EDITORIAL

Qualitative versus Quantitative Evaluation of Scientists' Impact: A Medical Toxicology Tale

REZA AFSHARI[1,2,*], SEYED MOSTAFA MONZAVI[1]

[1] Addiction Research Centre, Mashhad University of Medical Sciences, Mashhad, Iran
[2] Division of Basic Medical Sciences, Academy of Sciences of Islamic Republic of Iran, Tehran, Iran

Abstract

Evaluation of scientists working in a specific area of science is necessary, as they may strive for same limited resources, grants and academic promotions. One of the most common and accepted methods of assessing the performance and impact of a scientist is calculating the number of citations for their publications. However, such method suffer from certain shortcomings. It has become more and more obvious that evaluation of scientists should be qualitative in addition to quantitative. Moreover, the evaluation process should be pragmatic and reflective of the priorities of an institution, a country or an intended population. In this context, a scoring scale called "360-degree researcher evaluation score" is proposed in this paper. Accordingly, scientists are evaluated in 5 independent domains including (I) science development, (II) economic impact, (III) policy impact, (IV) societal impact and (V) stewardship of research. This scale is designed for evaluation of impacts resulted from research activities and thus it excludes the educational programs done by a scientist. In general, it seems necessary that the evaluation process of a scientist's impact moves from only scintometric indices to a combination of quantitative and qualitative indices.

Keywords: Comparative Effectiveness Research; Cost-Benefit Analysis; Employee Performance Appraisal; Journal Impact Factor; Social Change

How to cite this article: Afshari R, Monzavi SM. Qualitative versus Quantitative Evaluation of Scientists' Impact: A Medical Toxicology Tale. Asia Pac J Med Toxicol 2014;3:134-140.

INTRODUCTION

Evaluation of scientists working in a specific area of science is necessary, as they may strive for same limited resources, grants and academic promotions. Nevertheless, determining the performance parameters of a scientist has always been under debate (1). One of the most common and accepted methods of assessing the performance and impact of a scientist is calculating the number of citations for their publications (1,2). In that sense, the h-index proposed by Hirsch in 2005 has gained popularity (3). However, such method suffer from certain shortcomings. For example, the value of h-index is time dependent and relies upon the database used for calculating the number of citations (1). In addition, many scientists from the developing countries speak and produce science in non-English languages which are published in journals in their own language. These journals are less likely to be taken into account in major citation systems.

For some scientists, success resides within publishing more articles, gaining more citations and having higher h-index (4). However, it should be kept in mind that the impact a research leading to a successful change in health system policy or a new invention may have on improvement of public health, a highly cited paper may not. Besides, creating a meaningful change in system policies or invention of a new device takes more time and involves more people than studying on a specific subject matter and publishing its results. Therefore, it has become more and more obvious that evaluation of scientists should be qualitative in addition to quantitative (2). Moreover, it should be pragmatic and reflective of the priorities of an institution, a country or an intended population (1,2). In this context, in addition to scientometric analysis, other aspects for evaluation of scientist's impact have been suggested to be considered including research applicability and integrity, research economical rate of return, assistance in peer review of research papers and grant applications, benchmarking and capacity building (1,2,5-7). In the present paper, a scoring scale has been proposed which can be used as a prototype for evaluation of scientist's impact both qualitatively and quantitatively.

Details of the scoring scale for evaluation of scientist's impact

According to the proposed scoring scale, which can be called "360-degree researcher evaluation score", scientists are evaluated in 5 independent domains including (I) science development, (II) economic impact, (III) policy impact, (IV) societal impact and (V) stewardship of research (Table 1). For each domain the maximum score is determined to be 20. The domains contain many items that are categorized under subdomains. The allocated scores for items are researcher-driven, and further investigations are needed to be standardized.

