SUPPORT Tools for evidence-informed health Policymaking (STP)

8. Assessing the applicability of systematic reviews

John N Lavis
Andrew D Oxman
Nathan Mendes Soua
Simon Lewin
Russell L Gruen
Atle Fretheim

1. Centre for Health Economics and Policy Analysis, Department of Clinical Epidemiology and Biostatistics, and Department of Political Science, McMaster University, 1200 Main St. West, HSC-2D3, Hamilton, ON, Canada L8N 3Z5. Email: lavisj@mcmaster.ca
2. Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs plass, N-0130 Oslo, Norway. Email: oxman@online.no
3. Health Research Methodology PhD Program and Department of Clinical Epidemiology and Biostatistics, 1200 Main St. West, HSC-2D1 area, Hamilton, ON, Canada, L8N 3Z5. Email: souzanm@mcmaster.ca
4. Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs plass, N-0130 Oslo, Norway and Health Systems Research Unit, Medical Research Council of South Africa. Email: simon.lewin@nokc.no
5. The Alfred Hospital, Monash University, Level 4, 89 Commercial Rd, Melbourne, VIC, Australia 3004. Email r.gruen@alfred.org.au
6. Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs plass, N-0130 Oslo, Norway. Email: atle.fretheim@nokc.no

Corresponding author:
John N Lavis
1200 Main St. West, HSC-2D3
Hamilton, ON, Canada, L8N 3Z5

Email: lavisj@mcmaster.ca
Abstract

**Background:** This article is number 8 in a series of 21 articles on tools for evidence-informed health policymaking. Differences between health systems may often result in a policy or programme option not being feasible or acceptable in another setting. Or these may result in a programme or policy not working in the same way in another setting, or even achieving different impacts in another setting. A key challenge that policymakers and those supporting them must face therefore is the need to understand whether research evidence about policies and programmes can be applied to their setting. Systematic reviews make this task easier by summarising the evidence from studies conducted in a variety of different settings. However, many systematic reviews do not provide adequate descriptions of the features of the actual settings in which the original studies were conducted.

**Objectives:** In this article we suggest five questions to guide those assessing the applicability of the findings from systematic reviews to specific settings.

**Key messages:**
- The following five questions can be used to guide assessments of the applicability of findings from systematic reviews to specific settings:
  1. Were the studies included in a systematic review conducted in the same setting or were the findings consistent across settings or time periods?
  2. Are there important differences in on-the-ground realities and constraints that might substantially alter the feasibility and acceptability of a policy or programme option?
  3. Are there important differences in health system arrangements that may mean an option could not work in the same way?
  4. Are there important differences in the baseline conditions that might yield different absolute effects even if the relative effectiveness were the same?
  5. What insights can be drawn about scaling up, implementation, and monitoring and evaluation?
- Even if there are reasonable grounds for concluding that the impacts of policy and programme options might differ in a specific setting, insights can almost always be drawn from a systematic review about possible options, approaches to the scaling up and implementation of options, and approaches to monitoring and evaluation.
Background

This article is number 8 in a series of 21 articles on tools for evidence-informed health policymaking. It is also the 2nd of 6 articles in this series about characterising the costs and consequences of potential policy and programme options. Its purpose is to suggest questions to guide those involved in assessing the applicability of the findings from systematic reviews to specific settings.

Commonalities in human biology often mean that a clinical procedure or drug will work the same way in different patients. This is not always the case, however, and questions have therefore been developed to help healthcare providers to assess the applicability of research evidence to their patients (see Box 1 for a list of questions related to applicability criteria for clinical practice). Differences between health systems often mean that a policy or programme option may not be feasible or acceptable in another setting. These differences may also mean that a policy or programme may not work the same way in another setting, or that it may achieve different impacts in another setting [1, 2]. A key challenge that policymakers and those supporting them must face, therefore, is understanding whether research evidence about the impacts of a policy or programme are applicable to their setting.

