

The health and health system impacts of China's comprehensive primary health care reforms: a systematic review

Running title: China's PHC reforms: health and health system impacts

MA Chang Cai^{1*}, MSc Shangzhi Xiong^{2,3}, PhD Christopher Millett^{1,4}, PhD Jin Xu⁵, PhD Maoyi Tian^{2,6}, PhD Thomas Hone¹

¹ Public Health Policy Evaluation Unit, School of Public Health, Imperial College London, UK.

² The George Institute for Global Health, Faculty of Medicine and Health, University of New South Wales, Sydney, Australia.

³ Global Health Research Centre, Duke Kunshan University, Jiangsu, China.

⁴ Public Health Research Centre and Comprehensive Health Research Centre, NOVA National School of Public Health, NOVA University Lisbon, Portugal.

⁵ China Center for Health Development Studies, Peking University China Center for Health Development Studies, Beijing, China

⁶ School of Public Health, Harbin Medical University, Harbin, China.

*Corresponding author

Name: Chang Cai

Email: c.cai20@imperial.ac.uk

Address: Reynolds Building, St Dunstan's Road, London W6 8RP

Country: United Kingdom

ORCID: 0000-0003-0883-5021

Word counts: 4644

Keywords

Primary health care, China, health reforms, policy evaluations

Key messages

- Existing evaluations have mostly focused on single-component primary health care (PHC) interventions and evidence from Asia is lacking. Comprehensive system-wide PHC reforms are limited to few low- and middle-income countries (LMICs) and there is a dearth of evaluations on their synergetic impacts. There is no synthesised understanding of the health system and health impacts of China's comprehensive PHC reforms starting in 2009.
- China's comprehensive PHC reforms since 2009 have produced modest impacts including increased primary care utilisation, improved the health outcomes of people with non-communicable diseases (NCDs) and disproportionately benefited vulnerable, high-risk populations, and less-

resourced regions. The reform impacts on quality of primary care were underexamined, with only a few studies assessing impacts on antibiotic use, satisfaction, and perception of care quality among service users. Findings on health service costs were mixed. Substantial evidence gaps remain on the reform impacts on financial protection, general population health and equity.

- PHC-oriented reforms can increase primary care utilisation, reduce inequality, improve the health of people with NCDs, and should be a priority for LMICs to advance Universal Health Coverage (UHC). Other health and health system impacts remain poorly studied.
- PHC research in LMICs would benefit from robust studies, adequate data on care quality, and strengthened knowledge base of holistic PHC reforms, to inform policy-making for UHC.

Author contributions

CC, TH and CM contributed to study conceptualisation and the development of the review protocol. CC executed the data search. CC, XSZ and TH performed the title, abstract and full-text screening. CC and XSZ undertook the data extraction. CC and TH undertook the quality assessment. CC wrote the first draft of the manuscript. TH and CM made substantial edits to the draft. All authors contributed to the revision and edits of subsequent drafts, including the analysis and interpretation. All authors approved the final manuscript.

Reflexivity Statement

The authors include one female and five males and have various levels of seniority. The author group for this paper includes researchers from various institutions in both low and middle-income countries (LMICs) and high-income countries, including China, UK, and Australia. All the authors have expertise in health system research and policy evaluations in LMICs with a focus on primary health care.

Data availability

All data collected and generated in this systematic review is available from the corresponding author upon request.

Ethical Approval

Ethical approval is not required by for this type of article.

Funding

There is no funding for this study.

Conflicts of interest

None declared.

The health and health system impacts of China's comprehensive primary health care reforms: a systematic review

Abstract

China's comprehensive primary health care (PHC) reforms since 2009 aimed to deliver accessible, efficient, equitable and high-quality health care services. However, knowledge on the system-wide effectiveness of these reforms is limited. This systematic review synthesizes evidence on the reforms' health and health system impacts. In August 2022, 13 international databases and three Chinese databases were searched for randomised controlled trials, quasi-experimental studies and controlled before-after studies. Included studies assessed large-scale PHC policies since 2009; had a temporal comparator and a control group; and assessed impacts on expenditures, utilisation, care quality, and health outcomes. Study quality was assessed using ROBINS-I and results were synthesized narratively. From 49,174 identified records, 42 studies were included - all with quasi-experimental designs, except for one randomised control trial. Nine studies assessed as at low risk of bias. Only five low to moderate quality studies assessed the comprehensive reforms as a whole and found associated increases in health service utilisation, whilst the other 37 studies examined single-component policies. The National Essential Medicine Policy (N=15) and financing reforms (N=11) were the most studied policies, whilst policies on primary care provision (i.e., family physician policy and the National Essential Public Health Services) were poorly evaluated. The PHC reforms were associated with increased primary care utilisation (N=17) and improved health outcomes in people with non-communicable diseases (N=8). Evidence on healthcare costs was

unclear and impacts on patients' financial burden and care quality were understudied. Some studies showed disadvantaged regions and groups accrued greater benefits (N=8). China's comprehensive PHC reforms have made some progress in achieving their policy objectives including increasing primary care utilisation, improving some health outcomes, and reducing health inequalities. However, China's health system remains largely hospital-centric and further PHC strengthening is needed to advance Universal Health Coverage.

ACCEPTED MANUSCRIPT

INTRODUCTION

Primary health care (PHC) is the most effective and efficient approach to building strong health systems, and achieving universal health coverage (UHC) and the health-related Sustainable Development Goals (SDGs) (Hone et al., 2018; Kruk and Pate, 2020; Starfield et al., 2005; World Health Organization, 2018). However, PHC strengthening in low- and middle-income countries (LMICs) has been challenging due to limited resources, fragmented health systems, and poor-quality care (Kruk et al., 2018). These deficits straddle multiple building blocks of health systems and require system-wide reforms (Kruk et al., 2018; Lancet Global Health, 2018). Although policymakers in LMICs are aware of this imperative (Kruk and Pate, 2020; Thapa et al., 2019), the introduction of comprehensive PHC reforms is scarce and there is a dearth of robust evaluations and synthesized evidence (Bitton et al., 2019; Lancet Global Health, 2018; Rule et al., 2014).

China had a hospital-centric health system where public hospitals were the main health service providers and were allocated with a disproportionate share of resources, whilst the PHC sector was under-developed. In 2008, hospitals in China accounted for 53.1% of total health expenditures, employed 54.3% of doctors, and delivered 53.4% of outpatient care (Ministry of Health, 2009). In an attempt to change the hospital-dominated system, China initiated wide-reaching health reforms in 2009 aiming to shift to a PHC-oriented health system and deliver accessible, efficient, equitable, and quality primary care services to achieve UHC (Liu et al., 2017; Meng and Tang, 2013; The State Council of the People's Republic of China, 2009). From the perspective of PHC strengthening, this on-going health reform can

be divided into two phases: the first phase (between 2009 and 2014) focused on establishing a PHC-based health system and increasing access to primary care services. This included insurance expansion, investment in PHC infrastructure, education and training for medical students and primary care physicians, as well as the introduction of the National Essential Medicine Policy (NEMP) and the National Essential Public Health Services (NEPHS) (Yip et al., 2019). However, low utilisation and poor-quality primary care services persisted with little alleviation of the overused hospital system (Meng et al., 2019; Yip et al., 2019). To address these issues, the second phase of the reform starting in 2015 focused on restructuring health service delivery, including establishing a hierarchical, integrated delivery system with partial gatekeeping, and the introduction of a primary care model based on family physicians (Table 1) (Fang et al., 2019; Ma et al., 2019; Meng et al., 2019; Yuan et al., 2019). It is worth noting that this PHC reform is part of wider reforms, including those in the public hospital system (Liu et al., 2011; Yip et al., 2012).

While China's PHC reforms have been presented and debated extensively, there is no systematic synthesis of their health system and health outcome impacts (Liu et al., 2011; Liu et al., 2017; Yip et al., 2019). This systematic review aims to examine the impacts of China's PHC reforms since 2009 on the health system and health outcomes based on evaluations which use robust study designs, including randomised controlled trials, quasi-experimental studies, and controlled before-after studies.

METHODS

This systematic review follows a registered protocol (PROSPERO registration number CRD42021239991).

Search strategy

Thirteen international databases and three Chinese databases were searched in August 2022, including MEDLINE, Embase, ScienceDirect, Scopus, Web of Science, CINAHL, Cochrane, Ecolit, Jstor, HMIC, World Bank Library, WHOLIS, Opengrey, CNKI, Wanfang and cqvip.

The terms 'primary health care', 'family physicians', 'ambulatory care' and their synonyms were searched in titles or abstracts, along with relevant MeSH terms. Some reform-specific terms were also used, such as "zero mark-up" and "the National Essential Medicine Policy". These terms were linked by OR Boolean operators and the search was further restricted using AND Boolean operators and the words "reform", "China" and their synonyms. No language restriction was applied in the search. For studies in Mandarin, similar terms were searched along with specific policy titles, such as 'Yilianti' (part of the integrated delivery system, also known as Medical Alliance in Chinese), 'Fenjizhenliao' (partial gatekeeping, also known as Hierarchical Medical System in Chinese) and 'Xinnonghe' (abbreviation of New Rural Cooperative Medical Insurance in Chinese). More details on the search strategy can be found in Appendix 1. Grey literature was searched in Opengrey. When full-text articles were not available, corresponding authors of the identified studies were contacted for publication information. References of included articles were also screened for additional studies.

Study selection

Studies were eligible for inclusion if they: 1) studied Chinese citizens; 2) evaluated PHC reforms from 2009 described in Table 1 and were implemented at the equivalent of city/county level or above; 3) included both a temporal comparator and a control group with no exposure to the reforms or exposed to alternative interventions (e.g. capitation and pay-for-performance (P4P) for provider payment reforms); 4) examined outcomes that were PHC-related, including any relevant measures of system and individual health costs, health service utilisation, quality of primary care, health outcomes or health inequalities; 5) used study designs that were either randomised control trials (RCTs) or quasi-experimental studies with controls.

Studies were excluded if they: 1) assessed interventions at community or village level and clinical trials; 2) did not have both temporal comparators and control groups; 3) examined outcomes which only included hospital care or did not measure the actual primary care delivery, such as using simulated clients and performance examinations for primary care providers; 4) were qualitative studies, commentaries, reviews, cross-sectional studies, uncontrolled before-after studies and uncontrolled interrupted time series studies.

