Metrics: what they are and how to use them

By David Jenkins and Chris Biggs

Open the following in advance:

- *Library website*
- *Web of Science*
- *Scopus*
- *Google Scholar*
- https://github.com/ethanwhite/CV/blob/master/CV.md
- http://www.jonathansilvertown.com/about/
- ERC grant application form

Introductions (Chris & David)

- Introduce ourselves
- Get audience to introduce themselves
  - Acknowledge variety of roles/backgrounds in audience

Caveats (Chris & David)

- We appreciate there will be different levels of existing knowledge
- We will give some basics and background just to make sure we all have the same base level of knowledge
- Hopefully, there will be something in the session for everyone

Why you should care about metrics/the aim of this session (David)

- If you understand how they work, where to get them from and what their pros and cons are then they can help you achieve your goals
  - For example, you can increase your citation counts and show these figures
- They are somewhat controversial and problematic
- But because they are often asked for, the more familiar you are with them, the more time and hassle you'll save
- We aren't here to sell you on metrics but to explore the issues with you and equip you to make your own decisions
- We will look at

- What metrics are
- Where you can get them from
- How you can use them
- What their strengths and weaknesses are
- Recent developments
    - attitudes towards metrics
    - metrics and the REF

Defining metrics (David)

Bibliometrics

- The quantitative analysis of publication and citation data
- For example:
    - Citation counts
    - H-index
    - Journal impact factors
- Bibliometrics are well established
- You can get them form sources such as:
    - Scopus
    - WoS
    - Google Scholar

Altmetrics

- Look at the mentions/uses of research (or anything else) that are not captured by traditional bibliometrics
- In practice, involves looking at how many times research is mentioned, used, saved etc. on
    - Blogs
    - Social media
    - Social bookmarking services (e.g. Mendeley, Connotea, CiteULike) etc.
- It also involves looking at views and no. of downloads
- Altmetrics are newer but are increasingly widespread and increasingly valued
- You can see them in
    - Standalone services
        - Impact Story
        - Altmetric
        - Plum Analytics
    - Embedded in databases
        - Scopus (uses Altmetric)
        - Biomed Central (uses Altmetric)
        - PLOS
        - NPG (uses Altmetric)

Background questions (Chris)

- Who asks you for metrics?
    - E.g. publishers, funders, departments
- How are you using metrics?
- What do you think about metrics?
- What do you want to get out of this session?

Examples of using metrics (David)

- CVs
    - https://github.com/ethanwhite/CV/blob/master/CV.md
- Personal websites
    - http://www.jonathansilvertown.com/about/
- Bids
    - ERC application form

The differences between databases (David)

- Demo search live:
    - Fawcett, Tom. "An introduction to ROC analysis." Pattern recognition letters 27.8 (2006): 861-874.
    - "Higher or lower"?
        - Web of Science - 3505
        - Scopus - 4398
        - Google Scholar – 7054
- Higher citation counts in Google Scholar
    - This is because it indexes more
        - Including theses, websites, blogs, conference proceedings, grey literature, scholarly social networking sites (e.g. Academia), articles in press
    - There are also accusations that the way it counts citations is less accurate because it relies more on automation
        - It counts multiple versions of the same article as separate citation sources
        - There are metadata errors
        - Search for "title of paper"
    - This means it can be gamed

Journal Impact Factor (David)

- JIF is proprietary
    - Made by Thomson Reuters
    - Only found in their products such as JCR
- It is calculated thus:

    - JIF for a journal in 2015 = no. of citations received in 2015 by articles published in the journal during 2013 and 2014
    - No. of articles published in the journal during 2013 and 2014
- Designed to allow you to see how highly cited the average article in a journal is
- It is still widely used

- JIF distributes the citations earned by one article across the whole journal
    - Acta Crystalagraphica A 2008 vs 2009
    - Re: SHELX – in 2008 it had over 3500 citations and the next nearest paper had 4
    - "When Bill Gates walks on a train, everybody is a millionaire on average until he gets off"

  Example and points taken from "The Strange World of Bibliometric Numbers: Implications for Professional Practice" by Dr Ian Rowlands, http://www.slideshare.net/samgray/the-strange-world-of-bibliometric-numbers-implications-for-professional-practice

- JIF was never originally meant to be used to designate quality or impact
- Journal level metrics as a whole are increasingly seen as inadequate and people are calling for a move towards article-based metrics

H-index (David)

- H-index is biased towards more established researchers
- H-index does not take into account subject variations

- H-index can be seen to measure consistency rather than outright impact
- It is possible for people to have different h-index scores when they have the same number of papers and citations

