Health policy 134 (2023) 104576

Contents lists available at ScienceDirect

Health policy

journal homepage: www.elsevier.com/locate/healthpol

A critical review of methodologies used in pharmaceutical pricing policy analyses

Iris R. Joosse[a], David Tordrup[a], Lisa Bero[b], Aukje K. Mantel-Teeuwisse[a], Hendrika A. van den Ham[a,*]

[a] Utrecht WHO Collaborating Center for Pharmaceutical Policy and Regulation, Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Universiteitsweg 99, CG, Utrecht 3584, the Netherland
[b] Center for Bioethics and Humanities, University of Colorado Anschutz Medical Campus, Colorado, United States

ARTICLE INFO

Article history:
Received 31 May 2021
Revised 28 September 2021
Accepted 9 March 2022

Keywords:
Study designs
Pricing policies
Policy analysis
Interrupted time series
Risk of bias
Methodological issues

ABSTRACT

Robust evidence from health policy research has the potential to inform policy-making, but studies have suggested that methodological shortcomings are abundant. We aimed to identify common methodological weaknesses in pharmaceutical pricing policy analyses. A systematic review (SR) of studies examining pharmaceutical pricing policies served as basis for the present analysis. We selected all studies that were included in the SR ($n = 56$), and those that were excluded from the SR due to ineligible study designs only ($n = 101$). Risk of bias was assessed and specific study design issues were recorded to identify recurrent methodological issues. Sixty-one percent of studies with a study design eligible for the SR presented with a high risk of bias in at least one domain. Potential interference of co-interventions was a source of possible bias in 53% of interrupted time series studies. Failing to consider potential confounders was the primary cause for potential bias in difference-in-differences, regression, and panel data analyses. In 101 studies with a study design not eligible for the SR, 32% were uncontrolled before-after studies and 23% were studies without pre-intervention data. Some of the methodological issues encountered may be resolved during the design of a study. Awareness among researchers on methodological issues will help improve the rigor of health policy research in general.

1. Introduction

Evidence from health policy research has the potential to be translated into effective and appropriate strategies, policies and interventions [1,2]. Indeed, health policy research is considered essential in advancing health systems' performances with the ambition of achieving universal health coverage (UHC) and the health-related Sustainable Development Goals (SDG) [2]. Since the 1990s, increased importance has been placed on evidence-based healthcare policy making [3,4].

Randomized controlled trials (RCTs) remain the gold standard in clinical research for generating robust evidence with a considerable certainty [5]. However, waiting for this same level of certainty in generating evidence on healthcare policies would paralyze the policy-making process [6,7]. Particularly as conducting RCTs in policy research may be unfeasible or even undesirable [8,9]. To establish a measure of effect in health policy research that is both unbiased and feasible to produce, certainty of evidence and pragmatism need to be balanced [7]. This is by no means straightforward.

Recognizing this problem, the World Health Organization (WHO) published the *Health Policy and Systems Development: An Agenda for Research* in 1996 [10]. With this technical document the WHO provided researchers with guidance on identifying general research approaches that are potentially appropriate in studying health policies. This document states that, to achieve real advancement in the field of health policy research, policy assessments should move towards measuring the direct and indirect effects of policies on prespecified outcomes. It is emphasized that there is a need for both qualitative and quantitative policy assessments. The need for robust evidence on health policies was further stressed in the *WHO Handbook for Guideline Development* in 2012. This Handbook states that systematic reviews used to inform WHO guidelines are to be developed according to the standards outlined by the Cochrane Collaboration [11]. With that, only RCTs or observational study designs associated with a low risk of bias should qualify for systematic reviews used to support WHO guidelines, beside additional qualitative evidence.

* Corresponding author.
E-mail address: h.a.vandenham@uu.nl (H.A. van den Ham).

https://doi.org/10.1016/j.healthpol.2022.03.003

However, little explicit guidance exists on what research designs or methods best inform quantitative health policy analyses and how to perform these [12]. Instead, mostly general recommendations have been presented in the literature over the years [13]. For one, multiple types of outcomes should be adopted in health policy analyses including both unintended and unexpected consequences of a policy intervention [14]. To facilitate the identification of such consequences, the study should encompass a sufficiently long time span [12]. Additionally, an appropriate comparator or counterfactual is necessary to interpret the results [14]. Finally, a comprehensive and well-specified description of the intervention and the contextual factors is required as it may help to explain the success or failure of an intervention [6,14]. Although these general recommendations provide some direction, more concrete guidance is lacking.

The present study was inspired by the experiences from an extensive systematic review (SR) of studies evaluating ten pharmaceutical pricing policies, used for the development of the 2020 WHO Guideline on Country Pharmaceutical Pricing Policies [15]. We observed that a large proportion of studies were excluded during the review process due to ineligible study designs [16]. Additionally, we noted that many of the studies that did meet the eligibility criteria had methodological shortcomings. The frequent use of biased or weak study designs in pharmaceutical policy analyses has previously been reported in systematic reviews, each expressing that some of the shortcomings in study design may be preventable [17,18]. Insight into common weaknesses can provide concrete starting points for improving methodologies used in pharmaceutical pricing policy analyses specifically and health policy research in general. Accordingly, we aimed to identify some of the gaps and methodological weaknesses in pharmaceutical pricing policy analyses.

2. Methods

We conducted an extensive SR in 2019 that served as the basis for the present study [16]. The SR focused on the effects of ten pharmaceutical pricing policies, with the aim to identify which policies are effective in managing pharmaceutical prices. For the present study, search results of the SR at the full-text level were our primary source of data. Studies were selected if they (1) had been included in the original SR or (2) had been assessed for eligibility on full-text level but had subsequently been excluded from the SR due to an ineligible study design (but met all inclusion criteria otherwise). Studies were excluded from this analysis if there were other reasons for exclusion from the SR, such as an ineligible intervention, ineligible outcomes or an unsuitable publication type.

2.1. Description of data source

The original SR was undertaken according to the principles of systematic reviewing embodied in the Cochrane Handbook and guidance document published by the Center for Reviews and Dissemination (CRD) [19,20]. A literature search was performed in a number of databases, including but not limited to Ovid MEDLINE, Ovid Embase, Social Science Citation Index, EconLit, and NHS Economic Evaluation Database. Database searches were supplemented by grey literature searches and the reference lists of relevant articles were searched manually.

