Objective and numerical method of finding importance of medical research and prioritizing grants and publications

Jerzy Dyczkowski, Ph.D.

Independent consultant in data science

Kappeliweg 31,

4132 Muttenz, Switzerland

email: jd.datascience@gmx.ch

CLASSIFICATION

Social Sciences: Social sciences

Biological Sciences: Medical Sciences

KEYWORDS

research, public health, innovation, decision making, review

SIGNIFICANCE

Putting worth on research and selection of studies by importance are crucial in medical innovation. Practical applications include choosing personal study topics, publication review, study grant selection, and decisions of spending or misspending billions in public health. Multiple studies raised alarm that current methods perform poorly in reproducibility, prediction of best research and objectivity. I propose using the metrics how much disease burden is reduced and calculating objective, numerical research value. The concept is that worth of medical research is not subjective but can be reproducible and numerically quantified. The method increases transparency by giving decision makers an externally accountable proof, and frees peer reviewers to check scientific integrity. Its numerical form can capture small differences important in competition between studies.

ABSTRACT

Finding value and selecting knowledge by importance are crucial in medical innovation. Applications include individuals designing research, funding organizations selecting grants, journals – publications, institutions – priorities in public health and health policy, and decision makers spending or misspending billions of research funds. Currently finding value of knowledge is done by peer review together with checking scientific integrity. Multiple studies raised alarm that it performs poorly in prediction of highest citations, bias, transparency and quality. The resulting problems include perception of slow medical progress and wasting funds and time. I introduce a standard, objective and numerical method for finding value of medical research. It measures disease burden prevented by new knowledge contained in a study or a publication. In its simple form, it is calculated by multiplying disease prevalence, disease burden, and efficacy of the therapy. It can be modified for risk of failure, multi-disease effect and for ethical considerations. The process is described step-by-step in terms common in medical practice. A quick estimate is often sufficient. The advantage is objectivity, since it is calculated from real world data. This gives transparency and externally accountability of decision making. The second advantage is a numerical form. This can measure small differences in research value which, in sharp competition, determine which studies are selected. A researcher can calculate the value of own future effort. Institutions might ask to provide it at submission. The method is also applicable to broad policy analysis, objective evaluation of scientific achievement and bibliometric studies.

ARTICLE

Introduction

Important task in medical innovation is finding value of medical knowledge or selecting more versus less important research (1-8). Researchers do it whenever they choose a study topic. Further practical applications are selecting manuscripts in academic publishing (9,10), selecting grants during science funding review (2-6,11-14), and in choosing research priorities and trying to reduce waste in public health and health policy (1). The results are decisions to spend or misspend tens of billions of research funding annually.

Currently finding value of medical knowledge is partially subjective. Institutions and journals use peer review, which also checks for scientific integrity. However, multiple studies in the last two decades raised alarm that peer review is remarkably inefficient in estimating value of research (5,6,8,10,11-15). Quality of prediction which research becomes most valuable, by number of publications, citations (2,3,6,14) or several other metrics (13,14) is low (2,3,6,10,13,14). Journals frequently reject papers which become groundbreaking after publication elsewhere (10). Reproducibility of opinions between reviewers and journals is low, too (5,8,10,11-13). Common accusations are gender bias (13, 16), bias against novel topics and directions of research (5,9,13), and lack of transparency (1,3,8,13,15). Resulting problems include journals struggling to improve content, perception of wasting medical spending and science funding (1,3,13,15), and time at every level, from an individual researcher to nation-wide policies, perception of too slow progress in medicine (1,9,13), and poor image of medical research both within the research community and in the public (9).

Finding value of medical research additionally lacks precision (6). Sharp competition results in that accepted are only 10-15% of eligible research grants (6,18,19), and 4-10% manuscripts in top journals (20-23). Later, one accepted publication can spawn a multi-million grant, and a one grant or a publication can make or break a career of a researcher (14,24,25). Because small differences in quality determine which studies fall into the small minority which becomes successful (6,14), a

quality of the evaluation which is only better than random on average (6) to moderate (4,7) is not sufficient (1,7,9).

The weakness of peer review became well documented during two decades, but it persists because no better solution was found. Suggested modifications included changing selection or motivation or reviewers (1,7), reviewing by a larger group or a community (26), or modifying human reviews by scoring methods or algorithms (27). None of these became universally accepted. They all share the inherent weakness: they are based on personal opinions (8,13,15), which suffer from biases, including unconscious ones (17), and therefore are prone to be challenged (13). This also means that a branch of business which manages multi-billion funds, which medical research also is, has relatively poor main metrics (1,13). Other branches of business developed more detailed and stricter standards (e. g. 28).