*Correspondence to: Reza Afshari; MD, PhD, MPH. Associate Professor; Addiction Research Centre (ADRC), Imam Reza Hospital, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.
Tel/Fax: +98 51 3852 5315, E-mail: afsharir@mums.ac.ir

Table 1. Scoring scale for evaluation of scientist's impact: "360-degree researcher evaluation score"

Domain	Item	Definition	Max. Score	How to measure
Science Development (Scientific Output)			**20**	
	H-Index	Number of papers of a given scientist with a number of citations equal or greater than h	3	Rank of the h-index of a scientist working in the target institution in an intended period of time
	Number of citations	Total number of citations for published articles of a scientist	1	Rank of the number of citations of a scientist working in the target institution in an intended period of time
	Mean number of citations minus self-citations	Mean number of citations minus self-citations per paper	1	Rank of the mean of number of citations minus self-citations per paper of a scientist working in the target institution in an intended period of time
	Reciprocal value of standard deviation (SD) of number of citations minus self-citation	1 divided by the SD of mean of citations minus self-citations per paper	0.5	Rank of the reciprocal value of SD of number of citations minus self-citation of a scientist working in the target institution in an intended period of time
	Number of downloads of articles	Number of downloads of published articles	0.5	Rank of the total number of downloads of articles published by a scientist working in the target institution in an intended period of time
	Invention / Patent	Inventions that received patent number	3	For each patent 1 score can be allocated
	Number of publications	Any trackable publication		
	Book	Textbooks published by an academic publication or a well-known international publisher		
	Editor	Editor of a textbook	3	For each textbook 1 score can be allocated
	Author of a chapter	Author of a chapter in a textbook	2	For each textbook 0.5 score can be allocated
	Journal articles	Papers accepted or published in peer reviewed journals		
	Meta-analysis	Articles published in the format of meta-analysis or systematic review	2	For each meta-analysis or systematic review, the first or corresponding author receives 1 + journal impact factor (IF)/30 and other authors receive 0.25 + journal IF/30 score
	Review / Original	Articles published in the format of narrative review or original article	2	For each original or narrative review article, the first or corresponding author receives 0.5 + journal IF/30 and other authors receive 0.125 + journal IF/30 score
	Case report/ Editorial/ Letter to editor	Articles published in the format of case report, editorial, letters to editor	1.5	For each case report or editorial or letter to editor, the first or corresponding author receives 0.25 + journal IF/30 and other authors receive 0.063 + journal IF/30 score
	Conference presentations	Oral or poster presentation in internationally supported conferences	0.5	1 oral presentations or 5 poster presentations receive 0.25 score
Economic impact			**20**	
	Research grant productivity	Mean of (citations - self citations) / grant funding received from the institution	3.5	Rank of the research grant productivity of a scientist working in the target institution in an intended period of time
	Cost benefit	Updated methods and interventions that result in fulfillment of the objectives of health administration or training healthcare staff or preventing, diagnosis, treatment, and monitoring health conditions at a cost lower than the historical cost or the projected cost		Rank of money saved (or possibly can be saved) in health administration, training healthcare staff or preventing, diagnosis, treatment, and monitoring health conditions resulted from the methods and interventions developed by a scientist working in the target institution in an intended period of time assessed subjectively by the expert committee
	Administration	Health or healthcare administration	1.25	
	Education	Medical education	1.25	
	Prevention	Preventive measures in health	1.25	