Studies published after 1 January 2004 and up to October 2019 were eligible for inclusion. Eligible interventions were:

1. Cost-plus pricing
2. Policies promoting the use of generic and biosimilar medicines
3. Policies regulating mark-ups across the pharmaceutical supply and distribution chain
4. Pooled procurement
5. Price discounts for single source pharmaceuticals
6. (External and internal) reference pricing
7. Tax exemptions or tax reductions for pharmaceuticals
8. Tendering and negotiation
9. Policies promoting price transparency
10. Value-based pricing

Studies were eligible if they included at least one of the following outcomes: price (or expenditure as a proxy), volume, availability or affordability of pharmaceuticals. Studies that compared interventions to at least one comparator or counterfactual and that included pre-intervention data were eligible for inclusion in the SR. Eligible study designs were: randomized trials and non-randomized or quasi-experimental designs, e.g. controlled before-after studies, difference-in-differences (DID), interrupted time series (ITS), non-randomized controlled trials (nRCT), and repeated measures (RM). As the study label did not always represent the actual study design, studies were classified according to the features of a study's design rather than the label mentioned in the paper by the authors. Beside study types primarily aiming to prevent confounding at the design level, studies using techniques intended to correct for confounding during analysis (e.g. regression analyses, panel data analyses) were also eligible. All eligible study types either included a direct control or were able to correct for absence of a control, increasing the certainty of the evidence. Definitions and categorizations of study designs and analysis techniques as applied in the SR are shown in Table 1. Searches were conducted without language restriction.

Two reviewers independently assessed study eligibility and possible types of bias based on risk bias criteria as suggested by Cochrane Effective Practice and Organization of Care (EPOC) [21]. The quality of the evidence was assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) methodology. We did a narrative summary of the evidence describing the relationship between studies and patterns discerned in the data. The methodology and detailed search strategies have been published elsewhere [22].

2.2. Data extraction and analysis

For the present study the following information was extracted for all studies: inclusion status in the SR, type of publication, WHO region of study location, and type of intervention (according to one of ten pharmaceutical pricing policies as used in the SR). The year of publication and income setting of the study location (as designated by the World Bank for 2019–20) were extracted to test the hypothesis that there could be a relationship between these features and the studies with an ineligible study design.

For studies that had been included in the SR information on study design features and risk of bias as extracted for the SR was used. Risk of bias assessment criteria as suggested by EPOC had been adapted to study design (randomized trials, non-randomized trials and controlled before-after studies were assessed on nine criteria; ITS and RM studies were assessed on eight criteria; and a set of four assessment criteria applied to all other study types; Table S1 in appendix). For articles not included in the SR the specific design issues were recorded. The risk of bias was not assessed for this group.

We used descriptive statistics to identify recurrent methodological issues per type of study design. Some examples from the SR were selected to illustrate the issues encountered.

3. Results

We identified 32,011 publications in our initial literature search for the SR. After removal of the obviously irrelevant records we

Table 1
Definitions of study designs and analysis techniques as applied in the systematic review.

Design strategies	Characteristics	Results of analysis	Examples*
		Randomized designs	
Randomized trial	An experimental study in which subjects are allocated to different interventions using methods that are random [19].	Generally represented by an absolute or relative difference compared to the pre-intervention time period, usually corrected for the effect in the control group	Bhargava et al. [55]
		Non-randomized and quasi-experimental designs	
Controlled before-after	An observational study that uses observations from few time points (<3) before and after the implementation of an intervention, both in a group that receives the intervention and in a control group that does not [19].	Generally represented by an absolute or relative difference compared to the pre-intervention time period, usually corrected for the effect in the control group.	Adesina et al. [51]
Difference-in-Differences	Quasi-experimental design and analysis technique in which observations are made at multiple time points before and after an intervention, both in a group that receives the intervention and in a control group that does not.	An estimate for a regression coefficient, that signifies a difference in changes over time between the intervention and control group. May also be presented as a percentage difference.	Ghislandi et al. [52]
Interrupted Time Series	A quasi-experimental study designs that uses observations at multiple time points (≥3) before and after an intervention (the 'interruption'). The design attempts to detect whether the intervention has had an effect greater than any underlying trend over time [19].	Estimates for regression coefficients corresponding to two effects: a change in level (the difference between the observed level at the first post-intervention time point and that predicted by the pre-intervention time trend) and a change in trend (the difference between pre- and post-intervention slopes) before and after the intervention.	Yoo et al. [53]
Non-randomized trial†	An experimental study in which subjects are allocated to different interventions using methods that are not random [19].	Generally represented by an absolute or relative difference compared to the pre-intervention time period, usually corrected for the effect in the control group.	NA
Repeated Measures	An interrupted time series design where measurements are made in the same individuals at each time point [19].	Dependent upon analysis method, the results may be represented as estimates for regression coefficients (e.g. regression methods) or F-ratios (repeated measures ANOVA).	Ben-Aharon et al. [36]
		Analytic strategies	
		Conventional confounding correction methods	
Regression analysis	An analysis technique that examines the influence of one or more independent variables on a dependent variable. In (health) policy analysis, the technique usually uses longitudinal data.	Estimates for regression coefficients, representing the isolated effect of each independent variable on the dependent variable.	Kaiser et al. [56]
		Confounding correction methods for multidimensional data	
Panel data analysis	Analysis technique that uses multidimensional data (cross-sectional time-series data) and allows for variation along individual and time dimensions†.	Estimates for regression coefficients, representing the isolated effect of each independent variable on the dependent variable.	Von der Schulenburg et al. [54]

*Examples as encountered in the systematic review. Please note that these examples are not necessarily free of bias, see systematic review for details [16]. Non-randomized trials were not encountered in the systematic review. † If there are only two groups and two measurements, this model is equivalent to the difference-in-differences design.

assessed 1000 records for eligibility at full-text level [16]. Only 56 studies were deemed eligible at the time, meeting all requirements including an eligible study design (hereafter called 'eligible study designs group'). Important reasons for exclusion from the review were study design issues (n = 316), ineligible interventions (n = 241), ineligible outcomes (n = 181), and insufficient data (n = 161). Upon re-inspection of the 316 records that had been excluded for design issues, 215 studies were also ineligible for reasons other than design issues. This left 101 studies for the current analysis (hereafter called 'ineligible study designs group'). With that, 157 studies fulfilled the inclusion criteria of the present study (Fig. 1). The general characteristics of the studies are shown in Table 2.

A total of 144 (92**%**) of 157 included studies were published as an original research article and 102 (65%) were published in the past eight years (2012–19) (Table 2). Additionally, most studies were conducted in European countries (39**%**), followed by a fair number of studies originating in the WHO region for the Americas (15**%**, mainly the United States) and the WHO region for the Western Pacific (22**%**, mainly the China). This distribution is also reflected in the income setting with very little evidence from low-income countries (<1**%**). Internal reference pricing was the most researched pricing policy (24%), whilst cost-plus pricing and tax reductions were not the subject of any studies. Overall, the distribution of both groups of studies follows a similar pattern for all characteristics.

