Keynote Presentations

The Fundamentals of Heterodox Evaluative Scientometrics

Fourth international conference on science and technology metrics (STMet 2023)

Dubai, 16-17 November 2023

Giovanni Abramo
Laboratory for Studies in Research Evaluation
Institute for System Analysis and Computer Science
National Research Council of Italy

About myself and research group

Giovanni Abramo and Ciriaco Andrea D'Angelo are prominent contributors to the scientometric literature. Anyone familiar with their extensive oeuvre will have noticed that some of their ideas differ fundamentally from mainstream scientometric thinking.

Ludo Waltman (2016)

https://www.sciencedirect.com/science/article/pii/S1751157716300864

Mainstream scientometric thinking and practice

- Research productivity is measured by the number of publications
- Research impact is measured by the number of (normalized) citations
- Research performance measurements do not account for input

Outline

1. How to measure (and compare!) individual research performance
2. How to measure research performance at the aggregate level (discipline, institution, country)
3. How not to measure research performance
4. Ranking distortions when using invalid indicators
5. The research performance of Gulf countries
6. Conclusions and recommendations

Research evaluation problems

- ✓ Proliferation of (invalid) performance indicators
- ✓ Doubtful assessment methods
- ✓ Abundance of non theory-based rankings
- ✓ Media fanfare for (wrong!) world institutions rankings
- ✓ Do-it-yourself practices
- ✓ Poor strategic and policy perspectives

The Fractional Scientific Strength (FSS) in short

It measures total impact per dollar spent on research:

- ✓ It counts all publications of a scientist in a period of time
- ✓ It divides each publication by the number of authors (it weights their contribution based on their position in the byline, if not alphabetically ordered)
- ✓ It measures their value by a weighted combination of citations and IF, each scaled by a factor accounting for field and year of publication
- ✓ It divides total impact by the research expenditures

Individual research performance indicator

The Fractional Scientific Strength (FSS)

$$FSS_R = \frac{1}{(w_R + k)} \cdot \frac{1}{t} \sum_{i=1}^{N} c_i f_i$$

Where:

N = number of publications of the researcher in the period under observation

c_i = weighted combination of normalized citations and impact factor associated to publication i*

f_i = fractional contribution of the researcher to publication i

w_R = average yearly salary of the researcher

k = average yearly capital used by the researcher

t = number of years of work of the researcher in the period under observation

* Abramo, G., D'Angelo, C.A., & Felici, G. (2019). Predicting long-term publication impact through a combination of early citations and journal impact factor. *Journal of Informetrics,* 13(1), 32-49.

Data and method

- ✓ Output source: Web of Science (WoS) core collection
  - ✓ Publications indexed in over 21,000 peer-reviewed journals
- ✓ Input source: Italian Ministry of Research academic database: (prof. name, field classification, academic rank, gender, affiliation)
  - ✓ Professors field classification: 370 scientific disciplinary sectors (SDS); 14 areas (UDA)
- ✓ Tool: Disambiguation algorithm, assigning publications to each author:
  - ✓ Affiliation unification
  - ✓ Authors' name disambiguation

Comparing individual research performance

Name	John Doe	Jane Doe
Discipline	Clinical medicine	Clinical medicine
Indicator	Score	Score
O	6.6	3.6
FO	1.44	1.22
MNCS	2.02	0.95
FSS	**1.22**	**0.69**
h-index	12	5
g-index	19	7

The importance of researchers' field classification

Name	John Doe		Jane Doe	
Field	MED/15 (Blood diseases)		MED/22 (Vascular surgery)	
Indicator	Score	**Rank(η)**	Score	**Rank(η)**
O	6.6	**67.4**	3.6	**90.5**
FO	1.442	**68.4**	1.220	**95.2**
MNCS	2.021	**78.9**	0.595	**89.6**
FSS	1.228	**78.4**	0.692	**91.3**
h-index	12	**76.4**	5	**79.6**
g-index	19	**77.0**	7	**80.4**

