringos DS, Romaniello A, *et al*. Primary care efficiency measurement using data envelopment analysis: a systematic review. *J Med Syst* 2015;39:156.
37. Zhang X, Xiong Y, Ye J, *et al*. Analysis of government investment in primary healthcare institutions to promote equity during the three-year health reform program in China. *BMC Health Serv Res* 2013;13:114.
38. Ozcan YA. *Health Care Benchmarking and Performance Evaluation*. 2nd edn. New York: Springer, 2014.
39. Barber SL, Borowitz M, Bekedam H, *et al*. The hospital of the future in China: China's reform of public hospitals and trends from industrialized countries. *Health Policy Plann* 2014;29:367–78.
40. Kounetas K, Papathanassopoulos F. How efficient are Greek hospitals? A case study using a double bootstrap DEA approach. *Euro J Health Econ* 2013;14:979–94.
41. Löthgren M, Tambour M. Bootstrapping the data envelopment analysis Malmquist productivity index. *Appl Econ* 1999;31:417–25.
42. Simar L, Wilson PW. Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. *Manage Sci* 1998;44:49–61.
43. Simar L, Wilson PW. A general methodology for bootstrapping in non-parametric frontier models. *J Appl Stat* 2000;27:779–802.
44. Simar L, Wilson PW. Estimation and inference in two-stage, semi-parametric models of production processes. *J Econometrics* 2007;136:31–64.
45. Atkinson SE, Wilson PW. Comparing mean efficiency and productivity scores from small samples: a bootstrap methodology. *J Prod Anal* 1995;6:137–52.
46. Efron B. Bootstrap methods: another look at the jackknife. *Ann Stat* 1979;7:1–26.
47. Simar L, Wilson PW. Estimating and bootstrapping Malmquist indices. *Eur J Oper Res* 1999;115:459–71.
48. Ramírez-Valdivia MT, Maturana S, Salvo-Garrido S. A multiple stage approach for performance improvement of primary healthcare practice. *J Med Syst* 2011;35:1015–28.
49. Dong SP, Zuo YL, Tao HB. Study on DEA-based Chinese hospital efficiency and applied indicators. *Chinese Health Policy Res* 2014;7:40–5.
50. Wilson PW. FEAR: a software package for frontier efficiency analysis with R. *Socio-Econ Plan Sci* 2008;42:247–54.
51. Gerdtham UG, Lothgren M, Tambour M, *et al*. Internal markets and health care efficiency: a multiple-output stochastic frontier analysis. *Health Econ* 1999;8:151–64.
52. Varabyova Y, Schreyogg J. International comparisons of the technical efficiency of the hospital sector: panel data analysis of OECD countries using parametric and non-parametric approaches. *Health Policy* 2013;112:70–9.
53. Ineveld MV, Oostrum JV, Vermeulen R, *et al*. Productivity and quality of Dutch hospitals during system reform. *Health Care Manag Sci* 2016;19:279–90.
54. Torabipour A, Najarzadeh M, Arab M, *et al*. Hospitals productivity measurement using data envelopment analysis technique. *Iran J Public Health* 2014;43:1576–81.
55. Li H, Dong S. Measuring and benchmarking technical efficiency of public hospitals in Tianjin, China: a bootstrap-data envelopment analysis approach. *Inquiry* 2015;52:pii: 0046958015605487.
56. Deidda M, Lupiáñez-Villanueva F, Codagnone C, *et al*. Using data envelopment analysis to analyse the efficiency of primary care units. *J Med Syst* 2014;38:122.
57. Wang CX, Yue JL, Zhou F. Dynamic change on the productivity of township health centres after China's health care system reform: based on the nonparametric Malmquist index in Guangzhou, Guangdong Province, 2009–2013. *Chin J Health Policy* 2014;7:49–54.
58. Sun Q, Wu WB, Zheng XH, *et al*. Using data envelopment analysis method to explore the efficiency of township hospitals' efficiency in four provinces. *Mod Hosp Manag* 2012;10:11–9.
59. Ming YF, Huang R, Zhang YC. Operation efficiency of rural health centres in Hami Prefecture after health care reform at basic level. *Chin Gen Pract* 2015;18:759–63.
60. Marschall P, Flessa S. Efficiency of primary care in rural Burkina Faso. A two-stage DEA analysis. *Health Econ Rev* 2011;1:5.
61. Cordero Ferrera JM, Cebada EC, Murillo Zamorano LR. The effect of quality and socio-demographic variables on efficiency measures in primary health care. *Eur J Health Econ* 2014;15:289–302.
62. Li C. *Evaluating Efficiency of Township Hospital in Anhui Province Based on Data Envelopment Analysis*. Jinan: Shandong University, 2013.
63. Zheng J, Li J, Jiang X, *et al*. Sustaining health workforce recruitment and retention in township hospitals: a survey on 110 directors of township hospitals. *Front Med* 2015;9:239–50.
64. Babiarz KS, Miller G, Yi H, *et al*. New evidence on the impact of China's New Rural Cooperative Medical Scheme and its implications for rural primary healthcare: multivariate difference-in-difference analysis. *BMJ* 2010;341:c5617.
65. Fang PQ, Xu QH. The main problems and strategies regarding "separating revenue from expenditures" policy in township hospitals. *Chin Rural Health Serv Admin* 2007;27:794–7.
66. Ma PY. The study on performance evaluation of separating revenue from expenditures in rural health clinics of Beijing. *Chin Health Serv Manag* 2013;10:765–8.
67. Yang LP, Liu CJ, Ferrier JA, *et al*. The impact of the National Essential Medicines Policy on prescribing behaviours in primary care facilities in Hubei Province of China. *Health Policy Plann* 2013;28:750–60.
68. Li ZG, Shu DF, Xia M, *et al*. The assessment on impact of essential drugs policy on primary health care system in rural areas of Shandong Province policy and regulation division of the Health Department of Shandong Province. *Technol Health Care* 2015;23: S169–76.
69. Fang Y, Wagner AK, Yang S, *et al*. Access to affordable medicines after health reform: evidence from two cross-sectional surveys in Shaanxi Province, western China. *Lancet Global Health* 2013;1: E227–37.
70. Hu S. Essential medicine policy in China: pros and cons. *J Med Econ* 2013;16:289–94.
71. Yang L, Cui Y, Guo S, *et al*. Evaluation, in three provinces, of the introduction and impact of China's National Essential Medicines Scheme. *Bull World Health Organ* 2013;91:184–94.
72. Yang L, Liu C, Ferrier JA, *et al*. Organizational barriers associated with the implementation of National Essential Medicines Policy: a cross-sectional study of township hospitals in China. *Soc Sci Med* 2015;145:201–8.