Table 1. *(Continued)*

Domain	Item	Definition	Max. Score	How to measure
	Treatment	Medicines, therapeutic guidelines (protocols), rehabilitation methods, therapeutic measures or instruments	1.25	
	Diagnosis	Diagnostic guidelines (protocols), methods and instruments	1.25	
	Food and environment	Improvement in the quality and safety of food or environment	1.25	
Rate of return		Time to achieve a profitable impact from a research	0.75	Rank of the rate of return (or possible rate of return) of a cost beneficial research performed by a scientist working in the target institution in an intended period of time assessed subjectively by the expert committee
Cost effectiveness		Effectiveness of updated methods and interventions that result in fulfillment of the objectives of health administration or training healthcare staff or preventing, diagnosis, treatment, and monitoring health conditions in association with the related costs		Rank of cost effectiveness in health administration, training healthcare staff or preventing, diagnosis, treatment, and monitoring health conditions resulted from the methods and interventions developed by a scientist working in the target institution in an intended period of time assessed subjectively by the expert committee
	Administration	Health or healthcare administration	1.25	
	Education	Medical education	1.25	
	Prevention	Preventive measures in health	1.25	
	Treatment	Medicines, therapeutic guidelines (protocols), rehabilitation methods, therapeutic measures or instruments	1.25	
	Diagnosis	Diagnostic guidelines (protocols), methods and instruments	1.25	
	Food and environment	Improvement in the quality and safety of food or environment	1.25	
Rate of impact		Time to achieve an effective impact from a research	0.75	Rank of the rate of impact (or possible rate of impact) of a cost effective research performed by a scientist working in the target institution in an intended period of time assessed subjectively by the expert committee
Policy impact			**20**	
Science to policy translation		Implementation of a sustained policy in health administration, training healthcare staff or preventing, diagnosis, treatment, and monitoring health conditions (clinical excellence, clinical governance, guideline or protocol development) based on a scientific research		Rank of the policy implemented to improve health administration, medical education, prevention, diagnosis, treatment, or monitoring health conditions by a scientist working in the target institution in an intended period of time assessed subjectively by the expert committee
	Administration	Health or healthcare administration	3	
	Education	Medical education	3	
	Prevention	Preventive measures in health	3	
	Treatment	Medicines, therapeutic guidelines (protocols), rehabilitation methods, therapeutic measures or instruments	3	
	Diagnosis	Diagnostic guidelines (protocols), methods and instruments	3	
	Food and environment	Improvement of the quality and safety of food or environment	3	
Rate of return		Time to achieve a profitable impact from a policy change resulted from research	1	Rank of the rate of return (or possible rate of return) of a policy implemented by a scientist working in the target institution in an intended period of time assessed subjectively by the expert committee

Table 1. *(Continued)*

Domain	Item	Definition	Max. Score	How to measure
Maintaining intellectual properties		Maintaining policies to protect peoples' specific intellectual properties	1	Rank of the policies developed to maintain intellectual properties by a scientist working in the target institution in an intended period of time assessed subjectively by the expert committee
Societal impact			**20**	
	Improvement of public health literacy	Improving health literacy of society through research (action research)	3	Rank of the improved public health literacy by a scientist working in the target institution in an intended period of time
	Improvement of public perception of science/health	Improving public perception of science or health of society through research (action research)	2	Rank of the improved public perception of science/health by a scientist working in the target institution in an intended period of time
	Improvement of lifestyle	Improving lifestyle of society through research	4	Rank of the improved lifestyle of the society by a scientist working in the target institution in an intended period of time
	Increase in life expectancy	Improving life expectancy of society through research	4	Rank of the increased life expectancy by a scientist working in the target institution in an intended period of time
	Disease/harm reduction	Reducing morbidity and mortality of a disease or harm through research	2	Rank of the reduced morbidity and mortality of a disease by a scientist working in the target institution in an intended period of time
	Disease eradication	Eradication of a disease from society through research	5	Rank of the eradicated disease by a scientist working in the target institution in an intended period of time
Stewardship of research			**20**	
Capacity building		Expansion and improvement of working environment, funding and organizational capabilities		
	Establishment of research centers	Foundation and establishment of research centers		Rank of the research center established by a scientist working in the target institution in an intended period of time certifiable by the letter of acknowledgment. For uncertified research centers, half of the score can be allocated
	As founder		1	For each research center 1 score can be allocated
	As co-founder		0.5	For each research center 0.25 score can be allocated
	Peer-reviewed journal	Maintaining and development of a peer-reviewed journal		
	Establishment of a journal	Starting up a peer-reviewed journal	1	For each journal 1 score can be allocated. For scientists working in the target institution the scores can be ranked according to the journal indexing
	Editorial activities	Member of a journal editorial board		
	Editor-in-Chief		1	For each journal 1 score can be allocated. For scientists working in the target institution the scores can be ranked according to the journal indexing
	Editorial member		0.5	For each journal 0.25 score can be allocated. For scientists working in the target institution the scores can be ranked according to the journal indexing
	Reviewer		0.25	For each journal 0.125 score can be allocated. For scientists working in the target institution the scores can be ranked according to the journal indexing
	Thesis supervision	Supervision of thesis/dissertation approved by the institution		