One author performed the searches. The title and abstracts of all identified records were screened by two reviewers independently using Covidence (Veritas Health Innovation, 2021). Conflicts over study inclusion were discussed between the two reviewers and resolved by a third author. Full-text screening was undertaken by two

reviewers independently, and disagreements were discussed and resolved with a third author. Data from eligible studies were extracted by the same two reviewers separately and cross-checked for errors. The quality of included studies was assessed by two reviewers.

Data analysis

Information on the interventions, study settings, data sources, study designs, outcomes, effect directions, effect sizes, statistical significance and subgroup analysis was collected using a pre-designed standardised table. All relevant reported results were extracted. Outcomes were grouped into five categories: health expenditures (for both patients and providers), health service utilisation, quality of care, health outcomes and health inequalities.

The quality of included studies was assessed using the Cochrane Collaboration's tool for assessing Risk of Bias In Non-randomized Studies of Interventions (ROBINS-I) (Sterne JAC et al., 2016). Quality was assessed based on bias across seven domains: confounding, selection of participants into the study, classification of intervention (i.e., whether the interventions are clearly defined and the possibility of misclassification), deviation from intended interventions, missing data, measurement of outcomes, and selection of the reported results. An overall risk of bias rating was generated. Each of the risk of bias domains was graded into one of the five categories: low, moderate, serious, critical and no information (Schünemann et al., 2019). Compared with other quality assessment tools, ROBINS-I offers a comprehensive framework to identify weaknesses in intervention evaluations, but its

reported risk-of-bias is often higher and with a smaller range of variety (Losilla et al., 2018) since only studies comparable to “a well-performed randomized trial” can be judged as low risk of bias (Schünemann et al., 2019). To address the limitation, the ROBINS-I tool in this review was adjusted whereby the quality of studies was upscaled one level if they balanced control groups using matching or weighting, included more than four years of observation, ruled out co-interventions, or had random policy assignment (e.g., RCTs).

Heterogeneity in assessed policies, study populations and outcomes precluded meta-analysis, and a narrative approach was adopted. Studies were grouped by evaluated policies and outcome measures. A harvest plot was used to narratively summarise the heterogeneous effects of the PHC policies on each of the five outcome groups (Ogilvie et al., 2008). Heterogeneous effects examined in the included studies were extracted without any pre-specification and were grouped into three types for the analysis based on what was identified in studies: 1) geographical; 2) socio-demographic; 3) people with and without non-communicable diseases (NCDs), including cardiovascular diseases, cancers, chronic respiratory diseases and diabetes (World Health Organization, 2022).

RESULTS

The search identified 49,174 records. After removing duplicates, 30,053 studies were title and abstract screened and 347 studies were full-text screened (Figure 1). In total, 42 reports from 41 studies were eligible for inclusion, of which 31 were in English and 11 in Mandarin. Publication years ranged from 2012 to 2022. Eleven studies

were nationwide, and 31 were based on cities/counties within provinces, except for two across provinces (Wei et al., 2015; Yin et al., 2016). There were 18 studies from eastern China, eight from central China (Jiang et al., 2016; Liu et al., 2014; Liu et al., 2016; Miao et al., 2018; Miao et al., 2019; Tang et al., 2018; Yang et al., 2013; Yang et al., 2017) and five from western China (Miao et al., 2016; Powell-Jackson et al., 2015; Shen et al., 2021; Tan et al., 2015; Yao et al., 2020). Nearly half (N=18) studied rural populations, whilst 15 studies included both rural and urban populations. Eleven studies focused exclusively on people with NCDs, specifically with hypertension or diabetes. Study duration was less than five years for 36 studies, of which 13 had a duration of two years or less (Table 2).

ACCEPTED MANUSCRIPT

All included studies were quasi-experimental studies, except for one matched-pair cluster-randomised control trial (Liu et al., 2016). The majority of the studies employed difference-in-difference (DID) approaches (N=38), with seven combining this with propensity score matching/weighting (Ding and Wu, 2015; Hu et al., 2021; Miao et al., 2018; Miao et al., 2019; Wang and Liu, 2022; Zhang et al., 2014a; Zhang et al., 2017), eight using fixed effect panel regression (Duan et al., 2020; Pan and Yang, 2022; Shen et al., 2021; Sun et al., 2016b; Xu et al., 2020a; Yang et al., 2017; Yi et al., 2015; Zhou et al., 2021), three adopting dynamic DID models (also known as event-study analysis) (Shen et al., 2020a; Yao et al., 2020; Yuan et al., 2021), and four without any statistical tests (Chen et al., 2013; He et al., 2014; Ma et al., 2014; Wang et al., 2014). Four studies were controlled ITS studies (Liang et al., 2014; Sun et al., 2016a; Tang et al., 2018; Yang et al., 2013). Most studies used untreated populations in similar cities/counties in the province (N=38), except for one based on populations from a different province (Yin et al., 2016). Three studies used people with different intensities of the treatments (Shen et al., 2021; Yi et al., 2015) or alternative treatments (i.e., different combinations of provider payments and primary care delivery) across provinces (Wei et al., 2015). There were 17 ecological studies at facility level, whilst 22 studies used individual-level data. Most studies (N=35) used secondary data, of which 30 used routinely collected data, such as administrative data from PHC facilities and health insurance data (Table 2).

The 42 studies covered all of the major PHC reforms, of which 37 assessed single-component reforms. Of the five studies that evaluated the impacts of comprehensive PHC reform as a whole (Liang et al., 2014; Liu et al., 2014; Pan and Yang, 2022; Wei et al., 2015; Zhou et al., 2021), three focused on the first stage (2009-2014)

(Liang et al., 2014; Liu et al., 2014; Wei et al., 2015), whilst two examined the second stage (2015 onwards) (Pan and Yang, 2022; Zhou et al., 2021). The National Essential Medicine Policy was the most investigated PHC policy (N=15), followed by financing reforms. Of the 11 studies on financing reforms, six focused on demand-side insurance reforms (e.g. reducing co-payment for visits to PHC facilities or NCD-related outpatient visits) (Jiang et al., 2016; Miao et al., 2018; Miao et al., 2019; Shen et al., 2020b; Shen et al., 2020a; Zhang et al., 2014a), two focused on provider payment reforms, such as P4P and capitation (Sun et al., 2016b; Tan et al., 2015), two on the combination of these two approaches (Powell-Jackson et al., 2015; Sun et al., 2016a), and one on governmental subsidies (Shen et al., 2021). Of the five studies on reorganising the health service delivery system, one assessed gatekeeping (Xu et al., 2020a) and four examined the integrated delivery system (Duan et al., 2020; Hu et al., 2021; Miao et al., 2016; Yuan et al., 2021). Policies on primary care provision were the least studied, with three studies on the family physician policy (Wang and Liu, 2022; Yin et al., 2016; Zhu et al.) and one study on the National Essential Public Health Services (Zhang et al., 2017). Additionally, two studies investigated two micro-level interventions on providers: 1) PHC workforce training (Yao et al., 2020); and 2) public reporting on PHC facilities' performance (Figure 2) (Liu et al., 2016).

Health expenditures were the most studied outcome (26 studies). Three studies investigated out-of-pocket expenditures (OOPE) (Ding and Wu, 2015; Miao et al., 2018; Miao et al., 2019), whilst the rest examined NCD-related service or drugs costs in both PHC facilities and hospitals (N=6) (Hu et al., 2021; Liu et al., 2014; Miao et al., 2018; Miao et al., 2019; Shen et al., 2020b; Zhang et al., 2014a) or the costs in PHC facilities alone (N=19). Health service utilisation was examined by 25 studies, including visits to PHC facilities (N=20), essential health service usage (N=5) (Liu et al., 2014; Tan et al., 2015; Wang et al., 2014; Yang et al., 2017; Zhang et al., 2017), and hospitalisation (N=9) (He et al., 2014; Jiang et al., 2016; Liu et al., 2014; Miao et al., 2018; Miao et al., 2019; Powell-Jackson et al., 2015; Shen et al., 2020b; Shen et al., 2020a; Yi et al., 2015). Quality of care was examined by 13 studies from three aspects: antibiotic use (N=8) (Chen et al., 2014; Gong et al., 2016; Jin et al., 2013; Liang et al., 2014; Liu et al., 2016; Sun et al., 2016b; Sun et al., 2016a; Yang et al., 2013), satisfaction and perception of quality of care (N=5) (Duan et al.; Liu et al., 2014; Wei et al., 2015; Yin et al., 2016; Zhu et al.), and the delivery of essential medicines (N=1) (Yang et al., 2017). Ten studies investigated changes in health status using biomarkers (N=8) (Duan et al.; Hu et al., 2021; Jiang et al., 2016; Miao et al., 2016; Miao et al., 2018; Miao et al., 2019; Yuan et al., 2021; Zhang et al., 2017), mortality (N=1) (Duan et al.), self-reported health status (N=2) (Jiang et al., 2016; Zhu et al.), health-related quality of life (N=2) (Miao et al., 2016; Wang and Liu, 2022). Only one study assessed the reform effects on income-related health inequality (Pan and Yang, 2022) (Table 2). The effect sizes reported in the included studies were modest (see Appendix 3).

The risk of bias of the included studies ranged from low to serious, with nine studies assessed as high-quality (Ding and Wu, 2015; Gong et al., 2016; Hu et al., 2021; Liu et al., 2016; Powell-Jackson et al., 2015; Shen et al., 2020a; Xu et al., 2020a; Yang et al., 2017; Yao et al., 2020) and 24 with moderate risk of bias (Table 2). The most common potential bias was confounding from baseline differences and not accounting for concurrent interventions (N=31) (see Appendix 2). The nine studies at serious risk of bias did not adjust for any potential confounders (Chen et al.; Jiang et al., 2016; Jin et al., 2013; Li et al.; Liu et al., 2014; Sun et al., 2016a; Tan et al.; Yin et al., 2016; Zhu et al.), whilst one had a serious risk of bias from missing pre-intervention data and an inconsistent outcome measure (Sun et al., 2016a). Missing data and deviation from the intended interventions were generally not discussed in the studies (14 and 23 out of 42 respectively) (see Appendix 2).

