3rd International Conference on Public Policy (ICPP3)
June 28-30, 2017 – Singapore

Panel T07P06 Session 1
Evidence-based Policy Making: Evaluation and Methodological Challenges

Title of the paper

A Systematic Review of Empirical Studies of Essential Medicines Policy in China: Implications for Evidence-Based Policy Making in Developing Countries

Author(s)

Lili, Li, Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore, li.lili89@u.nus.edu
Jiwei, Qian, East Asian Institute, National University of Singapore, Singapore, jiwei.qian@nus.edu.sg
Xun, Wu, Division of Social Science & Division of Environment, Hong Kong University of Science and Technology, China, wuxun@ust.hk

Date of presentation

29th June, 2017

A Systematic Review of Empirical Studies of Essential Medicines Policy in China: Implications for Evidence-Based Policy Making (EBPM) in Developing Countries

Lili Li[1], Jiwei Qian[2] and Xun Wu[3, 1 *]

[1] Lee Kuan Yew School of Public Policy, National University of Singapore, 469E Upper Bukit Timah Road, 259774, Singapore
[2] East Asian Institute, National University of Singapore, Tower Block, 469A Bukit Timah Road, 259770, Singapore
[3] Division of Social Science & Division of Environment, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

Email: Lili Li, a0123860@u.nus.edu, Jiwei Qian, jiwei.qian@nus.edu.sg, Xun Wu, wuxun@ust.hk

*Correspondence to: Xun Wu wuxun@ust.hk, Tel (852) 2358 7836

Key words: evidence-based policy making, meta-analysis, research methods, essential medicine policy, China

Abstract

Although evidence-based policy making (EBPM) in health care has gained currency in developing countries, their analytical capacity for producing sound evidence for policy making cannot be taken for granted. The systematic review presented here covers 79 empirical studies published in Chinese academic journals 2010–2015 concerning the country's essential medicines policy, a measure introduced beginning in 2009 to control cost escalation in health expenditures, and 65 empirical studies were selected as eligible for meta-analysis. Results show that most of these studies relied on simple before-and-after comparisons to assess the effects of policy measures, and that few used control groups to deal with validity concerns. The analytical methods used in studies related to the estimated effect sizes concerning China's essential medicines policy. Our conclusions point to the critical importance of adequate analytical capacity in producing reliable evidence for policy making.

Introduction

Evidence-based Policy Making (EBPM) has made some major strides in recent years toward improving effectiveness of health policy interventions in developing countries (WHO 2012). In Mexico, for example, the use of best available evidence has contributed to development of a new health insurance scheme (Seguro Popular) providing universal access to health care (Frenk 2006). In China, empirical research on outcomes of market-based reforms has led to drastic policy changes since the mid-2000s (Liang et al. 2014; Wang and Jin 2011).

The greater progress made in applying EBPM in the health policy arena has been attributed to the widespread adoption and practices of empirically based medicine (EBM) in last few decades. Black (2001) argues that adoption of EBM led some clinicians to begin challenging managers and policy makers to apply the same analytical principles in their decision making. Jiang, Zhang, and Shen (2013) similarly have observed that EBM inspired a movement toward evidence-based public health policy in China.

However, EBPM differs from EBM in ways that could have critical implications for its progress in the future. First, evidence for EBM is often drawn from randomized control trials (RCTs), whereas evidence derived from quasi-experimental designs and observational studies plays a more dominant role in EBPM (Victora, Habicht, and Bryce 2004). In addition, EBPM may involve many more stakeholders in the policy process, and their values and interests must be reflected in ultimate results (Marston and Watts 2003). The production and interpretation of evidence in EBPM are thus subject to both technical and nontechnical considerations. Moreover, evidence highly relevant to EBPM tends to be context-specific (Behague et al. 2009).

The analytical capacity of EBPM to produce rigorous evidence will play a key role in determining its future in the health sector in developing countries. Today many developing countries are undergoing significant transformations of their health systems, and need reliable evidence for policy decisions (Gilson and Raphaely 2008; Hyder et al. 2011). Meanwhile, there is a global shortage of policy experts and researchers adequately trained in research methodologies for rigorous policy research. As a result, the quality of research presently producing evidence for EBPM cannot be taken for granted.

The analysis reported here presents a systematic review of empirical studies regarding the impacts of China's essential medicines policy, with the objective of highlighting potential obstacles to their usefulness to EBPM. Beginning in the 1990s, the Chinese government gradually retreated from the health sector, reducing fiscal subsidies for regional and local hospitals from more than 60% in the 1980s to about 10% by 2000s (Duckett 2010; Qian and Blomqvist 2014). A further, unintended result of these reforms was that doctors in hospitals and clinics found strong incentives to overprescribe medicines (Currie, Lin, and Zhang 2011; Currie, Lin, and Meng 2014; Ramesh, Wu and He 2014; Yip and Hsiao 2014). By the early 2000s, medical prescriptions accounted for about 40% of revenues in public hospitals (Ministry of Health, various years). Since 2009, the National Essential Medicine System (NEMS) has been implemented to reduce this rise in prescription drug expenses. Essential medicines are defined as cost-effective drugs that should be accessible to all patients and

serve basic medical needs (Ministry of Health, various years).