RESULTS, MATERIALS AND METHODS

The concept

I propose to evaluate biomedical knowledge using a metrics how much disease burden it prevents. This comes from the prime objective of medicine: protecting human life and health. The result is that value of knowledge for medicine is not a subjective human opinion, but becomes grounded in objective medical facts. Further, it can be measured quantitatively and potentially precisely.

The concept is perhaps most intuitively understood by an example. Imagine that a medical doctor reads two reports about two new therapies. The therapy A allows saving lives of 100 people. The therapy B allows saving lives of 200 people. This is a simplified example and all other factors are equal. The doctor could reasonably say that the piece of knowledge B is exactly twice as valuable as

A. This shows, that the value of knowledge for medicine can be based on facts in the real world and that it can be measured in exact numbers.

The definition of objective, numerical research value

For semantic clarity: the research value is not the same as monetary value. 'Knowledge' means knowledge contained in research studies, grant applications, publications, and manuscripts. All these are pieces of knowledge or future knowledge, which are evaluated by science reviewers using broadly similar criteria. I propose to derive objective and numerical research value I. In its simplest form:

$$I_t = p_t * b_t * e_t$$

where *I* - Numerical research value of new knowledge *t, p* - disease prevalence, *b* - disease burden, *e* - efficacy (established or expected) of improved therapy due to the new knowledge *t*.

Extension to risk of failure and multiple fields of research

Numerical research value can also include the possibility of failure of a study. For example, clinical trials in large majority do not result in an approved therapy (29-31). Therefore the research value of a clinical trial should be multiplied by the probability of success of clinical trials at a given stage. This way, the risk of failure can be quantified or at least estimated. Preclinical research, pure science research and one-off reports have even smaller probability of becoming an approved therapy. Probability of a drug candidate to pass to a phase I of trial is on average 0.29–0.35 (32-34), varying from 0.23–0.7 between disciplines (34). I did not find a number how many pure science and one-off reports result in a practical therapy. As an interim solution I suggest calculating value of a

preclinical research as 0.3 of a phase I clinical trial in the same discipline. This, however, is likely an overestimate.

Numerical research value could be also modified to studies applicable to multiple diseases or fields of medicine. For example, drugs for one type of cancer are often active against other types. Authors can accommodate a well founded expectation that one study can be applicable to several diseases. I did not found the numbers in literature how often such extensions occur in general. An ideal general modifier here could be a number of therapies found extensible between the disease areas, divided by all attempts (including the number of therapies tried and failed).

$$I_t = \Sigma_{1..n} (r_{tdi} * p_{tdi} * b_{tdi} * e_{td\,i})$$

where I - Numerical research value of a new knowledge t, $d_{1..}d_n$ – diseases 1 to n, r – chance of successful passing to a practical therapy, p - disease prevalence, b - disease burden, e_{tdi} - efficacy (established or expected) of the therapy in disease d_i due to new knowledge t. When knowledge helps in one disease, and the therapy is proved to be successful (probability = 1), the formula simplifies to the previous formula.

In many practical situations, a quick estimate of research value is enough for a decision. For example, often a researcher needs to choose between two options. Then, quick estimate of key factors, and knowing that some are very much bigger in one case, is sufficient for a decision.

In some cases, exact numbers are impossible to get and can be replaced by estimates. Nevertheless, such narrowing down the uncertainty in the well understood way is still helpful for decision.

Examples of use

The objective, numerical research value formalizes the general understanding that some medical discoveries are more important for human health by themselves, rather than by subjective appreciation of doctors. It allows to quantify and precisely describe many issues in medicine. Progress in common and serious diseases (large *p* and *b*) is valued more, exactly in proportion to their real world occurrence and severity. In the same way, therapies with more efficacy *e* are given more value. Discoveries applicable to many ailments *n* are valued more, too.

The concept addresses the problem of neglected diseases, and on the other hand the problem of throwing too much resources on popular topics. When more research deals with a disease, the disease burden falls, and less common diseases become more urgent. The numerical value can precisely quantify how much a neglected disease is under-funded and when a medical problem becomes over-hyped.

Effort already put into research can be included, too. A researcher entering a popular field faces a risk that competitors will solve the problem. Unless the researchers expect a synergy, the law of diminishing returns applies. In this case, the expected research value might be divided between the working groups in the field, or by estimated probability of success. This avoids the possible mistake that 100% researchers and funds would go to the single most common disease.