Systematic reviews make this task easier by offering a single summary of studies from different settings. The word settings, in this instance, refers to political/country jurisdictions (e.g. Canada or Cameroon), but settings can also include sectors (e.g. primary care or hospital care), and locales (e.g. urban or rural). Systematic reviews can also assist with the process of making judgements about the applicability of the evidence in specific settings by providing a framework and, when available, research evidence that can be used to identify those factors that are essential for a policy or programme to work – or that might modify its impacts. A systematic review of pharmaceutical policies (i.e. referencing pricing, other pricing, and purchasing policies), for example, provided a summary of the factors that could influence the impacts of referencing pricing, as well as the rationale for each [3]. These factors included the equivalence of the drug, incentives, exemptions, the availability of drugs, price levels, and electronic information systems.

Unfortunately, many systematic reviews do not do the following:

- Highlight the features of the settings in which studies were conducted, particularly those features that might modify the impacts of a policy or programme
- Provide a framework for identifying potential modifying factors, or
- Provide research evidence about modifying factors

In these cases, policy analytic articles or narrative reviews may provide more helpful frameworks that could be used to inform judgements about the applicability of the evidence in a systematic review.

A framework for corruption in the health sector, for example, highlighted how health system arrangements (e.g. governance arrangements that limit monopolies, require transparency, and support enforcement) and other factors, influenced opportunities and ‘pressure’ to abuse, as well as the rationalisation for abuse, and how this, in turn influenced the abuse of power for private gain [4]. But, ideally, a systematic review about the impacts of anti-corruption efforts would also have described the relevant health system arrangements in the settings where the studies were conducted. Those features that might influence opportunities and ‘pressure’ to abuse and the rationalisation of such behaviour, would be of particular interest.
Applicability considerations are equally, if not more important, for other types of systematic reviews, including those of administrative database studies and of community surveys that help to place problems in comparative perspective. Such considerations are equally important for reviews of observational studies that help to characterise an option’s likely harms, and reviews of qualitative studies that assist in: understanding the meanings that individuals or groups assign to particular problems, how and why particular options work, and the views and experiences of stakeholders related to particular options.

Questions to consider

The following five questions can guide how to assess the applicability of the findings from systematic reviews to specific settings.

1. Were the studies included in a systematic review conducted in the same setting or were the findings consistent across settings or time periods?
2. Are there important differences in on-the-ground realities and constraints that might substantially alter the feasibility and acceptability a policy or programme option?
3. Are there important differences in health system arrangements that may mean an option could not work in the same way?
4. Are there important differences in the baseline conditions that might yield different absolute effects even if the relative effectiveness was the same?
5. What insights can be drawn about scaling up, implementation, and monitoring and evaluation?

1. Were the studies included in a systematic review conducted in the same setting or were the findings consistent across settings or time periods?

If the studies included in a systematic review were conducted in the same setting where policymakers are based, or else in very similar settings, there may be little reason to be concerned about the applicability of the findings. Similarly, if the findings have been shown to be consistent across settings or time periods, then similar impacts might be expected. On-the-ground realities and constraints, health system arrangements and baseline conditions, are likely to differ across settings and change over time, so consistent findings in these circumstances are likely to mean the findings are broadly applicable (These three issues are the focus of the three next questions that follow in this section).

The following information in systematic reviews can be used by policymakers to inform judgements related to such issues:

- Information about the settings of studies and specifications regarding the time periods over which the studies were conducted. This can typically be found in sections entitled ‘Characteristics of included studies’
- Information about the consistency of findings can typically be found in the abstract or in the ‘Results’ section

When information about settings and time periods is lacking in a systematic review, policymakers will need to retrieve the original studies to locate this information (if the issue is of sufficiently high priority, and if resources and time allow).
Research comparing mortality rates in not-for-profit hospitals with mortality rates in for-profit hospitals provides an example of how such data can be used. This research has been conducted over several decades in the United States during which time the health system has changed dramatically [5]. The research has demonstrated remarkable consistency over time in the significant survival advantage of being treated in not-for-profit hospitals. Based on these data, a policymaker from Canada might then conclude that a similarly consistent finding would be seen in a Canadian setting. And this conclusion might lead them to avoid the introduction of for-profit hospitals into the current system which consists only of not-for-profit hospitals (or at least to avoid using health benefits as a justification for doing so).

2. Are there important differences in on-the-ground realities and constraints that might substantially alter feasibility and acceptability of a policy or programme option?

