Journal of Informetrics 11 (2017) 389–406

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Uncovering fine-grained research excellence: The global research benchmarking system

Peter Haddawy [a,*], Saeed-Ul Hassan [b], Craig W. Abbey [c], Lee Inn Beng [d]

[a] Faculty of ICT, Mahidol University, Nakhon Pathom 73170, Thailand
[b] Information Technology University, Lahore, 54770, Pakistan
[c] Office of Institutional Analysis, University at Buffalo, Buffalo, NY 14260 United States
[d] Elsevier, Singapore

ARTICLE INFO

Article history:
Received 10 April 2015
Received in revised form 6 February 2017
Accepted 8 February 2017
Available online 28 February 2017

Keywords:
Research benchmarking
University ranking
Bibliometrics
Fine-grained subject categories

ABSTRACT

Since few universities can afford to be excellent in all subject areas, university administrators face the difficult decision of selecting areas for strategic investment. While the past decade has seen a proliferation of university ranking systems, several aspects in the design of most ranking systems make them inappropriate to benchmark performance in a way that supports formulation of effective institutional research strategy. To support strategic decision making, universities require research benchmarking data that is sufficiently fine-grained to show variation among specific research areas and identify focused areas of excellence; is objective and verifiable; and provides meaningful comparisons across the diversity of national higher education environments. This paper describes the Global Research Benchmarking System (GRBS) which satisfies these requirements by providing fine-grained objective data to internationally benchmark university research performance in over 250 areas of Science and Technology. We provide analyses of research performance at country and university levels, using the diversity of indicators in GRBS to examine distributions of research quality in countries and universities as well as to contrast university research performance from volume and quality perspectives. A comparison of the GRBS results with those of the three predominant ranking systems shows how GRBS is able to identify pockets of excellence within universities that are overlooked by the more traditional aggregate level approaches.

1. Introduction

Universities are widely viewed as playing a central role in the economic competitiveness of modern knowledge economies. This is particularly the case in areas of science and technology where universities are seen as engines of innovation and sources of high quality talent for growth of high-tech industries. This perceived role has led to efforts in high and middle income countries to increase university research activity, with a resultant increase in competition for research funding and top research talent nationally and internationally. Since few universities have the resources to be excellent in all subject areas, research administrators at university and government levels face the difficult decision of selecting areas

* Corresponding author.
E-mail addresses: peter.had@mahidol.ac.th (P. Haddawy), saeed-ul-hassan@itu.edu.pk (S.-U. Hassan), cwabbey@buffalo.edu (C.W. Abbey), IB.Lee@elsevier.com (I.B. Lee).

http://dx.doi.org/10.1016/j.joi.2017.02.004
1751-1577/

for strategic investment (Salmi, 2009). To effectively do so requires first being able to identify the distribution of strengths in relation to competitors globally.

The past decade has seen a proliferation of university ranking systems, many of which claim to provide information to help universities benchmark their performance. While a number of rankings exert great influence over universities with many universities even formulating aspects of their strategies specifically to improve their standing, several aspects in the design of most ranking systems make them inappropriate to support formulation of effective research investment strategy. First, all existing rankings operate at the institutional and broad subject levels. Thus by design they mask variation in quality within universities and overlook focused pockets of excellence, information that is crucial to make effective strategic research investment decisions. Second, they focus primarily on the largest and most comprehensive universities, thus missing the important contributions being made by a myriad of more narrowly focused institutions. Finally, many of the indicators used by some of the most prominent ranking systems are subjective or exceedingly retrospective and are thus not appropriate as sources of benchmarking data to support management decisions.

Universities and government research funding agencies require research benchmarking data to support strategic decision making that is sufficiently fine-grained to identify focused areas of excellence; is actionable; is objective and verifiable; allows tracking of performance over time; and provides meaningful comparisons across the diversity of national higher education environments and university structures. In this paper we describe the Global Research Benchmarking System.[1] (GRBS) which satisfies these requirements by providing objective data to internationally benchmark university research performance in areas of Science and Technology. All GRBS data are freely available on the GRBS website[2] GRBS supports identification of fine-grained subject areas in which universities can excel; to make rational strategic and resource allocation decisions; to identify university research partners with complementary strengths; and to publicize program strengths. By covering 251 fine-grained subject areas and selecting universities for inclusion based on their performance in these areas, GRBS is able to shed light on variation within a single institution as well as to highlight the performance of universities with particular focused strengths.

2. University ranking and benchmarking systems

International university ranking systems fall into two broad categories: those that seek to cover a broad range of university activity and those that focus exclusively on research. Here we discuss six prominent ranking systems, three in each category.

2.1. Rankings covering multiple dimensions of university activity

QS (2017) launched its world university rankings (QS-WUR) in 2004 in collaboration with Times Higher Education (THE). In 2010 the partnership ended with QS continuing its ranking and THE establishing a new ranking. QS currently publishes eight different types of rankings: faculty rankings, subject rankings, graduate employability rankings, regional rankings, higher education system strength rankings, a ranking for universities under 50 years old, a best student cities ranking, and their original world university ranking. In addition, QS publishes a university stars rating for which they charge an audit fee. The indicators used vary among the rankings. Their world university ranking indicators cover academic reputation (40%), reputation among graduate employers (10%), international faculty ratio (5%), international student ratio (5%), student/faculty ratio (20%), and citations per full-time faculty (20%). Citation data is taken from Elsevier's Scopus database. The 2016–2017 ranking covers over 900 universities. Their 2016–2017 subject ranking covers universities in each of 42 subject areas. Indicators include academic reputation, reputation among graduate employers, citations per paper, and h-index. The weights applied to the indicators vary across the different subject areas.

THE publishes seven different types of rankings: a world university ranking, a BRICS & emerging economies ranking, a US college ranking, two regional rankings, a world reputation ranking, and a ranking of universities under 50 years old. The 2016–2017 release of THE world university ranking (THE 2017) covers 980 universities using 13 indicators grouped into five areas: International Outlook (7.5%), Research (volume, income, reputation) (30%), Citations (30%), Industry income (2.5%), and Teaching (reputation, staff-to-student ratio, doctorate-to-bachelor's ratio, number of doctorates awarded to academic staff, institutional income) (30%). The indicators include reputation surveys for research (18%) and for teaching (15%) for a total of 33% of the weight. The THE subject ranking covers 100 universities in each of six broad subject areas using the same indicators as for the world university ranking but with weights varying among the subject areas. Times Higher Education does not publish the process by which universities are selected for inclusion in their rankings.

U-Multirank (2017), launched in 2014, provides university performance evaluation for institutions overall and in 13 fields: Biology, Business Studies, Chemistry, Computer Science Programmes, Electrical Engineering, History, Mathematics,

[1] GRBS was initiated by the United Nations University International Institute for Software Technology and the Center for Measuring University Performance. Contributing organizations include: Arizona State University, Institute for Scientific and Technical Information of China, Korean Academy of Science and Technology, Ministry of Higher Education of Malaysia, National Assessment and Accreditation Council of India, National Institute for Informatics (Japan), National Institution for Academic Degrees and University Evaluation of Japan, ProSPER.Net, University of Melbourne, and University of Pisa. The governance structure of the initiative included an International Advisory Board providing expertise in university performance evaluation, bibliometrics, and Sustainable Development, and representing diverse regional and stakeholder perspectives.

[2] www.researchbenchmarking.org The currently available data is from the 2012 release, which uses Scopus data from 2008 – 2011.