3.1. Studies with an eligible study design

When focusing on the eligible study designs group only, these applied different designs with the majority being ITS (n = 17) and DID studies (n = 13) (Table 3). There is no apparent association between the type of intervention and the study designs that are used to study them (Fig. S1 in appendix). Most studies reported on multiple types of outcomes related to drug pricing and expenditure but none reported on the outcomes availability and affordability. Information on contextual factors or on the implementation of the intervention, both of which could help explain the failure or success of an intervention, was provided in 57% and 45% of the eligible studies, respectively. Thirty-four (61%) studies scored high

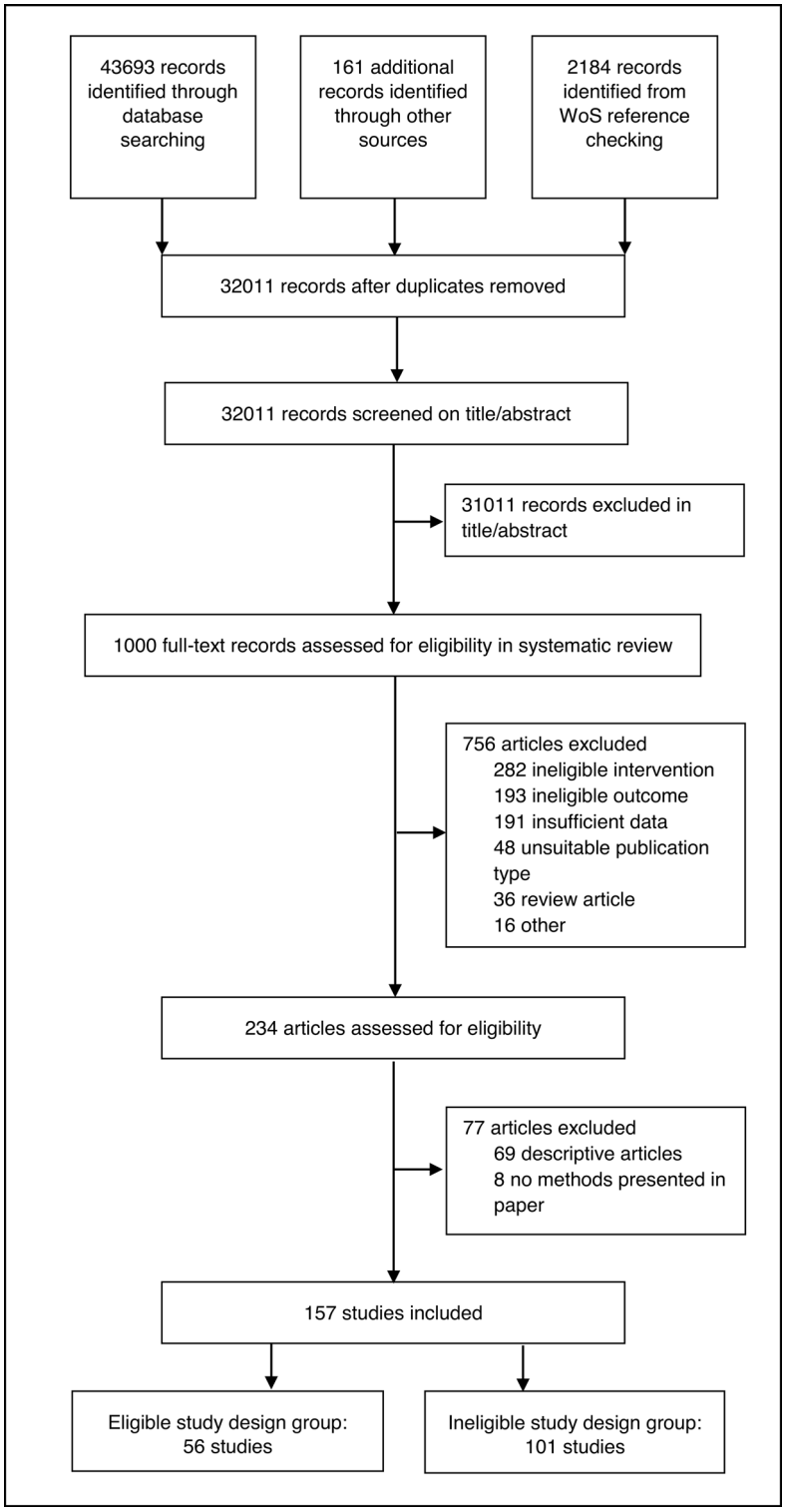

Fig. 1. Flow chart of study selection. WoS=Web of science.

risk of bias in at least one domain (Table 3). Only four studies (7%) were associated with a low risk of bias across all domains. Notably, 86% of the studies were considered to have an unclear risk of bias in the domain 'incomplete outcome data' (Fig. 2a).

Fifty-three percent (53%, $n = 17$) of ITS studies were associated with a risk of bias due to potential interference from co-interventions (see Fig. 2b), the effects of which could not always be discerned from the intervention of interest. Although the short time between successive interventions was often acknowledged but did not allow for a satisfactory separate analysis [23–27], several other studies disregarded the influence of co-interventions completely within their analysis. An example is the study by Kwon et al. [28] The authors noted the possible impact of two co-interventions that were implemented 17 and 21 months after the main intervention of equal medicine pricing (EMP), but did not introduce these as separate segments in their regression analysis.