Impacts of comprehensive PHC reforms

Of the five studies that examined the comprehensive reform as a whole package, three evaluated the first phase between 2009 and 2014 (Liang et al., 2014; Liu et al., 2014; Wei et al., 2015), and two assessed the second phase since 2015 (Pan and Yang, 2022; Zhou et al., 2021). There was no consensus on impacts and the quality of evidence was low. One study in eastern China found drug expenditures and antibiotic use among children decreased following the 2009 reforms (Liang et al., 2014), whilst a low-quality study from central China found essential health service and hospital utilisation and satisfaction increased, but outpatient visits (both at PHC facilities and hospitals) were unaffected (Liu et al., 2014). A third study compared two PHC models introduced in 2009 in eastern China and reported that full

government funding, better-trained health workforces, and services tailored to local health needs were associated with a higher perceived quality of care (Wei et al., 2015). Both studies on the second reform phase were nationwide and of moderate quality, with one study reporting an increase in primary care utilisation among urban populations, but not their rural counterparts (Zhou et al., 2021), with the other study showing that the reform decreased income-related health inequality (Table 3) (Pan and Yang, 2022).

Impacts of the National Essential Medicine Policy

There were 15 studies on the National Essential Medicine Policy. Impacts were mixed, except for a consistent decrease in drug costs. Health outcomes were not studied. Seven ecological studies observed a reduction in costs for drugs at PHC facilities after NEMP introduction (Chen et al., 2014; Gong et al., 2016; Li et al.; Ma et al.; Tang et al., 2018; Yang et al., 2013; Zhang et al., 2014b), but changes were insignificant in two national studies (Chen et al., 2014; Zhang et al., 2014b). One high-quality study on urban populations from eastern China found no impacts on drug costs or OOPE in PHC facilities (Ding and Wu, 2015). Three studies reported a decrease in healthcare costs in PHC facilities after the NEMP (Chen et al.; Han et al.; Ma et al.), whilst four studies found no impacts on healthcare costs (Ding and Wu, 2015; Li et al.; Yi et al., 2015; Zhang et al., 2014b). Of the six studies on health service utilisation, two found an increase in the utilisation of primary care services among rural populations in eastern China (Han et al., 2016; He et al., 2014), one study with no statistical test observed a decrease (Wang et al., 2014), and three studies found no significant effects following NEMP introduction (Ding and Wu, 2015; Li et al.; Yi et al., 2015), including one national study (Yi et al., 2015). One study

without statistical tests found vaccination rates in rural areas increased (Wang et al., 2014), whilst two studies found increasing inpatient visits in rural areas (He et al., 2014; Yi et al., 2015). Among the four studies on antibiotic use, one high-quality national study found decreased prescription of antibiotics in urban PHC facilities (Gong et al., 2016), two intermediate quality studies found no significant impacts (Chen et al., 2014; Yang et al., 2013), and one low-quality study found an increase in antibiotic use (Jin et al., 2013) after the NEMP (Table 3).

Impacts of financing policies

The 11 studies on financing policies showed a general increase in primary care utilisation across various regions and population groups, but the quality of evidence varied. Of the six studies on demand-side financing reforms, five were NCD-focused interventions (Jiang et al., 2016; Miao et al., 2018; Miao et al., 2019; Shen et al., 2020b; Zhang et al., 2014a), and one high-quality study assessed incentives for using PHC facilities (Shen et al., 2020a). All the six studies found increasing the charges for outpatient services provided by both hospitals and PHC facilities following a reduction in copayments (Miao et al., 2018; Miao et al., 2019; Shen et al., 2020b; Shen et al., 2020a; Zhang et al., 2014a), except for one low-quality study (non-significant effects) (Jiang et al., 2016). However, the six studies did not distinguish who covered these increased costs – either by insurers or individual out-of-pocket payments. Both studies on OOPE found decreased total OOPE after reducing copayments (Miao et al., 2018; Miao et al., 2019), but with a significant decline only in inpatient OOPE (Miao et al., 2018). All three studies on primary care utilisation found increased visits to PHC facilities following co-payment reduction

(Miao et al., 2018; Miao et al., 2019; Shen et al., 2020a). All five studies on hospitalisations showed decreased inpatient care visits after copayment reductions for visits to PHC facilities (Miao et al., 2018; Miao et al., 2019; Shen et al., 2020b; Shen et al., 2020a), except for one low-quality study reporting no effects (Jiang et al., 2016). The high-quality study further showed that outpatient visits to hospitals decreased while primary care utilisation increased (Shen et al., 2020a). Three studies examining health outcomes were all based on people with NCDs and suggested that reducing copayment was associated with decreased diastolic blood pressure (Miao et al., 2018; Miao et al., 2019) and blood glucose (Jiang et al., 2016), and increases in self-rated health (Jiang et al., 2016) (Table 3). The health impacts of the demand-side financing reforms among the general population were not studied.

Four studies investigated the health system impacts of provider payment reforms (i.e., capitation or P4P) with (Powell-Jackson et al., 2015; Sun et al., 2016a) or without (Sun et al., 2016b; Tan et al., 2015) combining demand-side incentives. Three of the four studies found visits to PHC facilities (Powell-Jackson et al., 2015; Sun et al., 2016a; Tan et al., 2015) and use of essential health services (Tan et al.) increased. No significant changes in inpatient care use were found (Powell-Jackson et al., 2015). One high-quality study also found that a combination of both demand and supply-side financing reforms showed no greater effects on visits to PHC facilities than the sole demand-side insurance reform did (Powell-Jackson et al., 2015). No studies found provider payment reforms had significant effects on system costs (Sun et al., 2016b; Sun et al., 2016a), antibiotic use (Sun et al., 2016b), or intensity of treatment (Powell-Jackson et al., 2015), except for one low-quality study

reporting decreases in the use of both antibiotics and injections (Sun et al., 2016a). One moderate quality study using continuous DID showed that increasing government subsidies to PHC facilities had an adverse impact on visits to PHC facilities, suggesting higher subsidies were insufficient to incentivise primary care practitioners to deliver additional services beyond the set goals (Shen et al., 2021) (Table 3). The effects of provider payment reforms on health outcomes were not examined.

Impacts of health service delivery policies

Regarding policies restructuring the service delivery system, one study assessed gatekeeping (Xu et al., 2020a) and four examined integration of delivery systems (Duan et al., 2020; Hu et al., 2021; Miao et al., 2016; Yuan et al., 2021). The high-quality study on gatekeeping found the use of primary care services increased, accompanied with a decrease in outpatient visits to hospitals, but with no impact on health costs in PHC facilities (Xu et al., 2020a). Of the four studies which examined integrated care, one high-quality study found reducing total health costs in all medical institutions (Hu et al., 2021), and two moderate-quality found increased visits to PHC facilities (Duan et al., 2020; Yuan et al., 2021). All three studies were based on people with hypertension or diabetes and showed service integration increased control rates for both of the diseases (Hu et al., 2021; Yuan et al., 2021) and increased self-reported health (Miao et al., 2016). One moderate-quality study on the general population found no significant impacts on essential health service use, health services costs, satisfaction, self-reported health or mortality (Duan et al.) (Table 3).

Four studies examined policies on primary care at community level, including family physicians (Wang and Liu, 2022; Yin et al., 2016; Zhu et al., 2017) and the National Essential Public Health Services (NEPHS) (Zhang et al., 2017). Two of the three studies on the family physician policy were low-quality and from eastern China (Yin et al., 2016; Zhu et al., 2017), whilst the other moderate quality study used evidence from a national survey (Wang and Liu, 2022). One low-quality study using self-reported measures found reduced healthcare costs and satisfaction, and increased visits to PHC facilities, but no changes in self-reported health (Zhu et al., 2017). By contrast, two studies found improved perceived quality of care (Yin et al., 2016) and health-related quality of life (Wang and Liu, 2022) among people with NCDs. One moderate-quality national study showed the NEPHS increased uptake of antihypertensive drugs and blood pressure examinations among people with hypertension and increased blood pressure control rates (Zhang et al., 2017) (Table 3).

The two micro-level provider interventions, workforce training (Yao et al., 2020) and performance reporting (Liu et al., 2016), were assessed by two high-quality studies, from western and central China respectively. One found PHC workforce training increased primary care utilisation (Yao et al., 2020), whilst one found public reporting on the performance of PHC facilities reduced drug costs and the use of combined antibiotics (Liu et al., 2016) (Table 3).

Heterogeneous impacts of the reforms

Heterogeneous impacts were investigated in 11 studies. Eight studies investigated differential effects across regions (i.e., urban/rural, or eastern/western/central China) (Chen et al., 2014; Powell-Jackson et al., 2015; Shen et al., 2020a; Yang et al., 2013; Yang et al., 2017; Yao et al., 2020; Zhang et al., 2017; Zhou et al., 2021). Two policies were found to benefit rural/western/central regions more in terms of health service utilisation and health outcomes. The National Essential Public Health Services was associated with larger increases in medication use, physical examinations and hypertension control in western/central China compared to eastern China (Zhang et al., 2017). PHC workforce training increased visits to PHC facilities in rural regions but not in urban areas (Yao et al., 2020). By contrast, the NEMP had no significant impacts on antibiotic use in either rural or urban areas (Chen et al., 2014; Yang et al., 2013), but reduced delivery rates of essential medicines in rural PHC facilities (Yang et al., 2017). One study on the second stage of the PHC reform (Zhou et al., 2021), as well as one on copayment reductions (Shen et al., 2020a), found increasing use of PHC facilities among urban areas but not rural areas. One study on financing reforms found people living closer to village clinics had greater increases in clinic utilisation (Powell-Jackson et al., 2015) (Figure 3).

Among six studies examining heterogeneous reform impacts among people with different socio-demographic characteristics (Chen et al., 2014; Liu et al., 2016; Powell-Jackson et al., 2015; Shen et al., 2020b; Shen et al., 2020a; Wei et al., 2015) and with/without NCDs (Chen et al., 2014; Shen et al., 2020b; Wei et al., 2015), one from eastern China assessed the first stage of the comprehensive reform and found that people from poorer households or with NCDs had higher ratings of primary care quality after the reform (Wei et al., 2015). Regarding single-component policies, four

studies from various regions found that financing interventions disproportionately benefited females (Powell-Jackson et al., 2015; Shen et al., 2020a), the elderly (Shen et al., 2020a), people from poor households (Powell-Jackson et al., 2015), and people with NCDs (Chen et al., 2014; Shen et al., 2020b) in terms of primary care utilisation. One study on performance reporting also found greater improvement in rational use of health services among the elderly (Liu et al., 2016) (Figure 3).