The first essential medicine list, including 307 drugs for primary care providers, was released in 2009. In 2012, the list was expanded to 520 (Qian and Blomqvist 2014). Each province has the discretion to add or remove drugs from this list according to local conditions. On average, 236 medicines were added to provincial level essential medicine lists in the 2012 update (Barber et al. 2013). All essential medicines are covered by social health insurance plans such as the Basic Medical Insurance plan for urban employees and the New Rural Cooperative Medical Scheme for rural residents. The reimbursement rate for essential medicines is also set substantially higher than the reimbursement rate for other drugs.

A great number of empirical studies have been conducted to evaluate the effects of China's NEMS since it was introduced in 2009, and their findings exhibit substantial disparities. Some report that medicine costs as a share of total expenses per treatment episode among urban and rural primary care providers have been reduced, others indicate that the share of drug revenue in total health expenditures has also decreased. But other researchers have found that despite apparent savings, total drug revenues per health facility increased.

Several characteristics of empirical studies investigating the effects of China's essential drug policy provide a good opportunity to conduct a meta-analysis of the quality of such research for producing evidence suitable for EBPM. The first advantage is the sheer volume of research papers on the topic, which permits the use of advanced analytical techniques such as meta-analysis for a systematic review of the literature. The second advantage is the diverse range in research methodologies employed across this great volume of studies, which enrich the variety of prospects for meta-analysis. The third advantage is the complexity involved in the implementation of NEMS in practice, which impose challenges to attempts at assessing the program's impacts.

With collected quasi-experimental studies evaluating China's NEMS since its implementation in 2009, our systematic review and analysis explore how research methodologies employed in

different empirical studies can affect the observed effects concerning the use of medicines in primary-care providers. There are limited systematic reviews of empirical studies on the effects of health policy interventions in developing countries. Among few studies using systematic reviews, there is no prior work focusing on the research methodologies used in producing evidence. Most of the studies in our research sample did not use control groups and simply compared differences in measures before and after policy interventions, which were subject to serious internal validity concerns. The flawed evidence from such empirical methodologies not only may mislead policymakers and practitioners, but also may weaken prospects for further progress in research and applications for EBPM in the heath sector.

Methodology

Data

Our data sources were articles on the impact of NEMS published in leading medical journals in China from 2010 through 2015. We searched Google Scholar, NUS Library databases and the China Knowledge Resource Integrated Database (CNKI) with a combination of key words such as "NEMS," "impact assessment," and "empirical studies" both in English and Chinese language, and then narrowed our search to include only articles focused on impact assessment specifically of NEMS and adopting quantitative analysis. The search work by 2016 obtained the raw sample of 86 published articles with 7 English articles and 79 Chinese articles. As the English sample was not sufficient for a sub-analysis, the resulting sample of 79 Chinese articles were selected for the systematic review. We assigned each study a unique ID number so that it could keep through subsequent analysis.

Data extraction was done independently by two authors following pre-agreed indicators and confirmed by the third author. Data extracted cover the use of research methodologies, the observed policy effects and other general information including authorship, number of references, number of pages, journal information, research funding, publication time, spatial scope, and

location of the study areas

Effect Size

We focused our analysis on four key indicators: outpatient drug cost per prescription (ODCP), average number of drug types per prescription (ANDTP), proportion of prescriptions involving antibiotics (PPA), and proportion of prescriptions involving injections (PPI).

Measures of effect sizes can take the form of mean differences or regression coefficients based on the information reported in individual studies (Lipsey and Wilson 2001a, 2001b; Borenstein and Cooper 2009). To make effect sizes comparable in scale, we computed a standardized mean difference for each individual study. For studies with a control group, the mean difference is $d_i = \overline{\mathrm{X}}_T - \overline{\mathrm{X}}_C$, , where $\overline{\mathrm{X}}_T$ and $\overline{\mathrm{X}}_C$ are respectively the sample means in the treatment group and the control group; for studies with no control group, the mean difference is $d_i = \overline{\mathrm{X}}_{post} - \overline{\mathrm{X}}_{pre}$, where $\overline{\mathrm{X}}_{pre}$and $\overline{\mathrm{X}}_{post}$are respectively the sample means of observations before and after the policy intervention. We computed effect sizes from regression coefficients (Borenstein 2009) for studies based on regression analysis. We also computed overall effect size and estimated the consistency of effect sizes (i.e. I^2 statistic) across studies according to methodologies reported in Lipsey and Wilson (2001b) and Borenstein (2009).

Meta-regression

Meta-analysis has been increasingly applied in sociology, medicine, and psychology to synthesize existing studies on a given topic (Ringquist 2005). A meta-regression analysis defines the effect sizes as the dependent variable and investigates the impacts of study characteristics on effect sizes (Thompson and Higgins 2002).

A relevant study was deemed suitable for meta-analysis if it (1) examined changes in ODCP, ANDTP, PPA, and PPI in primary care providers before and after the policy intervention and (2) contained sufficient statistical details that its results could be included in meta-analysis. Of our overall sample of 79 studies, 65 were found suitable for meta-analysis.

The independent variables analyzed were, in particular, use of DID (difference in difference) method, use of a simple before-and-after comparison method, use of primary data, use of random sampling, authorship affiliated with government/health care providers, the availability of research funds, and whether an article was published in a "core" Chinese journal[1].