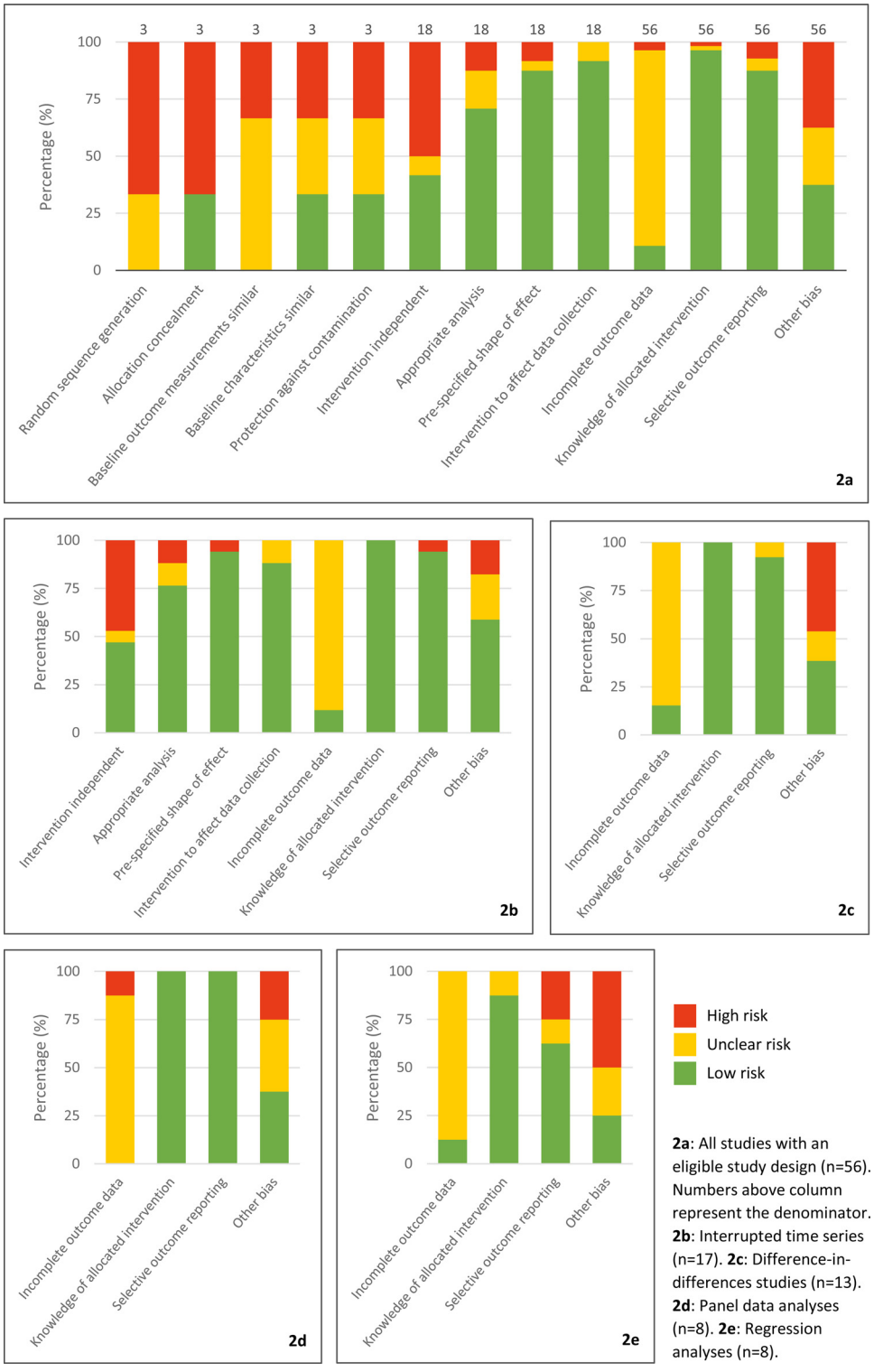

Fig. 2. The risk of bias of studies in the eligible study design group.

Additionally, a third co-intervention 18 months before the intervention was not mentioned in this publication but was described in another study examining the EMP [29]. The immediate effect of the intervention (presented as the change of intercept in an ITS analysis) and the long-term effects (presented as a change in slope) may thus have been influenced by co-interventions. In contrast, in the study by Langley et al. [30] the effects of the introduction of new treatment guidelines 6 months after the implementation of a transparency measure was separated using a different segment.

Sixty-six percent (66%, $n = 29$) of DID studies, regression analyses and panel data analyses were associated with a high or unclear risk of bias in the domain 'other bias' (Fig. 2c–e). In the majority of cases the lack of relevant confounding factors in the empirical model resulted in this assessment. In some of these studies the authors described several factors as potential confounders, but

Table 2
General characteristics of included studies.

	Eligible study designs	Ineligible study designs
***n* (%)**	56	101
Year of publication		
2004–2007	7 (13)	12 (12)
2008–2011	9 (16)	27 (27)
2012–2015	21 (38)	36 (36)
2016–2019*	19 (34)	26 (26)
WHO Region		
Africa	2 (4)	5 (5)
Americas	9 (16)	14 (14)
South-East Asia	2 (4)	4 (4)
Europe	23 (41)	38 (38)
Eastern Mediterranean	0 (0)	6 (6)
Western Pacific	15 (27)	20 (20)
Global	5 (9)	14 (14)
Setting		
Low-income	0 (0)	1 (1)
Lower-middle income	2 (4)	7 (7)
Upper-middle income	11 (20)	28 (28)
High-income	39 (70)	49 (49)
Multiple income settings	4 (7)	16 (16)
Publication type		
Dissertation	0 (0)	1 (1)
Guidelines	0 (0)	1 (1)
Original research article	49 (88)	95 (94)
Report	7 (13)	3 (3)
Other	0 (0)	1 (1)
Type of intervention		
Cost-plus pricing	0 (0)	0 (0)
Promoted use of generic and biosimilar medicines	10 (18)	15 (15)
Price setting and mark-up thresholds	10 (18)	15 (15)
Pooled procurement	6 (11)	12 (12)
Price discounts for single source pharmaceuticals	0 (0)	0 (0)
External reference pricing	0 (0)	6 (6)
Internal reference pricing	18 (32)	19 (19)
Tax exemptions or tax reductions	0 (0)	0 (0)
Tendering and negotiation	1 (2)	12 (12)
Policies promoting price transparency	3 (5)	2 (2)
Value-based pricing	2 (4)	2 (2)
Multiple interventions	6 (11)	18 (18)

*Data was included until October 2019.

Table 3
Additional characteristics for studies with an eligible study design.

	Eligible study design group (n = 56)
Study design	
Design strategies	
Controlled before-after	2 (4)
Difference-in-differences	13 (23)
Interrupted time series	17 (30)
Randomised trial	1 (2)
Repeated measures	1 (2)
Other	6 (11)
Analytic strategies	
Regression analysis	8 (14)
Panel data analysis	8 (14)
Information on contextual factors provided	32 (57)
Information on implementation method provided	25 (45)
Risk of bias	
No domains with a risk of bias	4 (7)
≥1 domain with a high risk of bias	34 (61)

Table 4
Issues with studies with an ineligible study design (n = 101).