Mechanical Engineering, Medicine, Physics, Psychology, Social Work and Sociology. The ranking uses over 50 indicators covering five broad dimensions: teaching and learning, research, knowledge transfer, international orientation, and regional engagement. The research indicators are: research publications (number of research publications indexed in Web of Science), size normalized research publications (number of research publications relative to number of students), publication output (number of all research publications included in the institution's publications databases), top cited papers (proportion of publications by field in the top 10% of most cited publications), interdisciplinary publications, post-doc positions (number of post-doc positions relative to fte academic staff), external research income (corrected for purchasing power parity and expressed per fte acadmic staff), art-related output (number of scholarly outputs of the creative and performing arts relative to the number of fte academic staff), citation rate (the average number of times the department's research publications from the period 2010–2013 are cited in other research published in 2010–2015) and research orientation of teaching (the degree to which the education is informed by research in the field). U-Multirank does not combine indicators into a single composite indicator but rather rates universities along each indicator separately in five performance categories: "very good" to "weak". The website can sort universities either alphabetically, based on their score on a particular indicator, or based on the number of high scores ("similar to the Olympic medal table"). Data for U-Multirank comes from the publications that are indexed in the CWTS-licensed edition of the Web of Science (WoS) database, the PATSTAT patent database, student surveys from participating institutions, and from the institutions themselves. Universities register to participate in U-Multirank and the system currently includes 1300 institutions. According to the website, the ranking includes 300 universities that have never before appeared in any of the global ranking systems.

2.2. Rankings focusing on research

Shanghai Jiao Tong University launched their Academic Ranking of World Universities (ARWU, 2017) in 2003. In addition to the original ARWU, Shanghai Jiao Tong publishes a field ranking covering five fields: Natural Sciences and Mathematics, Engineering/Technology and Computer Science, Life and Agricultural Sciences, Clinical Medicine and Pharmacy, and Social Sciences; as well as a subject ranking covering five subject areas: Mathematics, Physics, Chemistry, Computer Science, and Economics/Business. The ARWU ranks universities using a composite of six indicators: Alumni of an institution winning Nobel Prizes and Fields Medals (10%), Staff of an institution winning Nobel Prizes and Fields Medals (20%), Highly cited researchers in 21 broad subject categories (20%), papers published in Nature and Science (20%), papers indexed in Science Citation Index-expanded and Social Science Citation Index (20%), and the weighted scores of the above five indicators divided by the number of full-time equivalent academic staff (10%). For the field and subject rankings, the indicator for number of papers published in Nature and Science is replaced with an indicator measuring the percentage of papers published in the top 20% journals of that field or subject. In addition, for the Engineering/Technology and Computer Science fields the indicators for Nobel Prizes and Fields Medals are replaced with an indicator for the total engineering-related research expenditures. The three rankings rank 1200 universities selected based on whether they have any Nobel Laureates, Fields Medalists, Highly Cited Researchers, or papers published in Nature and Science. In addition, universities with a significant number of papers indexed by Science Citation Index-Expanded and Social Science Citation Index are included. The ARWU publishes the ranking of only the top 500 universities, while the field and subject rankings publish only the top 200.

The CWTS Leiden (2017) publishes university overall and subject area rankings covering seven broad subject areas. The rankings do not use a composite indicator but rather rank universities based any of five indicators. The indicators are grouped into two broad categories: impact indicators and collaboration indicators. Except for publication volume, the indicators are all available in size dependent and size independent forms. By default, size independent indicators are used. The impact indicators include the volume and proportion of publications in the top 1%, 10% and 50% most frequently cited, total citation score (TCS), mean citation score (MCS), total normalized citation score (TNCS), and mean normalized citation score (MNCS). The collaboration indicators are proportion of inter-institutional collaborative publications, proportion of international collaborative publications, proportion of collaborative publications with industry, proportion of short distance collaborative publications, and proportion of long distance collaborative publications. All indicators are calculated using data from Thomson Reuters' Web of Science. The 2016 version of the ranking covers 842 universities worldwide selected by identifying the universities with the largest scientific output published in "core journals".

Scimago (2017) publishes an institutional ranking covering the aspects of research performance, innovation outputs and societal impact. The 2016 release ranks 5147 institutions, including 2894 higher education institutions, 1203 government institutions, 824 health institutions, and 174 private institutions. The sole criterion for selecting institutions is that they must be research institutions with over 100 published works indexed in the Scopus database during the last year of the time window of consideration. Their research ranking uses a composite indicator that combines the following ten size-dependent and two size-independent indicators: publication output, international collaboration, item oriented field normalized citation score average, high quality publications (ratio of publications in the top 25% of journals according to the Scimago Journal Rank), excellence (an institution's scientific output that is included in the top 10% of the most cited papers), scientific leadership (percentage of output for which the corresponding author is from that institution), excellence with leadership (number of publications in the top 10% of most cited papers in that field for which the main contributor author is from that institution), scientific talent pool (total number of authors from that institution), number of scientific publications cited in patents, percentage of the publications cited in patents, web pages containing institute's URL as per Google, and the number

of links (incoming) to institute domains as indexed by ahrefs (https://ahrefs.com). All bibliometric indicators are computed using data from Elsevier's Scopus database.

2.3. Critique of the rankings

Ranking systems including those discussed above have been heavily criticized in the academic literature. The critique falls primarily along the lines of methodological flaws and negative effects. Examples of the latter include a benefiting of resource intensive universities and an associated norming effect on higher education (Altbach, 2015; Hazelkorn, 2014; Marginson, 2014).

In terms of methodology, the aspect that has perhaps been criticized most heavily is the aggregation of indicators of various dimensions of university performance into a single composite indicator, resulting in publication of a league table. It has been argued that the weights assigned to indicators are arbitrary and reflect subjective decisions (Baty and Hong, 2012; Billaut et al., 2010) and that the arbitrary combination of indicators of different aspects of university performance prevents a clear interpretation of the aggregate indicator (Waltman et al., 2012). From the perspective of multi-criteria decision making, Billaut et al. (2010) liken this exercise to coming up with the concept of best car or best wine in the world, which is meaningless unless we identify a group of stakeholders, structure their objectives, and study ways in which indicators can measure attainment of the objectives. Finally, it has been pointed out that assigning a single score to a university does not accommodate the different missions of universities (Altbach, 2015; Waltman et al., 2012) and that the composite score "can mask real excellence in specific fields or areas of performance" (Baty and Hong, 2012).

Moed (2016) presents a critical comparative analysis of five ranking systems (ARWU, Leiden, THE, and QS). He shows that several pairs of very similar or identical indicators from the different ranking systems rank correlate only moderately, especially those based on student and faculty numbers. He further shows that normalizations applied to some of the indicators (e.g. THE research and teaching performance scores) have "severe implications", causing the rankings of some institutions to be much more affected by the indicator than others.

With the exception of Scimago, all the above ranking systems that publish their university selection methodology, select universities based on overall institutional performance and thus have a bias toward the most comprehensively strong universities (Rauhvargers, 2013). Such an approach to selection can miss world class programs or centers of excellence that exist in universities that do not as a whole rank in the top few percent worldwide. It is, indeed, well recognized that great differences exist in performance among researchers and programs within individual institutions (Kehm, 2014).

Another major methodological issue is the use of opinion surveys by QS-WUR and THE-WUR. It has been argued that the way in which academic reputation surveys are organized leads to selection of elite universities only (Rauhvargers, 2013). Van Raan (2005) argues that it is questionable whether experts involved in the surveys can be regarded as knowledgeable experts in all parts of the evaluated entities that matter. In a study of the THE-WUR, Bowman and Bastedo (2011) have shown that the published ranking itself affects the perceptions of surveyed academics in subsequent years, even to the extent that "academics' perceptions of the top institutions in their field are affected by the world rankings that do not differentiate among fields". An additional methodological element that THE-WUR, QS-WUR, ARWU, and U-Multirank share is the use of self-reported data. Such data is problematic because of lack of internationally standardized definitions and vulnerability to manipulation (Rauhvargers, 2013; Waltman et al., 2012).

