Developing a research evaluation framework

There is growing demand internationally for research evaluation, due to an increasing emphasis on governance and accountability in both the public and private sectors. There is also greater understanding that policymaking must be based on objective evidence, and therefore a need for explicit and transparent evaluation methods.

The traditional approaches to research evaluation are summative, assessing, for example, outputs such as the quality and number of papers published, as measured with bibliometrics, or comparing institutions' past performance. These examine what has happened in the past but do not tell us why. Wider and more specific measures of research success, such as payback frameworks, are now being applied, and these extend the range of summative evaluations; there are also new, formative, evaluation methods, based on learning, that contribute to improving the way research is done in the future.

There are many different evaluation frameworks to choose from, and many tools for gathering appropriate data to use within frameworks, but no single approach works in all contexts and for all purposes. We have therefore constructed a decision tree for funders, policymakers and researchers who want to evaluate research and need practical guidance on how to choose the appropriate approach. In the full report, we present a comprehensive list of tools and techniques, and discuss when each might be most useful; examine the advantages and disadvantages of different approaches; and look at the context in which each of 14 different frameworks has been used before.

The examples provided have been collected internationally and many are multidisciplinary, so that while the resource is tailored towards the biomedical research arena, lessons can be drawn more widely.

The decision tree overleaf sets out the questions to ask and what needs to be considered in selecting a research evaluation framework. It shows that the first question to resolve is: what is the purpose of your evaluation? Different types of information will be required from an evaluation, depending on what it is trying to achieve, and this will determine what methods are most suitable. There may be one or more purposes, and it is important to be clear which has primacy, due to tensions between the different purposes and their appropriate methods. The main purpose will usually be described by one of the '4 As':

Key findings:

- In an environment of increasingly scarce resources, economic and other accountability pressures require that research funders, policymakers and researchers themselves need to evaluate research

- There is no single evaluation method that suits all contexts, so we have assembled a survey of the options available internationally, and describe when these are best used

- The first step is to decide on the purpose of the evaluation in terms of the '4 As': Analysis, Accountability, Advocacy and Allocation

- We present a decision tree to help develop a research evaluation framework that suits the purpose

- Further details of trade-offs, advantages and disadvantages, and previous applications of 14 research evaluation frameworks from six continents can be found in the full report: *Measuring research: A guide to research evaluation frameworks and tools*

Analysis: for example, what funding is most effective in terms of different outputs and outcomes, including the impact of research?

Accountability: for example, for those distributing public funds who need to show they are doing the right thing. Likewise, funding organisations need to demonstrate impact to donors.
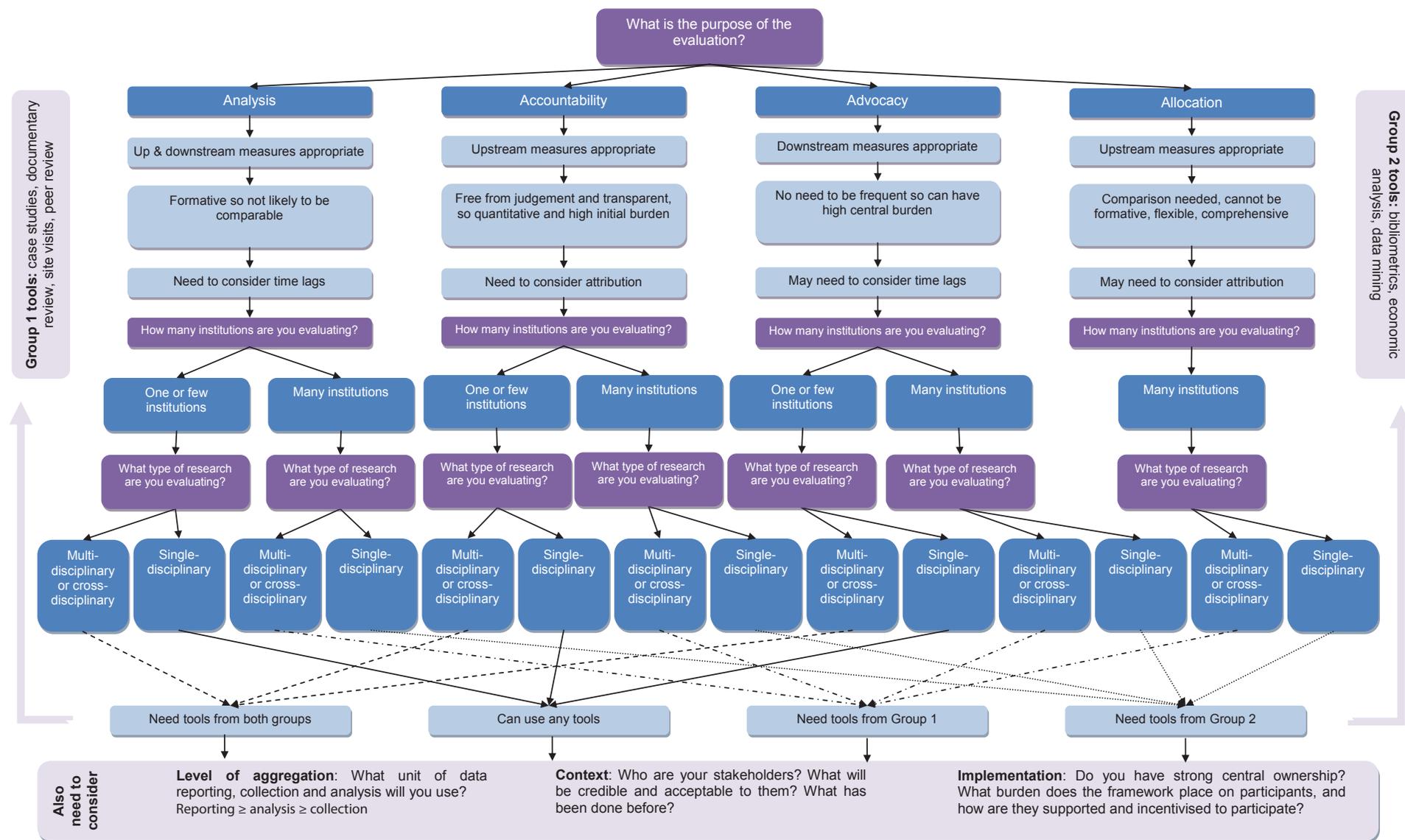
Advocacy: for example, how the research benefits society; this would help funders wanting evidence to support their decisions, or advocates seeking evidence for their cause.