Ineligible designs and design issues	n (%)
Ineligible study designs	
Cross-sectional studies	12 (12)
Descriptive study without statistical analysis	17 (17)
Theoretical study	16 (16)
Uncontrolled before-after study	32 (32)
Design issue	
Lack of pre-intervention data	23 (23)
Other design issue	1 (1)

were unable to control for these elements because the data was unavailable to them [31–34]. In other studies potential explanatory factors did not seem to have been considered at all. For example, in the study by Wu et al. [35] the characteristics of the medicines included in the study were not described. Because interventions are often specific for certain products, factors such as formulation and pack size may have influenced the results and could have been taken into account.

A problem observed across study designs is related to the timing of the intervention. As time is an important co-variate in all longitudinal policy analyses in which one expects to see changes over time, an exact definition of the timing and the correct analysis thereof is crucial. However, in several studies the exact timing of the intervention was either not described [36] or difficult to establish [37–39]. In two other studies the authors did not apply the point of intervention as point of analysis [23,40]. To illustrate, in an ITS study by Hsiao et al. [40] the quarter in which the intervention occurred (Jan–Mar 2003) was regarded in the analysis as 'pre-intervention' even though the policy was implemented in March 2003. This makes interpretation of immediate changes in usage patterns difficult.

Instead, authors should consider the use of a phase-in period, also when the implementation of an intervention has been gradual or when there may have been an anticipatory response to implementation of a policy. To allow for this possibility, Leopold et al. [26] considered a four-month transition period prior to implementation of the policy and excluded these data points from analysis.

3.2. Studies with an ineligible study design

Studies that were ineligible for the original SR due to design issues can be subdivided in roughly six categories, among which four ineligible study designs and two design issues: (1) cross-sectional study, (2) descriptive study without statistical analysis, (3) theoretical study, (4) uncontrolled before-after study, (5) lack of pre-intervention data, (6) other design issue (see Table 4). Although a study design that is prone to bias, uncontrolled before-after studies were nonetheless abundant (32 of 101 ineligible studies). This design's sensitivity to bias becomes clear in the study by Law et al. [41] in which quarterly data from 2010 was used to estimate the potential savings of a policy reducing generic drug prices. Data from quarters 1 and 2 was used to derive a counterfactual, which was then compared to the observed data from quarters 3 and 4. An analysis based on so few datapoints risks seasonal variation or randomly deviating datapoints being incorporated and leading to biased conclusions. This is likely the case in this example: an aberrant datapoint in quarter 2 resulted in an upward counterfactual trend pre-policy that was not prolonged in the observed data post-policy. The inclusion of either a control group or more timepoints before and after the intervention would allow correction for random or seasonal variations.

Studies without pre-intervention data were also encountered frequently (23 of 101 ineligible studies). Without a pre-policy baseline the effectiveness of an intervention cannot be determined, as any trend could be pre-existing and not due to the intervention. To illustrate, Adriaen et al. [42] aimed to examine pricing strategies in Belgium, including internal reference pricing. The Belgian reference-pricing system itself was introduced in June 2001, yet

data was collected beginning at July 2001. With that, the authors were able to report on factors that influence pricing strategies, but not on the effectiveness of the pricing strategy. Inclusion of data from before the intervention would have allowed for this.

No apparent association between the design issues and the interventions of interest was found (Fig. S2 in appendix).

4. Discussion

We found that methodological weaknesses in pharmaceutical pricing policy analyses were multifold. Our results show that three out of five studies that met the eligibility criteria of the original SR were associated with a high risk of bias in at least one domain. In ITS studies this was predominantly due to the potential effects of co-occurring interventions. In DID studies, regression analyses, and panel data analyses, the failure to account for potential confounders often resulted in a high or unclear risk of bias. Establishing an exact timing of a policy intervention was problematic across all study designs, and information on contextual factors or implementation methods of the policy was often limited. Finally, a large absolute number of studies were excluded from the original SR for study design issues alone.

The large proportion of studies ineligible due to design issues suggests that there is a mismatch between the type of evidence generated by researchers and that required to make evidence-informed decisions. It is worth noting, however, that research evidence is not the only input that is considered in policy-making. Other components such as politics, social culture, financial concerns and timing impact policy decisions as well, as suggested by the term 'evidence-informed' rather than 'evidence-based' policy-making [43]. Understanding the motives and perspectives of researchers on one hand and policy-makers on the other may be an important step in aligning evidence generation with policy-making in practice. Nonetheless, the prevalent use of study designs that are highly vulnerable to bias, and the limited attention in scientific research to specific pricing policy topics creates an important evidence gap. The building of an encyclopedia to map evidence and impacts has been proposed as a way to identify such gaps, with the ultimate goal of enhancing current efforts and furthering future research [44].

Some of the methodological issues that we encountered could probably be resolved without much difficulty. For instance, the analysis of co-interventions as separate segments in ITS should be considered when the time of implementation of co-interventions is known. Likewise, careful selection of potential confounders in empirical models could markedly reduce the risk of confounding bias in pharmaceutical pricing policy analyses. And although not considered a methodological flaw, the reporting of contextual information on interventions can often be improved to facilitate interpretation of results. Lastly, we acknowledge that some of the issues may be borne from a lack of data and require a more fundamental solution. Both researchers and policy-makers could play an important role in collecting the required data for adequate monitoring of implemented policies.

Similarly, some methodologies with high associated risks of bias that were not eligible for inclusion in the SR can be modified in such a way to make them more rigorous. To illustrate, uncontrolled before-after studies are very sensitive to bias, because the number of datapoints before and after the intervention is insufficient to distinguish an effect that is different from random variations or a preexisting trend [45,46]. The addition of another intervention group and multiple control groups could tackle this flaw [21,46]. However, including a suitable control is complicated by variations in health systems, disease burden or demographics that may result in different effects in different countries or regions following implementation of the same intervention. Due to these complications, the use of a control alone is oftentimes insufficient unless a highly similar setting can be identified or a control from within the same setting (such as a different medication group). Yet a control group still does not address the issue of preexisting trends, which can only be elucidated if historical data is available. Including data from before the intervention is therefore imperative in these policy analyses, but we often observed it to be missing. When including pre-intervention data, the study should preferably include multiple timepoints before and after the intervention to permit correction for preexisting trends [47].

Not only do our results indicate a relatively low awareness of more robust observational study designs, it is also suggested that pharmaceutical pricing policy analyses have remained challenging even in recent years and that these challenges are experienced in all regions of the world. Concretely, we hypothesized that there would be a relationship between the studies with an ineligible study design and the year of publication or income-setting of the country, but substantial differences between studies with eligible and ineligible study designs were not found. In addition to this, studies from low-income settings were widely missing. More awareness on rigorous study designs among both researchers, journal editors and policy makers could help encourage the generation of higher quality evidence that can be used to inform policy-making, as noted by others [17,44]. This also provides opportunities for capacity-building in low-income economies, which could further contribute to strengthening methods used in the field.