A critique specific to the ARWU ranking is its heavy reliance on Nobel Prizes and alumni with Nobel Prizes as proxies for research and teaching excellence. Since the Nobel Prize is typically given toward the end of a researcher's career, it has been pointed out that use of these indicators provides at best a historical perspective on university performance that may not relate much to current performance (Billaut et al., 2010; Van Raan, 2005). An addition problem is that the indicator based on number of Nobel Prize winners uses the affiliation at the time the award was received, not the time the research was done, and thus may reflect better on the ability of a university to attract strong researchers rather than the actual research environment. It has also been argued that use of these indicators downplays social sciences and humanities (Altbach, 2015) as well as other fields (Billaut et al., 2010) in which the Nobel Prize is not awarded.

2.4. Benchmarking systems

In addition to the numerous university ranking systems, a number of public and commercial benchmarking systems have become available in recent years. The Center for World-Class Universities (CWCU) of Shanghai Jiao Tong University launched the freely available Global Research University Profiles (GRUP, 2017) project in 2012. GRUP provides the ARWU data on the 1200 universities analyzed for the ranking as well as more detailed data on students, academic staff, and resources for participating universities. In 2015, 551 universities provided such data. The system provides benchmarking functions using 40 indicators: 13 about students, 9 about faculty, 13 about resources, and the 5 ARWU indicators.

Two notable research benchmarking exercises carried out in the context of national block grant exercises are the UK's Research Excellence Framework (REF, 2017) and Australia's Excellence in Research for Australia (ERA, 2017). The REF 2014 is carried out over 36 subject areas or "units of assessment" and based on expert peer review with one expert panel for each unit of assessment. Australia's ERA analyses research performance across 127 fields of research grouped into 22 broad areas. ERA is based on the principle of expert review informed by indicators. The indicators in ERA include volume and activity,

publishing profile, citation analysis (based on the Scopus database), peer review of a sample of research output, esteem measures, research income, and applied measures.

There are a number of commercial research benchmarking tools and services. These include Elsevier's SciVal (2017) which uses Scopus data, Thomson Reuters' InCites (2017) which uses Web of Science data, and Academic Analytics (2017). Of these SciVal is closest to our work since it uses Scopus data. But while GRBS selects universities for inclusion based on fine-grained subject area publication volume, SciVal selects universities based on overall volume of publications as well as based on request. In addition there are some differences in definition of metrics. For example, GRBS uses a 4-year h-index to show change over time, while SciVal uses an h-index that is computed over all years in the database.

In addition, some consortia of universities run their own benchmarking exercises. For example Australia's Group of Eight publishes benchmarking data for the member universities across a wide range of indicators and provides the online Group of Eight Executive Dashboard (2017).

3. Methodology

3.1. Vetting of affiliations

When benchmarking university research performance the data used must represent similar, meaningful entities in order for comparisons to be of value. There would be little sense, for example, in comparing the research output of a single university with that of an entire university system. This is a complex issue and internationally there is currently no standard for identifying what is and what is not a component part of a university. For the purposes of university research benchmarking, we use the term "university" to apply to a single institution that has substantially independent academic decision-making authority over its entities that may be geographically separated (Abbey, Capaldi, & Haddawy, 2011). GRBS considers a paper to belong to a university if at least one affiliation in the paper contains one of its entities. The situation is complex for affiliated institutions like academic hospitals. Under the judgement of country experts, those that are owned or controlled by a university are considered part of that university.

Using this definition, the Scopus data for all universities covered in GRBS was vetted in order to clean the affiliations. The Scopus database consists of affiliation IDs for institutional sub-entities that are then grouped with the main institutional affiliation. The GRBS vetting process consisted of collecting the list of all sub-entities that could possibly be associated with a given university based on string and location matching. The GRBS Advisory Board then organized experts[3] in higher education from each country covered (typically, at least one per country) to examine the proposed labeling and aggregation of entities, guided by the above definition of university. The final results were then used to procure Scopus publication data for each university covered in GRBS.

3.2. Database and subject area coverage

GRBS covers 24,963 source titles of type journal, conference proceeding, and book series from Elsevier's Scopus database. From among these source titles, the publication types included are journal articles, reviews, and conference papers. Bibliometric data for the 2012 release covers the 4-year window 2008–2011. Data in GRBS is organized into a 3-level subject hierarchy. Disciplinary categories are structured using the All Science Journal Classification (ASJC), which maps source titles in a structured hierarchy of disciplines and sub-disciplines. ASJC classifies all Scopus source titles into a two-level hierarchy. The top level contains 27 subject areas including a Multidisciplinary category and the second level contains 309 subject areas. Because GRBS currently only covers Science and Technology, only 23 of the top level subject areas and 251 of the sub-areas are used in the system. In order to work with broad subject areas of approximately the same size, GRBS combines some of the 23 top level ASJC subject areas together, resulting in the 15 broad subject areas at the top level of the hierarchy as shown in Table 1.

3.3. Country coverage and selection of universities

GRBS covers universities in three regions: North America, Europe, and Asia Pacific. North America is taken to be the USA and Canada; Europe is taken to be the European Union plus Norway and Switzerland; and Asia Pacific is taken to be the ASEAN + 3 nations, SAARC, the Pacific Islands Forum, and fourteen other countries and economies. A total of 41 countries are covered as shown in Table 2. The choice of country coverage was driven largely by the availability of experts to help with the labor intensive affiliation vetting process described above. The strategy was to expand regional coverage over time. While the selection criteria for inclusion of universities were applied to all countries in this list, not all countries contained universities meeting the selection criteria. In particular, the Asia Pacific region includes universities from eleven countries and economies out of all those analyzed. In total, GRBS covers 1343 universities: 241 in North America, 614 in Europe, and 488 in Asia Pacific.

[3] Find the list of domain experts at the following URL: http://www.researchbenchmarking.org/people.php

Table 1
Broad subject areas.

1	Agriculture and Biological Sciences
2	Biochemistry, Genetics and Molecular Biology
3	Chemistry
4	Computer Science
5	Earth and Planetary Sciences
6	Economics and Business Sciences
7	Engineering
8	Environmental Sciences
9	Health Professions and Nursing
10	Material Sciences
11	Mathematics
12	Medicine
13	Other Life and Health Sciences
14	Physics and Astronomy
15	Multidisciplinary

Table 2
Countries covered per region.

Asia Pacific	Europe	North America
Australia	Austria	Canada
China	Belgium	United States
Hong Kong	Bulgaria	
India	Cyprus	
Japan	Czech Republic	
Malaysia	Denmark	
New Zealand	Estonia	
Singapore	Finland	
South Korea	France	
Taiwan	Germany	
Thailand	Greece	
	Hungary	
	Ireland	
	Latvia	
	Lithuania	
	Luxembourg	
	Netherlands	
	Norway	
	Poland	
	Portugal	
	Romania	
	Slovakia	
	Slovenia	
	Spain	
	Sweden	
	Switzerland	
	United Kingdom	

Universities were selected for inclusion in GRBS by examining research output in the 4-year window 2008–2011 at two levels. First in each of the 251 fine-grained subject areas the universities with the highest number of publications were identified for inclusion. For Asia Pacific the top 50 were taken, for the US & Canada the top 40 were taken and the Europe the top 100 were taken. The different numbers are due to the differences in the sizes of the regions. In each category a minimum cut-off of 50 publications was applied for statistical reasons so that universities with fewer than 50 publications in the 4-year window in that subject area are not included in the list. The cutoff of 50 publications is necessary because some of the indicators lose their meaning if the volume is too low. Second, to ensure inclusion of universities with significant research activity in a broad subject area, but lacking particular fine-grained area strength, the 200 universities in each region with the highest number of publications in each of the 15 broad subject areas were identified for inclusion. Again, a cutoff of 50 publications was applied. The set of universities included in GRBS is then the union of all these resulting subject area lists. Any university that appears in at least one list is included in GRBS and analyzed in all subject areas for which it satisfies the statistical requirement for the minimum number of publications. In this way GRBS is able to recognize universities that have particular focused strengths.

