RAND EUROPE

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Skip all front matter: Jump to Page 1 ▼

Support RAND

Browse Reports & Bookstore

Make a charitable contribution

For More Information

Visit RAND at www.rand.org

Explore RAND Europe

View document details

Limited Electronic Distribution Rights

Measuring research

A guide to research evaluation frameworks and tools

Susan Guthrie, Watu Wamae, Stephanie Diepeveen, Steven Wooding and Jonathan Grant

Prepared for the Association of American Medical Colleges

RAND EUROPE

The research described in this report was prepared for the Association of American Medical Colleges.

RAND Europe is an independent, not-for-profit research organisation whose mission is to improve policy and decision making for the public good. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

Executive summary

Interest in and demand for the evaluation of research is increasing internationally. Several factors account for this: agendas for accountability and good governance and management, and fiscal austerity in many countries. There is a need to show that policymaking is evidence based and, in the current economic climate, to demonstrate accountability for the investment of public funds in research. This is complemented by a shift in the type of evaluation needed: from the traditional, summative, assessment to more formative evaluations and those covering wider outputs from research.

Given this growing need for effective and appropriate evaluation, it is increasingly important to understand how research can and should be evaluated in different contexts and to meet different needs. The purpose of this study was to understand the challenges and trade-offs in evaluating research, by looking at examples of frameworks and tools for evaluating research used internationally. The frameworks and tools we explored in detail are summarised in Annex A. Our key findings are as follows.

Designing a research evaluation framework requires trade-offs: there is no silver bullet.
By investigating some of the characteristics of research evaluation frameworks, and the relationships between them, we have identified some trade-offs in developing a framework to evaluate research:

- Quantitative approaches (those which produce numerical outputs) tend to produce longitudinal data, do not require judgement or interpretation and are relatively transparent, but they have a high initial burden (significant work may be required at the outset to develop and implement the approach).
- Formative approaches (those focusing on learning and improvement rather than assessing the current status) tend to be comprehensive, evaluating across a range of areas, and flexible, but they do not produce comparisons between institutions.
- Approaches that have a high central burden (requiring significant work on the part of the body organising the evaluation process) tend not to be suitable for frequent use.
- Approaches that have been more fully implemented tend to have a high level of central ownership (by either the body organising the evaluation, or some other body providing oversight of the process).
- Frameworks that place a high burden on participants require those participants to have a high level of expertise (or should provide capacity building and training to achieve this).

To be effective, the design of the framework should depend on the purpose of the evaluation: advocacy, accountability, analysis and/or allocation.
Research evaluation aims to do one, or more, of the following:

- advocate: to demonstrate the benefits of supporting research, enhance understanding of research and its processes among policymakers and the public, and make the case for policy and practice change
- show accountability: to show that money and other resources have been used efficiently and effectively, and to hold researchers to account

- analyse: to understand how and why research is effective and how it can be better supported, feeding into research strategy and decisionmaking by providing a stronger evidence base
- allocate: to determine where best to allocate funds in the future, making the best use possible of a limited funding pot.

The aim(s) of the evaluation process should be decided at the outset, as this will influence most other decisions in the development of the research evaluation approach. For example, our analysis suggests that frameworks that are used to inform allocation often require comparisons to be made between institutions, and are therefore summative rather than formative, and are unlikely to be comprehensive. By contrast, analysis typically requires a more formative approach in order to show 'how' and 'why' the research is effective, rather than just provide a summative measure of performance. This suggests that frameworks for allocation are not likely to be well suited to use for analysis, and vice versa. To achieve both aims effectively would require two parallel, though potentially connected, evaluation processes. Alternatively, allocation, in particular accountability purposes, calls for transparency. This suggests that quantitative approaches may be most appropriate, and that there may be a high initial burden in setting up the framework.

Research evaluation tools typically fall into one of two groups, which serve different needs; multiple methods are required if researchers' needs span both groups.
The two groups are:

- formative tools that are flexible and able to deal with cross-disciplinary and multi-disciplinary assessment
- summative tools that do not require judgement or interpretation, and are quantitative, scalable, transparent, comparable and suitable for high frequency, longitudinal use.

The units of aggregation used for collecting, analysing and reporting data will depend on the target audience(s), and the tool(s) used.
Developing a framework also requires careful selection of units of aggregation for the collection, analysis and reporting of data. The units selected for each evaluation phase are interrelated. The unit of reporting depends on the needs of the audience(s), as well as privacy requirements related to the unit being assessed. The unit of data collection will depend on feasibility, burden and the selection of tools, but must be at a lower (or the same) level of aggregation as reporting. The unit of analysis must be between (or the same as) those of collection and reporting.

There are some perennial challenges to research evaluation that need to be addressed.
The importance of these challenges varies depending on the aim(s) of the framework. Attribution and contribution are important when the evaluation is for accountability or allocation, but the fact that there may have been multiple inputs to downstream outcomes and impacts tends to be less important in advocacy, and may not be important in analysis depending on the purpose and context. Time lags between research being conducted and final outcomes are important when frameworks are using downstream measures of performance, such as outcomes and impacts. Being able to discriminate levels of performance at a fine-grained level is only really significant where the output of the evaluation process includes some kind of score or comparative measure between institutions (or groups, researchers and so on), typically for allocation purposes.

Research evaluation approaches need to suit their wider context.
Different approaches may be acceptable and credible in different environments, and it is important to consider this when developing a framework. The history, politics and wider social and economic context need to be taken into account. For example, bibliometrics as a methodology to evaluate research is credible and acceptable to the research community in some countries (e.g. Australia) but has met with

greater hostility in others (e.g. the UK). Credibility and acceptability can be improved by considering the risks of discrimination against specific groups (e.g. early career researchers, part-time workers and minority groups) through the implementation of the framework, and by carefully considering the unexpected consequences that may result. Moreover, the evaluation of emerging areas of research (for example, implementation science) presents special challenges and opportunities for new approaches, as traditional evaluation tools and metrics may be wholly unsuitable.

Finally, implementation needs ownership, the right incentives and support.
Implementation is a key consideration, whether participation in the framework is compulsory or voluntary. Where compulsory, the challenge is to obtain support from the academic and wider community. Where participation is voluntary, incentives need to be in place to promote and sustain uptake. In both cases, participants need to be given the skills necessary for the process, through simplicity, training or a toolkit. In all cases, strong central ownership is needed for effective large-scale implementation.