RAND EUROPE

AAMC research evaluation initiative: Panel workshop report

Cross-panel analysis

Interest in and demand for the evaluation of research is increasing internationally.[1] This increase is linked to a growing emphasis on accountability, driven by a focus on good governance and management alongside fiscal austerity in a number of countries. This produces a need to show that policymaking is evidence-based and, particularly in the current economic climate, to demonstrate accountability for the investment of public funds in research. In this context, the AAMC has launched a research evaluation initiative, with the support of RAND Europe, with the aim of helping medical schools to introduce more comprehensive approaches to research evaluation. These approaches would go beyond the traditional metrics of publications and grant awards to address issues that matter to wider stakeholders outside of the academic community. To do this, the AAMC has engaged three panels focusing on different areas of research to consider some of the key challenges and questions relating to the evaluation of research in their areas.

- **Clinical outcomes improvement research:** Research focused on benefitting the clinical care mission of the institution at which the research is being conducted
- **Health equity research:** Research focusing on removing the disparities in health and healthcare provision
- **Basic research:** Fundamental biomedical research

Each of the three panels convened between September and November 2013 for an initial workshop to think about some of the challenges and opportunities in research evaluation in their field. Each of the two-day workshops explored areas in which research evaluation is occurring and how it can be improved and developed. The workshops were interactive and participative, offering the panel members the opportunity to share their experience of research evaluation and their perspectives on the current state of evaluation of research. The aim of the workshops was to start the process of putting together a toolkit or approach that AAMC and RAND, together with the panels, can provide to support medical schools in evaluating their research.

Separate reports are available for each panel, summarizing the views of the panels and the issues raised in discussion as well as providing a detailed summary of the content and process of the workshops. This document compares the findings of the different panels and identifies some of the common themes and threads emerging across the three workshops to inform our thinking in the next stage of the initiative.

[1] Examples include the Research Excellence Framework in the UK (see http://www.ref.ac.uk/) or Excellence in Research for Australia (see http://www.arc.gov.au/era/). More examples of international research evaluation efforts can be found in Guthrie et al. (2013). "Measuring research: A guide to research evaluation frameworks and tools". RAND Report MG-1217-AAMC. Available at http://www.rand.org/pubs/monographs/MG1217.html

CURRENT APPROACHES TO RESEARCH EVALUATION

All of the workshops started with some discussion of existing approaches to research evaluation at the different institutions represented on the panels, to identify commonalities and differences in their understanding and experience of research evaluation. A common theme across the panels was that the institutions that the panelists represented all used the traditional research evaluation metrics such as number of publications or research grants, or number of citations received. However, all panels also had examples of the use of novel approaches at individual institutions, demonstrating the potential for learning across institutions.

The health equity and the clinical outcomes improvement research panels talked about the importance of the community in their research endeavours, as both participants in and beneficiaries of research. Both of these groups also talked about the importance of their research being able to contribute to the wider goals of their institutions. To meet the needs of researchers in these two areas, research evaluation approaches will need to be able to address these issues. For the basic science panel, research funding was a key issue and there was some discussion around the ways in which evaluation can support this.

All the panels discussed the challenges and importance of measuring 'outcomes' which are of importance to wider stakeholders, although each did this in a different context. However, what were considered outcomes differed significantly between these three groups. For the clinical outcomes improvement research panel, 'outcomes' referred to health outcomes at their institution and amongst their local patients and populations. This was also important to the health equity research panel, but they had a greater emphasis on equity of health outcomes as well as their general improvement. For the basic research panel, wider elements around capacity building, including research training, and outreach were important. Health outcomes were considered important but it was felt that linking these directly to research might be too challenging.

All panels were interested in making the case for the importance of their particular type of research to their institutions, and felt evaluation offered an opportunity to do that, by reaching a shared understanding of the intended outcomes of their research with institutional leadership and other stakeholders, and by illustrating some of the benefits of their research through the evidence that the evaluation could provide.

Developing a shared understanding of what is meant by research evaluation also formed an important part of these early discussions at the workshops. One way of conceptualizing research evaluation across the various stages of the research and translation pathway is through a logic model of the form presented in Figure 1. Although this is a simplification of the research process, panelists across all three panels found this useful in thinking about the various elements and stages of research that could be measured as part of an evaluation and this will form part of our thinking as we develop an approach which can be used by medical schools.