Table 3
Indicators used in GRBS.

Performance Indicators	Description
Total Publications	Total number of publications during the 4-year time window.
Percentage Publications in top 25% SNIP	Percentage of Total Publications published in source titles that are within top 25% of that subject area, based on the Source Normalized Impact Per Paper (SNIP) value of the last year in the time window.
Total Citations	Total number of citations within the 4-year time window to papers published in that time window. All citation counts used in GRBS exclude author self-citations
Percentage publications in top 25% highly cited publications	Percentage of publications within the 4-year time window that are within top 25% highly cited publications of that subject area. Citations are counted within the 4-year time window.
4 –year H-Index	A university having 4-year h-index of X means that at least X of their publications (during that 4-year window) have no less than X publications citing them (during that window). H-index is computed for each subject area.
Internationality Indicators	**Description**
Internationally Co-authored Publications	Number of publications in the 4-year window that are written in collaboration with at least one researcher from outside the country in which university in located.
Percent International Collaboration	Percentage of Total Publications with international co-authorship
Total Citations of Internationally Co-authored Publications	Total number of citations received within the 4-year window by internationally co-authored publications in the 4-year window
International Citations	Citations received from papers authored (only) by researchers from outside the country in which a given university is located. This is a strong definition of international citations so that citations from internationally co-authored publications are not counted.
Percent International Citations	Percentage of International Citations relative to Total Citations
International Citations of Internationally Co-authored Publications	International citations received by Internationally Co-authored Publications.
International Impact	Measure of the impact a university's research is having outside the country in which it is located. It is defined as the ratio of International Citations to the total number of references made by the papers in a given field which are authored (only) by researchers from outside the country in which the university is located. The denominator normalizes for the size of the market of citations outside the country.

3.4. Indicators

To provide for a comprehensive evaluation of research performance, GRBS utilizes the twelve indicators shown in Table 3. The first five in the table are indicative of research performance while the other seven measure internationality of research activity and impact, which while not traditional measures of research performance, are considered important with respect to the mission of an increasing number of universities. All indicators may be used to compare university research performance and activity without aggregation.

When using indicators that depend on citation counts, an appropriate time window needs to be chosen. The window must be sufficiently wide to result in an accurate count of citations but at the same time must not be so wide as to make the indicators insensitive to changes in performance over time. Similar to the 2016 version of Leiden Ranking (Waltman et al., 2012), a 4-year window has been chosen for use in GRBS. In addition, the Higher Education Funding Council for England (HEFCE, 2009) published the results of an analysis to determine what timeframe would give a reasonable balance between allowing enough time for citation counts to grow and being unduly retrospective. The study found that a four-year publication window for citation analysis provides a measurable benefit over a two-year window in producing results that are not significantly different from those that would be obtained by using a six-year window.

We discuss the performance indicators in more detail since they are used in the analyses in the subsequent sections of this paper. These five indicators were chosen to reflect volume, quality, output, and scholarly impact of research. Volume indicators include total publications, total citations, and h-index which measures both publications and citations. All citations in GRBS exclude author self-citations and all indicators are computed within the 4-year time window. A 4-year H-index is computed for each university in each broad and fine-grained subject area. The 4-year H-index uses the same formula proposed by Hirsch (2005) for individual researchers, but limited to publications in a 4-year time window. GRBS includes two measures of scholarly quality. The percentage of publications that appear in the top 25% of source titles within each broad level field is intended to measure the quality of output and has the advantage of being measurable in a much shorter time than direct citation-based metrics. Source titles (journals, conference proceedings and book series) are ranked based on their Source Normalized Impact per Paper (SNIP) (Moed 2010) values in each of the 15 top level GRBS categories. The SNIP measures the citation impact of scientific journals using a source-normalized approach which corrects for differences in citation rates across fields. Recent studies by Ahlgren and Waltman (2014) and Haddawy, Hassan, Asghar, and Amin (2016) show that SNIP has stronger agreement with expert judgement of journal quality than journal impact factor or raw impact per paper. Waltman et al. (2012) argue that the original SNIP indicator, proposed by Moed (2010), has some counterintuitive properties, and they introduce a revised SNIP indicator that they show does not have these properties. Though some systematic differences can be observed between the original SNIP indicator and the revised one, empirical

Table 4
Number of fine-grained areas in top 10 percent for each indicator for all countries covered in GRBS.

Region	Country	Publications	Percent Publications in Top 25% SNIP	Citations	Percent Publications in Top 25% Highly Cited	4-Year H-Index	Fine-Grained Areas in top 10 percent in all indicators
North America	United States	1601	1956	2171	2184	2612	195
	Canada	207	162	223	138	292	10
Asia Pacific	China	1170	134	597	166	601	3
	Japan	278	30	147	32	140	0
	South Korea	203	97	146	56	156	0
	Australia	138	157	136	123	172	0
	Taiwan	102	162	82	79	109	5
	Singapore	82	44	91	38	102	9
	Hong Kong	41	78	62	95	80	13
	India	20	14	17	26	20	0
	Malaysia	10	7	7	9	10	0
	Thailand	4	8	2	9	3	0
	New Zealand	3	11	3	11	5	0
Europe	United Kingdom	296	511	411	628	503	22
	Netherlands	202	294	238	220	266	19
	Germany	184	182	190	294	243	1
	Sweden	84	88	83	81	100	0
	Italy	78	152	65	113	84	2
	France	70	196	56	153	70	0
	Switzerland	60	122	97	153	131	6
	Denmark	52	66	61	48	67	1
	Belgium	45	88	58	64	71	0
	Spain	31	182	35	85	55	0
	Finland	24	58	20	31	19	0
	Greece	18	25	10	32	14	0
	Poland	16	10	0	4	0	0
	Portugal	12	29	6	16	9	0
	Austria	11	26	13	26	18	0
	Czech Republic	10	4	2	1	3	0
	Norway	9	55	11	25	18	1
	Romania	4	1	4	3	2	0
	Ireland	3	25	3	28	7	0
	Slovenia	1	3	1	7	1	0
	Luxembourg	0	0	0	1	0	0
	Cyprus	0	1	0	0	0	0
	Lithuania	0	3	2	5	2	0
	Bulgaria	0	0	0	0	0	0
	Latvia	0	0	0	0	0	0
	Estonia	0	4	0	0	0	0
	Hungary	0	1	0	0	0	0
	Slovakia	0	0	0	1	0	0

Note: Countries within each region are ordered by the number of fine-grained areas in top 10 percent of publications.

differences between them turn out to be relatively small. We use the original version of the SNIP in our study. The SNIP of the last year in the time window is used, so for the window 2008–2011, the values from 2011 are used. The other measure of quality is the percentage of publications in the top 25% highly cited publications within each field – both for 15 broad subject areas and 251 fine-grained subject fields. Note that since percentages are being used rather than absolute numbers, the quality measures are independent of volume of publications and citations. In this way GRBS is able to recognize the contributions of smaller institutions or departments that may lack volume yet produce consistently high quality research.