Our results also indicate that the Cochrane tool for grading the risk of bias may not be sufficiently tailored to the study types that we see in the field of pharmaceutical policy evaluation [48]. For one, the large number of studies that were associated with an unclear risk of bias in the domain 'incomplete outcome data' is striking. According to the tool, complete data should not be assumed unless specifically stated. However, where missing data may suggest a problem in a randomized drug trial, the study designs that we encounter here mostly use periodically collected and validated data from databases. Hence, it is reasonable that most of the studies did not specify whether data was missing. Along the same lines, the domain 'Knowledge of allocated intervention' is in clinical randomized trials related to the blinding of researchers, but in this context regarded as the objectivity of outcomes. This may be a less relevant sign of bias in this field of study because pharmaceutical pricing policy analyses are predominantly based on objective outcomes such as unit prices. Thirdly, the large proportion of studies associated with a high risk of bias in the domains 'random sequence generation', 'allocation concealment' and 'random sequence generation' is misleading. Not only does the small denominator overstate the pattern, but more importantly is the use of non-randomized studies penalized. Indeed, non-randomized or controlled before-after studies are always scored as high risk according to the EPOC guidelines, even if performed well. The pattern that is now shown in Fig. 2a is thus the result of the choice for these study design themselves and not the methodological choices within these studies. Fourthly, the domain 'other bias' was often assessed to have an unclear or high risk of bias because relevant and common issues could not be captured under the other domains. A simpler, empirically based tool could possibly provide more accurate measures of risk of bias and study quality in pharmaceutical pricing policy analysis [49]. As many tools for assessing risk of bias already exist [50], the pros and cons of these tools can be assessed to guide development of an empirical tool. Furthermore, we encourage the modification of existing tools in the development of empirically based tools that match the specific characteristics of health policy research. Particularly, biases that are typically found in policy research – such as confounding bias – should be addressed in a new tool. Joint efforts of the research commu-

nity and the Cochrane Collaboration should be made to develop a tool appropriate for assessing bias in health policy analyses.

This study maps the methodological weaknesses of studies that have been published in the field of pharmaceutical pricing policies, and intends to encourage researchers, journal editors, policy makers and other relevant stakeholders to increase both the supply of and demand for high quality observational research on pharmaceutical and other health policies. A strength of this study is that it includes a representative sample of studies, not only due to an extensive literature search for the original SR but also the inclusion of multiple interventions within one field of study. Another strength is that gaps in reporting could be identified through literature complementing each other. We encountered several cases in which the exact same intervention was studied in different settings or focusing on different outcomes. Information in one study then enabled us to make better assessments in another study.

Our study has several limitations. The first limitation is that generalizability to other fields of study may be limited. The present study only included evidence from a SR on ten pharmaceutical pricing policies and may not accurately reflect the issues that are encountered in other areas of health policy. However, the methodological designs that were identified in this study are not unique to this field of study and issues identified are equally important to consider in other areas of health policy analysis. Additionally, the purpose of the present work is to illustrate some of the gaps and methodological weaknesses, which could be informative for researchers outside of pharmaceutical pricing policy analysis. Another limitation is the inconsistent naming of the study design used in included studies, if declared at all. In many cases, the method of analysis was presented as a study design. In others, neither the analysis nor design method was described in the paper. Both instances required the classification of study designs to be made based on the methods as presented. This could have introduced misclassification in the present work. A third limitation is the small number of studies that was included per study type. Saturation of possible recurrent methodological shortcomings may not have been achieved with this sample.

5. Conclusion

We have described that study design issues occur often in pharmaceutical pricing policy analyses and lead to a reduction in the volume of evidence that can be effectively used for policy-making. The common issues identified in the present study might be indicative of similar issues within other fields of health policy analysis and should be used as starting point for improving commonly applied methodologies in the field. Our results also indicate that a more tailored tool is needed for the assessment of the quality and risk of bias of health policy analyses. Ultimately, the generation of more robust evidence should go hand in hand with aligning the efforts of researchers and policy-makers to bridge the existing gap between generating evidence and policy-making in practice.

Role of the funding source

The systematic review of ten pharmaceutical pricing policies was commissioned and funded by the World Health Organization (WHO), under a grant from the United Kingdom Department for International Development (DFID), as part of the process for developing the 2020 WHO Guideline on Country Pharmaceutical Pricing Policies. The WHO secretariat and its advisors provided technical supports for formulating the systematic review protocol, performing specific searches on government websites, and translating several non-English publications.

The present study further builds upon the literature collected for the systematic reviews, but was not commissioned nor funded by a third party. WHO, its advisors and DFID had no role in the data analysis or interpretation of the systematic reviews or the present study.

Declaration of Competing Interest

The authors report receiving funding from the World Health Organization for performing a systematic literature review on ten pharmaceutical pricing policies, which served as primary data source for the present work.

Acknowledgments

The authors wish to thank colleagues at York Health Economics Consortium (YHEC) who contributed to the initial search, document management and screening process of studies: Julie Glanville, Eleanor Kotas, Ross Birtles, Mick Arber, Chris Bartlett, and James Mahon. The authors are also grateful to the other members of our systematic review team for their role in extracting the data: Christine Leopold, Lizanne Arnoldy, Lynn Al-Tayara, Tom Buijs and Daniela Moye Holz. Finally, we would like to thank YingYao Chen and Kiu Tay-Teo for their assistance in extracting and translating information from articles in Chinese.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.healthpol.2022.03.003.

References

[1] Xiu-xia L, Ya Z, Yao-long C, Ke-hu Y, Zong-jiu Z. The reporting characteristics and methodological quality of Cochrane reviews about health policy research. Health Policy 2015;119:503–10. doi:10.1016/j.healthpol.2014.09.002.

[2] World Health Organization Strengthening health research and evidence-based decision making, Geneva: World Health Organization; 2020. Available from https://www.who.int/westernpacific/activities/strengthening-health-research-and-evidence-based-decision-making [accessed 19/02/].

[3] World Health Organization Guidelines for WHO guidelines, Geneva: World Health Organization; 2003. Available from: https://apps.who.int/iris/bitstream/handle/10665/68925/EIP_GPE_EQC_2003_1.pdf?sequence=1 [accessed 19/02/2020].

[4] Niessen LW, Grijseels EW, Rutten FF. The evidence-based approach in health policy and health care delivery. Soc Sci Med 2000;51(6):859–69. doi:10.1016/s0277-9536(00)00066-6.