The importance of researchers' field classification

SDS	Obs	Unproductive	Average output
MED/15 - Blood Diseases	187	3.2%	5.84
MED/16 - Rheumatology	114	8.8%	5.68
MED/11 - Cardiovascular Diseases	265	6.0%	5.11
MED/01 - Medical Statistics	108	4.6%	5.01
MED/26 - Neurology	414	4.8%	5.00
MED/13 - Endocrinology	258	7.0%	4.46
MED/08 - Pathological Anatomy	316	3.5%	4.43
MED/12 - Gastroenterology	175	6.3%	4.09
MED/03 - Medical Genetics	144	2.8%	3.86
MED/06 - Medical Oncology	132	9.8%	3.73
…	…	…	…
MED/28 - Odonto-Stomalogical Diseases	431	17.2%	1.74
MED/42 - General and Applied Hygiene	366	16.7%	1.67
MED/30 - Eye Diseases	274	20.1%	1.52
MED/33 - Locomotory Diseases	223	20.2%	1.51
MED/48 - Neuropsychiatric and Rehabilitation Nursing	15	13.3%	1.49
MED/32 - Audiology	60	20.0%	1.33
MED/45 - General. Clinical and Pediatric Nursing	34	20.6%	1.33
MED/43 - Legal Medicine	265	36.6%	0.98
MED/02 - History of Medicine	37	45.9%	0.73
MED/47 - Nursing and Midwifery	7	42.9%	0.46

The performance of single researchers

The national percentile ranking of researchers of the Biopathology Dept of Institution "X"

Name	Ac. rank	SDS	O	FO	SS	FSS
John Doe 1	Ass.	MED/03	37	25	22	**23**
John Doe 2	Full	MED/08	75	59	61	**58**
John Doe 3	Full	MED/15	42	23	23	**27**
John Doe 4	Full	MED/30	52	37	39	**41**
John Doe 5	Res.	MED/36	23	13	6	**11**
John Doe 6	Ass.	BIO/14	50	36	38	**38**
John Doe 7	Ass.	MED/08	83	72	70	**64**
John Doe 8	Full	FIS/07	74	56	62	**55**
John Doe 9	Res.	MED/15	54	35	40	**44**
John Doe 10	Ass.	BIO/14	25	23	18	**20**
John Doe 11	Res.	MED/15	28	25	27	**22**
John Doe 12	Res.	MED/30	38	22	20	**21**
John Doe 13	Res.	FIS/07	27	25	15	**17**
John Doe 14	Res.	MED/36	83	70	70	**67**
John Doe 15	Res.	MED/36	31	13	13	**13**
John Doe 16	Full	BIO/13	86	72	69	**75**
John Doe 17	Full	MED/30	95	83	75	**77**

Research performance at the aggregate level

- ✓ Premise: To measure research institution performance, one needs to know the identity, output, and field of research of each research staff member
- ✓ Statement: Research institutions are not homogenous in terms of number and size of research fields
- ✓ Proposition: Individual performance is absolutely required to measure performance at organizational level

The Fractional Scientific Strength (FSS)

aggregate level

Productivity of research units (e.g. field, discipline, department, institution) based on FSS_R

$$FSS_D = \frac{1}{RS} \sum_{j=1}^{RS} \frac{FSS_{R_j}}{\overline{FSS_R}}$$

Where:

RS = research staff of the unit, in the observed period

FSS_{Rj} = productivity of researcher *j* in the research unit

$\overline{FSS_R}$ = average productivity of all national productive researchers in the same SDS as researcher *j*

The performance in each field (SDS)

The fields within the UDA "Medicine" of Institution "*X*"

	O			FSS		
Field	Score	Rank	Rank(η)	Score	Rank	Rank(η)
MED/09-Internal medicine	0.739	6 out of 12	55	0.435	8 out of 12	36
BIO/14-Pharmacology	0.457	25 out of 37	33	0.287	25 out of 37	33
MED/38-General and specialised pediatrics	0.524	33 out of 42	22	0.460	28 out of 42	34
MED/40-Gynaecology and obstetrics	0.816	5 out of 22	81	0.242	5 out of 22	81
MED/42-General and applied hygiene	1.103	8 out of 52	86	1.000	14 out of 52	75
MED/07-Microbiology and clinical microbiology	1.525	3 out of 23	91	2.077	2 out of 23	95
BIO/13-Applied biology	0.425	37 out of 42	12	0.225	38 out of 42	10
MED/08-Pathological anatomy	0.667	28 out of 40	31	0.485	26 out of 40	36
MED/11-Cardiovascular diseases	1.023	8 out of 27	73	1.053	9 out of 27	69