Calculations can be also applied to the choice of several not exclusive options. An institution which can pick several or a portfolio (3) of projects, could use methods similar to constructing an optimal portfolio of options (35).

METHODS: A practical step-by-step guide to objective research value of a manuscript or a study

Note, as said above, that a quick estimate is often enough in practice.

1. Choose the metrics.

An easy mistake would be to trying to compare the incomparable, for example lifetime cases worldwide with cases per year in the USA. The metric should be common to all compared cases, appropriate to the topic and objective. The metric can include impact of the disease (mortality, quality-adjusted life years, disability-adjusted life years, financial cost, other), time frame (lifetime occurrence, occurrence per year, or other) and geographical scope (worldwide, in a country etc).

A researcher can define the metrics for oneself. A journal or a grant committee could select the metric for the field of study and ask the submissions to use it, unless a good reason is given.

Generally, the criteria of a metric mentioned earlier in the above lists are preferable to latter ones, for example mortality is a better metrics than life quality. However, the latter criteria are better in some disciplines, for example for non-lethal diseases. The appropriate recalculations could be used there. However, in some cases, a simple recalculation may be inappropriate for the impact of the disease, because the opinion in society is that life has no absolute priority over quality of life. Here the metrics partially depends on subjective ethical considerations.

2. Obtain the burden of the disease in concern according to the metrics. If an effect is measured over the existing therapy, a correction is needed.

3. Calculate the numerical research value itself – the disease burden lowered. Multiply the burden of the disease by efficacy of the therapy (actual or expected). For example, if mortality was reduced twofold, divide the disease burden by two.

4. For the early stage research, estimate probability of passing to the therapy. In case of clinical trials, divide by a rate of failure of trials appropriate for the discipline and type of study (29-31). For a pre-clinical research or single case study, the worth is even lower (32-34). If no data are available, I suggest an estimate of 0.3 of a phase I of a trial but it can be an overestimate.

5. Theoretically, the research value could be modified by applications outside the field of study. For example, drugs for one type of cancer are often active against other types. If there is a well-founded belief that it is possible, value of other areas can be added. However, I did not found the numbers in literature how often such extensions occur. An ideal modifier here would be a number of therapies found extensible between the disease areas, divided by all attempts (including the number of therapies tried and failed).

The above guide could be modified appropriately to a medical area.

Caution is needed especially because raw values, e.g. efficacy of a therapy are often imprecise and contain a margin of error. Therefore researchers should take special care of data quality, and be aware that small differences of value can be artifacts. The numerical research value potentially allows complex calculations with great precision, for example an optimal research strategy in multiple medical fields, but the suitably precise raw data may not be available. However, even when exact calculation is impossible or impractical, the method can provide brackets of estimation, or a relative choice of one option before another.

The caution should be taken also in case of non-clinical research, which has no easily measurable effect on curing disease. I propose two solutions. One is comparing research value only between studies in preclinical research, and avoiding comparing preclinical with clinical research. An

alternative is carefully formulating how exactly a non-clinical research results in lower disease burden. For example, a new diagnostic procedure can be more precise, and lower number of medical errors, which measurably reduces disease burden. Or a new diagnostic procedure can be faster, so allow faster therapy and increased survival.

Ethical considerations could warrant modification. Note, however, that the ethical principle of lowering disease burden is already the core basis of the method. Particular ethical considerations can often be helped by calculations of research value. As an example, it is possible to calculate which patients are neglected by the current medical research. To calculate what help already is available for a particular disease, concentrate on the point 2. above. Calculate the burden of the disease if no therapy was used. Calculate separately what proportion of this burden is lowered by the existing therapies. This shows objectively which diseases are neglected and how much. Particular effort should be directed at these. Such efforts might fit into the general medicine portfolio (see above, 3, 35) as a low-hanging fruit, where small, targeted effort can produce big return.

Discussion

Advantages of the concept

The concept that the value of a medical discovery can be independent of a personal opinion, and can be a precise number can be surprising to some readers. The method itself is a scientific novelty. However, there has always been an understanding in society that some medical discoveries are objectively very important, and some not so. For example, it would be difficult to find a person who thinks that cure for cancer is less important than cure for dandruff. The objective research value validates and formalizes this imprecise understanding.

The practical advantage of this approach is objectivity. It uses real world medical facts which are externally verifiable and have relatively limited scope for interpretation. The value is reproducible, in the sense than every researcher using the same raw data on disease burden, chance of success of clinical trials etc. should arrive at the same number. This makes decisions explainable, justifiable and externally accountable.