GRBS also provides for normalization by a measure of the number of active researchers for the size dependent indicators, which is all indicators except those that represent percentages. This is done by dividing the indicator value by the number of authors in the Scopus database affiliated to that institution who published at least one paper in the 4-year time window.

4. Results and discussion

4.1. Country level analysis

We begin by examining research performance at the country level. For each of the five performance indicators, Table 4 shows the number of fine-grained areas in which universities are performing within the top 10%. Note that a university may excel in several areas and several universities in the same country may excel in any given area. All such occurrences are counted. The last column shows the number of areas in which universities in each country are performing within the top

Table 5
Distribution of fine-grained areas by research performance levels across all indicators for each country covered in GRBS.

Country	Region	Distribution (%)				Total Research Active Areas
		Top 10%	Top 25%	Top 50%	Top 75%	
North America	United States	1.64	11.11	39.27	71.57	11912
	Canada	0.46	4.27	26.36	63.62	2155
Asia Pacific	Singapore	4.21	24.30	59.35	82.24	214
	Hong Kong	2.71	11.48	37.16	69.31	479
	Taiwan	0.42	2.89	17.06	49.75	1178
	China	0.05	0.58	4.58	20.59	5998
	Japan	0.00	0.15	4.01	25.47	3239
	South Korea	0.00	1.45	9.21	34.98	1998
	Australia	0.00	3.21	25.16	60.13	1590
	India	0.00	0.25	7.19	24.09	793
	New Zealand	0.00	0.00	9.06	43.40	265
	Malaysia	0.00	0.87	2.16	9.09	231
	Thailand	0.00	0.00	1.79	26.01	223
Europe	Netherlands	1.75	16.56	52.25	80.40	1087
	Switzerland	0.97	10.88	40.42	70.94	616
	United Kingdom	0.67	7.17	33.18	65.93	3264
	Norway	0.30	1.50	20.42	57.06	333
	Denmark	0.25	5.60	33.84	70.48	393
	Italy	0.08	0.68	10.19	46.73	2493
	Germany	0.03	1.96	19.69	57.86	3372
	France	0.00	1.65	15.32	47.39	1762
	Spain	0.00	1.03	11.99	45.00	1651
	Sweden	0.00	3.57	31.28	66.26	812
	Poland	0.00	0.00	0.29	5.44	699
	Belgium	0.00	6.43	38.80	69.71	482
	Greece	0.00	0.64	10.00	42.77	470
	Finland	0.00	3.50	19.04	58.21	457
	Austria	0.00	1.83	17.20	55.50	436
	Portugal	0.00	0.94	9.67	42.22	424
	Czech Republic	0.00	0.00	0.34	11.74	298
	Ireland	0.00	1.09	9.49	44.16	274
	Hungary	0.00	0.00	0.00	17.88	179
	Romania	0.00	0.00	0.66	4.64	151
	Slovenia	0.00	0.00	6.32	34.74	95
	Slovakia	0.00	0.00	2.63	5.26	76
	Estonia	0.00	0.00	1.92	21.15	52
	Lithuania	0.00	0.00	2.33	9.30	43
	Bulgaria	0.00	0.00	0.00	4.55	22
	Cyprus	0.00	0.00	0.00	31.58	19
	Latvia	0.00	0.00	0.00	0.00	11
	Luxembourg	0.00	0.00	0.00	0.00	214

Notes: Countries are ordered by the percentage of fine-grained areas in top 10% across all indicators.

10% across all indicators. This shows consistent excellent performance along all five dimensions. Along the three volume indicators (Publications, Citations, 4-Year H-index) the top three countries are United States, China, and UK. The ranking of countries looks different for the two quality indicators. For percent publications in top 25% source titles, the top three are United States, UK, and Netherlands; while for percent highly cited papers, the top three are United States, UK, and Germany. The United States dominates along all the dimensions measured. Along the dimension of publications China is a close second behind the United States. Looking at the number of fine-grained subject areas in the top 10% across all indicators, the United States is far ahead with 195 areas, followed by UK with 22 and then Netherlands with 19, Hong Kong with 13, and Canada with 10. Within Europe, the UK and the Netherlands are significantly ahead of the remaining countries. Hong Kong's strong standing globally and the fact that it tops the list in Asia is remarkable given its relatively small size.

Since the countries analyzed vary greatly in size, it is useful to examine the distribution of research performance for each county, which provides insight into the focus on excellence. Table 5 shows the distribution of research performance among the areas in which each country is research active (last column). A university is considered to be active in a research area if it produces more than 50 publications in that area within the 4-year window. Columns three through six show the percentage of fine-grained areas in the top 10%, 25%, 50%, and 75% across all indictors. Table 5 is related to Table 4 in that the value in the last column of Table 4 is equal to the percentage in the top 10% column of Table 5 multiplied by the total research active areas. Singapore stands out in this list with a remarkably high 4.21% of its research active areas in the top 10%. This is followed by the Hong Kong at 2.71%, the Netherlands at 1.75%, and United States at 1.64%. Among these four countries, Singapore and Hong Kong have a relatively small number of research active areas, while the Netherlands has a moderate number, and United States has the largest number. The fact that many countries have no areas in the top 10% across all

Table 6a
Universities with at least five fine-grained areas in Top 10% across all indicators (total 16 universities).

Rank	Scope	University	Country	Top 10% in all indicators	Top 25% in all indicators [Rank]
1	Medium	University of California – San Francisco	United States	24	45 [7]
2	Broad	Harvard University	United States	22	62 [1]
3	Broad	Stanford University	United States	16	59 [2]
3	Broad	Massachusetts Institute of Technology	United States	16	48 [5]
5	Broad	University of California, Berkeley	United States	12	40 [10]
6	Broad	National University of Singapore	Singapore	9	31 [15]
6	Medium	Erasmus University Rotterdam	Netherlands	9	28 [22]
8	Broad	The University of Oxford	United Kingdom	8	50 [3]
8	Medium	City University of Hong Kong	Hong Kong	8	15 [46]
10	Broad	The University of Cambridge	United Kingdom	7	42 [9]
11	Broad	University of California – San Diego	United States	6	45 [7]
11	Broad	Columbia University in the City of New York	United States	6	38 [11]
11	Broad	University of Pennsylvania	United States	6	37 [12]
11	Medium	Washington University in St. Louis	United States	6	31 [15]
11	Broad	Duke University	United States	6	29 [21]
16	Narrow	University of Texas – M. D. Anderson Cancer Center	United States	5	13 [52]

Note: Universities are ordered by the number of fine-grained areas in top 10 percent across all indicators.

indicators shows how difficult this is to achieve. In Asia Pacific the counties with highest concentration of research active areas in the top 10% are Singapore, Hong Kong, Taiwan, and China. In Europe the countries with the highest concentration in the top 10% are the Netherlands, Switzerland, and the UK. Looking at percentage of research active areas in the top 25%, again Singapore leads with 24.30%, followed by the Netherlands (16.56%), Hong Kong (11.48%), United States (11.11%), and Switzerland (10.88%).

4.2. University level analysis

We now drill down to examine research performance at the university level. To facilitate comparison of universities of varying breadth of research activity, universities are grouped into three categories: broad scope (active in at least 110 fine-grained subject areas), medium scope (active in 60–110 areas), and narrow scope (active in fewer than 60 areas). The definition of active research area is as given in the previous section. Of the 1336 universities covered in GRBS, 97 are classified as broad, 214 as medium, and 1025 as narrow.