DISCUSSION

China's PHC reforms since 2009 have made some progress towards the stated policy objectives of increasing primary care utilisation and improving the health of people with NCDs. The PHC reforms disproportionately benefited vulnerable populations (i.e., women, the elderly and lower-income populations) and under-resourced regions (i.e., rural areas and central/western China), suggesting some progress in reducing inequalities. However, the overall impacts were relatively modest. The reform impacts on care quality, financial protection, and general population health were understudied, and evidence on health care system costs was mixed.

Comprehensive system-wide PHC reforms have been identified as a prerequisite for high-performing health systems and UHC achievement (Kruk et al., 2018; Kruk and Pate, 2020; Lancet Global Health, 2018), yet few countries have introduced ambitious reforms such as those in China. Comprehensive PHC reforms which involve both demand and supply sides are more desirable than selective, single-component approaches to build strong health systems (Kruk et al., 2018; Kruk and Pate, 2020) and achieve UHC (Atun, 2015) since single-component policies are

unlikely to achieve system-wide changes or synergetic effects. Policy change across multiple health system functions aligns with and can accelerate progress to a “whole-of-society approach” and multisector actions for health (World Health Organization, 2018). Additionally, the synergistic effects of comprehensive reforms can adapt to the dynamics and interconnected parts of health systems. Comprehensive PHC reforms have been implemented in a few middle-income countries, such as Brazil (Paim et al., 2011), Mexico (González-Pier et al., 2006; Knaul et al., 2012), Bolivia (Alvarez et al., 2016), Turkey (Atun et al., 2013; Atun, 2015), and Iran (Ghasemyani et al., 2022; Heshmati and Joulaei, 2016). These countries have well-established PHC sectors and share a paradigm of comprehensive PHC reforms – centering on family physician care models and financing reforms (Bitton et al., 2019). Although China pioneered the “barefoot doctor programme” in the 1960s to expand the coverage of primary care services in rural areas, this programme was unsuccessful in substantially changing the hospital domination of the health system due to scarce funds and unskilled workforces (Xu et al., 2020b; Zhang and Unschuld, 2008). China’s hospital-centric health system remained dominated by large public hospitals (Liu et al., 2017; Xu et al., 2020b), with a drastically imbalanced health resource distribution and poor coordination between hospital and PHC sectors (Yip et al., 2012). In this context, China’s PHC reforms have aimed to not only revive and strengthen the PHC system, but also improve integration between the PHC and hospital sector, making it a valuable case study for other settings.

Existing evaluations of PHC reforms in LMICs have mainly focused on single-component PHC policies (Angell et al., 2019; Bastos et al., 2017; Bitton et al., 2019; Hone et al., 2017; Qin et al., 2019), whilst there is a dearth of evidence on system-

wide, comprehensive PHC reforms. Turkey's comprehensive PHC reforms were found to increase infant and maternal health service utilisation and improve health outcomes (Atun et al., 2013). Similar impacts were found in Bolivia (Alvarez et al., 2016) and other Latin American countries (Ramírez et al., 2011). Comparably, this review found some evidence of increased use of maternal and general health services in China. However, there were no studies examining how comprehensive PHC reforms in the country affected health system costs, health outcomes and equity. The small number of studies and low-quality evidence identified in this review highlight a missed opportunity to examine the synergetic effects of China's ambitious PHC reforms.

Regarding single-component PHC policies, our finding of increased primary care utilisation in China aligns with evidence from reforms in other LMICs (Hone et al., 2016; Hone et al., 2017; Qin et al., 2019). This review found that deductible copayments for primary care services (Powell-Jackson et al., 2015; Shen et al., 2020a), free-of-charge essential public health services (Zhang et al., 2017), and the introduction of family physicians (Zhu et al.) all increased primary care utilisation in China. For people with hypertension or diabetes, these three interventions increased primary care service utilisation and medication and treatment adherence, reducing hospitalisations and resultant OOPE (Miao et al., 2018; Miao et al., 2019; Zhang et al., 2017), and improving their health outcomes (Miao et al., 2018; Miao et al., 2019; Wang and Liu, 2022; Zhang et al., 2017). The management of other NCDs, such as cardiovascular diseases, was not investigated. There should be a cautious interpretation of these increases in primary care utilisation. These increases might indicate increased access to care since unmet healthcare needs and undertreated

NCDs are prevalent in China (Wang et al., 2021; Yip et al., 2019). It could also suggest a shift in patient flows from hospitals to PHC facilities – supported by two studies reporting PHC increases were concomitant with reducing hospital utilisation (Shen et al., 2020a; Xu et al., 2020a). However, the increases in primary care utilisation could also be explained by overutilisation as the appropriateness of this usage was not ascertained. Notably, short-term increases in PHC utilisation following the reforms contradicts with nationwide decreases in the preference for using PHC facilities (Ta et al., 2020; Wan et al., 2021; Zhang et al., 2020) and the dwindling patient share of visits to PHC facilities (Liu et al., 2021b). This contradiction suggests the PHC reforms may have been ineffective in achieving enduring impacts, outpacing the growth of the hospital sector, or reversing the hospital-centric orientation of the health system. Detailed understanding of the nature of primary care utilisation, including appropriateness and efficiency, the long-term effectiveness of the reforms, and relative improvement in the PHC system compared with the hospital sector is needed.

This review is unable to draw firm conclusions on reform impacts on service quality given relevant evidence is scarce and of poor quality, and comprehensive measurements of care quality were missing. The mixed evidence on antibiotic use in this review, a useful indicator of quality, is consistent with a recent review from China (He et al., 2019). Evidence on quality perception was mixed and the reform impacts on process quality were not studied. This substantial evidence gap underscores the lack of knowledge on quality improvement and comprehensive quality-oriented performance measurements in China (Li et al., 2020; Xiong et al., 2022) - a common

impediment in LMICs for the achievement of a high-performance PHC system and UHC (Kruk and Pate, 2020; Lancet Global Health, 2018).

The finding of health improvements in this review were concentrated in high-risk populations (i.e. people with NCDs), and the few studies on general population health identified in this review found no significant health impacts. Evidence from Latin America and other Asian countries showed general population health improvements identified following PHC reforms (Bastos et al., 2017; Hone et al., 2020; Kruk et al., 2010). This difference may stem from three reasons. First, China's PHC reforms prioritised NCD management (Xiong et al., 2022), and some of the assessed copayment reductions (Jiang et al., 2016; Miao et al., 2018; Miao et al., 2019; Shen et al., 2020b; Sun et al., 2016a) and service delivery models (Miao et al., 2016) were targeted towards people with NCDs. This priority is reasonable since NCDs are poorly managed (Li et al., 2020) and predicted to remain an obstacle in China's path toward UHC (Liu et al., 2021b). Second, this review highlights the scarce studies and lack of quality measurements of the health of general populations. Reported health outcomes were selective for hypertension and diabetes, and health improvements among general population may not be captured or have longer lag times to show. More sensitive measurements for population health (e.g. infant and maternal health, mental health, and other cardiovascular diseases) were not assessed. Third, PHC policies that benefit the wider population may have been hampered by the high dependency on hospital sectors, the lack of system-level changes, and the knowledge gaps on quality improvement mentioned above. China's prioritisation of people with NCDs aligns with the global commitment and efforts in NCD control (World Health Organization, 2022). Nevertheless,

comprehensive system-wide PHC reforms remain imperative for LMICs to deliver benefits to wider populations and to achieve UHC (Kruk et al., 2018; Kruk and Pate, 2020).

Impacts of China's PHC reform on health care costs were unclear, with financial protection infrequently studied. This review finds that the NEMP reduced drug costs at PHC facilities, but did not affect OOPE – similar conclusions to a previous systematic review (Liu et al., 2021a). The limited impacts on OOPE may stem from primary care providers compensating lost drug mark-up income with other health service charges (Ding and Wu, 2015), offsetting intended financial protection. This plausibly reflects an underfunded PHC system and potentially explains unchanged provider behaviours found in this review, such as continued overuse of antibiotics (Chen et al., 2014; Sun et al., 2016b; Yang et al., 2013), overtreatment (Powell-Jackson et al., 2015), and continued unnecessary hospitalisation (Yi et al., 2015). This finding hints at the interactions of financing with other health system building blocks and reiterates the necessity of a system perspective on PHC financing to curb OOPE (Hanson et al., 2022).

The PHC reforms in China showed some pro-equity features which is consistent with findings from other LMICs (Atun et al., 2013; Lankaran et al., 2017; Ramírez et al., 2011) and the declining trend of inequalities in health service utilisation in China (Liu et al., 2021b). All assessed PHC policies disproportionately benefited vulnerable populations, including the elderly, females, lower-income populations, and high-risk populations (Chen et al., 2014; Ding and Wu, 2015; Liu et al., 2016; Powell-Jackson et al., 2015; Shen et al., 2020b; Shen et al., 2020a; Wei et al., 2015). This finding is

anticipated since China's PHC reforms have aimed to narrow the gaps in health service accessibility between different population groups and regions to achieve UHC. Additionally, these population groups are more likely to use PHC facilities (Sang et al., 2021; Tao et al., 2021; Zhang et al., 2020), and likely more responsive to improved access brought about by PHC reforms than their counterparts. Some evidence in this review shows that essential public health services and workforce training delivered more benefits to rural areas and middle/central China in terms of primary care utilisation and health outcomes (Yao et al., 2020; Zhang et al., 2017), whilst financial interventions failed to do so. This may suggest that China's PHC reforms in under-resourced regions faced challenges beyond affordability in increasing primary care utilisation, such as low trust in primary care providers (Duckett et al., 2016) and low accessibility of PHC facilities (Tao et al., 2021). These challenges indicate the needs for improving the accessibility and quality of primary care services among under-resourced areas, which can be addressed by essential public health services and workforce training but not sole financing incentives.