If the studies included in a systematic review were conducted in settings with largely similar resource and capacity constraints, and largely similar perspectives amongst health system stakeholders and levels of influence, policymakers might reasonably expect that an option would be both feasible and acceptable in their own setting. However, policymakers will rarely be able to find information about resource and capacity constraints and stakeholder influence in a systematic review. Instead, they will find a description of the policy or programme that was studied. They will typically be sufficiently familiar with the resources, capacity, and stakeholder influence in their own setting to enable them to judge the feasibility and acceptability of the policy or programme.

Policymakers in a setting with very significant resource and capacity constraints when implementing an option will have to think twice about the feasibility of an option [6]. Some settings, for example, may face a shortage of nurse practitioners, and therefore any option requiring a significant role for this category of healthcare provider might not be feasible in the short-term [7]. Similarly, some settings have such limited financial resources that an option shown to have significant impacts, such as artemisinin-based combination therapies (ACT) to treat malaria, might not be considered feasible on a large scale without significant donor support [8]. Some settings may have sufficiently over-stretched health systems that could not accommodate the increase in demand that may accompany the introduction of conditional cash transfers (i.e. providing money to households on the condition that they comply with certain health and healthcare-seeking behaviours) [9]. Or settings may lack the capacity within government or among managers, healthcare providers and consumers to support the widespread use of a particular option. Audit and feedback (i.e. the provision of healthcare providers with data about their performance), for example, might not be feasible in settings where routinely collected data are unreliable. Article 13 in this series addresses resource and capacity constraints in more detail, albeit specifically related to the context of scaling up policies and programmes [10].

Policymakers in a setting in which stakeholders are opposed to an option and have significant influence on practice and policy, may have to assess the likely acceptability of an option particularly carefully. Healthcare provider associations such as nurses association, for instance, may resist the introduction or expansion of a lay health worker programme if they thought that the income or status of nurses might be affected [11]. Civil society organisations, too, may actively oppose changes that would reduce prescription drug use among consumers, particularly for life-sustaining drugs, and drugs that are important in treating chronic conditions [12]. Such changes could include the introduction of caps (i.e. consumers are
reimbursed up to a set maximum number of prescriptions), co-insurance (i.e. consumers pay a percentage of the price of the prescription drug), and co-payments (i.e. consumers pay a fixed amount per prescription drug).

This broader question of whether differences in on-the-ground realities and constraints might substantially alter the feasibility and acceptability of a policy or programme option is very similar to the focus of the more general question – “Will it work?” – which has been raised in the context of the testing of clinical interventions [13]. It is also similar to the two questions raised in Box 1 later in this paper, about compliance with treatment requirements, that can be used by healthcare providers in assessing the applicability of research evidence to individual patients [14].

Significantly, on-the-ground realities and constraints can be addressed over time: nurse practitioner training programmes, for example, can be scaled up. And donors can subsidise the cost of ACT. Governments can improve the quality of routinely collected data. And healthcare provider associations and civil society organisations can be engaged in a series of negotiations or dialogues.

3. Are there important differences in health system arrangements that may mean an option could not work in the same way?

If the studies included in a systematic review were conducted in settings with largely similar health system arrangements, particularly those that might substantially alter the potential impacts of an option, policymakers might reasonably expect similar relative effectiveness in their setting. Deciding whether health system arrangements might alter an option’s impacts requires an understanding of how and why an option might work. Within a systematic review policymakers may find a framework and research evidence that will identify those factors essential for a policy or programme to work – or that might modify its impacts. Policymakers may also find a summary of the features of the settings in which studies were conducted that might modify the impacts of a policy or programme.

If a systematic review does not provide the information necessary to determine whether particular health system arrangements might mean an option could not work in the same way, policymakers could look for:

- Policy analytic articles or narrative reviews that include helpful frameworks that could be used to identify factors that might modify the impacts of an option, and
- Detailed descriptions of the health system arrangements, specifically those that might substantially alter the potential impacts of an option, in the settings where the studies were conducted

The European Observatory on Health Systems and Policies publishes, and periodically updates, profiles of the health systems of a large number of middle- and high-income countries. These ‘Health in Transition’ (HiT) profiles can be found online (http://www.euro.who.int/observatory/hits/20020525_1) and downloaded free of charge. The Health Policy Monitor provides a searchable online database of key health system features in some of the same countries (http://www.hpm.org/index.jsp).