4.2.1. Fine-grained and broad area performance

Table 6a shows the sixteen universities with at least five fine-grained areas in the top 10% across all five performance indicators. At the top of the list is UC – San Francisco with 24 areas, followed closely by Harvard with 22 and then Stanford and MIT with 16 each. The universities in the list come from five countries with eleven from N. America (USA), two from Asia Pacific (Singapore and Hong Kong), and three from Europe (UK 2 and Netherlands 1). Among the universities in this list, eleven or 69% are broad scope, four are medium, and one is narrow. Notably, the university at the top of the list, UC – San Francisco, is a medium scope university.

If the threshold is relaxed to count areas in the top 25% as shown in the last column, the top ranked universities are now Harvard with 62 areas, Stanford with 59 areas, Oxford with 50 areas, University of Michigan – Ann Arbor with 49, and MIT with 48 areas. Michigan has only four areas in the top 10% and so does not appear in the table.

To illustrate how choice of subject level granularity affects the measurement of university research excellence, we carry out the same analyses of university research performance as in Table 6a but using the fifteen broad subject categories listed in Table 1. Table 6b shows the 15 universities with at least five broad areas in the top 10% across all five performance indicators. The list is now dominated by USA with eleven universities and the UK with three. The only other country in the list is Singapore. Out of the fifteen universities, fourteen or 93% are broad scope. Seven universities that appeared in Table 6a do not appear in 6b, two of which are broad scope, four medium scope and one narrow scope. Table 6c lists the fine-grained areas in the top 10% for the three highest ranked universities in Table 6a that do not appear in Table 6b: University of California – San Francisco, Erasmus University, and City University of Hong Kong. The broad areas under which the fine-grained strengths fall are also shown in the table. In terms of broad area strength, UC – San Francisco is in the top 10% in four areas: Medicine (all); Biochemistry, Genetics and Molecular Biology (all); Other Life and Health Sciences;

Table 6b
Universities with at least five broad areas in top 10% across all indicators (total 15 universities).

Rank	Scope	University	Country	Top 10% in all indicators	Top 25% in all indicators [Rank]
1	Broad	Stanford University	United States	13	13 [7]
2	Broad	Harvard University	United States	11	15 [1]
2	Broad	Massachusetts Institute of Technology	United States	11	14 [2]
4	Broad	University of California, Berkeley	United States	10	13 [7]
5	Broad	University of California – Los Angeles	United States	9	13 [7]
6	Broad	Imperial College	United Kingdom	7	14 [2]
6	Broad	University Michigan – Ann Arbor	United States	7	14 [2]
6	Broad	The University of Cambridge	United Kingdom	7	12 [14]
9	Broad	University of California – San Diego	United States	6	14 [2]
9	Broad	University of Washington – Seattle	United States	6	14 [2]
11	Broad	The University of Oxford	United Kingdom	5	13 [7]
11	Broad	National University of Singapore	Singapore	5	11 [16]
11	Broad	University of Texas – Austin	United States	5	10 [20]
11	Broad	Columbia University in the City of New York	United States	5	9 [30]
11	Medium	Princeton University	United States	5	9 [30]

Note: Universities are ordered by the number of broad areas in top 10 percent across all indicators.

and Multidisciplinary. Erasmus University is also in the top 10% in four areas: Medicine (all), Biochemistry, Genetics and Molecular Biology (all), Other Life and Health Sciences (all), and Health Professions and Nursing (all). City University of Hong Kong is in the top 10% in two broad areas: Computer Science (all), and Mathematics (all). The fine-grained analysis provides important information missing from the broad area analysis alone. For example, although UC – San Francisco is in the top 10% in the area of Medicine, it has fine-grained strength at the top 10% in only 11 of the total 48 areas under Medicine. And conversely, although the university is in the top 10% in two fine-grained areas under Health Professions and Nursing (all), it is not in the top 10% in that broad area. Similarly, although City University is in the top 10% in three fine-grained areas under Engineering (all), it is not in the top 10% in the broad area. Finally, comparing UC – San Francisco and Erasmus University, although they are both in the top 10% in the broad area of Medicine (all), their specific profiles differ significantly, with only three fine-grained areas of excellence in common in this category. These points illustrate that the two granularities offer complementary perspectives and it is important to include both to provide comprehensive insightful analysis of university research performance. These three medium scope universities are also ranked quite low by the traditional ranking systems, as shown in Table 8, which further shows how universities that are performing at a world class level in terms of fine-grained subject areas can be completely overlooked if one takes a broad area perspective and how the broad area analysis can disproportionally exclude universities with more narrow focus.

4.2.2. Distribution of research strength

Similar to the country-level analysis shown in Table 5, Table 7 shows the distribution of research strength among the fine-grained areas in which each university is research active and provides an indication of the extent to which each university has a focus on research excellence. The universities in the list are ordered by the percentage of research active areas rated in the top 10% across all five performance indicators. The top three universities in the list are the same as in Table 6a: UC San Francisco, Harvard, and MIT. UC San Francisco has a remarkable concentration of 25.6% of its research active areas in the top 10%. In comparison, the next university, Harvard, has 15.07%. UC San Francisco also has the highest percentage (47.37%) in the top 25%. City University of Hong Kong and Erasmus University are ranked 4 and 5, respectively, although their percentages differ by a very small amount and Erasmus University has, in fact, a significantly higher percentage than City University of Hong Kong in the top 25%. It is also noteworthy that University of Texas – M. D. Anderson Cancer Center is ranked 7th in this list, compared with 16th in Table 6a. We might expect more narrowly focused universities to have a higher percentage of their research active areas in the top 10% and this is, in fact, the case. Among the top five universities in this table, three are medium scope. We would also expect to have some correspondence between the country level results in Table 5 and the university level results in Table 7. This too is the case. The top four countries with high percentages of their fine-grained areas in the top 10% (Singapore, Hong Kong, Netherlands, and United States) all have universities (National University of Singapore, City University of Hong Kong, Erasmus University, UC San Francisco) with correspondingly high percentages in the top 10% in Table 7. In the case of Singapore, the National University of Singapore contributes all of the areas in the top 10% for the country.

Table 6c
Top three ranked universities in terms of fine-grained areas in top 10% that do not appear in list of top universities by broad subject areas. The table shows the list of fine-grained areas in top 10%, broad areas under which they are classified, and number of fine-grained areas in top 10% out of total in that broad area.

University of California – San Francisco		Erasmus University		City University of Hong Kong	
Broad Area Classification	Fine-grained Areas in top 10% (24 total)	Broad Area Classification	Fine-grained Areas in top 10% (9 total)	Broad Area Classification	Fine-grained Areas in top 10% (8 total)
Medicine (all) (11 fine-grained areas of 48 total)	Clinical Neurology	Medicine (all) (7 fine-grained areas of 48 total)	Clinical Neurology	Computer Science (all) (3 fine-grained areas of 12 total)	Computer Science Applications
	Oncology		Oncology		Software
	Immunology & Allergy		Immunology & Allergy		Artificial Intelligence
	Infectious Diseases		Hematology		Information Systems
	Surgery		Public Health, Environmental and Occupational Health	Engineering (all) (3 fine-grained areas of 29 total)	Electrical and Electronic Engineering
	Pharmacology (medical)		Obstetrics and Gynecology		Control and Systems Engineering
	Endocrinology, Diabetes, and Metabolism		Reproductive Medicine		Civil and Structural Engineering
	Geriatrics and Gerontology	Other Life and Health Sciences (1 fine-grained area of 20 total)	Immunology	Mathematics (all) (1 fine-grained area of 14 total)	Applied Mathematics
	Pathology and Forensic Medicine	Biochemistry, Genetics and Molecular Biology (all) (1 fine-grained area of 15 total)	Genetics		
	Microbiology (medical)				
	Histology				
Biochemistry, Genetics and Molecular Biology (all) (8 fine-grained areas of 15 total)	Molecular Biology				
	Cell Biology				
	Biochemistry				
	Cancer Research				
	Physiology				
	Developmental Biology				
	Endocrinology				
	Clinical Biochemistry				
Other Life and Health Sciences (3 fine-grained areas of 20 total)	Immunology				
	Pharmacology				
	Drug Discovery				
Health Professions & Nursing (all) (2 fine-grained areas of 39 total)	Oncology (nursing)				
	Care Planning				

4.2.3. Comparison with predominant ranking systems

Table 8 shows a comparison of research ratings using the GRBS methodology with the three predominant global university ranking systems: Academic Ranking of World Universities (ARWU), Times Higher Education World University Ranking (THE-WUR), and QS World University Ranking (QS-WUR).[4] The table compares fine-grained subject area performance of the 79

[4] For a comparison of the rankings of these three ranking systems see the work of Moed (2016).