This study has limitations. Our conclusions are constrained by the sparse evidence, varying study quality, and the observational nature of studies. Heterogeneity across studies prevented meta-analysis. The exclusion of small-scale studies, including experimental studies, is a trade-off between internal and external validity of evidence and may overlook some relevant evidence. Well-conducted RCTs can produce robust causal inferences but may not produce generalisable knowledge needed to inform large-scale policies (Deaton and Cartwright, 2018; Pritchett and Sandefur, 2015). This review has key strengths. This is the first systematic review on the topic, and addresses an evidence gap on the impact of comprehensive PHC reforms in

Asia and globally (Bitton et al., 2019; Macinko et al., 2009). Comprehensive search strategies and broad search terms were used to minimise the possibility of omitting eligible studies. Studies were searched for in English and Mandarin. The review only focused on robust evidence by restricting eligible study designs to RCTs and controlled before-after studies. The evidence on China's PHC reforms mirrors wider weaknesses in PHC research in LMICs (Adam et al., 2011; Alvarez et al., 2016; Bitton et al., 2019). This includes considerable dearth of evaluations with rigorous study designs, inadequate data and limited measurement of quality of primary care and health outcomes which hampers robust and insightful research. Evidence gaps in PHC research often overlap with policy gaps in PHC reforms and hinder effective evidence-based policy making, creating a vicious cycle. The concentrated evidence on financing and primary care utilisation found in this review aligns with the identified "major policy initiatives" of China's health reform which were considerably strengthened over time (Xiong et al., 2022). By contrast, the scarce evidence on care quality highlighted in this review resonates with the policy gaps in performance improvements (Xiong et al., 2022). Most of the studies in this review did not discuss how well the reforms were implemented or deviations from the intended interventions. Due to the lack of detailed process evaluations, we cannot conclude whether the mixed results found in this review were due to poor design or poor implementation of the reforms. Existing studies mostly focused on single-component PHC interventions, whereas synthesized evidence and evaluations on the comprehensive reforms as a whole are rare. Understanding of long-term health system and health impacts and financial protection remains limited. Policy makers and PHC researchers in LMICs should work together and closely to address these challenges in the future.

CONCLUSION

China's comprehensive PHC reform since 2009 has made some progress in increasing primary care utilisation, improving equity, and improving some health outcomes for people with hypertension or diabetes. Progress regarding other objectives of the reforms is not clear. A question remains as to whether China's PHC reforms have been sufficiently wide-reaching and transformative to achieve a truly PHC centric health care system given evidence of continued dominance of the hospital sector (Liu et al., 2021b). While there remains a need to create a stronger knowledge base of comprehensive system-wide PHC reforms globally, including evaluations with more robust study designs and quality impacts, the mixed evidence from China shows that evidence-based comprehensive reforms and strong PHC systems are essential for altering hospital domination and achieving UHC, providing crucial lessons to inform and advance PHC policy initiatives globally.

ACCEPTED MANUSCRIPT

REFERENCE

- Adam T, Ahmad S, Bigdeli M *et al.* 2011. Trends in health policy and systems research over the past decade: still too little capacity in low-income countries. *PloS one*, 6 **11**: e27263.
- Alvarez FN, Leys M, Mérida HER *et al.* 2016. Primary health care research in Bolivia: systematic review and analysis. *Health policy and planning*, 31 **1**: 114-28.
- Angell B, Dodd R, Palagyi A *et al.* 2019. Primary health care financing interventions: a systematic review and stakeholder-driven research agenda for the Asia-Pacific region. *BMJ global health*, 4 **Suppl 8**: e001481.
- Atun R. 2015. Transforming Turkey's health system—lessons for universal coverage. *New England Journal of Medicine*, 373 **14**: 1285-9.
- Atun R, Aydin S, Chakraborty S *et al.* 2013. Universal health coverage in Turkey: enhancement of equity. *The Lancet*, 382 **9886**: 65-99.
- Bastos ML, Menzies D, Hone T *et al.* 2017. The impact of the Brazilian family health on selected primary care sensitive conditions: a systematic review. *PloS one*, 12 **8**: e0182336.
- Bitton A, Fifield J, Ratcliffe H *et al.* 2019. Primary healthcare system performance in low-income and middle-income countries: a scoping review of the evidence from 2010 to 2017. *BMJ Global Health*, 4 **Suppl 8**: e001551.
- Chen M, Mao W, Chen W *et al.* 2013. Impact of the essential medicines policies on operation and expenditure burden at primary health care institutions. *Chinese Health Resources*, 16 **2**: 93-4.
- Chen M, Wang L, Chen W *et al.* 2014. Does economic incentive matter for rational use of medicine? China's experience from the essential medicines program. *Pharmacoeconomics*, 32 **3**: 245-55.
- Deaton A and Cartwright N. 2018. Understanding and misunderstanding randomized controlled trials. *Social science & medicine*, 210: 2-21.
- Ding L and Wu J. 2015. The Impact of China's National Essential Medicine Policy and Its Implications for Urban Outpatients: A Multivariate Difference-in-Differences Study. *Value in Health*, 18 **3**: A72-A3.
- Duan H, Zhang Y, Hou Y *et al.* 2020. Research on the Influence of the Policy of Medical Alliance on the Service Ability of Primary Community Medical and Health Institutions. *Journal of Gansu Administration Institute*, **4**: 4-16.
- Duckett J, Hunt K, Munro N *et al.* 2016. Does distrust in providers affect health-care utilization in China? *Health policy and planning*, 31 **8**: 1001-9.
- Fang H, Eggleston K, Hanson K *et al.* 2019. Enhancing financial protection under China's social health insurance to achieve universal health coverage. *bmj*, 365 **12378**: 1-4.
- Ghasemyani S, Raofi S, Hamidi H *et al.* 2022. Iran's Health Transformation Plan; Main Issues and Opportunities for Improvement: A Systematic Review. *Iranian Journal of Public Health*, 51 **9**: 1977-89.
- Gong Y, Yang C, Yin X *et al.* 2016. The effect of essential medicines programme on rational use of medicines in China. *Health Policy Plan*, 31 **1**: 21-7.

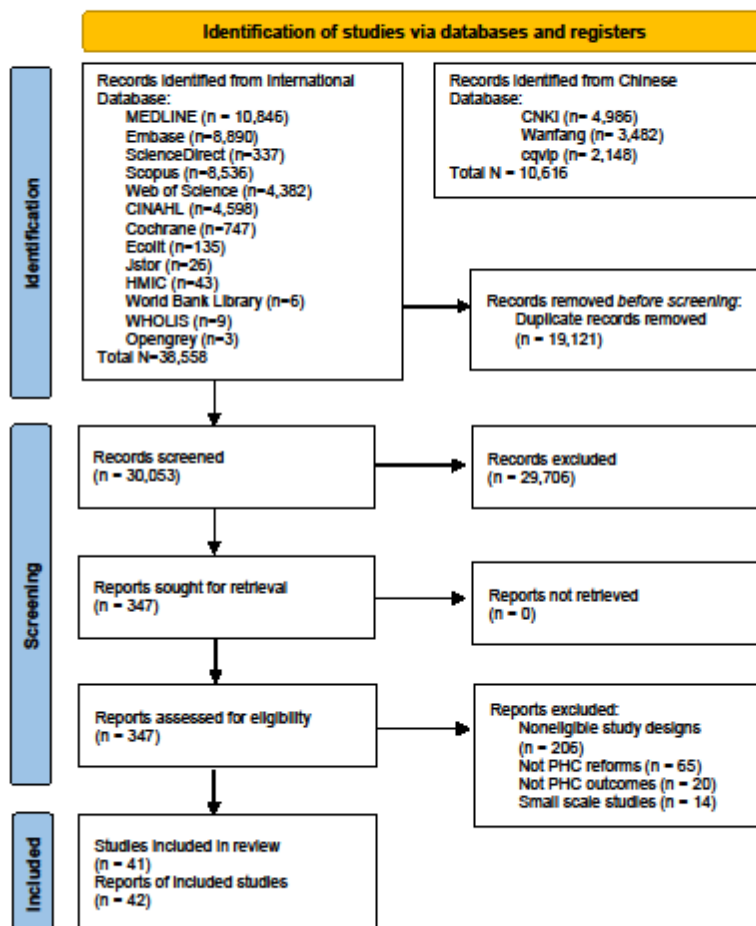
- González-Pier E, Gutiérrez-Delgado C, Stevens G *et al.* 2006. Priority setting for health interventions in Mexico's System of Social Protection in Health. *The Lancet*, 368 **9547**: 1608-18.
- Han C, Li C, Li J *et al.* 2016. Effects of Essential Drug Zero-profits Sales to Village Clinic Based on the Evaluation Method of Difference-in-Differences. *Chinese General Practice*, 19 **22**: 2647-50.
- Hanson K, Brikci N, Erlangga D *et al.* 2022. The Lancet Global Health Commission on financing primary health care: putting people at the centre. *The Lancet Global Health*, 10 **5**: e715-e72.
- He P, Sun Q, Shi L *et al.* 2019. Rational use of antibiotics in the context of China's health system reform. *BMJ*, 365: 14016.
- He Y, Shen Y, Wang Y *et al.* 2014. Comparative Analysis on the Operation Status of Township Hospitals before and after the Implementation of Essential Medicine System. *Chinese Health Economics*, 33 **11**: 54-7.
- Heshmati B and Joulaei H. 2016. Iran's health-care system in transition. *The Lancet*, 387 **10013**: 29-30.
- Hone T, Macinko J, and Millett C. 2018. Revisiting Alma-Ata: what is the role of primary health care in achieving the Sustainable Development Goals? *The Lancet*, 392 **10156**: 1461-72.
- Hone T, Lee JT, Majeed A *et al.* 2017. Does charging different user fees for primary and secondary care affect first-contacts with primary healthcare? A systematic review. *Health policy and planning*, 32 **5**: 723-31.
- Hone T, Gurol-Urganci I, Millett C *et al.* 2016. Effect of primary health care reforms in Turkey on health service utilization and user satisfaction. *Health Policy and Planning*, 32 **1**: 57-67.
- Hone T, Saraceni V, Medina Coeli C *et al.* 2020. Primary healthcare expansion and mortality in Brazil's urban poor: A cohort analysis of 1.2 million adults. *PLOS Medicine*, 17 **10**: e1003357.
- Hu H, Liang H, and Wang H. 2021. Longitudinal study of the earliest pilot of tiered healthcare system reforms in China: Will the new type of chronic disease management be effective? *Soc Sci Med*, 285: 114284.
- Jiang Q, Jiang Z, Xin Z *et al.* 2016. Evaluation of the impact of a chronic disease scheme reimbursing medical costs of patients with diabetes in Anhui province, China: a follow-up study. *BMC Public Health*, 16: 980.
- Jin C, Yang H, Liang X *et al.* 2013. The Impact Evaluation of Essential Medicine Policies on Joint Usage of Antibiotics, Steroids and Intravenous Infusion in Township Health Care. *China Health Economics*, 32 **7**: 53-5.
- Knaul FM, González-Pier E, Gómez-Dantés O *et al.* 2012. The quest for universal health coverage: achieving social protection for all in Mexico. *The Lancet*, 380 **9849**: 1259-79.
- Kruk ME and Pate M. 2020. The Lancet global health Commission on high quality health systems 1 year on: progress on a global imperative. *The Lancet Global Health*, 8 **1**: e30-e2.
- Kruk ME, Porignon D, Rockers PC *et al.* 2010. The contribution of primary care to health and health systems in low- and middle-income countries: a critical review of major primary care initiatives. *Social science & medicine*, 70 **6**: 904-11.
- Kruk ME, Gage AD, Arsenault C *et al.* 2018. High-quality health systems in the Sustainable Development Goals era: time for a revolution. *The Lancet global health*, 6 **11**: e1196-e252.