  Example and points taken from "The Strange World of Bibliometric Numbers: Implications for Professional Practice" by Dr Ian Rowlands, http://www.slideshare.net/samgray/the-strange-world-of-bibliometric-numbers-implications-for-professional-practice

- Depends on accuracy and amount of a person's papers that are listed

Altmetrics (David)

- Altmetrics can help address the "evaluation gap"
    - Give a fuller picture of the use of research
    - Citation-based metrics focus on traditional scholarly use
    - Altmetrics can capture non-scholarly use and cover things like evidence-based policy and decision documents
- Should be used alongside traditional bibliometrics
- Can be used e.g. by early career researchers who may not have a high h-index yet
- How meaningful are altmetrics?
    - Use "Evidence for a limit to human lifespan" as example
    - Social media sources display Western bias
    - What does a retweet mean? Look at **AmelS44Uameron**
    - Some papers get high metrics for non-research related reasons

- e.g. Obama paper

The importance of an accurate profile (Chris)

- If you don't have an accurate profile, then your data suffers

1. Author Profile systems – here are some:

2. The benefits of Author Profile Systems:

- To make your research known
- To increase chance of citation
- To correct attributions
- To ensure research is counted in research assessment
- To increase chance of new collaboration
- To increase chance of funding

3. Some profiles already exist, some need to be created, all need to be curated.

- Some profiles are readymades – they exist without your intervention. Other require your intervention to create them – some require more intervention than others.

4. Readymade profiles/OU People profile – these profiles exist even if you don't create them – and herein lies the first problem of Zombie profiles. OU People profile – automatically created profiles can become zombie profiles

5. Readymade profiles in commercial datasets like Scopus. These profiles automatically assign papers by algorithm to readymade author profiles. The algorithm is not always right. Scopus profiles – automatically created based on algorithms. Need to be managed and merged. And papers associated to any profile weeded and potentially added to. 7.3% journal coverage indexed by Scopus.

6. ORCID profiles – need to be created and populated. By creating the right connecting a can be auto-populated with all outputs that have your ORCID assigned to it.

7. Google Scholar – need to be created and managed – relying on automatic updates won't work if you have a common name like John Smith.

   You can add your own papers if they haven't already been indexed by Google Scholar. Note the i10 index metric an example of a home grown metric

   http://ideophone.org/some-things-you-need-to-know-about-google-scholar/

8. ResearchGate – needs to be created and maintained. Great for making connections but RG and Acdaemia.edu are commercial operations and will monetize their network once they work out how!

   Note the ResearchGate score which is a home grown RG metric – RG Score "A good example of a bad metric… ResearchGate thus creates a sort of black-box evaluation machine that

keeps researchers guessing, which actions are taken into account when their reputation is measured." http://blogs.lse.ac.uk/impactofsocialsciences/2015/12/09/the-researchgate-score-a-good-example-of-a-bad-metric/

9. Mendeley – needs to be created and maintained – like RG and Academia.edu sits in a walled garden. Note H-index and citations that appear to be similar/identical to the scores in Scopus.
10. DIY – nice! But probably takes a lot more time to create and curate than some of the out od box models.
11. Different functions have created different systems
12. Different systems have different modus operandi

Uses and Abuses of Metrics (Chris)

- A shortcut to evaluation
- A shortcut to evaluation "Their success is not related to the quality of the information they provide, but more on the facility: a saving of the time necessary to make a real evaluation." (Bibliometric evaluation of individual researchers: not even right... not even wrong!, p.29)
- Funders might use metrics to evaluate bids: Charles Jennings has reported the following example, the Italian Association for Cancer Research compels grant applicants to calculate the impact factor for each journal in which they have published for the preceding five years – cumulative impact factor. (Trends in the Usage of ISI Bibliometric Data: Uses, Abuses, and Implications., p.114))
- Funders might use metrics to allocate monies: e.g. REF
- Metrics were used to support REF 2014 11 of 36 sub panels used citation data to inform their review of outputs. Citation data was sourced from Scopus. Panels and HEIs were provided with contextual data to support them in interpreting citation counts.
- On the REF, the Higher Education Green Paper... controversially calls for "making greater use of metrics and other measures to 'refresh' the REF results and capture emerging pockets of research excellence in between full peer review". THE, November 6th, 2016
- Metrics and the Stern Review: Recommendation 4: Panels should continue to assess on the basis of peer review. However, metrics should be provided to support panel members in their assessment, and panels should be transparent about their use.
- Metrics and a light touch REF – discussions in and around government late last year around a metrics led REF. THE article on an Elsevier analysis; Government Green paper for Higher Education; HJEFCE/BIS Contract for provision of citation indicators for research outputs 2008-14
- Metrics and the Stern Review - has been informed by The Metric Tide report (Independent Review of Metrics for Research Assessment) which emphasised the need for peer review. One recommendation of the Metric Tide was "In assessing research outputs in the REF, it is not currently feasible to assess research outputs or impacts in the REF using quantitative indicators alone." http://www.hefce.ac.uk/pubs/rereports/Year/2015/metrictide/Title,104463,en.html
- Universities might use metrics to recruit and promote: In Spain, a university determines promotion by multiplying the number of articles published with the impact factor. (Trends in