The performance of Institution '*X*' in each discipline (UDA)

	O		SS		FO		FSS	
UDA*	Score	Rank(η)	Score	Rank(η)	Score	Rank(η)	Score	Rank(η)
2	1.231	81	1.246	76	0.988	69	1.087	76
3	1.031	72	0.973	63	1.111	86	1.092	79
5	1.031	65	0.853	45	1.064	72	0.865	47
6	1.033	74	1.033	67	1.080	76	1.115	79
7	0.775	46	0.643	39	0.845	54	0.734	46
9	0.741	26	0.763	43	0.763	33	0.766	39

** 2, Physics; 3, Chemistry; 5, Biology; 6, Medicine; 7, Agricultural and veterinary sciences; 9, Industrial and information engineering*

Key performance indicators

- Individual level:
  - Productivity (FSS) and its components
  - Highly-cited articles per researcher
- Institution level:
  - Productivity (FSS)
  - Share of unproductive staff
  - Share of top scientists
  - Effectiveness of recruitment
  - Rate of institution-industry collaborations

How not to measure research performance

- ✓ The MNCS
- ✓ The h-index
- ✓ The Shanghai ranking and the like

Research-based (?!) Leiden rankings

http://www.leidenranking.com/ranking.aspx

Mean normalized citation score (MNCS). The average number of citations of the publications of a university, normalized for field differences, publication year, and document type. An MNCS value of two for instance means that the publications of a university have been cited twice above world average.

Validity of the most popular indicators

- ***Mean normalized citation score*** (MNCS): The average number of field-normalized citations of the publications of a university …

Univ. A = (10) => MNCS = 10

Univ. B = (10, 10, 10, …, 9) => MNCS < 10

Distortion of rankings by the Leiden's new crown indicator (MNCS)

AREA	Percentage of Q1 scientists by FSS not included in the same set by MNCS
Mathematics and computer science	31
Physics	57
Chemistry	42
Earth sciences	40
Biology	44
Medicine	46
Agricultural and veterinary science	42
Civil engineering	26
Industrial and information engineering	35
Total	42

Validity of the most popular indicators

- ✓ The h-index: the maximum number *h* of works by a scientist that have at least *h* citations each

John Doe = (4, 4, 4, 4)

Jane Doe = (400, 400, 400, 400, 4, 4, …, 4)

John Doe	Jane Doe
h = 4	h = 4

Distortion of universities rankings by h and g indexes

UDA	Percentage of Q1 universities by FSS not included in the same set by h	g
Mathematics and computer science	45	47
Physics	48	51
Chemistry	49	46
Earth sciences	42	35
Biology	42	36
Medicine	40	35
Agricultural and veterinary science	41	33
Civil engineering	28	26
Industrial and information engineering	40	35
Total	*42*	*38*

The Shanghai ranking

ARWU
by Shanghai Jiao Tong University

http://www.shanghairanking.com/ARWU2018.html

Sapienza, Padua: 151-200

Academic Ranking of World Universities 2018

World Rank	Institution*	By location: All	National/Regional Rank	Total Score	Score on: Alumni
1	Harvard University		1	100	100
2	Stanford University		2	75.6	44.5
3	University of Cambridge		1	71.8	82.3
4	Massachusetts Institute of Technology (MIT)		3	69.9	70.9
5	University of California, Berkeley		4	68.3	65.6
6	Princeton University		5	61	55.8
7	University of Oxford		2	60	50.8
8	Columbia University		6	58.2	62.8
9	California Institute of Technology		7	57.4	53.5
10	University of Chicago		8	55.5	59.2
11	University of California, Los Angeles		9	51.2	29.2
12	Cornell University		10	50.7	43.1
12	Yale University		10	50.7	47.1
14	University of Washington		12	50	24.9
15	University of California, San Diego		13	47.8	19