- Lancet Global Health. 2018. Adding quality to primary care. *The Lancet Global health*, 6 **11**: e1139.
- Lankaran KB, Khankeh HR, Zarei N *et al.* 2017. Toward equity under health system reform; a systematic review. *Shiraz E-Medical Journal*, 18 **11**.
- Li K, Sun Q, Zuo G *et al.* 2012. Study of the Impact of Essential Medicine System on the Patient Visits and Cost in Township Hospitals: Based on the Evaluation Method of Difference in Difference. *Chinese Health Economics*, 31 **4**: 62-4.
- Li X, Lu J, Hu S *et al.* 2017. The primary health-care system in China. *The Lancet*, 390 **10112**: 2584-94.
- Li X, Krumholz HM, Yip W *et al.* 2020. Quality of primary health care in China: challenges and recommendations. *The Lancet*, 395 **10239**: 1802-12.
- Liang X, Xia T, Zhang X *et al.* 2014. Governance structure reform and antibiotics prescription in community health centres in Shenzhen, China. *Fam Pract*, 31 **3**: 311-8.
- Liu C, Zhang X, Wang X *et al.* 2016. Does public reporting influence antibiotic and injection prescribing to all patients? A cluster-randomized matched-pair trial in china. *Medicine (Baltimore)*, 95 **26**: e3965.
- Liu GG, Vortherms SA, and Hong X. 2017. China's health reform update. *Annual review of public health*, 38: 431-48.
- Liu Q, Wang B, Kong Y *et al.* 2011. China's primary health-care reform. *The Lancet*, 377 **9783**: 2064-6.
- Liu W, Hsu C, Liu T *et al.* 2021a. Systematic Review of the Effect of a Zero-Markup Policy for Essential Drugs on Healthcare Costs and Utilization in China, 2015–2021. *Frontiers in Medicine*, 8: 618046.
- Liu X, Wang Z, Zhang H *et al.* 2021b. Measuring and evaluating progress towards Universal Health Coverage in China. *Journal of global health*, 11.
- Liu Y, Yuan Z, Liu Y *et al.* 2014. Changing community health service delivery in economically less-developed rural areas in China: impact on service use and satisfaction. *BMJ Open*, 4 **2**: e004148.
- Losilla J-M, Oliveras I, Marin-Garcia JA *et al.* 2018. Three risk of bias tools lead to opposite conclusions in observational research synthesis. *Journal of clinical epidemiology*, 101: 61-72.
- Ma A, Li C, Shen Y *et al.* 2014. Study on the Changes of Hospitalization Costs for Patients in Township Hospitals before and after the Implementation of Essential Medicine System. *Chinese Health Economics*, 33 **11**: 31-4.
- Ma X, Wang H, Yang L *et al.* 2019. Realigning the incentive system for China's primary healthcare providers. *BMJ*, 365 **12406**: 8-11.
- Macinko J, Starfield B, and Erinoshio T. 2009. The impact of primary healthcare on population health in low-and middle-income countries. *The Journal of ambulatory care management*, 32 **2**: 150-71.
- Meng Q and Tang S. 2013. Universal coverage of health care in China: challenges and opportunities. *Procedia-Social and Behavioral Sciences*, 77: 330-40.
- Meng Q, Mills A, Wang L *et al.* 2019. What can we learn from China's health system reform? *BMJ*, 365 **12349**: 1-5.
- Miao Y, Gu J, Zhang L *et al.* 2018. Improving the performance of social health insurance system through increasing outpatient expenditure reimbursement ratio: a quasi-experimental evaluation study from rural China. *Int J Equity Health*, 17 **1**: 89.

- Miao Y, Yuan X, Gu J *et al.* 2019. Constructing a value-based healthcare system for hypertensive patients through changing payment mode: evidence from a comparative study in rural China. *J Med Econ*, 22 **3**: 245-51.
- Miao Y, Zhang L, Sparring V *et al.* 2016. Improving health related quality of life among rural hypertensive patients through the integrative strategy of health services delivery: a quasi-experimental trial from Chongqing, China. *Int J Equity Health*, 15 **1**: 132.
- Ministry of Health. 2009. *China Health Statistical Yearbook*, Beijing: Peking Union Medical College Press.
- Ogilvie D, Fayer D, Petticrew M *et al.* 2008. The harvest plot: a method for synthesising evidence about the differential effects of interventions. *BMC medical research methodology*, 8 **1**: 1-7.
- Paim J, Travassos C, Almeida C *et al.* 2011. The Brazilian health system: history, advances, and challenges. *The Lancet*, 377 **9779**: 1778-97.
- Pan C and Yang J. 2022. The Impact of the Implementation of Hierarchical Medical Policy on Health Inequality among the Chinese Elderly. *Social Security Studies*, **1**: 49-60.
- Powell-Jackson T, Yip WC, and Han W. 2015. Realigning demand and supply side incentives to improve primary health care seeking in rural China. *Health Econ*, 24 **6**: 755-72.
- Pritchett L and Sandefur J. 2015. Learning from experiments when context matters. *American Economic Review*, 105 **5**: 471-75.
- Qin VM, Hone T, Millett C *et al.* 2019. The impact of user charges on health outcomes in low-income and middle-income countries: a systematic review. *BMJ Global health*, 3 **Suppl 3**: e001087.
- Ramírez NA, Ruiz JP, Romero RV *et al.* 2011. Comprehensive primary health care in South America: contexts, achievements and policy implications. *Cadernos de Saúde Pública*, 27 **10**: 1875-90.
- Rule J, Ngo DA, Oanh TTM *et al.* 2014. Strengthening primary health care in low- and middle-income countries: generating evidence through evaluation. *Asia Pacific Journal of Public Health*, 26 **4**: 339-48.
- Sang H, González-Vallejo C, Zhao J *et al.* 2021. Is low cost really conducive to primary care utilisation: An empirical analysis of community health centers in China. *Health & Social Care in the Community*, 29 **5**: e163-e73.
- Schünemann HJ, Cuello C, Akl EA *et al.* 2019. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *Journal of clinical epidemiology*, 111: 105-14.
- Shen C, Zhou Z, Lai S *et al.* 2021. Whether high government subsidies reduce the healthcare provision of township healthcare centers in rural China. *BMC Health Serv Res*, 21 **1**: 1184.
- Shen M, He W, and Li L. 2020a. Incentives to use primary care and their impact on healthcare utilization: Evidence using a public health insurance dataset in China. *Soc Sci Med*, 255: 112981.
- Shen M, He W, Yeoh EK *et al.* 2020b. The association between an increased reimbursement cap for chronic disease coverage and healthcare utilization in China: an interrupted time series study. *Health Policy Plan*, 35 **8**: 1029-38.
- Starfield B, Shi L, and Macinko J. 2005. Contribution of primary care to health systems and health. *The milbank quarterly*, 83 **3**: 457-502.

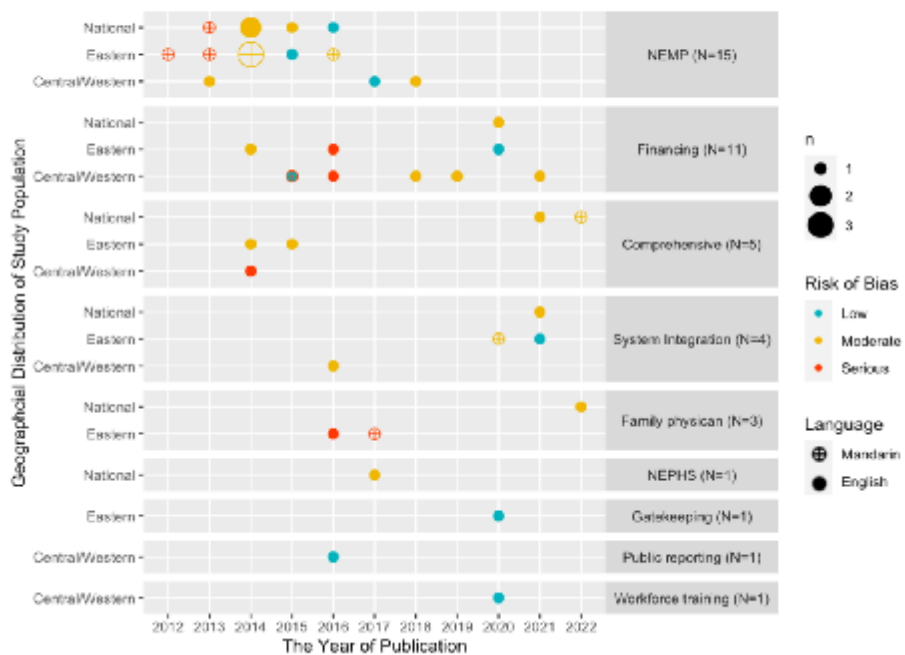
- Sterne JAC, Hernán MA, Reeves BC *et al.* 2016. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. *BMJ* 355 **i4919**: 1-7.
- Sun J, Zhang X, Zhang Z *et al.* 2016a. Impacts of a new insurance benefit with capitated provider payment on healthcare utilization, expenditure and quality of medication prescribing in China. *Trop Med Int Health*, 21 **2**: 263-74.
- Sun X, Liu X, Sun Q *et al.* 2016b. The Impact of a Pay-for-Performance Scheme on Prescription Quality in Rural China. *Health Econ*, 25 **6**: 706-22.
- Ta Y, Zhu Y, and Fu H. 2020. Trends in access to health services, financial protection and satisfaction between 2010 and 2016: Has China achieved the goals of its health system reform? *Social Science & Medicine*, 245: 112715.
- Tan H, Chen F, and Jing S. 2015. Analysis on The Performance Related-pay Impact on Effect of Basic Public Health Services - An Empirical Study Based on the DID. *Chinese Health Service Management*, 1: 68-70.
- Tang Y, Liu C, Zhang Z *et al.* 2018. Effects of prescription restrictive interventions on antibiotic procurement in primary care settings: a controlled interrupted time series study in China. *Cost Eff Resour Alloc*, 16: 1.
- Tao T, Hou L, and Shao R. 2021. Factors associated with the choice of primary treatment at the community level among the middle-aged and elderly in China. *The International Journal of Health Planning and Management*, 36 **2**: 442-58.
- Thapa G, Jhalani M, García-Saisó S *et al.* 2019. High quality health systems in the SDG era: country-specific priorities for improving quality of care. *PLoS medicine*, 16 **10**: e1002946.
- The State Council of the People's Republic of China. *Opinions on Deepening Health System Reform in China*. http://www.china.org.cn/government/scio-press-conferences/2009-04/09/content_17575378.htm accessed 16 April 2022.
- Veritas Health Innovation. 2021. *Covidence systematic review software*, Melbourne: Veritas Health Innovation.
- Wan G, Wei X, Yin H *et al.* 2021. The trend in primary health care preference in China: a cohort study of 12,508 residents from 2012 to 2018. *BMC health services research*, 21 **1**: 1-15.
- Wang L and Liu W. 2022. Effects of Family Doctor Contract Services on the Health-Related Quality of Life Among Individuals With Diabetes in China: Evidence From the CHARLS. *Front Public Health*, 10: 865653.
- Wang L, Peng W, Zhao Z *et al.* 2021. Prevalence and Treatment of Diabetes in China, 2013-2018. *JAMA*, 326 **24**: 2498-506.
- Wang X, Wang C, Sang X *et al.* 2014. Study on the Impact of Essential Medical System on the Medical Services of Township Hospitals. *Chinese Health Economics*, 33 **11**: 42-5.
- Wei X, Li H, Yang N *et al.* 2015. Changes in the perceived quality of primary care in Shanghai and Shenzhen, China: a difference-in-difference analysis. *Bull World Health Organ*, 93 **6**: 407-16.
- World Health Organization. 2018. *Declaration of Astana*, Geneva: World Health Organization.
- World Health Organization. 2022. *Noncommunicable diseases progress monitor*, Geneva: World Health Organization.
- Xiong S, Cai C, Jiang W *et al.* 2022. Primary health care system responses to non-communicable disease prevention and control: A scoping review of national policies in Mainland China since the 2009 health reform. *The Lancet Regional Health-Western Pacific*: 100390.