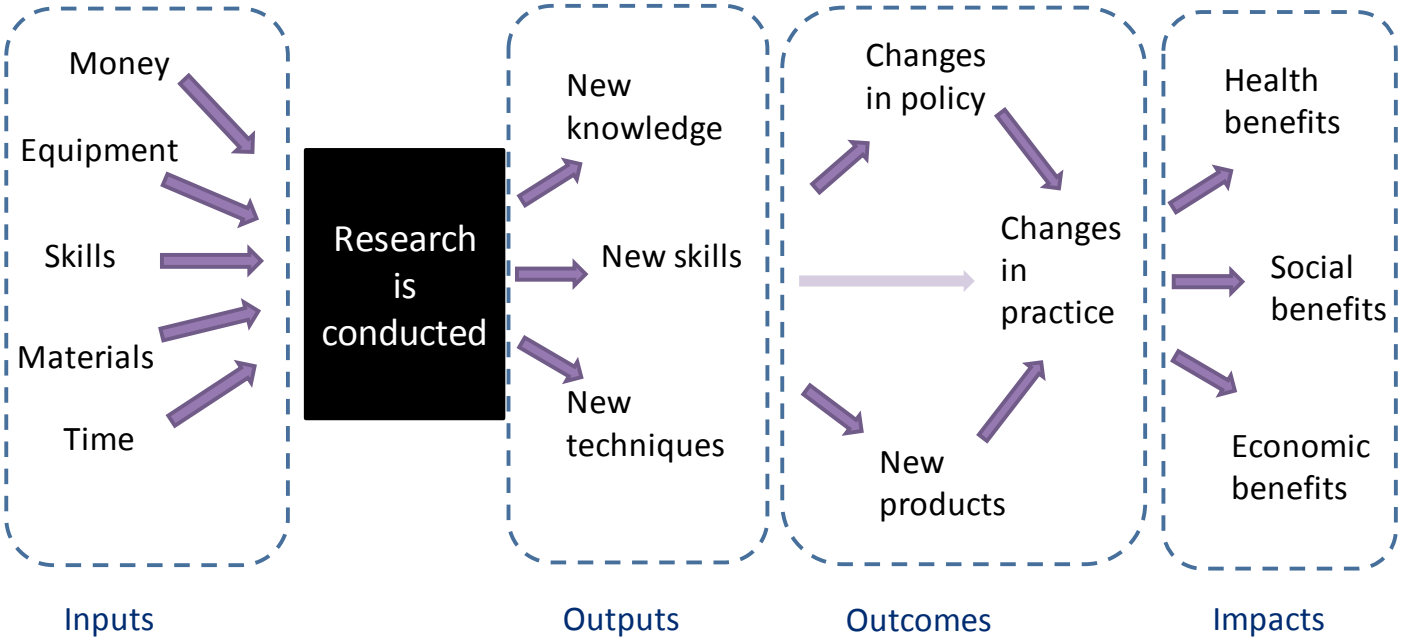

Figure 1. Slide showing a simple model for the research process and examples of the different types of measures (input, output, outcome and process) that could be captured. This is a simplification of the true research process which does not always follow such a linear model and would include many feedback loops.

PURPOSES FOR RESEARCH EVALUATION

Different reasons for undertaking evaluations were discussed based on four key purposes for research evaluation:

- Analysis: To answer questions about 'what works' and make improvements based on data
- Advocacy: To demonstrate the benefits of research to government and society
- Allocation: To support the development of effective investment strategies
- Accountability: To promote responsible management of research funds.

Based on the discussion at the workshops there is not one clear winner in terms of which purpose is most important for medical schools. All the panels had at least one person who valued each of these possible purposes. For the clinical outcomes improvement research panel, and the health equity research panel, the workshop participants were evenly spread across these different purposes, depending on their differing institutional contexts. For the basic research panel, there were many participants who sat between purposes, feeling their institutions had needs that fell into three or all four of these categories.

In addition, the health equity research panel identified a potential additional reason for research evaluation. They suggested that developing a research evaluation framework which places importance on health equity could also play a role in sending the message that health equity is important to that institution and in setting out their aspiration to improve health equity. This use of the evaluation criteria and metrics to communicate the importance of particular areas of research or particular outcomes may also have relevance to other research areas.