The Shanghai ranking criteria

Metodology: total score

Criteria	Indicator	Weight
Quality of Education	Alumni of an institution winning Nobel Prizes and Fields Medals	10%
Quality of Faculty	Staff of an institution winning Nobel Prizes and Fields Medals	20%
	Highly cited researchers in 21 broad subject categories	20%
Research Output	Papers published in Nature and Science	20%
	Papers indexed in SCI-E and SSCI (Web of Science)	20%
Per Capita Performance	Per capita academic performance of an institution	10%

90% of the score is size dependent!

The scientific productivity of the Gulf countries

We assess the scientific productivity of the Gulf countries, overall and in 221 subject categories (SCs)*, in the 2015-2019 period.

We identify the research staff of each country**, and classify each one in the prevalent SC of their publications.

* WoS classification schema (all SCs but *Art & Humanities)*

** Through the Caron & van Eck author name disambiguation algorithm, on WoS data (Caron, E., & van Eck, N. J. (2014). Large scale author name disambiguation (AND) using rule-based scoring and clustering. In E. Noyons (Ed.), *Proceedings of the Science and Technology Indicators Conference 2014 Leiden* (pp. 79–86). Leiden: Universiteit Leiden—CWTS.)

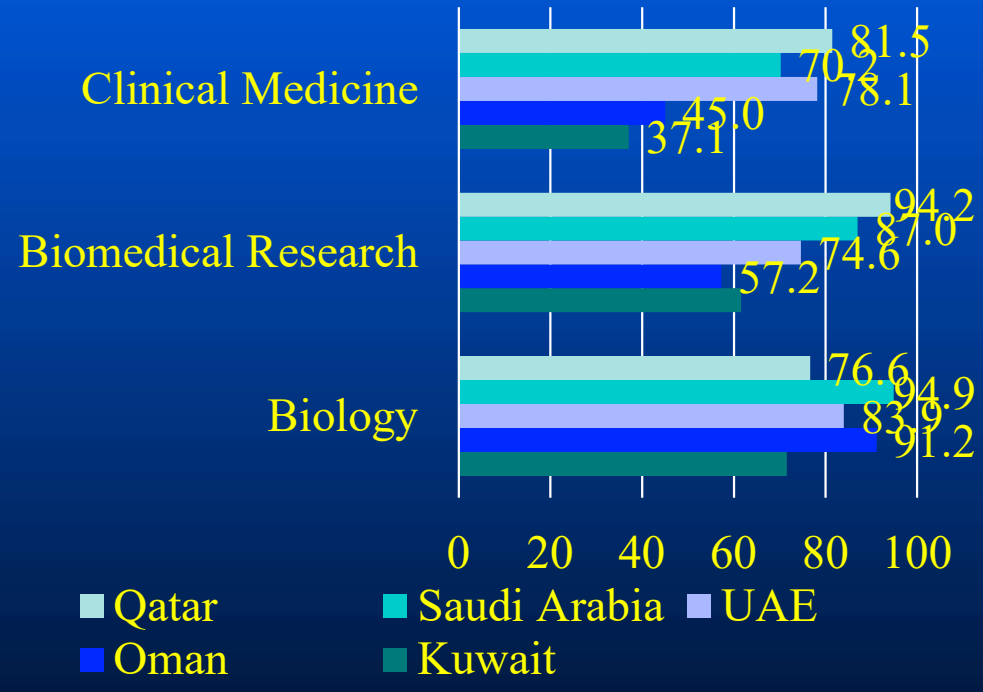
The research productivity at overall and area level

Overall level

Country	Obs	FSS_A	Rank (out of 146)*
Qatar	1,142	1.322	3
Saudi Arabia	8,079	1.205	9
UAE	1,737	1.041	19
Oman	540	0.767	43
Kuwait	655	0.550	75