Table 1. *(Continued)*

Domain	Item		Definition	Max. Score	How to measure
		PhD thesis		2	Rank of the theses supervised by a scientist working in the target institution in an intended period of time
		MD / MPH / MSc thesis		1.5	Rank of the theses supervised by a scientist working in the target institution in an intended period of time
	Organizing congresses		Organizing scientific congresses		
		International		1	Chair of a congress receives 1 score per event and head of scientific or organizing committee receives 0.5 score
		National		0.25	Chair of a congress receives 0.25 score per event and head of scientific or organizing committee receives 0.125 score
	Attracting external research funding and sponsorship		Attracting funds and sponsorship from external sources out of institution	2	Rank of the external funds and sponsorship attracted by a scientist working in the target institution in an intended period of time
Teamwork in research			Performing research in a team rather than individually		
	Number of co-authors		Number of co-authors in the published papers	0.5	Rank of the number of co-authors of a scientist working in the target institution in an intended period of time
	Outside institution authors		Having compatriot coauthors with different affiliations in the published papers	1	Rank of the number of outside institution co-authors of a scientist working in the target institution in an intended period of time
	International co-authors		Having international coauthors in the published papers	1.5	Rank of the international co-authors of a scientist working in the target institution in an intended period of time
Pursuing a "research line"			Pursuing and promoting a set of interrelated projects on a specific topic of research. This can be mainly assessed by number of articles published by a scientist with identical keywords	2	Rank of a scientist working in the target institution in an intended period of time based on pursuing a "research line"
Board member of research or health related societies/committees			Being the board member of societies or committees related to research or healthcare development		
	International administrative/advisory position			2.5	Head of an international society or committee receives 2.5 score and board member of an international society or committee receives 1.25 score
	National administrative/advisory position			1	Head of a national society or committee receives 1 score and board member of an national society or committee receives 0.5 score
	Institutional administrative/advisory position			0.5	Head of an institutional committee receives 0.5 score and board member of an institutional committee receives 0.25 score

It is noteworthy that the scale is designed for evaluation of impacts resulted from research activities and thus it excludes the educational programs done by a scientist. For some items the score that can be allocated to a given scientist is based on the rank of that scientist among the scientists working in a specific (target) institution (or organization). The important details of the domains and items are explained hereunder:

1. H-index can be calculated from Web of Science, Scopus, Google Scholar or some software designed for this purpose (8). Other indexing databases may also be applicable for calculating the h-index. However, Google scholar seems to be superior for this purpose as it is more inclusive, especially regarding the language diversity of articles included compared to other databases. Nevertheless, Google scholar includes many non-peer reviewed articles. It is therefore upon the policy of the target institution to opt for which database.

2. For calculation of the mean number of citations minus self-citations, the number of citations minus self-citations of all articles published by a scientist should be calculated and then the mean of them should be measured. The standard

deviation (SD) of this mean will be used for the item "reciprocal value of SD of number of citations minus self-citation".

3. In order to calculate the score of published articles of a scientist, an additional score from the impact factor of the journal that published the article divided by 30 should be added to maximum score for each article.

4. For a same amount of grant that two scientists receive, the scientific productivity of the research done should be compared by the number of citations minus self-citations of articles produced.

5. Although research, in its nature, is useful to develop science, if it only remains in papers and does not translate into policy or change in society may not be influential. Two domains of policy impact and societal impact are added in this scale to encourage scientists not just focusing on increasing their publications.

Comparison of cost effectiveness and cost benefit analysis

Comprehension of these two sets of scales is somehow difficult as they have close meanings. Cost benefit is a measure of total money return (or saving) expected from a new medical method per unit of money spent. On the other hand, in cost effectiveness, the effectiveness and applicability of alternative methods, strategies and interventions to achieve a specific set of results is compared. Cost effectiveness of research is referred to the impacts of research which led to improved quality in healthcare approaches or decrease in morbidity or mortality but currently may not be translated to economical values due to lack of supporting evidence or needing more time for evaluation (9). A cost beneficial method is undoubtedly a cost effective method that is also shown to be economical.