Results and discussion

Descriptive Statistics

[TABLE 1 ABOUT HERE]

Table 1 shows descriptive statistics of general information extracted from our sample articles. About half of the articles in our sample (51.90%) were written by 4 to 6 authors. The majority of authors were from academic institutions, but government officials and health care practitioners are also represented. In general, the length of articles is considerably shorter than one would expect in leading medical journals in English: a little more than (54.43%) ran to less than 3 pages. As categorized by a recognized Chinese journal ranking system (Peking University Library 2014), 40.51% of studies in our sample appeared in core journals. Half of all studies (49.37%) were supported by state-level or provincial-level research funds. A large majority of studies (85%) covered only a single province, and a greater proportion of studies (about 50%) focused on the eastern part of the country rather than other regions.

Research Methodologies

[TABLE 2 ABOUT HERE]

Only 12 studies in our sample (15.19%) used control groups; the majority (67 articles, 84.81%) did not include a control group. Of the 12 studies with control groups, 11 applied DID methods in

[1] "Core" Chinese journals are assessed by Peking University Library (Peking University Library 2014). These Chinese journals publish higher-quality articles with greater impact than "non-core" journals in the library's ranking system.

estimating policy effects, but none used random assignment of treatment groups and control groups.

Selection bias, which occurs when control groups and treatment groups are not assigned randomly, was particularly pronounced in our sample articles, given the complexity in policy implementation for NEMS. For instance, systematic differences may have existed between the control group and the treatment group before the policy intervention, because it is possible that health care providers in different areas were faced, even before policy intervention, with different fiscal issues, operational conditions, and patients with varying economic capacity (Tian, Hou, and Dong 2010). Six studies in the sample attempted to control for selection bias due to absence of random assignment using matching techniques.

Among the 67 studies without control groups, 58 (73.42%) involved both pre-tests and post-tests, before and after the policy intervention, respectively. Most of these studies (55 articles, 69.62%) simply computed the differences in measures before and after the policy intervention. The estimated policy effects with simple before-and-after comparison would be subject to a host of threats to internal validity. For instance, ODCP may change over time because of socioeconomic interventions other than NEMS. Three studies employed alternative quasi-experimental techniques, such as interrupted time series analysis and regression discontinuity.

The remaining nine studies used neither control groups nor pre-tests. Without pre-tests, it is hard to tell whether changes have occurred with NEMS implementation. Without control groups, it is hard to demonstrate what might have happened without the policy intervention. To show that changes did occur, five studies conducted surveys among citizens/doctors/patients after NEMS in order to gauge their perceptions of changes before and after the policy intervention. The other four studies (5.06%) without control groups and without pre-tests simply focused on observations after NEMS implementation.

The units of analysis frequently used in studies were "primary care providers" (in 49.37% of

studies) and "prescriptions" (in 36.71%). About half of the studies (48.10%) selected samples randomly. In terms of data collection, the majority of studies (70.89%) were based on primary data.

Effect Size

Table 3 presents the overall mean differences in our effect-size measures before and after implementation of NEMS. Overall mean difference was computed from raw mean differences in individual studies and averaged across studies, to indicate the measures' overall policy effects even though they might be flawed by validity threats. We use *t*-tests to detect the significance of policy effects.

[TABLE 3 ABOUT HERE]

The overall mean difference in ODCP across studies is –5.82 yuan per prescription, a statistically significant difference. Similar reductions are also found in other outcome indicators, such as ANDTP, PPA, and PPI, indicating not only that outpatients' drug costs decreased significantly following implementation of NEMS but also that the number of drug types per prescription and the percentage of prescriptions using antibiotics or injections all experienced downward trends.

To ensure comparability across different studies given their diversity in research methodologies, we standardized the raw mean differences of measures by using the random effect model. Figure 1 display the standardized effect sizes for individual studies and their 95% confidence intervals, and overall effect size by measures. The horizontal lines (rows) center on the effect sizes of individual studies and display their 95% confidence intervals. The bottom horizontal line (last row) displays the 95% confidence interval of the overall effect size.

[FIGURE 1 ABOUT HERE]

The standardized effect sizes and I^2 statistics show that effect sizes on each measure are highly heterogeneous across studies. From Figure 1, we can also see that, for each measure, the effect

sizes for individual studies dispersed to either side of the horizontal line. The overall effect sizes of NEMS on ODCP and on ANDTP are all significantly negative at the 5% level. In Figure 1(a), the overall effect size on ODCP is –0.51, averaged over 26 studies, with a 95% confidence interval of (–1.34, 0.06). The overall effect size on ANDTP is –0.30 computed from 16 studies, with individual effect sizes ranging from –0.58 to –0.04. However, after standardization, the overall effect sizes on PPA or PPI become insignificant—different outcomes from the results reported in Table 3. In Figure 1(b), the overall effect size on PPA is –0.32 averaged over 18 studies, but not significant ($p > 0.10$). The overall effect size on PPI is –0.44, with a 95% confidence interval of (–1.00, 0.12).

