Where the rubber meets the road: some practical thoughts about analysing evaluation data

Nick Smith, Cancer Research UK
Who are Cancer Research UK?

- The largest fundraising medical research charity in the world
- The largest funder of cancer research in Europe
- The second largest global funder of cancer research
- We are almost exclusively funded through public donations
- The money we raise is spent on research, information, advocacy and public policy
- We fund 45% of cancer research activity in the UK
About me & the CRUK evaluation team

• A team of 4.5 people, which sits alongside the wider strategy team

• I have worked with the team for the last year, running the Researchfish submission and analysis and leading the production of strategic progress dashboards

• Bring a human and sociological perspective to research evaluation data

• Today I want to speak with you about my experience “where the rubber meets the road” in evaluation at CRUK
Interested in near-to-medium term outputs

Publications, collaborations, engagement

IP, spin-outs, medical products

“Lives saved”
Research Strategy Evaluation Dashboards

- Framework outlining strategy (need good strategy for good evaluation)
- Supports evaluation of progress (need “data culture”)
- Uses input, activity and outcome data as measures of progress
- Summary dashboard = quick comparison of strategy areas
- Complements expert review
Existing guides to analysing research evaluation data are numerous but have limitations

- ISRIA
- Metric Tide/ Responsible Metrics
- Snowball Metrics
- Leiden Manifesto
- REWARD
- Funder initiatives
- AMRC & Wellcome events
- Technical literature, e.g. Cronin and Sugimoto’s edited volume, *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*
Schematics can be over-complicated (but also not complex enough)

Theory of Change Perspective on Agricultural Development Interventions
There is an opportunity for evaluators to move beyond performance

• As evaluators working with evaluation data alone, we risk treating researchers as ciphers

• How does the way researchers work, as people, affect research outcomes?

• A social scientific approach to the data helps unlock meaning and value and move beyond performance and towards understanding complexity
The added value of a social scientific approach

• Can handle complex, takes account of the human dimension

• Helps us move beyond evaluating purely for performance

• Informs the generation of meaningful questions and hypotheses

• Hypotheses help us generate nuanced findings about the dynamics of research that are relevant to the concerns and aims of our organisations
Normalised bibliometric data is fairer, but trickier to interpret

**CRUK**

**Articles by Influence Bands**

- Below
- Average
- Highly Cited
- Very Highly Cited

**“Average” for all scientific publications**

**Articles by Influence Bands**

- Below
- Average
- Highly Cited
- Very Highly Cited
There is a lot more in the publications (and collaborations) data than performance

• Finding more than the good, the bad and the indifferent means asking more probing questions

• Can we find evidence for social scientific theories and models about the mechanics of science in our datasets?

• E.g. do we under-value “average” papers? Can funders measure their researchers slow and steady progress in a field? Can we identify paradigm-shifting papers?
Conclusion

• Move beyond counts and volume-based measures of performance

• Engage social science literature on the sociology and anthropology of science to develop organisationally relevant hypotheses to investigate in the data

• Design new measures and combinations of measures that speak to a more complex understanding of the scientific system

• Produce a rich bank of information on our impact on the research landscape and how to modify our impact to achieve organisational priorities
Appendix
Existing guides

**Leiden Manifesto**

1) Quantitative evaluation should support qualitative, expert assessment
2) Measure performance against the research missions of the institution, group or researcher
3) Protect excellence in locally relevant research (allow for variation across academic cultures, e.g. English vs. other language publications)
4) Keep data collection and analytical processes open, transparent and simple
5) Allow those evaluated to verify data and analysis.
6) Account for variation by field in publication and citation practices (normalize)
7) Base assessment of individual researchers on a qualitative judgement of their portfolio
8) Avoid misplaced concreteness and false precision
9) Recognize the systemic effects of assessment and indicators (don’t create perverse incentives)
10) Scrutinize indicators regularly and update them

**The Metric Tide**

- Robustness: basing metrics on the best possible data in terms of accuracy and scope
- Humility: recognising that quantitative evaluation should support – but not supplant – qualitative, expert assessment
- Transparency: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results
- Diversity: accounting for variation by field, and using a variety of indicators to support diversity across the research system
- Reflexivity: recognising systemic and potential effects of indicators and updating them in response.
Challenges

- Time lags

- Attribution and contribution

- Understanding high and low impact when the differences are small and there isn’t consensus on what good looks like

- Ensuring evaluation offers added value

- Identifying the correct unit of assessment when research is multi-disciplinary and has impact in a variety of fields

- Scale: what is the level at which a particular mode of assessment is appropriate.

Solutions

- mixed methods and multi-data sources

- the responsible selection of indicators and metrics

- ISRIA suggests triangulating data sources, using multiple or baskets of data points to highlight a finding.