Another practical advantage is that the value is quantitative, so potentially very precise. The value can contain a margin of error, which can be also quantified. Currently decisions are made using qualitative adjectives like 'important', 'very important', 'breakthrough', or essentially appeals to majority: 'of big general interest', 'considered important' etc. Compared to these, numerical research value allows making decisions with much more clarity.

Potential applications

Objective research value can especially help an individual researcher, who wants a personal guide to which study is worth undertaking and which publications to read. Entering a vast field of medicine, with only too many possible research topics and more published papers than time to read, a researcher can make decisions with unusual clarity.

Further applications are publication review and science funding review of research grants. The method greatly increases objectivity and reduces human bias, providing decision makers with transparency and external accountability of their actions, important for example in justifying medical spending. Journals and funding institutions could calculate the research value by themselves, or ask submissions to include the pre-calculated research value and raw numbers used. These bodies could also decide to produce guidelines to the authors containing, for example,

preferred metric to use, e.g. saved mortality annually worldwide, or quality-adjusted life years in the U.S. This would reduce the burden on peer reviewers who are notoriously overworked (7,10,13). In such setup, peer reviewers could check faster whether the research value was calculated properly, and would focus on scientific integrity: whether the research is scientifically valid, methods will produce the results, results support conclusions etc.

The method is applicable for studies trawling medical research for the purpose of data science in metaresearch, bibliometrics, scientometrics and science of science. Interestingly, it is possible to compute such a thing as medical value of all knowledge in human health. It becomes also possible to directly compare research from very different branches, for example cardiology with oncology, as long as the metric is the same. The formula is therefore ideal for calculating and analyzing contribution to medical progress of different branches of medicine, countries, methods, time periods etc. The value can be also used as an objective metrics in general research: in cost analysis and more broad science pricing and research pricing in health economics, in public health and health policy. The value might also be adopted, together with citations ranks, impact factor etc., for research quality assessment, evaluating personal scientific output, academic productivity and scientific achievement.

References

1. I. Chalmers, M. B. Bracken, B. Djulbegovic, S. Garattini, J. Grant, A. M. Gülmezoglu, D. Howells, J. P. Ioannidis, S. Oliver, How to increase value and reduce waste when research priorities are set. *Lancet*, 383, 156–165 (2014).

2. N. Danthi, C. O. Wu, P. Shi, M. Lauer, Percentile ranking and citation impact of a large cohort of National Heart, Lung, and Blood Institute-funded cardiovascular R01 grants. *Circ. Res,* 114, 600–606 (2014).

3. S. A. Gallo, A. S. Carpenter, D. Irwin, C. D. McPartland, J. Travis, S. Reynders, L. A. Thompson,

S. R. Glisson. The validation of peer review through research impact measures and the implications for funding strategies. *PLoS One*. 9, e106474 (2014).

4. D. Li, L. Agha, Research funding. Big names or big ideas: do peer-review panels select the best science proposals? *Science*. 348, 434–438 (2015).

5. K. J. Boudreau, E. C. Guinan, K. R. Lakhani, C. Riedl, Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science. *Manage Sci*. 62, 2765–2783 (2016).

6. F. C. Fang, A. Bowen, A. Casadevall, NIH peer review percentile scores are poorly predictive of grant productivity. *ELife*, 5, e13323 (2016).

7. S. Copiello, On the money value of peer review. *Scientometrics*. 115, 613–620 (2018).

8. E. L. Pier, M. Brauer, A. Filut, A. Kaatz,J. Raclaw, M. J. Nathan, C. E. Ford, M. Carnes, Low agreement among reviewers evaluating the same NIH grant applications. *Proc Natl Acad Sci U S A*. 115, 2952–2957 (2018).

9. B. Alberts, M. W. Kirschner, S. Tilghman, H. Varmus, Rescuing US biomedical research from its systemic flaws. *Proc Natl Acad Sci U S A*. 111, 5773–5777 (2014).

10. K. Siler, K. Lee, L. Bero, Measuring the effectiveness of scientific gatekeeping. *Proc Natl Acad Sci U S A*. 112, 360–365 (2015).

11. S. Cole, J. R. Cole, G. A. Simon, Chance and consensus in peer review. *Science*. 214, 881–886 (1981).

12. N. E. Mayo, J. Brophy, M. S. Goldberg, M. B. Klein, S. Miller, R. W. Platt, J. Ritchie, Peering at peer review revealed high degree of chance associated with funding of grant applications. *J Clin Epidemiol*. 59, 842–848 (2006).

13. S. Guthrie, I. Ghiga, S. Wooding. What do we know about grant peer review in the health sciences? *F1000Res*. 6, 335 (2017).