Multivariate meta-regression

Table 4 shows descriptive statistics for the variables included in the multivariate meta-regression described above. Some variables show only small variations over studies on ANDTP, PPA, and PPI. For instance, studies using DID account for only 6% of studies on ANDTP, 5% of studies on PPA, and zero studies on PPI. Small variations over studies render the impact of these study characteristics indeterminate through the subsequent regression analysis. Because of this concern we performed regressions only for ODCP. The effect sizes are in the format of raw mean differences.

[TABLE 4 ABOUT HERE]

Table 5 reports the results of the meta-regression. The first column regresses effect sizes on choices of analytic methods. Interrupted time series estimation, our benchmark analytic method, shows that observed effect sizes are significantly affected by the use of DID method. The same result, which can be drawn from regression models 2, 3, and 4, indicates that when using DID method, the implementation of NEMS seems to have been more effective, compared to results of studies that used the other two methods. The difference of the observed effect size is –21.66 yuan per prescription (i.e. -24.950-(-3.290)) between DID and a simple before-and-after comparison in regression model 1. However, effect size is not significantly different whether it is estimated by time series estimation or a simple before-and-after comparison.

[TABLE 5 ABOUT HERE]

Regression model 2 adds a variable representing the use of primary data. Results suggest that the use of primary data had a significant effect on the policy effects reported: if a study used primary data rather than secondary data, NEMS policy appears to have been more effective.

In addition to regression model 2, regression model 3 includes a dummy variable for random sampling, although whether a study used random sampling did not show a significant impact on the observed policy effect. Regression model 4 involves dummy variables for authorship affiliated with government units/health care providers, the availability of research funds, and whether the article was in a core Chinese journal. None of these shows a significant impact on the observed policy effect.

Overall, the use of DID method and use of primary data showed significant impacts on the estimated effect sizes even when controlling for other variables, with DID relatively larger in effect than use of primary data. These results suggest that using non-DID methodology or using secondary statistical data could underestimate the policy effects of NEMS.

Conclusion

The growing interest in EBPM has led to a surge in empirical research on various topics in health policy in China. The CNKI database lists a total of 23 journal articles with "empirical analysis" and "health" in their titles from 1996 through 2005, but the number increases to 650 from 2006 through 2015, a 30-fold rise. New journals have been launched to emphasize the importance of the producing high-quality evidence for policy making (Wang and Jin 2011; Jiang, Zhang, and Shen 2013).

But the quality of empirical research for policy making cannot be taken for granted. The majority of studies we surveyed from the years 2010–2015 had no control groups and simply compared differences in measures taken before and after policy intervention, and some studies

were based on post-tests only. Only a few studies used DID to estimate policy effects, and fewer still attempted to address selection bias when control or treatment group selection was not randomized.

The use of methodologies might be different for English publications. We had thorough search for published articles in order to provide sufficient articles for meta-analysis, but could only get access to few eligible English publications on the topic by 2016. In addition, Meta-regressions effectively work when there are adequate number of studies for meta-analysis. If studies covered in meta-analysis are only a few, many study characteristics could be not significant in meta-regressions, even though in actual they might have important impacts on the observed sizes of policy intervention effects (Higgins and Green 2008). We only performed the meta-regressions with 35 studies on ODCP, because there were not enough studies to do the same analyses for other indicators. As studies suitable for meta-regression have to meet certain requirements, many studies failed in the selection process, which happened in other meta-analysis studies as well.

Our analysis also offers insights into potential mechanisms for improvement. Academic journals should continue to serve as gatekeepers for best research by imposing high standards on quality of publications, including attention to research methodologies and to the analytical rigor of research. Systematic review of available evidence, as we have done here, can play a key role in effective use of evidence for policy making in developing countries, especially for learning beyond single studies. Analytical tools such as meta-analysis can also be used more extensively to explore the potential sources of disparities across different studies.

Contributors: All authors were involved in article collection and data extraction. Lili Li provided quantitative analysis skills. All authors contributed to the writing and subsequent revisions of the manuscript.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

References

Arnold, Matthias M., Andreas W. Rathgeber, and Stefan Stöckl. 2014. "Determinants of corporate hedging: A (statistical) meta-analysis." *Quarterly Review of Economics and Finance* 54 (4):443-458. doi: 10.1016/j.qref.2014.05.002.

Barber, Sarah L., Baobin Huang, Budiono Santoso, Richard Laing, Valerie Paris, and Chunfu Wu. 2013. "The reform of the essential medicines system in China: a comprehensive approach to universal coverage." *Journal of Global Health* 3 (1):010303. doi: 10.7189/jogh.03.010303.

Behague, Dominique, Charlotte Tawiah, Mikey Rosato, Télésphore Some, and Joanna Morrison. 2009. "Evidence-based policy-making: The implications of globally-applicable research for context-specific problem-solving in developing countries." *Social Science & Medicine* 69 (10):1539-1546. doi: 10.1016/j.socscimed.2009.08.006.

Black, Nick. 2001. "Evidence Based Policy: Proceed with Care." *BMJ: British Medical Journal* 323 (7307):275-278.

Borenstein, Michael. 2009. Introduction to meta-analysis. Vol. 1. Hoboken;Chichester, West Sussex, U.K;: John Wiley & Sons.