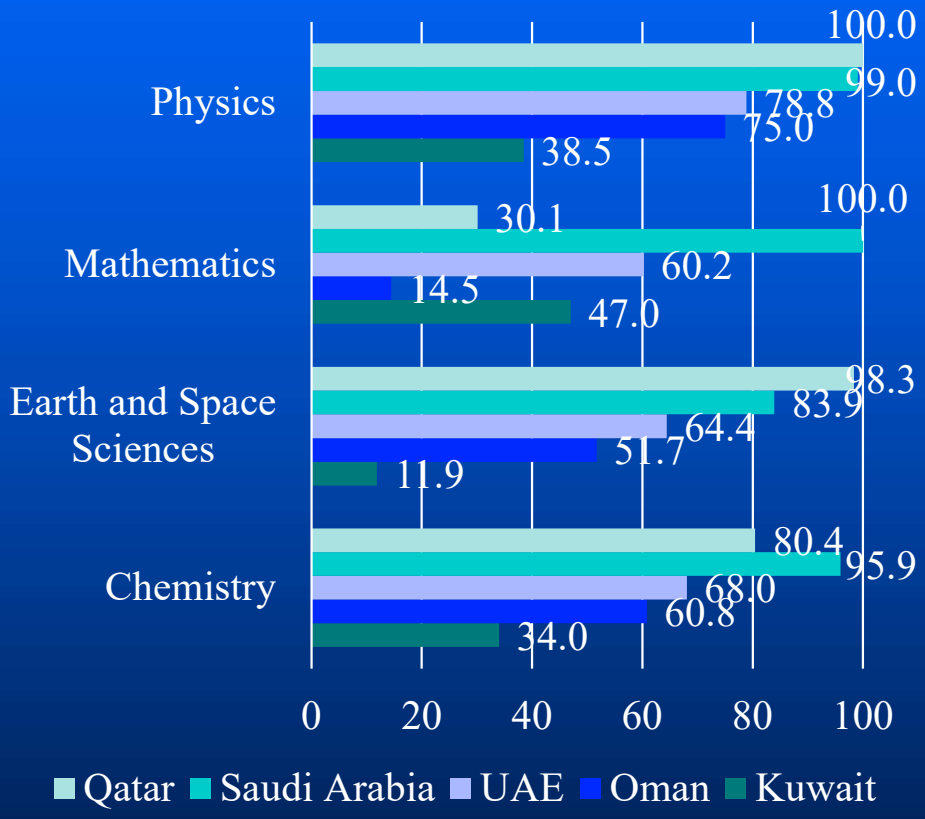
* *Countries with at least 100 productive researchers overall*

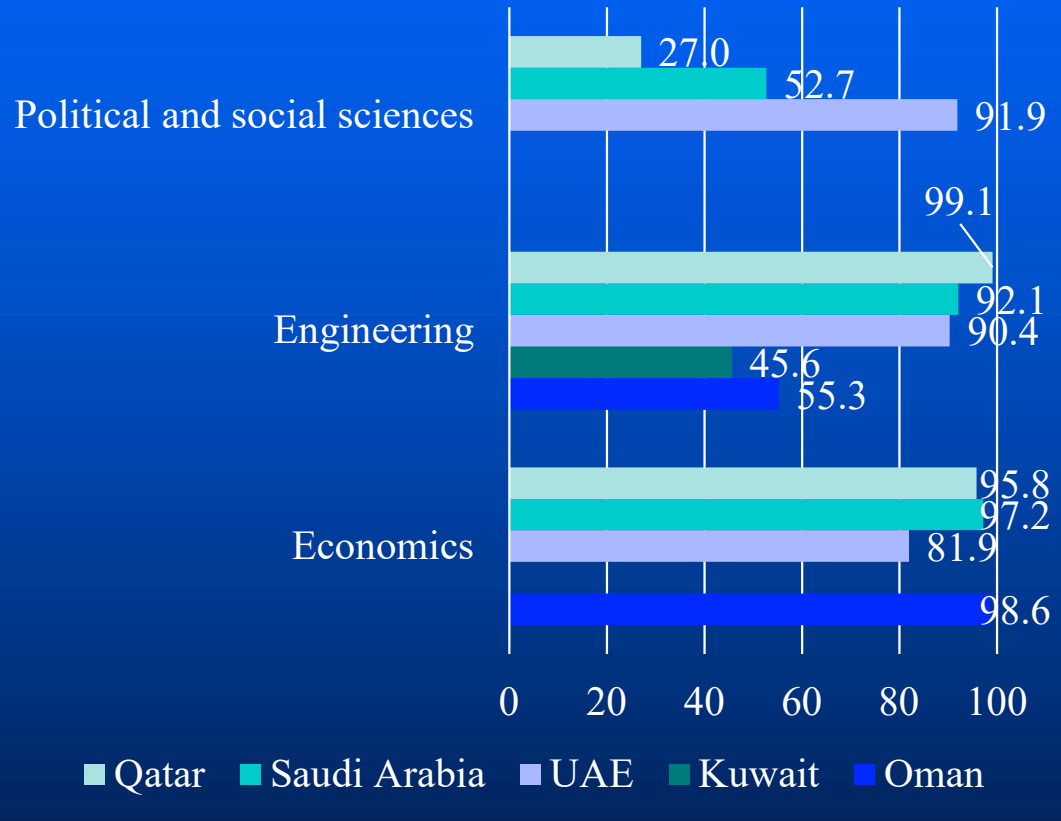
Area level – Life Sciences (FSS world percentile; 100, the best)**