Role of the expert committee policy

As for some items, the level of an individual scientist working in an intended institution should be ranked; an expert committee should be formed. It is better to be comprised of senior policy makers, members of the board of directorate of an institution, a group of experts in that field of science and lay people. This committee will allocate comparative ranks to the scientists.

For some items an exact numerical value for the impact or benefit of a research is not possible to be calculated, which necessitates the ranking of scientists to be subjectively assessed by the expert committee based on priorities, their experience and personal judgment.

DISCUSSION

Importance of evaluation scales in assessment of scientist's impact

All institutions or organizations that fund for research are inclined to support science and scientists that make a progress. But there is no fixed formula for identifying the exact value of research and researchers (1,2). Some institutions consider the amount of science produced or in a better word, measurement of the quantity of the science production. Recently we have shown that the productivity of the science of toxicology in Asia Pacific region has been expanding faster than Northern America (10). Potential reasons were given in phenomena so called "Ceiling effect of science production" in Northern America and the "the rise of the rest" in Asia Pacific region. As a result, the gap between these regions is narrowing down (10). Nevertheless, measurement of the quantity of the science production might not be insightful, as there may be papers with limited or redundant findings (11). As a reason, and quite rightfully, the number of citations of articles and h-index of individuals or institutions have been suggested to overcome this weakness (3). Hence, the majority of institutions and organizations are now trusting on the h-index; despite having its own limitations and shortcomings. For an organization, defining and prioritizing outcome of research topics requires value judgments and decision-making (1,2). Most directors are seeking for a researcher that creates a real advancement, not only in basic levels but also in improvement of health policies in the administrative levels. The evaluation scale proposed in this article can be helpful in this matter.

Some scientists are only eager to increase their publications and citations as they may be encouraged by their institution (6). This may create a one-dimensional scientist (Figure 1). Notwithstanding, an effective scientist is probably a multi-aspect one who opens new scientific fronts and connections to other disciplines (2), trains junior researchers to come up with him/her during the research, attempts to improve policies in healthcare, tries to build up improved capacities for research, helps other scientists write high quality papers (5), and is also eager to find effective and cost-beneficial approaches that strengthens healthcare and ultimately brings prosperity to the society.

In less affluent countries, there are scientists who dedicated their life to development of science in their society and in improvement of public health. However, they were not able to increase their citations as their knowledge has never been published or has only been disseminated in local journals. Probably a better metric than only h-index and scientometric analysis can provide a clearer image for their scientific impact. It seems that applicability, benchmarking and implementation of research are more important in underprivileged societies. Moreover, as in the developing world, the governmental support and non-governmental organizations (NGOs) are less developed; therefore, the independent impact of an individual scientist is vital. Perhaps these individuals should be defined as non-governmental individuals (NGIs). The value of their impact is heavily underrepresented if h-index alone is taken into account.

There has been an ongoing battle between basic and applied sciences. This article tries to favor applied sciences in scoring the research activities. Furthermore, it is indeed necessary to magnify the value of implementation of research-based principles. We have to set strategies in which scientists are also directly responsible for advocacy, science to policy translation and promoting public health literacy. This is where real effective scientists show their quality. Scientists with highly cited papers who spend most of their time in laboratories and libraries and are not inclined to change methods and implement policies have minor impacts on people's life.