Table 7
Distribution of fine-grained areas by research performance levels across all indicators for universities presented in Table 6a.

University [Rank from Table 6a]	Country	Distribution (%)				Number of Research Active Areas	Scope
		Top 10%	Top 25%	Top 50%	Top 75%		
University of California – San Francisco [1]	United States	25.26	47.37	70.53	90.53	95	Medium
Harvard University [2]	United States	15.07	42.47	73.97	91.78	146	Broad
Massachusetts Institute of Technology [3]	United States	13.22	39.67	75.21	90.91	121	Broad
City University of Hong Kong [8]	Hong Kong	11.76	22.06	55.88	76.47	68	Medium
Erasmus University Rotterdam [6]	Netherlands	11.25	35	71.25	85	80	Medium
Stanford University [3]	United States	10.19	37.58	73.25	92.36	157	Broad
University of Texas – M. D. Anderson Cancer Center [16]	United States	9.26	24.07	53.7	81.48	54	Narrow
University of California, Berkeley [5]	United States	8.63	28.78	65.47	87.05	139	Broad
National University of Singapore [6]	Singapore	6.87	23.66	54.2	78.63	131	Broad
The University of Oxford [8]	United Kingdom	5.63	35.21	65.49	87.32	142	Broad
Washington University in St. Louis [11]	United States	5.56	28.7	62.96	87.04	108	Medium
The University of Cambridge [10]	United Kingdom	4.76	28.57	63.27	85.03	147	Broad
Duke University [11]	United States	4.76	23.02	57.94	87.3	126	Broad
University of Pennsylvania [11]	United States	4.62	28.46	73.85	90.77	130	Broad
University of California – San Diego [11]	United States	4.26	31.91	66.67	90.78	141	Broad
Columbia University in the City of New York [11]	United States	4.23	26.76	64.79	80.99	142	Broad

Notes: Universities are ordered by the percentage of fine-grained areas top 10% across all indicators.

universities with at least one fine-grained area in the top 10% across all five GRBS performance indicators, showing their rankings within the different systems. The entries in the table are ordered first by the number of areas in the top 10% and ties on that measure are ordered by the number of areas in the top 25%. Table 9 shows the correlation between the various rankings. For universities given ranking ranges by the ranking systems, the midpoint was used. Universities not ranked in the other systems were not included in the computation of correlation. The correlation between world class fine-grained area ranking and the other three ranking systems shows approximately equal moderate correlation with all three: 0.580 with ARWU, 0.552 with THE-WUR and 0.537 with QS-WUR (1-tailed Spearman's rho, significant at the 0.01 level). With regard to the top ranked universities, the world class fine-grained area ranking has the closest agreement with ARWU: The universities ranked 2 through 5 according to world class fine-grained areas are ranked 1 through 4 in ARWU.

There are a number notable cases of universities that rank highly in terms of fine-grained area performance but are either ranked low in the other ranking systems or not ranked at all. The University of California – San Francisco which ranks first in terms of fine-grained performance in the top 10% ranks 18th in ARWU and does not appear in the THE-WUR or QS-WUR. Other examples are Erasmus University Rotterdam, City University of Hong Kong, University of Texas – M.D. Anderson Cancer Center, and University of California – Santa Cruz. All of these are medium or narrow scope universities. While differences between the GRBS fine-grained area ranking and the three ranking systems is not surprising given they are using different sets of indicators, the very low ranking or lack of coverage of some excellent narrow and medium scope universities in the three ranking systems indicates that the rankings may be systematically overlooking an important class of universities captured by the GRBS methodology.

5. Conclusions

The Global Research Benchmarking System provides university benchmarking data to support strategic decision making that is sufficiently fine-grained to show variation within a university and to identify focused areas of excellence; is actionable; is objective and verifiable; allows tracking of performance over time; and provides meaningful comparisons across the diversity of national higher education environments and university structures. By covering 251 fine-grained subject areas and selecting universities for inclusion based on their fine-grained area performance, GRBS is able to shed light on variation in a single institution as well as to highlight excellence within universities that may not be considered world class overall. Comparison between fine-grained and broad subject area analyses as well as comparison with the predominant rankings highlight the fact that performance analyses carried out only at the aggregate university level or even at the broad subject level miss significant pockets of research excellence, possibly creating the false impression of a lower level of research

Table 8
Universities with top 10% research performance across all indicators in at least one fine-grained area and their positions in selected international university rankings.

Rank	University	Country	Fine-Grained Areas in top 10%	Fine-Grained Areas in top 25% [Rank]	ARWU 2012	THE-WUR 2012	QS-WUR 2012	Scope
1	University of California – San Francisco	United States	24	45 [7]	18	N/A	N/A	Medium
2	Harvard University	United States	22	62 [1]	1	4	3	Broad
3	Stanford University	United States	16	59 [2]	2	2	15	Broad
3	Massachusetts Institute of Technology	United States	16	48 [5]	3	5	1	Broad
5	University of California, Berkeley	United States	12	40 [10]	4	9	22	Broad
6	National University of Singapore	Singapore	9	31 [15]	101–150	29	25	Broad
6	Erasmus University Rotterdam	Netherlands	9	28 [21]	151–200	72	99	Medium
8	The University of Oxford	United Kingdom	8	50 [3]	10	2	5	Broad
8	City University of Hong Kong	Hong Kong	8	15 [41]	201–300	182	95	Medium
10	The University of Cambridge	United Kingdom	7	42 [9]	5	7	2	Broad
11	University of California – San Diego	United States	6	45 [7]	15	38	70	Broad
11	Columbia University in the City of New York	United States	6	38 [11]	8	14	11	Broad
11	University of Pennsylvania	United States	6	37 [12]	14	15	12	Broad
11	Washington University in St. Louis	United States	6	31 [15]	31	44	84	Medium
11	Duke University	United States	6	29 [20]	36	23	20	Broad
16	University of Texas – M. D. Anderson Cancer Center	United States	5	13 [43]	151–200	N/A	N/A	Narrow
17	University Michigan – Ann Arbor	United States	4	49 [4]	22	20	17	Broad
17	University of Washington – Seattle	United States	4	46 [6]	16	24	59	Broad
17	University of California – Los Angeles	United States	4	34 [13]	12	13	31	Broad
17	Federal Institute of Technology Zurich	Switzerland	4	31 [15]	23	12	13	Broad
17	Boston University	United States	4	22 [27]	71	54	64	Medium
17	University of Chicago	United States	4	18 [36]	9	10	8	Medium
17	University of Pittsburgh	United States	4	17 [38]	58	76	98	Broad
24	Imperial College	United Kingdom	3	33 [14]	24	8	6	Broad