Allocation: for example, to prioritise which projects, people and institutions are given funding.

The next factors to consider include the number of institutions or specific research programs or areas to be evaluated, and whether the research is single- or multi-disciplinary. Working through the tree leads to recommendations about the types of tools that will best suit the situation. Tools are classified as either Group 1 (case studies, documentary review, site visits and peer review) or Group 2 (bibliometrics, economic analysis and data mining). For some evaluations, tools from both groups are warranted.

The full report and executive summary explore these dimensions in greater depth.

A decision tree for developing a research evaluation framework

What is the purpose of the evaluation?

Analysis
- Up & downstream measures appropriate
- Formative so not likely to be comparable
- Need to consider time lags
- **How many institutions are you evaluating?**
  - One or few institutions
    - What type of research are you evaluating?
      - Multi-disciplinary or cross-disciplinary
      - Single-disciplinary
  - Many institutions
    - What type of research are you evaluating?
      - Multi-disciplinary or cross-disciplinary
      - Single-disciplinary

Accountability
- Upstream measures appropriate
- Free from judgement and transparent, so quantitative and high initial burden
- Need to consider attribution
- **How many institutions are you evaluating?**
  - One or few institutions
    - What type of research are you evaluating?
      - Multi-disciplinary or cross-disciplinary
      - Single-disciplinary
  - Many institutions
    - What type of research are you evaluating?
      - Multi-disciplinary or cross-disciplinary
      - Single-disciplinary

Advocacy
- Downstream measures appropriate
- No need to be frequent so can have high central burden
- May need to consider time lags
- **How many institutions are you evaluating?**
  - One or few institutions
    - What type of research are you evaluating?
      - Multi-disciplinary or cross-disciplinary
      - Single-disciplinary
  - Many institutions
    - What type of research are you evaluating?
      - Multi-disciplinary or cross-disciplinary
      - Single-disciplinary

Allocation
- Upstream measures appropriate
- Comparison needed, cannot be formative, flexible, comprehensive
- May need to consider attribution
- **How many institutions are you evaluating?**
  - Many institutions
    - What type of research are you evaluating?
      - Multi-disciplinary or cross-disciplinary
      - Single-disciplinary

- Need tools from both groups
- Can use any tools
- Need tools from Group 1
- Need tools from Group 2

Also need to consider

Level of aggregation: What unit of data reporting, collection and analysis will you use? Reporting ≥ analysis ≥ collection

Context: Who are your stakeholders? What will be credible and acceptable to them? What has been done before?

Implementation: Do you have strong central ownership? What burden does the framework place on participants, and how are they supported and incentivised to participate?

---

www.randeurope.org

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.

Support RAND

Browse Reports & Bookstore

Make a charitable contribution

For More Information

Visit RAND at www.rand.org

Explore RAND Europe

View document details

Research Brief

This product is part of the RAND Corporation research brief series. RAND research briefs present policy-oriented summaries of individual published, peer-reviewed documents or of a body of published work.

Limited Electronic Distribution Rights