- Xu J, Powell-Jackson T, and Mills A. 2020a. Effectiveness of primary care gatekeeping: difference-in-differences evaluation of a pilot scheme in China. *BMJ Glob Health*, 5 8: e002792.
- Xu J, Gorsky M, and Mills A. 2020b. A path dependence analysis of hospital dominance in China (1949–2018): lessons for primary care strengthening. *Health Policy and Planning*, 35 2: 167-79.
- Yang L, Huang C, and Liu C. 2017. Distribution of essential medicines to primary care institutions in Hubei of China: effects of centralized procurement arrangements. *BMC Health Serv Res*, 17 1: 727.
- Yang L, Liu C, Ferrier JA *et al.* 2013. The impact of the National Essential Medicines Policy on prescribing behaviours in primary care facilities in Hubei province of China. *Health Policy Plan*, 28 7: 750-60.
- Yao Y, Liu GG, and Cui Y. 2020. Job training and organizational performance: Analyses from medical institutions in China. *China Economic Review*, 60.
- Yi H, Miller G, Zhang L *et al.* 2015. Intended and unintended consequences of China's zero markup drug policy. *Health Aff (Millwood)*, 34 8: 1391-8.
- Yin J, Wei X, Li H *et al.* 2016. Assessing the impact of general practitioner team service on perceived quality of care among patients with non-communicable diseases in China: a natural experimental study. *Int J Qual Health Care*, 28 5: 554-60.
- Yip W, Fu H, Chen AT *et al.* 2019. 10 years of health-care reform in China: progress and gaps in universal health coverage. *The Lancet*, 394 10204: 1192-204.
- Yip WC, Hsiao WC, Chen W *et al.* 2012. Early appraisal of China's huge and complex health-care reforms. *The Lancet*, 379 9818: 833-42.
- Yuan B, Balabanova D, Gao J *et al.* 2019. Strengthening public health services to achieve universal health coverage in China. *bmj*, 365 12358: 1-4.
- Yuan S, Fan F, and Zhu D. 2021. Effects of Vertical Integration Reform on Primary Healthcare Institutions in China: Evidence From a Longitudinal Study. *Int J Health Policy Manag*, 11 8: 1-9.
- Zhang A, Nikoloski Z, Albalá SA *et al.* 2020. Patient choice of health care providers in China: primary care facilities versus hospitals. *Health Systems & Reform*, 6 1: e1846844.
- Zhang D and Unschuld PU. 2008. China's barefoot doctor: past, present, and future. *The Lancet*, 372 9653: 1865-7.
- Zhang D, Pan X, Li S *et al.* 2017. Impact of the National Essential Public Health Services Policy on Hypertension Control in China. *Am J Hypertens*, 31 1: 115-23.
- Zhang L, Wang Z, Qian D *et al.* 2014a. Effects of changes in health insurance reimbursement level on outpatient service utilization of rural diabetics: evidence from Jiangsu Province, China. *BMC Health Serv Res*, 14: 185.
- Zhang X, Wu Q, Liu G *et al.* 2014b. The effect of the National Essential Medicines Policy on health expenditures and service delivery in Chinese township health centres: evidence from a longitudinal study. *BMJ Open*, 4 12: e006471.
- Zhou Z, Zhao Y, Shen C *et al.* 2021. Evaluating the effect of hierarchical medical system on health seeking behavior: A difference-in-differences analysis in China. *Soc Sci Med*, 268: 113372.
- Zhu Z, Zhao M, Sun G *et al.* 2017. Evaluation of Interventional Effects of Family Doctor Services on Community Residents in Beijing. *Chinese General Practice*, 20 13: 1549-53.



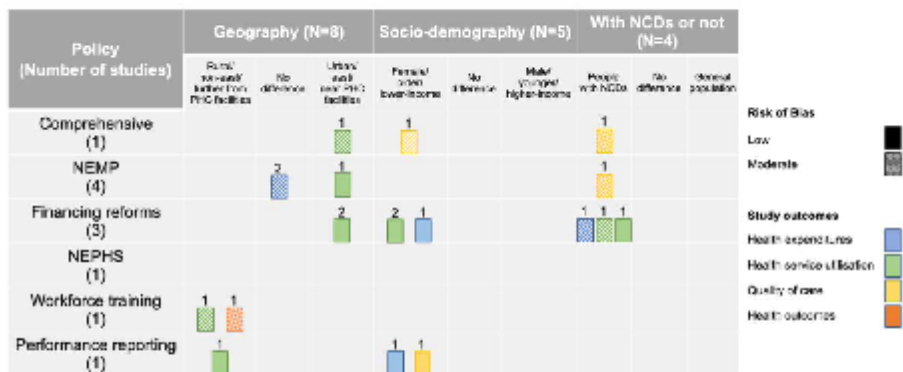
Notes: Small-scale studies were defined as studies that assessed Interventions Implemented within community or village level.

Figure 1 - PRISMA diagram for study selection process



Notes: 1. NEMP for the National Essential Medicine Policy; 2. NEPHS for the National Essential Public Health Services. 3. n for the number of studies.

Figure 2 - Language, publication year, geographical distribution and risk of bias of the 42 studies over time grouped by policies



Notes: 1. The matrix has six rows (one for each PHC policy) and three columns (one for each type of heterogeneity) 2. A bar in the cell indicates that the included study found greater policy effects among the corresponding population groups (defined by the columns) compared to their counterfactuals. The bar in the columns of "No difference" indicates that no differential policy effects were found; 3. The number annotated on the top of each of the bars indicates the number of studies; 4. NEMP for the National Essential Medicine Policy; 5. NEPHS for the National Essential Public Health Services; 6. NCD for non-communicable disease.

Figure 3 - Heterogeneous effects of China's PHC reforms

The first phase of the PHC reform (2009 - 2014)	
Policy	Content
Infrastructure strengthening (The State Council of the People's Republic of China, 2009b, 2009a, 2011) (March 2009)	<ul style="list-style-type: none"> • Provided new public funds to build new PHC facilities and refurbish and re-equip existing ones; • Introduced new governance models for PHC facilities, including government-regulated PHC facilities, whose funding and revenues are directly managed by local governments, and hospital-affiliated PHC facilities, whose funding and revenues are directly managed by public hospitals; • Established a GP system with standardized training programs for certificated/licensed GPs; • Provided subsidies and free training for medical students who are trained as GPs.
Performance reporting (Ministry of Health, 2010; The State Council of the People's Republic of China, 2009a) (June 2010)	<ul style="list-style-type: none"> • Introduced a list of performance indicators for PHC facilities, including antibiotic use, injection prescribing rates, and average expenditures for patients; • Required PHC facilities to record performance indicators in monthly reports and make monthly reports openly accessible to the public.
Financing reforms (July 2009)	<p>Health insurance reforms(Ministry of Health et al., 2009; Ministry of Human Resources and Social Security et al., 2009)</p> <ul style="list-style-type: none"> • Increased per capita subsidies for enrollees of basic health insurance to raise enrolment rates (from 120 CNY in 2010 to 610 CNY in 2022); • Funds for basic health insurance were pooled and managed at city level; • Expanded coverage of benefit packages in basic health insurance from hospitalisation to primary care services; • Reduced co-payment (to about 50%) for costs of outpatient care and medication for people with NCDs (with a cap).
	<p>Provider payment reforms(Ministry of Health et al., 2009; Ministry of Human Resources and Social Security et al., 2009)</p> <ul style="list-style-type: none"> • Introduced an evaluation scheme for the performance of PHC facilities, which affects the overall salary of the personnel in each facility. • Combined or replaced fee-for-service with P4P or capitation for the delivery of primary care services.
The NEMP (Ministry of Health, 2009; Ministry of Health et al., 2013) (August 2009)	<ul style="list-style-type: none"> • Created a National Essential Medicine List available at all PHC facilities; • Increased the reimbursement rate for all essential medicines; • Mandated unified medicine procurement and distribution at provincial level, and engaged local governments in price negotiation, drug bidding, procurement and quality assurance; • Removed the mark-up of essential medicines as revenue source.
The NEPHS (Ministry of Finance and National Population and Family Planning Commission, 2009) (July 2009)	<ul style="list-style-type: none"> • Provided the population with a free-of-charge, defined package of essential health services, including vaccination, health examinations, screening, health management, follow-up visits, prescription and health education (with a priority for people over 65-year-old, with NCDs, 0-6 years old children, and pregnant women); • Provided extra subsidies to PHC facilities for the delivery of the package based on per capita allocation; • Introduced individual electronic health records for all citizens at all PHC facilities.