Borenstein, Michael, and Harris M. Cooper. 2009. "Effect Sizes for Continuous Data." In The handbook of research synthesis and meta-analysis, edited by Harris M. Cooper, Larry V. Hedges and Jeff C. Valentine, 221-235. New York: Russell Sage Foundation.

Borenstein, Michael, Larry V. Hedges, Julian P. T. Higgins, and Hannah R. Rothstein. 2010. "A basic introduction to fixed-effect and random-effects models for meta-analysis." *Research Synthesis Methods* 1 (2):97-111. doi: 10.1002/jrsm.12.

Brownson, R. C., J. G. Gurney, and G. H. Land. 1999. "Evidence-based decision making in public health." *Journal of Public Health Management and Practice (JPHMP)* 5 (5):86-97.

Cheung, Mike W. L. 2015. Meta-analysis: a structural equation modeling approach. Chichester, West Sussex: John Wiley & Sons, Inc.

Currie, Janet, Wanchuan Lin, and Juanjuan Meng. 2014. "Addressing Antibiotic Abuse in China: An Experimental Audit Study." *Journal of Development Economics* 110:39-51. doi: 10.1016/j.jdeveco.2014.05.006.

Currie, Janet, Wanchuan Lin, and Wei Zhang. 2011. "Patient knowledge and antibiotic abuse: Evidence from an audit study in China." *Journal of Health Economics* 30 (5):933-949. doi: 10.1016/j.jhealeco.2011.05.009.

Duckett, Jane. 2010. The Chinese state's retreat from health: Policy and the politics of retrenchment. Abingdon, Oxon.: Routledge.

Frenk, Julio. 2006. "Bridging the Divide: Global Lessons from Evidence-Based Health Policy in Mexico." *The Lancet* 368 (9539):954-961.

Gilson, Lucy, and Nika Raphaely. 2008. "The terrain of health policy analysis in low and middle income countries: A review of published literature 1994-2007." *Health Policy and Planning* 23 (5): 294-307.

Gray, J. A. Muir. 2004. "Evidence Based Policy Making: Is about Taking Decisions Based on Evidence and the Needs and Values of the Population." *BMJ: British Medical Journal* 329 (7473):988-989.

Higgins, Julian, Sally Green, in collaboration with Cochrane and Wiley InterScience. 2008. Cochrane handbook for systematic reviews of interventions. Hoboken, NJ; Chichester, England: Wiley-Blackwell.

Higgins, Julian P. T., Simon G. Thompson, Jonathan J. Deeks, and Douglas G. Altman. 2003. "Measuring Inconsistency In Meta-Analyses." *BMJ: British Medical Journal* 327 (7414):557-560. doi: 10.1136/bmj.327.7414.557.

Hyder, Adnan A., Adrijana Corluka, Peter J. Winch, Azza El-Shinnawy, Harith Ghassany, Hossein Malekafzali, Meng-Kin Lim, Joseph Mfutso-Bengo, Elsa Segura, and Abdul Ghaffar. 2011. "National policy-makers speak out: Are researchers giving them what they need?" *Health Policy and Planning* 26 (1): 73-82.

Jiang, Fan, Jun Zhang, and Xiaoming Shen. 2013. "Towards evidence-based public health policy in China." *The Lancet* 381 (9882):1962-1964.

Lavis, John, Huw Davies, Andy Oxman, Jean-Louis Denis, Karen Golden-Biddle, and Ewan Ferlie. 2005. "Towards systematic reviews that inform health care management and policy-making." *Journal of Health Services Research & Policy* 10 (suppl. 1):35-48.

Liang, Wannian, Jipan Xie, Hongpeng Fu, and Eric Q. Wu. 2014. "The Role of Health Economics and Outcomes Research in Health Care Reform in China." *PharmacoEconomics* 32 (3):231-234. doi: 10.1007/s40273-014-0141-2.

Lin, Tengfei, Ming Hu, and Yuan Mi. 2014. "National Essential Medicines System in China: A Status-quo Analysis." *Chinese Journal of Evidence-Based Medicine* 14 (5):575-583. In Chinese.

Lipsey, Mark W., and David B. Wilson. 2001a. "The Way in Which Intervention Studies Have 'Personality' and why it is Important to Meta-Analysis." *Evaluation & the Health Professions* 24 (3):236-254. doi: 10.1177/01632780122034902.

Lipsey, Mark W., and David B. Wilson. 2001b. Practical meta-analysis. Vol. 49. Thousand Oaks, Calif: Sage Publications.

Marston, G., and R. Watts. 2003. "Tampering with the evidence: A critical appraisal of evidence-based policy-making." *Australian Review of Public Affairs Journal* 3 (3):143-163.

McMichael, Celia, Elizabeth Waters, and Jimmy Volmink. 2005. "Evidence-based public health: what does it offer developing countries?" *Journal of Public Health (Oxford)* 27 (2):215-221.

Ministry of Health., various years. China Health Statistical Yearbook. Beijing: China Union Medical University Press.

O'Dwyer, Lisel. 2004. A critical review of evidence-based policy making. Australian Housing and Urban Research Institute.

Peking University Library. 2014. A guide to Chinese core journals (Zhong Wen He Xin Qi Kan Yao Mu Zong Lan). Edited by Zhu Qiang, He Jun and Cai Rong. Peking University Press.