[5] Hariton E, Locascio JJ. Randomised controlled trials—the gold standard for effectiveness research. Int J Obstet Gynaecol 2018;125(13):1716. doi:10.1111/1471-0528.15199.

[6] Baicker K, Chandra A. Evidence-based health policy. N Engl J Med 2017;377(25):2413–15. doi:10.1056/NEJMp1709816.

[7] Oliver K, Pearce W. Three lessons from evidence-based medicine and policy: increase transparency, balance inputs and understand power. Palgrave Commun 2017;3(1):1–7.

[8] White H, Sabarwal S, de Hoop T. Randomized controlled trials (RCTs): methodological briefs - impact evaluation No. 7, Florence: UNICEF office of research; 2014. Available from https://www.unicef-irc.org/publications/752-randomized-controlled-trials-rcts-methodological-briefs-impact-evaluation-no-7.html [accessed 24/02/2020].

[9] WHO European Working Group on Health Promotion Evaluation & World Health Organization Health promotion evaluation: recommendations to policy-makers: report of the WHO European Working Group on Health Promotion Evaluation, Copenhagen: WHO Regional Office for Europe; 1998. Available from https://apps.who.int/iris/handle/10665/108116 [accessed 25/02/2020].

[10] Janovsky K. Health policy and systems developments - an agenda for research, Geneva: World Health Organization; 1996. Available from https://apps.who.int/iris/handle/10665/62987 [accessed 19/02/2020].

[11] World Health Organization WHO handbook for guideline development, Geneva: World Health Organization; 2001. Available from https://apps.who.int/iris/handle/10665/75146 [accessed 19/02/2020].

[12] Walt G, Shiffman J, Schneider H, Murray SF, Brugha R, Gilson L. 'Doing' health policy analysis: methodological and conceptual reflections and challenges. Health Policy Plan 2008;23(5):308–17. doi:10.1093/heapol/czn024.

[13] Bonell CP, Hargreaves J, Cousens S, Ross D, Hayes R, Petticrew M, Kirkwood BR. Alternatives to randomisation in the evaluation of public health interventions: design challenges and solutions. J Epidemiol Commun Health 2011;65(7):582–7. doi:10.1136/jech.2008.082602.