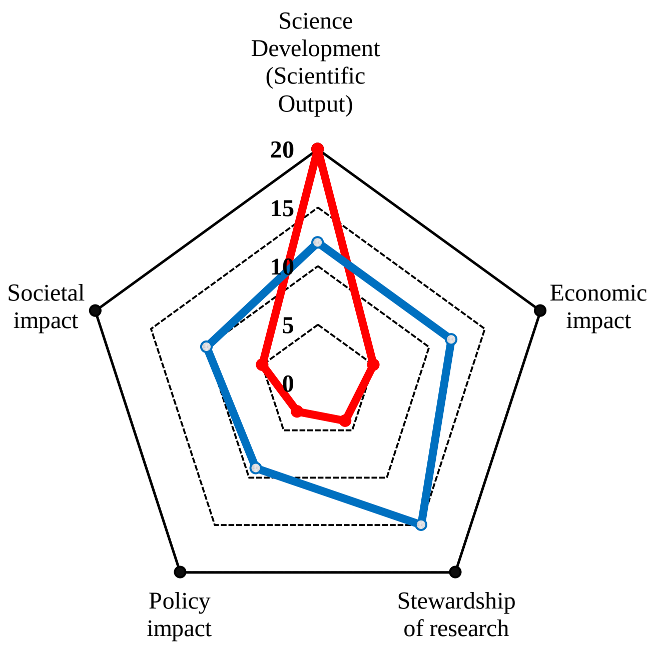

Figure 1. Template of a two potential scientists including one whose research impact is mainly coming out from production of science (red line) and the other one whose impact is coming out from wide range of research-related activities (blue line). It seems while science production is necessary to expand the edge of knowledge, it is probably more important to encourage scientists to work on other aspects as well.

A medical toxicology tale

It is a privilege to see that in recent decades, the number of scientists working in the field of medical toxicology especially in Asia Pacific region has been increasing (12). As a result, the productivity in this field of science has increased having an upward slope (10). Nevertheless, several questions rise when the amount of the production by a medical toxicologist is taken into account: How does this amount of research contributed to growth of medical toxicology? Has the performed research been effective and cost beneficial in development of newer diagnostic approaches, treatment protocols, improved therapeutic formulations, updated interventions and ultimately upgrading public health? Have these pieces of research resulted in implementation of newer diagnostic approaches, treatment protocols, improved therapeutic formulations and updated interventions? Has medical toxicologist been able to collaborate with other toxicologists or scientists in other disciplines? Has medical toxicologist been capable of making change in policies? Has medical toxicologist been able to prepare fresh young researchers for the future of this discipline? Indeed, an effective and multidimensional medical toxicologist is the one who can receive positive answers for the majority of these questions.

CONCLUSION

For evaluation of a scientist, in addition to the amount of their scientific production and the citations, the economic impact, policy impact and societal impact of their research as well as their ability in stewardship of research is necessary to be assessed and quantified. This strategy is more important in the developing world where the opportunities and resources are scarce. For this purpose, the evaluation process should be moved from only scintometric indices to a combination of quantitative (including scientometric) and qualitative indices as we called "360-degree researcher evaluation score". This proposed scale seems to be a good start. Its overall coverage, subdomains and proposed weights of scores need to be standardized. Your comments on this scale are highly welcomed to be published in the next issues of the journal.

Conflict of interest: None to be declared.
Funding and support: None.

REFERENCES

1. Kumar MJ. Evaluating Scientists: Citations, Impact Factor, h-Index, Online Page Hits and What Else? IETE Technical Rev 2009;26:165-8.
2. Van Houten BA, Phelps J, Barnes M, Suk WA. Evaluating scientific impact. Environ Health Perspect 2000;108:A392-3.
3. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A 2005;102:16569-72.
4. Afshari R. What is the “Best Research” for Low Income Countries? Asia Pac J Med Toxicol 2013;2:1.
5. Krause ET. Impact: Take peer review into account. Nature 2013;503:198.
6. Stroobants K, Godecharle S, Brouwers S. Research evaluation: Flanders overrates impact factors. Nature 2013;500:29.
7. Jamali HR, Asadi S, Sedghi S. Research Outcome and Impact Assessment of Medical Research. 1st ed. Tehran: Iranian Academy of Medical Sciences; 2012.
8. Harzing AW. Publish or Perish [Internet]. 2007 [Cited 2014 November 5] Available from: http://www.harzing.com/pop.htm
9. Cellini SR, Kee JE. Cost-effectiveness and cost-benefit analysis. In: Wholey JS, Hatry HP, Newcomer KE, editors. Handbook of Practical Program Evaluation. 3rd ed. San Francisco, USA: Jossey-Bass; 2010. p.493-530.
10. Afshari R. Scientometric Analysis of Toxicology in Asia Pacific Region: Signs of Growth. Asia Pac J Med Toxicol 2014;3:92-6.
11. Rawat S, Meena S. Publish or perish: Where are we heading? J Res Med Sci 2014;19:87-9.
12. Afshari R. Empowerment of Medical Toxicology in Asia Pacific Region. Asia Pac J Med Toxicol 2013;2:36.