Qian, Jiwei, and Ake Blomqvist. 2014. Health policy reform in China: A comparative perspective. Vol. 36. Hackensack, N.J; Singapore;: World Scientific Pub. Co.

Ramesh, M., Xun Wu, and Alex Jingwei He. 2014. "Health governance and healthcare reforms in China". *Health Policy and Planning* 29 (6): 663-672.

Ringquist, Evan J. 2005. "Assessing Evidence of Environmental Inequities: A Meta-Analysis." *Journal of Policy Analysis and Management* 24 (2):223-247. doi: 10.1002/pam.20088.

The Lancet. 2007. Editorial, "Reforming Research in China." *The Lancet* 369 (9565):880.

Thompson, Simon G., and Julian P. T. Higgins. 2002. "How should meta-regression analyses be undertaken and interpreted?" *Statistics in Medicine* 21 (11):1559-1573. doi: 10.1002/sim.1187.

Thompson, Simon G., and Stephen J. Sharp. 1999. "Explaining heterogeneity in meta-analysis: A comparison of methods." *Statistics in Medicine* 18 (20):2693-2708. doi: 10.1002/(SICI)1097-

0258(19991030)18:20<2693::AID-SIM235>3.3.CO;2-M.

Tian, Xiujuan, Jianlin Hou, and Zhumin Dong. 2010. "Evaluation on the Policy Effect of the New Rural Cooperative Medical Scheme." *Chongqing Social Science* (2):20–24.

Victora, Cesar G., Jean-Pierre Habicht, and Jennifer Bryce. 2004. "Evidence-Based Public Health: Moving beyond Randomized Trials." *American Journal of Public Health* 94 (3):400-405. doi: 10.2105/AJPH.94.3.400.

Wang, Jiyao, and Xuejuan Jin. 2011. "Translating evidence into policy in China: Opportunities and challenges." *Frontiers of Medicine* 5 (3):315-320. doi: 10.1007/s11684-011-0146-6.

Wei, Yan, Wenqiang Yin, Xin Ma, Zhongming Chen, Fei Wang, and Dongmei Huang. 2014. "Impact of the essential drug system on the proportion of drug sales to the total revenue: A meta-analysis" (Ji Ben Yao Wu Zhi Du Dui Ji Ceng Yi Liao Ji Gou Yao Zhan Bi Ying Xiang De Meta Fen Xi). *Chinese Journal of Health Statistics* 31 (2):291-292. In Chinese.

World Health Organization (WHO). 2012. Changing mindsets: Strategy on health policy and systems research. World Health Organization.

Wu, Xun, Yanling He, and Milan Tungwen Sun. 2013. "Public Administration Research in Mainland China and Taiwan: An Assessment of Journal Publications, 1998-2008." *Public Administration* 91 (2):261-280. doi: 10.1111/j.1467-9299.2011.01978.x.

Xu, Bingfa, Kai Qin, Lei Li, and Luyan Fan. 2012. "Influence of National Essential Drug System on the Application of Antibiotics in Primary Hospitals of Hefei." *China Pharmacy* 23 (16):1533-1535. In Chinese.

Yip, Winnie, and William Hsiao. 2014. "Harnessing the privatisation of China's fragmented health-care delivery." *The Lancet* 384 (9945):85-818.

Table 1. Basic features of relevant studies

	Number of studies (%)		Number of studies (%)
Descriptive statistics on authorship			
Number of authors		Authors' Affiliations	
1~3	24 (30.38)	Academic institutions [a]	42 (53.16)
4~6	41 (51.90)	Government or health care providers[b]	37 (46.84)
>6	14 (17.72)		
Number of references and number of pages: proxies for the quality of research			
Number of references		Number of pages	
0~5	22 (27.85)	1~3	43 (54.43)
6~10	47 (59.49)	4~6	35 (43.04)
>10	10 (12.66)	>6	5 (2.53)
Journal type			
Core journals [c]	32 (40.51)		
Other journals	47 (59.49)		
Publication time and research fund: proxies for attentions to the study field			
Publication Time		Research fund	
2011	15 (18.99)	State/provincial level research fund	39 (49.37)
2012	13 (16.46)	Other research fund[d]	16 (20.15)
2013	25 (31.64)	No research fund	24 (30.38)
2014	26 (32.91)		
Spatial information of study area			
Study scope [e]		Location	
City or County	47 (59.49)	Eastern China	40 (50.63)
Single province	20 (25.32)	Middle China	13 (16.46)
Region or state	12 (15.19)	Western China	11 (13.92)
		Cross regions	10 (12.66)
		Not mentioned	5 (6.33)

[a] Includes universities and academic institutions other than universities.

[b] Counts studies with at least one author from government units, or health care providers that contain primary care providers and hospitals above county-level.

[c] As assessed by Peking University Library (Peking University Library 2014). These Chinese journals publish higher-quality articles with greater impact than other journals in the library's ranking system.

[d] Studies without by state/provincial-level research funds, supported by funding from international or corporate sources.