Although participants expressed interest in all possible purposes for evaluation, and indeed the purposes are not mutually exclusive, it is important to note that there is some tension between them in some cases. For example, for advocacy you want to showcase the best examples of your research, and how it has supported evidence-based policy making. In contrast, for analysis it is important to look at both success and failure and to understand and learn from the differences. Given this tension and the desire to be able to access evaluation tools appropriate for all purposes, it is important that any research evaluation tool or approach developed for medical schools is flexible, allowing schools to tailor their approach to their particular needs.

STAKEHOLDERS

For one of the exercises during the workshops, the panel members explored the different stakeholders who were important audiences for the research evaluation. All three panels identified funders, the academic community, and institutional leadership/governance as key stakeholders. For the basic research panel, donors were also considered important. For the health equity research panel, the community or constituency which their institution served was also considered an important stakeholder as an audience for their evaluation, as were state and federal government.

Table 1 below indicates how many members of each panel indicated that a particular stakeholder was one of the top three stakeholders most important as an audience for research evaluation in their opinion.

Table 1. Important stakeholder audiences for research evaluation identified by each panel. The number indicates the number of panelists who indicated that the particular stakeholder was one of the top three most important stakeholders.

Stakeholder	Clinical outcomes improvement research	Health equity research	Basic research
Institutional leadership/ governance	6	8	5
Research funders	7	4	5
Academic community/ researchers	7	3	4
Policy decision makers (at state or federal level)	1	6	3
Constituents/ community	1	5	0
Patients and families	1	1	2
Donors	0	0	4
Hospitals/medical centres	1	0	2
Faculty	0	3	0
Learners/trainees	0	2	0
Promotion and tenure committees	0	1	0

It should be noted that although there is some consensus between panels here, the messages which the panels wish to communicate to the different stakeholders, as well as the definition of the different stakeholder groups, may differ between panels and individuals. For example, while the research funders for basic researchers would be primarily NIH, there might be a more diverse range of research funders for the other two panels. In particular, for the clinical outcomes improvement research panel, this is likely to include funding sources internal to the institution. Similarly, institutional leadership and governance will be defined in different ways for different medical schools, depending on the nature and structure of their institution. Within panels, there was some discussion to ensure definitions were shared, or at least that the diversity of definitions was understood, but differences in definition between panels remain.

CHARACTERISTICS OF RESEARCH EVALUATION APPROACHES

Panel members also undertook an exercise considering the characteristics which would be important in a research evaluation framework for medical schools. There was agreement on many of these within and across groups. All panels indicated that they wanted a framework which could handle cross- and multi-disciplinary research and that includes a mix of quantitative and qualitative approaches. All panels thought that collecting longitudinal data is important, and that participant burden needs to be low. All panels also indicated that transparency is important. Scalability was important to the health equity research panel and clinical outcomes improvement research panels, though opinion on this was more spread for the basic research panel. However for many characteristics, there was a spread of responses, perhaps representing differences in requirements depending on the purpose of the evaluation and the relevant stakeholders. There were few systematic differences between groups, indicating that the importance of particular characteristics of frameworks does not seem to depend significantly on the type of research.

The challenge here is that these characteristics are not independent: they are interlinked and trade-offs need to be made between them based on their relevant importance in a particular context. For example, looking at the characteristics which are common across panels as described above, there are already some challenges. Scalability and low participant burden are compatible, as is transparency, and all these characteristics point towards a relatively quantitative, summative approach. However, this type of approach has limitations in its ability to deal with cross- and multi-disciplinary research and developing fixed quantitative metrics for more downstream outcomes and impacts of research can be challenging, which was something in which all panels expressed an interest. It is also important to note that the relative importance of these different characteristics is likely to differ depending on the audience and purpose of the evaluation. For example, if producing information for advocacy to possible donors, scalability and the ability to collect longitudinal data is not likely to be important. However, if collecting ongoing monitoring data across the whole institutional portfolio for institutional leadership, these characteristics will be crucial. This illustrates that even for a particular institution, there is unlikely to be one framework or approach which is universally applicable. The right tools and the right measures, with the right characteristics, need to be considered in the context of the needs of the audience and the purpose of the evaluation.