** *Countries with at least 15 researchers in the area*

The research productivity at area level





*** Counties with at least 15 researchers in the area*

WORLD LEADERSHIP AT SC LEVEL

Country	SC	Obs	FSS	Rank*
Oman	Chemistry, Medicinal	4	4.554	1 out of 90
	Business, Finance	3	2.722	1 out of 49
Qatar	Optics	4	12.072	1 out of 85
	Computer Science, Artificial Intelligence	13	3.493	1 out of 87
	Immunology	15	2.142	1 out of 106
	Sport Sciences	42	2.028	1 out of 57
Saudi Arabia	Mechanics	44	4.468	1 out of 78
	Physics, Fluids & Plasmas	7	4.119	1 out of 57
	Mathematics, Applied	132	3.283	1 out of 87
	Marine & Freshwater Biology	36	2.721	1 out of 94
	Statistics & Probability	25	2.485	1 out of 69
	Psychology, Multidisciplinary	3	2.337	1 out of 60
	Ecology	9	2.121	1 out of 105
UAE	Physiology	7	1.723	1 out of 60

** Countries with at least 3 researchers in the SC*

The UAE excellence in Physiology

Name	Affiliation	FSS	FSS_scaled	Rank*	Perc.
Beegam, Sumaya		4.746	2.781	322	92.7
Yuvaraju, Priya		4.593	2.692	348	92.1
Yasin, Javed	College of Medicine and Health Sciences UAE University, Al Ain, Abu Dhabi	4.344	2.546	381	91.4
John, Anne		3.601	2.110	499	88.7
Howarth, Frank Christopher		2.810	1.647	697	84.2
Qureshi, Muhammad Anwar		0.488	0.286	2897	34.2
Jacobson, Michael	Higher Colleges of Technology, Abu Dhabi	0	0	4273	0

** Out of 4,401 physiology researchers in the world*

Conclusions

- ✓ Count only what counts and be aware of what you cannot count
- ✓ The most popular research performance indicators are invalid
- ✓ Field classification of scientists is absolutely required to compare performance at the individual level
- ✓ Research performance at the individual level is absolutely required to measure performance at organizational level
- ✓ World performance comparisons are not accurate … yet
- ✓ Avoid the "do-it-yourself" temptation
- ✓ The performance evaluation conducted in Italy can be replicated in any other institution or country, and the first might serve as a benchmark for comparison

Giovanni Abramo

Laboratory for Studies in Research Evaluation

at the Institute for System Analysis and Computer Science (IASI-CNR)

National Research Council of Italy

Via dei Taurini 19, 00185 Roma – ITALY

Tel. +39 339 4508461

giovanni.abramo@uniroma2.it

https://sites.google.com/view/laboratoriortt