Policymakers in a setting with very different health system arrangements, specifically arrangements that appear significant to an option functioning in the same way, should be
cautious about assuming that comparable relative effectiveness could be achieved. For example, in a review of reference drug pricing [3], six of the ten studies were conducted amongst older people/pensioners in British Columbia, Canada. Policymakers in other settings may conclude that they would not achieve comparable impacts to those seen in Canada if they have the following health system arrangements:

- Inadequate incentives for consumers, healthcare providers, pharmacists and pharmaceutical companies to comply with the reference drug price system, and
- An electronic processing system that lacks the capacity to realise the low administration costs associated with identifying, prescribing and dispensing the reference drugs and with handling exemptions

Similarly, other pricing policies examined in competitive pharmaceutical markets may yield a different relative effectiveness in markets with monopolies.

The question of whether there are important differences in health system arrangements that may result in an option not working in the same way, is again similar to the general question “Can it work?” raised in wider discussions of the testing of healthcare interventions [13]. It is also similar to the question, about whether biological results (e.g. sex, co-morbidities, age) that might modify the treatment response have been excluded, that can be used by healthcare providers in assessing the applicability of research evidence to individual patients (see Box 1) [14]. Unlike Question 2 above, which focused on on-the-ground realities and the possibility of associated change, there is less chance that health system arrangements could be modified. Health system arrangements are difficult to change and typically the rationale underpinning a change would need to be more compelling than only the possibility that it would enhance the impact of a single policy or programme.

4. Are there important differences in the baseline conditions that might yield different absolute effects even if the relative effectiveness were the same?

If the studies included in a systematic review were conducted in settings with largely similar baseline conditions, such as the use of a programme or policy coverage, policymakers might reasonably expect similar absolute effects in their setting (provided their answer to Question 3 above led them to expect similar relative effectiveness). Policymakers will often be able to find information about baseline conditions within systematic reviews in a section titled ‘Characteristics of included studies’ (or similar). Alternatively, they may have to retrieve the original studies included in the review in the hope that baseline conditions were better described in them. Policymakers will typically be able to find local evidence about baseline conditions in their own setting. (Article 6 in this series addresses how to find and use local evidence)[15].

Policymakers in a setting with different baseline conditions may expect different absolute impacts. The absolute impact of audit and feedback, for example, is likely to be larger than in instances where the baseline compliance to recommended practice is low [16]. Similarly, the absolute impact of a pay-for-performance initiative may have a bigger impact in low- and middle-income countries (because small financial incentives may be larger relative to wages) than in high-income countries [17].

This question highlighting the link between baseline conditions and absolute effects does not have an immediate parallel among those questions used to assess the applicability of research evidence to individual patients (see Box 1). Nevertheless, the question is highly relevant in
clinical settings where the relative effectiveness of a clinical intervention is often the same across patients but patients’ baseline risks may vary quite dramatically [18, 19]. The question is also highly relevant in public health settings where immunisation programmes, for example, may be introduced in countries with very different baseline conditions. Article 11 of this series discusses the use of balance sheets to summarise important impacts and provides further detail about relative effectiveness and absolute impacts [20].

5. **What insights can be drawn about scaling up, implementation, and monitoring and evaluation?**

Important lessons can still be drawn from systematic reviews even if the findings are not directly applicable to a given setting. Policymakers may be provided with an idea for a policy or programme option that they might not have considered otherwise. They may also gain insight into how policies and programmes have been scaled up or implemented in other settings. And they may be able to draw directly on the systematic review itself in developing a monitoring and evaluation plan. Policymakers, for example, may learn about a new approach to supporting team-based care, the importance of engaging both mid-level managers and front-line nurses in the implementation of an option, and what types of outputs and outcomes they should track as they monitor and evaluate the implementation of a new policy.