TOOLS FOR RESEARCH EVALUATION

There are a range of tools available for research evaluation. The majority of these were familiar to the panel members before the workshop, but the breadth of different ways in which they can be used in different contexts and to meet different needs made this a valuable discussion. The panels focused on thinking about how the tools can be applied in their particular research area and what the particular challenges and opportunities are in that context. The discussion also covered innovative applications of the tools that had been or could be attempted. These discussions are covered in more depth in the individual panel reports. Here we present a summary of the overall response to the tools discussed across the panels.[2]

Site visits: All three panels were interested in site visits being used formatively. There was also interest across panels in the possibility of inter-institutional use of site visits to promote shared learning and development, and provide some external objectivity.

Logic models: Panels shared the view that logic models were potentially useful in helping to develop a shared understanding of the goals of research, and of institutions more widely, and strategies for achieving those goals. This could be

[2] For a full description of these tools and their application in research evaluation see Guthrie et al. (2013) "Measuring research: A guide to research evaluation frameworks and tools". RAND Report MG-1217-AAMC. Available at http://www.rand.org/pubs/monographs/MG1217.html

particularly useful with institutional leadership and governance, an important stakeholder across the panels.

Case studies: All panels were fairly positive about the use of case studies. It was felt they were a powerful tool for some (though not all) stakeholders, and in particular could have political clout. However, it was felt that other tools were also needed to give the 'big picture'.

Document review: All panels thought this approach was potentially useful, but expressed concerns around the reliance on what has been recorded in the document. The method is reliant on the quality of that evidence, and it can be hard to know what has not been included in a document, and why.

Data mining: This seemed to be the tool of least interest to the panels, and was only discussed in detail by one of the three groups. The challenge of data quality was raised.

Data visualization: This appealed to the groups as a useful and potentially intuitive way to understand large data sets and make them available to a wider audience. However, it was stressed that the presentation needs to be clear and simple, and tailored to the relevant audience.

Economic analysis: This was considered to be a pragmatic tool across the panels, which is likely to be useful for some audiences, notably institutional leadership. However, concerns were expressed about the challenges in practical implementation, since some things are hard to measure reliably, and one panel questioned the credibility of such analyses.

Bibliometrics: The broad feeling across the panels was that bibliometrics are not going away as part of the research evaluation system, but that the question is how to use them better. All panels expressed an interest in novel ways of using bibliometrics, such as approaches that can incorporate team science, and measures of collaboration using social network analysis approaches.

All panels were given the opportunity to work in groups to start and put their knowledge into practice by developing an outline research evaluation framework which would be appropriate for use by a medical school. A range of interesting ideas and approaches as well as challenges were identified by the groups. One emerging theme across the panels was the importance of viewing the evaluation from the perspective of the relevant stakeholder. This importance of taking the audience for the evaluation into account emerged from all three groups, as well as the importance of engaging these stakeholders as the evaluation approach is developed, and we will use this to guide our development of a resource for medical schools in the next stage of the initiative.

NEXT STEPS

A clear finding across the panels is that there is no one approach or framework that will meet the needs of all medical medical schools. The best approach will differ depending on the institutional context, the stakeholder they are addressing, and the message they wish to communicate to that stakeholder. However, despite this diversity, there are commonalities. Medical schools share a common group of stakeholders and for each stakeholder there are some common messages which medical schools are likely to want to communicate. Given this, there is significant potential to compile and deploy a menu of tools and approaches which medical schools can use as a resource to draw upon. There are also opportunities for shared learning and some collaborative or centralized effort in particular circumstances to reduce the burden on individual medical schools.

Looking forward, AAMC together with RAND Europe will work with the panels in consultation with wider stakeholders amongst AAMC membership and beyond to start to bring together a menu of tools and approaches which can serve as a resource for medical schools in looking to develop and improve their approach to research evaluation. This resource will be piloted over the next year, with the results of the initiative intended to be published in late 2014. In addition, opportunities for medical schools to share experiences and best practice will also be supported.

Based on the findings across the three panels, the intention is that this menu of tools will take a stakeholder-based approach, allowing medical schools to identify one or more key stakeholders, and, considering the information that they want to communicate to that audience, suggested tools and approaches will be supplied based on the needs and interests of that stakeholder. This approach will be refined and developed alongside the panels and with the input of wider stakeholders over the coming months.

In addition, panel members across all three panels expressed a desire to start discussion about these issues at their own institutions. To support this dissemination, AAMC and RAND Europe intend to produce materials, in the form of a slide set and short briefing note, which will be supplied to panel members to brief their institutional leadership and senior faculty about the initiative and about research evaluation more widely. This will be made more generally available as part of the toolkit for institutions once tested and refined through use by panel members.