[14] World Health Organization Health policy and systems research - a methodology reader, Geneva: World Health Organization; 2012. Available from https://www.who.int/alliance-hpsr/resources/alliancehpsr_reader.pdf?ua=1 [accessed 19/02/2020].
[15] World Health Organization WHO guideline on country pharmaceutical pricing policies. 2nd ed. Geneva: World Health Organization; 2020. Available from https://www.who.int/publications/i/item/9789240011878 [accessed 18/02/2021].
[16] Web Annex A: Systematic reviews for the update of the WHO Guideline on country pharmaceutical pricing policies. Tordrup D., van den Ham R., Leopold C., Joosse I., Glanville J., Kotas E., Mantel-Teeuwisse A. Utrecht: Utrecht University, 2020. In: WHO guideline on country pharmaceutical pricing policies, 2nd ed.. Geneva: World Health Organization; 2020. Available from: https://www.who.int/publications/i/item/9789240011878 [accessed 18/02/2021]
[17] Gilson L, Raphaely N. The terrain of health policy analysis in low and middle income countries: a review of published literature 1994–2007. Health Policy Plan 2008;23(5):294–307. doi:10.1093/heapol/czn019.
[18] Jacobs TG, Robertson J, van den Ham HA, Iwamoto K, Bak Pedersen H, Mantel-Teeuwisse AK. Assessing the impact of law enforcement to reduce over-the-counter (OTC) sales of antibiotics in low- and middle-income countries; a systematic literature review. BMC Health Serv Res 2019;19(1):536. doi:10.1186/s12913-019-4359-8.
[19] Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. The Cochrane Collaboration; 2011. version 5.1.0 (updated March 2011)Available from https://training.cochrane.org/handbook/archive/v5.1/ [accessed 24/02/2020].
[20] Khan K., Ter Riet G., Glanville J., Sowden A., Kleijnen J.. Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews. CRD Report 4. York Cent Rev Dissem; 2001.
[21] Cochrane Effective Practice and Organisation of Care EPOC resources for review authors. Oslo: Norwegian Knowledge Centre for the Health Services; 2014. https://www.epocoslo.cochrane.org.proxy.library.uu.nl/epoc-specific-resources-review-authors.
[22] Tordrup D, van den Ham HA, Glanville J, Mantel-Teeuwisse AK. Systematic reviews of ten pharmaceutical pricing policies - a research protocol. J Pharm Policy Pract 2020;13:22. doi:10.1186/s40545-020-00228-0.
[23] Lee IH, Bloor K, Hewitt C, Maynard A. The effects of new pricing and copayment schemes for pharmaceuticals in South Korea. Health Policy 2012;104(1):40–9. doi:10.1016/j.healthpol.2011.09.003.
[24] Moodley R., Suleman F. The impact of the single exit price policy on a basket of generic medicines in South Africa, using a time series analysis from 1999 to 2014. 2019; 14(7): e0219690. DOI: 10.1371/journal.pone.0219690.
[25] Moodley R, Suleman F. Evaluating the impact of the single exit price policy on a basket of originator medicines in South Africa from 1999 to 2014 using a time series analysis. BMC Health Serv Res 2019;19(1):576. doi:10.1186/s12913-019-4403-8.
[26] Leopold C, Zhang F, Mantel-Teeuwisse AK, Vogler S, Valkova S, Ross-Degnan D, Wagner AK. Impact of pharmaceutical policy interventions on utilization of antipsychotic medicines in Finland and Portugal in times of economic recession: interrupted time series analyses. Int J Equity Health 2014;13:53. doi:10.1186/1475-9276-13-53.
[27] Puig-Junoy J. The impact of generic reference pricing interventions in the statin market. Health Policy 2007;84(1):14–29. doi:10.1016/j.healthpol.2007.02.010.
[28] Kwon HY, Bae S, Choi SE, Park S, Lee EK, Park S, Kim J. Easy cuts, easy rebound: drug expenditures with massive price cuts in Korea. Health Policy 2019;123(4):388–92. doi:10.1016/j.healthpol.2018.11.002.
[29] Suh HS, Kim JA, Lee IH. Effects of a price cut reform on the cost and utilization of antidiabetic drugs in Korea: a national health insurance database study. BMC Health Serv Res 2018;18(1):429. doi:10.1186/s12913-018-3255-y.
[30] Langley T, Lacey J, Johnson A, Newman C, Subramanian D, Khare M, Skelly R, Norwood M, Sturrock N, Fogarty AW. An evaluation of a price transparency intervention for two commonly prescribed medications on total institutional expenditure: a prospective study. Future Healthc J 2018;5(3):198–202. doi:10.7861/futurehosp.5-3-198.
[31] Yasaitis L, Gupta A, Newcomb C, Kim E, Newcomer L, Bekelman J. An insurer's program to incentivize generic oncology drugs did not alter treatment patterns or spending on care. Health Aff 2019;38(5):812–19. doi:10.1377/hlthaff.2018.05083.
[32] Baldi S, Vannoni D. The impact of centralization on pharmaceutical procurement prices: the role of institutional quality and corruption. Reg Stud 2017;51(3):426–38.
[33] Chu HL, Liu SZ, Romeis JC. Assessing the effects of drug price reduction policies on older people in Taiwan. Health Serv Manag Res 2011;24(1):1–7. doi:10.1258/hsmr.2010.010016.
[34] Stargardt T. The impact of reference pricing on switching behaviour and healthcare utilisation: the case of statins in Germany. Eur J Health Econ 2010;11(3):167–77. doi:10.1007/s10198-009-0172-3.
[35] Wu B, Zhang Q, Qiao X. Effects of pharmaceutical price regulation: China's evidence between 1997 and 2008. J Asia Pac Econ 2015;20(2):290–329.
[36] Ben-Aharon O, Shavit O, Magnezi R. Does drug price-regulation affect healthcare expenditures? Eur J Health Econ 2017;18(7):859–67. doi:10.1007/s10198-016-0832-z.
[37] Bhaskarabhatla A, Chatterjee C, Anurag P, Pennings E. Mitigating regulatory impact: the case of partial price controls on metformin in India. Health Policy Plan 2017;32(2):194–204. doi:10.1093/heapol/czw109.
[38] Balmaceda C, Espinoza MA, Diaz J. Impacto de una política de equivalencia terapéutica en el precio de medicamentos en Chile. Value Health Reg Issues 2015;8:43–8. doi:10.1016/j.vhri.2015.04.003.
[39] Sahay A., Jaikumar S. Does pharmaceutical price regulation result in greater access to essential medicines? Study of the impact of drug price control order on sales volume of drugs in India. Research and Publications. Ahmedabad; 2016. Report No.: 2016-02–01.
[40] Hsiao FY, Tsai YW, Huang WF. Price regulation, new entry, and information shock on pharmaceutical market in Taiwan: a nationwide data-based study from 2001 to 2004. BMC Health Serv Res 2010;10:218. doi:10.1186/1472-6963-10-218.
[41] Law MR, Ystma A, Morgan SG. The short-term impact of Ontario's generic pricing reforms. PLoS One 2011;6(7):e23030. doi:10.1371/journal.pone.0023030.
[42] Adriaen M, De Witte K, Simoens S. Pricing strategies of originator and generic medicines following patent expiry in Belgium. J Generic Med 2008;5(3):175–87.
[43] WHO Regional Office for Europe Evidence-informed policy-making, Copenhagen: WHO Regional Office for Europe; 2021. Available from https://www.euro.who.int/en/data-and-evidence/evidence-informed-policy-making [accessed 18/01/].
[44] Babar ZU. The need for an evidence-based encyclopaedia in health services research in pharmacy. Int J Environ Res Public Health 2020;17(7):2549. doi:10.3390/ijerph17072549.
[45] Bastuji-Garin S, Sbidian E, Gaudy-Marqueste C, Ferrat E, Roujeau JC, Richard MA, Canoui-Poitrine F. Impact of STROBE statement publication on quality of observational study reporting: interrupted time series versus before-after analysis. PloS One 2013;8(8):e64733. doi:10.1371/journal.pone.0064733.
[46] Soumerai SB, Ceccarelli R, Koppel R. False dichotomies and health policy research designs: randomized trials are not always the answer. J Gen Intern Med 2017;32(2):204–9. doi:10.1007/s11606-016-3841-9.
[47] Goodacre S. Uncontrolled before-after studies: discouraged by Cochrane and the EMJ. Emerg Med J 2015;32:507–8. doi:10.1136/emermed-2015-204761.
[48] Waddington H, Aloe AM, Becker BJ, Djimeu EW, Hombrados JG, Tugwell P, et al. Quasi-experimental study designs series-paper 6: risk of bias assessment. J Clin Epidemiol 2017;89:43–52. doi:10.1016/j.jclinepi.2017.02.015.
[49] Bero L, Chartres N, Diong J, Fabbri A, Ghersi D, Lam J, Lau A, McDonald S, Mintzes B, Sutton P, Turton JL, Woodruff TJ. The risk of bias in observational studies of exposures (ROBINS-E) tool: concerns arising from application to observational studies of exposures. Syst Rev 2018;7(1):242. doi:10.1186/s13643-018-0915-2.
[50] Taylor KW, Wang Z, Walker VR, Rooney AA, Bero LA. Using interactive data visualization to facilitate user selection and comparison of risk of bias tools for observational studies of exposures. Environ Int 2020;142:105806. doi:10.1016/j.envint.2020.105806.
[51] Adesina A, Wirtz VJ, Dratler S. Reforming antiretroviral price negotiations and public procurement: the Mexican experience. Health Policy Plan 2013;28(1):1–10. doi:10.1093/heapol/czs015.
[52] Ghislandi S, Armeni P, Jommi C. The impact of generic reference pricing in Italy, a decade on. Eur J Health Econ 2013;14(6):959–69. doi:10.1007/s10198-012-0442-3.
[53] Yoo KB, Lee SG, Park S, Kim TH, Ahn J, Cho MH, Park EC. Effects of drug price reduction and prescribing restrictions on expenditures and utilisation of antihypertensive drugs in Korea. BMJ Open 2015;5(7):1–10. doi:10.1136/bmjopen-2014-006940.
[54] von der Schulenburg F, Vandoros S, Kanavos P. The effects of drug market regulation on pharmaceutical prices in Europe: overview and evidence from the market of ACE inhibitors. Health Econ Rev 2011;1(1):18. doi:10.1186/2191-1991-1-18.
[55] Bhargava V. Addition of generic medication vouchers to a pharmacist academic detailing program: effects on the generic dispensing ratio in a physician-hospital organization. 2010; 16(6): 384–92. doi:10.18553/jmcp.2010.16.6.384.
[56] Kaiser U, Mendez SJ, Ronde T, Ullrich H. Regulation of pharmaceutical prices: evidence from a reference price reform in Denmark. J Health Econ 2014;36:174–87. doi:10.1016/j.jhealeco.2014.04.003.