[e] Scope of a study area, within a city/county, within a province, or regional/national

Table 2. Characteristics of research methods

Research designs	**Number of studies (%)**
Pretest-posttest design with control group, matching	6 (7.59)
Pretest-posttest design with control group, non-matching	5 (6.33)
Post-test only design with control group, non-matching	1 (1.27)
One-group pretest-posttest design	58 (73.42)
One-group posttest-only design	9 (11.39)
Data analysis methods	
DID	11 (13.92)
A simple before-and-after comparison [a]	55 (69.62)
Interrupted time series analysis	3 (3.80)
Others [b]	10 (12.66)
Unit of analysis	
Primary care providers [c]	39 (49.37)
Prescriptions	29 (36.71)
Others [d]	11 (13.93)
Sampling	
Random sampling [e]	38 (48.10)
Others	41 (51.90)
Data	
Primary data [f]	56 (70.89)
Secondary data	23 (29.11)

[a] Evaluates effects of policy intervention by simply computing the differences in measures before and after intervention.

[b] Three studies using factor analysis techniques, five studies that investigate on individual perceptions of the policy effects, and two studies that simply describe statistics of measures in post-tests.

[c] Primary care providers can be county level hospitals providing primary health care, township/community primary care providers, or village clinics; hospitals above county-level are not counted as primary care providers.

[d] "Others" includes other types of unit of analysis, such as counties, hospitals above county-level and individuals (citizens, doctors, or patients).

[e] Some studies use stratified sampling and their samples would be counted as "Random sampling" if they select samples randomly at one or more strata.

[f] Data (e.g., prescription data) collected from health care providers directly in fieldwork, in contrast to secondary data collected from statistical databases in other publications.

Table 3. Overall mean difference in measures

Measures	Mean [a]	Std. dev.	95% CI [b] lower	95% CI [b] upper	Number of studies (%)
ODCP [c]	-5.74**	16.69	-11.47	-0.01	35 (53.85)
ANDTP [d]	-0.21***	0.28	-0.36	-0.06	16 (24.61)
PPA [e]	-2.07*	5.22	-4.51	0.38	20 (30.77)
PPI [f]	-3.52**	5.96	-6.39	-0.65	19 (29.23)

[a] Overall mean difference in measures before and after the policy intervention across studies.

[b] Lower and upper limits of the 95% confidence interval.

[c] Yuan/prescription.

[d] 1 /prescription.

[e] Percentage.

[f] Percentage.

$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4. Descriptive characteristics of studies suitable for meta-analysis

Variables	**ODCP**	**ANDTP**	**PPA**	**PPI**
Total number of studies	35	16	20	19
Analytic method, data and sampling	**Mean** [a]	**Mean**	**Mean**	**Mean**
Method: DID, dummy	0.11	0.06	0.05	0.00
Method: A simple before-and-after comparison [b] dummy	0.80	0.94	0.95	1.00
Method: Interrupted time series analysis, dummy	0.09	0.00	0.00	0.00
Data: Primary data, [c] dummy	0.74	1.00	0.90	0.89
Sampling: Random sampling, [d] dummy	0.54	0.81	0.80	0.84
Basic information	**Mean**	**Mean**	**Mean**	**Mean**
Authorship: Government/health care providers, [e] dummy	0.46	0.38	0.30	0.32
Research fund, [f] dummy	0.71	0.69	0.75	0.74
Core journals, [g] dummy	0.34	0.19	0.25	0.21

[a] Mean value of each variable. Under each column, variables are averaged over studies that are suitable for meta-analysis and report the same measure. For each dummy variable, the mean value also indicates the percentage of studies that have a value of 1 for the dummy variable.

[b] Measuring policy implementation effects by imply computing the differences in measures before and after policy intervention.

[c] "Primary data" refers to data (e.g. prescription data) collected from health care providers directly in fieldwork.

[d] In studies using stratified sampling, a study has a value of 1 for this variable if it selects samples randomly at one or more strata.

[e] "Government/health care providers" has a value of 1 when a study has at least one author from government units, or from health care providers at hospitals above county-level.

[f] Studies supported by a research fund at any level, have a value of 1 for the variable.

[g] As assessed by the by Peking University Library (Peking University Library 2014), These journals publish articles of higher quality and greater impact in comparison to others in the library's ranking.

Table 5. Meta-regression of effect sizes of ODCP on study characteristics

	(1) Yuan/ prescription [a]	(2) Yuan/ prescription	(3) Yuan/ prescription	(4) Yuan/ prescription
Method: DID	−24.950**	−30.940**	−32.310**	−29.680**
	(11.870)	(11.720)	(12.090)	(12.270)
Method: a simple before-and-after comparison [b]	−3.290	−6.284	−8.704	−11.160
	(9.442)	(9.141)	(10.170)	(10.120)
Data: Primary data [c]		−11.970*	−13.050**	−12.050*
		(5.939)	(6.292)	(6.235)
Sampling: Random sampling [d]			3.348	2.583
			(5.871)	(5.986)
Authorship: Government/health care providers [e]				2.582
				(5.669)
Research fund [f]				−10.140
				(6.616)
Core Journals [g]				−2.091
				(5.897)
Constant	−0.257	11.720	12.790	20.910
	(8.973)	(10.430)	(10.710)	(12.910)
N	35	35	35	35
adj. R^2	0.133	0.209	0.191	0.225

[a] The dependent variable is $d_{ODCP,i}$, which is the effect size in i^{th} study measured in the form of raw mean difference; standard errors are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

[b] Measuring effects of policy implementation by computing the differences in measures before and after policy intervention.