(Continued) The second phase of the PHC reform (2015 - now)

Policy	Content
---------------	----------------

<p>Referral system Restructuring (The State Council of the People's Republic of China, 2015) (September 2015)</p>	<ul style="list-style-type: none"> • Introduced a three-tier health system, including PHC facilities, secondary hospitals, and tertiary hospitals, and defined the responsibilities for each level of health facilities; • Provided higher reimbursement rates for visits to PHC facilities and referrals from PHC facilities and reduced reimbursement rates or no reimbursement for direct visits to hospitals; • Introduced a new gatekeeping policy that requires referral letters from PHC facilities for hospital visits. This policy was not mandatory and implemented according to the local context;
<p>An integrated delivery system (The State Council of the People's Republic of China, 2017) (April 2017)</p>	<ul style="list-style-type: none"> • Established a standardized, dual referral system between PHC facilities and hospitals to integrate preventive-curative services; • Built provider networks to share health resources (e.g., equipment and workforce) and health technology across facilities at different levels; • Established an online platform to share electronic health records across facilities, as well as other health information.
<p>Family physician care model (The State Council of the People's Republic of China et al., 2016) (May 2016)</p>	<ul style="list-style-type: none"> • Introduced a new care model which featured family physician teams led by GPs or village doctors from PHC facilities, delivering primary care services at community level; • Provided subsidies to family physician teams for services delivered based on capitation; • Defined a list of basic health services delivered by family physician teams, including health education and examinations, prescription consulting, home visits, and personalised care based on patients' needs; • Introduced benefits to patients who are referred by family physicians, including priority to referrals and reservations and higher reimbursement rates for hospital visits.
<p>PHC workforce training (National Health and Family Planning Commission and National Administration of Traditional Chinese Medicine, 2015; The State Council of the People's Republic of China, 2015) (November 2015)</p>	<ul style="list-style-type: none"> • Provided subsidized job training programs for workforce at PHC facilities. • The provided training excluded medical certification/licenses and focused on medical treatment skills, nursing skills, using health information technology, leadership and teamwork.

Notes: 1. PHC for primary health care; 2. NEMP for National Essential Medicine Policy; 3. NEPHS for National Essential Public Health Services. 4. GP for general practitioner. 5. NCDs for non-communicable diseases; 6. P4P for pay-for-performance.

Table 1 - A summary for major China's primary health care policies since 2009 Reference

- Ministry of Finance, National Population and Family Planning Commission. *Opinions on Promoting the Gradual Equalisation of Essential Public Health Services*. http://www.gov.cn/ztl/ygzt/content_1661065.htm, accessed 05 May 2022.
- Ministry of Health *Notice of the Opinions on the Establishment of the National Essential Medicine Scheme*. http://www.gov.cn/ztl/ygzt/content_1661112.htm accessed 05 May 2022.
- Ministry of Health. *Regulations on Open Information on Medical Institutions*. <http://www.nhc.gov.cn/wjw/zcyj/201310/42a2c4febe8f4381b106dfd1214a9c0f.shtml>, accessed 05 May 2022.
- Ministry of Health, Ministry of Civil Affairs, Ministry of Finance, Ministry of Agriculture, National Administration of Traditional Chinese Medicine *Opinions on Strengthening and Developing the New Rural Cooperative Medical Insurance Scheme*. <http://www.nhc.gov.cn/jws/s6477/200907/e9b169b0ec274051b45fc646d336814f.shtml>, accessed 05 May 2022.
- Ministry of Health, National Development and Reform Commission, Ministry of Industry and Information Technology, Ministry of Human Resources and Social Security, Ministry of Commerce, National Administration of Traditional Chinese Medicine. *Opinions on strengthening and developing the National Essential Medicine Scheme and the New Operation Mechanism at Primary Level*. http://www.gov.cn/zwgk/2013-02/20/content_2335737.htm, accessed 05 May 2022.
- Ministry of Human Resources and Social Security, Ministry of Finance, Ministry of Health. *Guidance on Pooling Funds for Outpatient Care Services Under the Basic Medical Insurance for Urban Residents*. http://www.gov.cn/zwgk/2009-08/05/content_1384111.htm, accessed 05 May 2022.
- National Health and Family Planning Commission, National Administration of Traditional Chinese Medicine. *Guiding Opinions on Further Regulating the Management and Improving the Quality of Community Health Services*. <http://www.nhc.gov.cn/jws/s3581r/201511/1742007746a64005a16e32de00cc5fc5.shtml>, accessed 05 May 2022.
- The State Council of the People's Republic of China. *Opinions on Deepening Health System Reform in China*. http://www.china.org.cn/government/scio-press-conferences/2009-04/09/content_17575378.htm accessed 16 April 2022.
- The State Council of the People's Republic of China. *Healthcare System Reform Implementation Plan with Priority (2009 to 2011)*. http://www.gov.cn/zwgk/2009-04/07/content_1279256.htm, accessed 05 May 2022.
- The State Council of the People's Republic of China. *Guidance on the Establishment of General Practitioner Scheme*. http://www.gov.cn/zwgk/2011-07/07/content_1901099.htm, accessed 05 May 2022.
- The State Council of the People's Republic of China. *Guiding Opinions on Building a Hierarchical Health System*. http://www.gov.cn/zhengce/content/2015-09/11/content_10158.htm, accessed 05 May 2022.
- The State Council of the People's Republic of China. *Guiding Opinions on the Establishment and Development of Medical Alliance*. http://www.gov.cn/zhengce/content/2017-04/26/content_5189071.htm, accessed 05 May 2022.
- The State Council of the People's Republic of China, National Health and Family Planning Commission, National Development and Reform Commission,

Ministry of Civil Affairs, Ministry of Finance, Ministry of Human Resources and Social Security, National Administration of Traditional Chinese Medicine.

Guiding Opinions on Building the Family Physician Scheme

http://www.mohrss.gov.cn/SYrlzyhshbzb/shehuibaozhang/zcwj/yiliao/201606/t20160615_241854.html, accessed 05 May 2022.

ACCEPTED MANUSCRIPT

Characteristics of the included studies		Number of Studies
Language	Mandarin	11
	English	31
Year of publication	2012-2015	17
	2016-2021	25
Geographical region	Central	8
	Eastern	18
	Western	5
	National	11
Study Duration	Two years or less	13
	3-5 years	23
	More than 5 years	6
Data source	Primary Data	7
	Secondary Data	35
Data types	Cohort	8
	Panel	34
RCD data	No	12
	Yes	30
People with NCDs	No	31
	Yes	11
Urban/rural population	Rural	18
	Urban	9
	Both	15
Unit of analysis	Individual	22
	Aggregated	18
	Both	2
Evaluated Policies	Comprehensive	5
	NEMP	15
	Financing	11
	System Integration	4
	Gatekeeping	1
	Family Physician	3
	NEPHS	1
	Workforce training	1
Performance reporting	1	
Study outcomes	Health expenditures	26
	Health service utilisation	25
	Quality of care	13
	Health outcomes	10
	Health equity	1
Study quality	Low risk of bias	9
	Moderate risk of bias	24
	Serious risk of bias	9

Notes: 1. NEMP for the National Essential Medicine Policy; 2. NEPHS for the National Essential Public Health Services. 3. RCD for routinely collected data. 4. NCD for non-communicable disease.

Table 2 - Characteristics of the 42 included studies

Policy (Number of studies)	Health expenditures			Health service utilisation			Quality of care		Health outcomes		Health inequality
	Health services costs	Drug costs	O O P E	Visits to PHC facilities	Public health services	Hospitalisations	Antibiotic use	Satisfaction/Quality rating	Biomarkers/Mortality	SRH/HRQoL	Income-related
Comprehensive (3)		▼		⊖	▲	▲	▼	▲			▼
The NEMP (13)	▼▼▼▼ ⊖⊖⊖⊖	▼▼▼▼ ⊖⊖⊖⊖	⊖	▲▲▲ ⊖⊖⊖	▲	▲▲	▼ ⊖⊖ ▲				
Financing reforms (11)	▲▲▲▲ ⊖⊖ (NCD)	⊖	▼ ⊖ (NCD)	▲▲▲ ▲▲▲ ▲▲▲ ▼	▲	▼▼▼▼ (NCD) ⊖⊖ ⊖ (NCD)	▼ (NCD) ⊖		▼▼▼▼ (NCD)	▲ (NCD)	
Gatekeeping (1)	⊖			▲							
Integrated delivery system (4)	▼ (NCD) ⊖			▲ (NCD) ▲	⊖			⊖	▼▼▼▼ (NCD) ⊖	▲ (NCD)	
Family Physicians (3)	▼			▲				▲ (NCD) ▼		▲ (NCD) ⊖	
The NEPHS (1)					▲ (NCD)				▼ (NCD)		
Workforce training (1)				▲							
Performance reporting (1)		▼					▼				

Notes: 1. Each symbol represents one study on the corresponding assessed policies and outcomes, and studies assessing multiple outcomes have more than one symbol. 2. NEMP for National Essential Medicine Policy; 3. NEPHS for National Essential Public Health Services; 4. OOPE for individual out-of-pocket expenditures; 5. NCD in brackets indicates that the study population were people with NCDs; 6. ▲ for a statistically significant increase; ▼ for a statistically significant decrease; ⊖ for no statistically significant changes. 7. Red indicates the study was rated as serious risk of bias; yellow as moderate risk of bias; green as low risk of bias. 8. Studies that only compared with alternative interventions or did not report overall impacts were not included in Table 3. 9. Biomarkers including indicators from blood pressure and blood sugar tests, and control rates. A decrease symbol in biomarkers indicates better hypertension/diabetes control. 10. SRH for self-reported health, and HRQoL for health-related quality of life.

Table 3 - Estimated effects and certainty of evidence grouped by evaluated policies and outcomes