14. S. A. Gallo, S. R. Glisson, External Tests of Peer Review Validity Via Impact Measures. *Front. Res. Metr. Anal.* 3, 22 (2018).

15. Y. Bhattacharjee, Science funding. NSF's 'Big Pitch' tests anonymized grant reviews. *Science*. 336, 969–670 (2012).

16. Helmer M, Schottdorf M, Neef A, Battaglia D. Gender bias in scholarly peer review. *Elife*. 2017;6:e21718. Published 2017 Mar 21. doi:10.7554/eLife.21718

17. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. *BMC Med Ethics*. 2017;18(1):19. Published 2017 Mar 1. doi:10.1186/s12910-017-0179-8

18. European Commission. Horizon 2020 In Full Swing -Three Years On – Key facts and figures 2014-2016, Fig. 8. https://ec. europa. eu/programmes/horizon2020/en/horizon-2020-statistics (2018, accessed Jun 9, 2020).

19. Research Project Success Rates by NIH Institute for 2018. Published online. https://report. nih. gov/success_rates/Success_ByIC. cfm (accessed Jun 8, 2020).

20. Science. The *Science* Contributors FAQ https://www.sciencemag.org/site/feature/contribinfo/faq/#pct_faq (accessed Aug 8, 2020).

21. The Lancet. Information for Authors. https://els-jbs-prod-cdn.jbs.elsevierhealth.com/pb/assets/raw/Lancet/authors/tl-info-for-authors.pdf (accessed Jun 8, 2020).

22. Editorial criteria and processes. *Nature*. https://www.nature.com/nature/for-authors/editorial-criteria-and-processes (accessed Jun 9, 2020).

23. JAMA Network. About JAMA. https://jamanetwork.com/journals/jama/pages/for-authors (accessed Jul 8, 2020).

24. C. Bloch, E. E. Krogh Graversen, H. Skovgaard Pedersen, Competitive Research Grants and Their Impact on Career Performance. *Minerva* 52, 77–96 (2014).

25. T. Bol, M. de Vaan, A. van de Rijt, The Matthew effect in science funding. *Proc Natl Acad Sci U S A* 115, 4887–4890 (2018).

26. D. M. Herron, Is expert peer review obsolete? A model suggests that post-publication reader review may exceed the accuracy of traditional peer review. *Surg Endosc*. 26, 2275–2280 (2016).

27. J. Bollen, D. Crandall, D. Junk, Y. Ding, K. Börner, An efficient system to fund science: from proposal review to peer-to-peer distributions. *Scientometrics*. 110, 521–528 (2017).

28. International Accounting Standards Board. International Financial Reporting Standards 2007 (including International Accounting Standards (IAStm) and Interpretations as at 1 January 2007). LexisNexis, Available online at: https://www. iasplus. com/en/standards/standards (2007).

29. M. Hay, D. W. Thomas, J. L. Craighead, C. Economides, J. Rosenthal, Clinical development success rates for investigational drugs. *Nat Biotechnol*. 32, 40–51 (2014).

30. C. H. Wong, K. W. Siah, A. W. Lo, Estimation of clinical trial success rates and related parameters. *Biostatistics.* 20, 273–86 (2019) [published correction appears in *Biostatistics*. 20, 366 (2019)].

31. D. W. Thomas, J. Burns, J. Audette, A. Carroll, A. Dow-Hygelund, M. Hay, Clinical Development Success Rates 2006-2015. BIO industry analysis. https://www.bio.org/sites/default/files/legacy/bioorg/docs/Clinical%20Development%20Success%20Rates%202006-2015%20-%20BIO,%20Biomedtracker,%20Amplion%202016.pdf (2016, accessed Jun 9, 2020).

32. W. Bains, Failure rates in drug discovery and development: will we ever get any better? *Drug Discover World* 2004, Fall. https://www.ddw-online.com/business/p148365-failure-rates-in-drug-discovery-and-development:-will-we-ever-get-any-better. Html (2004, accessed Jun 6, 2020).

33. S. M. Paul, D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg, A. L. Schacht, How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov*. 9, 203-214 (2010).

34. T. Takebe, R. Imai, S. Ono, The Current Status of Drug Discovery and Development as Originated in United States Academia: The Influence of Industrial and Academic Collaboration on Drug Discovery and Development. *Clin Transl Sci*. 11, 597-606 (2018).

35. Z. S. Galis, W. K. Hoots, J. P. Kiley, M. S. Lauer, On the value of portfolio diversity in heart, lung, and blood research. *Am J Respir Crit Care Med*. 186, 575–578 (2012).