[c] Data (e.g. prescription data) collected from health care providers directly in fieldwork.

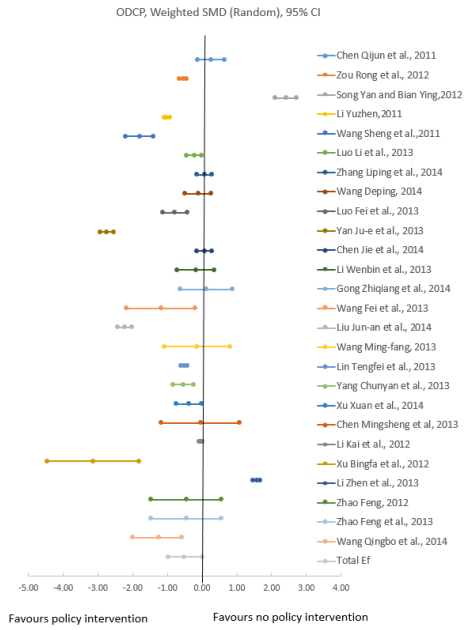
[d] In studies using stratified sampling, a value of 1 for this variable if it selects samples randomly at one or more strata.

[e] "Government/health care providers" has a value of 1 for at least one author from government units primary care providers in hospitals above county level.

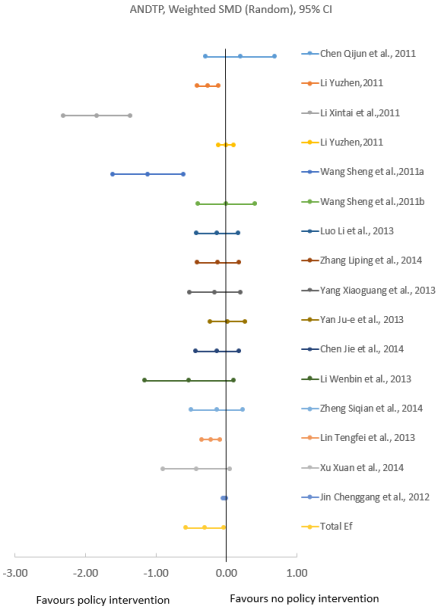
[f] When supported by research funding at any level, variable has a value of 1.

[g] As assessed by Peking University Library (Peking University Library 2014), publications of higher quality and greater impact than other journals in the Library's ranking.

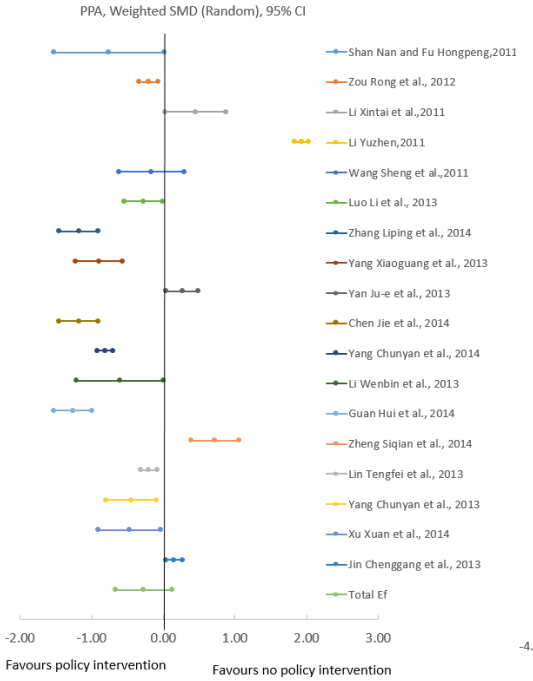
Figure 1. Forest plot for policy effects



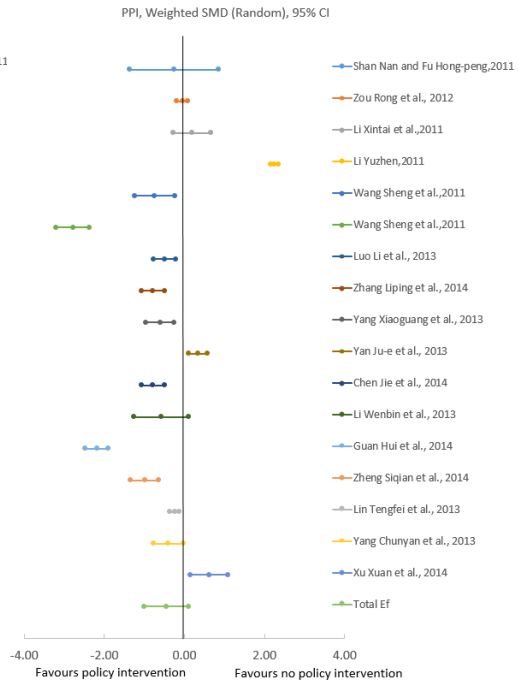
(a) ODCP, Yuan/prescription



(b) ANDTP, 1/prescription



(c) PPA, %



(d) PPI, %