the Usage of ISI Bibliometric Data: Uses, Abuses, and Implications., p. 114)

- Journals might encourage authors to increase reference lists to increase the impact factor of a particular journal "Impact Factor Wars"

- Individuals will seek to publish in Journals with high impact factors: The prestige of appearing in the major journals has led the Chinese Academy of Sciences to pay successful authors the equivalent of $30,000 (£18,000). Some researchers made half of their income through such "bribes", Schekman said in an interview.
  https://www.theguardian.com/science/2013/dec/09/nobel-winner-boycott-science-journals

- Guardian article "A paper can become highly cited because it is good science – or because it is eye-catching, provocative or wrong. Luxury-journal editors know this, so they accept papers that will make waves because they explore sexy subjects or make challenging claims. This influences the science that scientists do. It builds bubbles in fashionable fields where researchers can make the bold claims these journals want, while discouraging other important work, such as replication studies."

  What is the impact of this! It's suggested that this can lead to a change in research topics and hinder the development of minority research areas as the dominant research areas become defined by these journals with the highest impact factor.

- 5 years to get a paper published - Enrique Martin-Blanco, a principal investigator at the Molecular Biology Institute of Barcelona, claimed that anonymous peer reviews from "extremely opinionated physicists" tried to censor his paper. He added that the research provided an alternative to an established methodological analysis that has been used for years by some physicists, who did not like the idea."

- Individuals might seek to increase citations by self-citing and citation cartels: "To a certain extent, author self-citations are natural, as researchers usually build on their own previous research. However, in the context of research evaluation, where citations are used as a proxy for impact on the scientific community, self-citations are problematic as they do in fact not mirror influence on the work of other researchers and thus distort citation rates." (The use of bibliometrics for assessing research: Possibilities, limitations and adverse effects., p148)

  Citation cartels… where authors agree to cite each other's papers

- Individuals might seek to increase number of publications by salami slicing - what might be contained in a single research paper is spread out over several thereby artificially inflating the output of researchers.

- Individuals might seek to increase number of publications and citations as guest or gift authors.
  Guest author - those who do not meet accepted authorship criteria but are listed because of their seniority, reputation or supposed influence;
  Ghosts - ghost authors are those who meet authorship criteria but are not listed – this might obscure the participation of researchers who may have a conflict of interest
  Gifts - authors are those who do not meet accepted authorship criteria but are listed as a personal favour or in return for payment

- Your (real) impact factor http://www.phdcomics.com/comics/archive.php?comicid=1108

- The Metric Tide and Responsible metrics. Foreword of the report:
  "Too often, poorly designed evaluation criteria are "dominating minds, distorting behaviour and determining careers." At their worst, metrics can contribute to what Rowan Williams, the former Archbishop of Canterbury, calls a "new barbarity" in our universities. The tragic case of Stefan Grimm, whose suicide in September 2014 led Imperial College to launch a review of its

use of performance metrics, is a jolting reminder that what's at stake in these debates is more than just the design of effective management systems. Metrics hold real power: they are constitutive of values, identities and livelihoods."

- The Metric Tide and Responsible Metrics. "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 range of indicators to reflect and support a plurality of research and researcher career paths across the system. Reflexivity: recognising and anticipating the systemic and potential effects of indicators, and updating them in response."

- Responsible Metrics and The Leiden Manifesto – 10 principles to guide research evaluation http://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351

- Responsible Metrics and the San Franscisio Declaration on Research Assessment. "One central theme of DORA is : "the need to eliminate the use of journal-based metrics, such as Journal Impact Factors, in funding, appointment, and promotion considerations" http://www.ascb.org/dora/

Conclusion (Chris & David)

- What are your thoughts about metrics at this stage?
- What role do you think metrics should have in evaluating research?
- How would you like to go away and use them?
- What further help would you like?