Box 2 provides an example of an assessment of the applicability of a systematic review.
Resources

Useful documents and further reading

Links to websites
- SUPPORT Collaboration – Example of a source of policymaker-friendly summaries of systematic reviews that provides an assessment of the applicability of the research evidence (in this case to low- and middle-income countries), and that highlights the factors that policymakers need to bear in mind when assessing its applicability to their own setting http://www.support-collaboration.org/
- European Observatory on Health Systems and Policies – Example of a source of (Health in Transition) profiles of the health systems of a large number of middle- and high-income countries http://www.euro.who.int/observatory/hits/20020525_1
- Health Policy Monitor – Searchable online database of key health system features in a number of middle- and high-income countries http://www.hpm.org/index.jsp
Box 1: Comparing the applicability criteria for clinical practice and policymaking

The applicability criteria proposed here for policymaking have analogues in clinical practice. Four questions to assess the applicability of research evidence to individual patients have been proposed:

- Have biological results (e.g. sex, co-morbidities, age) that might modify the treatment response been excluded?
- Can consumers comply with the treatment requirements?
- Can healthcare providers comply with the treatment requirements?
- Are the likely benefits worth the potential risks and costs? [14]

The more general formulation of this approach focuses on three questions:

- Can it work?
- Does it work?
- Is it worth it? [13]

The second question in this list – Does it work? -- relates to the second and third questions, about compliance, in the above list.
Box 2: An assessment of the local applicability of a systematic review about home care (from the perspective of a Canadian policymaker)

Policymakers assessing the applicability of a 2005 review of home care could apply the following series of questions [21].

1. Were the studies included in the systematic review conducted in the same setting or were the findings consistent across settings or time periods?
   - 22 studies were included in the review
     - 9 from the United Kingdom
     - 3 from Australia
     - 1 from each of Italy, Norway, and the United States
     - 7 were not described in a way that identified the country in which the study was conducted
   - Findings were not consistent across settings
   - Two studies were published in 1978. The others were published from 1992 onwards (and many did not specify a time period), making it difficult to support the contention that the findings were consistent over time periods

2. Are there important differences in on-the-ground realities and constraints that might substantially alter the feasibility and acceptability of an option?
   - In Canada, nurses are in tremendous demand (particularly in hospitals) and are not used to the scope of practice required for care in home care settings. This means that many of the best nursing graduates might not embrace career opportunities in home care settings
   - In Canada, unlike in the United Kingdom where 9 of 13 identifiable studies were conducted, patients differ in whether they have supplementary coverage permitting more intensive home care. This means that relatively more wealthy people may get access to home care but the less well off may not
   - In Canada, unlike in the United Kingdom, home care recipients and their families may have to travel very long distances if they have to seek acute care. Some may therefore delay their discharge from hospital; others may suffer if a hospital transfer is difficult
   - In Canada, nurses may face a drop in pay if they move from hospitals to the community. Many of them may therefore actively oppose a shift from hospital care to home care
   - In Canada, as suggested earlier, caregivers may face a greater burden that is not covered by social services (at least outside the province of Quebec)

3. Are there important differences in health system arrangements that may mean an option could not work in the same way?
   - In Canada, there is even more of a separation between health and social services (at least outside the province of Quebec) than there is in the United Kingdom, which means that Canadian home care recipients and their families cannot rely on the same breadth of services available to those in the United Kingdom
   - In Canada, unlike in the United Kingdom, there is a governmental commitment to first-dollar coverage for hospital-based and physician-provided care but not for home care, which means that Canadian home care recipients and their families may face significant financial barriers to accessing home care
   - In Canada, unlike in the United Kingdom, most Canadians are not ‘attached’ to a (multi-disciplinary) primary healthcare practice, which means that some Canadian home care recipients would not even have a regular primary healthcare provider
4. Are there important differences in the baseline conditions that might yield different absolute effects – even if relative effectiveness were the same?
   - In Canada, home care is already well established for most types of care, which means that the benefits may be small in absolute terms, at least for those not facing financial barriers

5. What insights can be drawn about options or their scaling up, implementation, and monitoring and evaluation?
   - In Canada, admission-avoidance schemes may be a relatively unknown option compared to well established schemes such as the early discharge of elderly medical patients or patients following surgery or care of terminally ill patients

The review has now been updated and divided into two separate reviews, one of which deals specifically with admission-avoidance schemes and would be particularly germane to Canada [22].
References