Table 8 (*Continued*)

Rank	University	Country	Fine-Grained Areas in top 10%	Fine-Grained Areas in top 25% [Rank]	ARWU 2012	THE-WUR 2012	QS-WUR 2012	Scope
24	University of Illinois – Urbana-Champaign	United States	3	23 [25]	25	33	56	Broad
24	University of Texas – Austin	United States	3	22 [27]	35	25	68	Broad
24	University of Amsterdam	Netherlands	3	19 [34]	101–150	83	62	Broad
24	Princeton University	United States	3	16 [40]	7	6	9	Medium
24	Emory University	United States	3	13 [43]	101–150	79	122	Medium
24	University of Alberta	Canada	3	12 [45]	101–150	121	108	Broad
24	Hong Kong University of Science and Technology	Hong Kong	3	12 [45]	201–300	65	33	Medium
24	National Cheng Kung University	Taiwan	3	9 [57]	201–300	301–350	271	Broad
24	New York University	United States	3	9 [57]	27	41	43	Medium
34	Utrecht University	Netherlands	2	30 [19]	53	67	85	Broad
34	University of Minnesota – Twin Cities	United States	2	28 [21]	29	47	104	Broad
34	University of Toronto	Canada	2	27 [23]	27	21	19	Broad
34	University College London	United Kingdom	2	25 [24]	21	17	4	Broad
34	Yale University	United States	2	22 [27]	11	11	7	Broad
34	Northwestern University	United States	2	22 [27]	30	19	27	Broad
34	VU University Amsterdam	Netherlands	2	21 [31]	101–150	140	177	Broad
34	The University of British Columbia	Canada	2	19 [34]	39	30	45	Broad
34	University of Maryland – College Park	United States	2	17 [38]	38	97	117	Broad
34	University of California – Santa Barbara	United States	2	15 [41]	34	35	118	Medium
34	University of Texas Southwestern Medical Center	United States	2	12 [45]	48	N/A	N/A	Narrow
34	Pennsylvania State University – University Park	United States	2	12 [45]	49	61	101	Broad
34	Tufts University	United States	2	11 [49]	101–150	87	181	Medium
34	Hong Kong Polytechnic University	Hong Kong	2	11 [49]	201–300	251–275	159	Medium
34	Mount Sinai School of Medicine	United States	2	10 [53]	151–200		N/A	Medium
34	University of Waterloo	Canada	2	8 [59]	151–200	226–250	191	Medium
34	University of California – Santa Cruz	United States	2	5 [62]	N/A	N/A	N/A	Narrow

Table 8 (*Continued*)

Rank	University	Country	Fine-Grained Areas in top 10%	Fine-Grained Areas in top 25% [Rank]	ARWU 2012	THE-WUR 2012	QS-WUR 2012	Scope
51	University of Wisconsin – Madison	United States	1	31 [15]	19	31	38	Broad
51	University of North Carolina at Chapel Hill	United States	1	23 [25]	41	42	57	Broad
51	Cornell University	United States	1	21 [31]	13	18	14	Broad
51	University of California – Davis	United States	1	20 [33]	47	44	100	Broad
51	Leiden University	Netherlands	1	18 [36]	73	64	75	Medium
51	Baylor College of Medicine	United States	1	11 [49]	101–150	N/A	N/A	Medium
51	Technical University of Denmark	Denmark	1	11 [49]	151–200	149	132	Medium
51	Federal Institute of Technology Lausanne	Switzerland	1	10 [53]	101–150	40	29	Medium
51	Radboud University Nijmegen	Netherlands	1	10 [53]	101–150	127	136	Medium
51	Eindhoven University of Technology	Netherlands	1	10 [53]	301–400	114	158	Medium
51	The University of Sheffield	United Kingdom	1	6 [60]	101–150	110	66	Medium
51	University of Utah	United States	1	6 [60]	82	134	256	Broad
51	Iowa State University	United States	1	5 [62]	151–200	193	319	Medium
51	National Chiao Tung University Taiwan	Taiwan	1	5 [62]	301–400	251–275	238	Medium
51	University of Hawaii at Manoa	United States	1	4 [65]	101–150			Medium
51	McMaster University	Canada	1	4 [65]	92	88	152	Medium
51	Indiana University – Bloomington	United States	1	4 [65]	N/A	N/A	N/A	Medium
51	Tübingen University	Germany	1	4 [65]	N/A	N/A	N/A	Medium
51	University of Bern	Switzerland	1	3 [69]	151–200	151	149	Medium
51	Louisiana State University – Baton Rouge	United States	1	3 [69]	201–300	N/A	N/A	Medium
51	Michigan State University	United States	1	3 [69]	96	94	174	Broad
51	Peking University	China	1	2 [72]	151–200	46	44	Broad
51	London School of Hygiene and Tropical Medicine	United Kingdom	1	2 [72]	401–500	N/A	N/A	Narrow
51	Southeast University	China	1	2 [72]	401–500	N/A	551- 600	Medium
51	Norwegian School of Sport Sciences	Norway	1	2 [72]	N/A	N/A	N/A	Narrow

Table 8 (*Continued*)

Rank	University	Country	Fine-Grained Areas in top 10%	Fine-Grained Areas in top 25% [Rank]	ARWU 2012	THE-WUR 2012	QS-WUR 2012	Scope
51	National Taiwan University of Science and Technology	Taiwan	1	2 [72]	N/A	351–400	N/A	Narrow
51	University "La Sapienza"	Italy	1	1 [77]	101–150	301–350	216	Broad
51	Wuhan University	China	1	1 [77]	401–500	N/A	451- 500	Medium
51	University Vita-Salute San Raffaele	Italy	1	1 [77]	N/A	N/A	N/A	Narrow

Notes: Universities are ordered by the number of fine-grained areas in top 10 percent across all indicators. N/A denotes not ranked.

Table 9
Correlations (Spearman's rho) among rankings for those universities listed in Table 8.

	GRBS	ARUW	THE-WUR	QS-WUR	
GRBS	Correlation Coefficient	1.000	0.580**	0.552**	0.537**
	Sig. (1-tailed)		0.000	0.000	0.000
	N	63	63	63	63
ARUW	Correlation Coefficient	0.580**	1.000	0.882**	0.767**
	Sig. (1-tailed)	0.000		0.000	0.000
	N	63	63	63	63
THE-WUR	Correlation Coefficient	0.552**	0.882**	1.000	0.922**
	Sig. (1-tailed)	0.000	0.000		0.000
	N	63	63	63	63
QS-WUR	Correlation Coefficient	0.537**	0.767**	0.922**	1.000
	Sig. (1-tailed)	0.000	0.000	0.000	
	N	63	63	63	63

**. Correlation is significant at the 0.01 level (1-tailed).

performance at some universities than is the case. The GRBS analysis methodology also supports mission specialization rather than the normalized institutional model promoted indirectly by many of the rankings that operate at the institutional level.

The reported work has some limitations that need to be kept in mind. The current version of GRBS covers only fields of Science and Technology, so some universities with major strengths in humanities and social sciences may not appear in our lists of top universities by fine-grained area performance. The decision to cover Science and Technology was a conscious one since there is a strong consensus that bibliometric indicators are a valid way to measure performance in these areas. In addition, the methodology of counting number of fine-grained areas in which universities excel means that all fields are treated equally, despite the fact that some fields are larger than others. A possible alternative analysis methodology would examine weighted sums of fields with weights of fields varying based on a measure of their size.

Author contributions

Peter Haddawy: Conceived and designed the analysis; Contributed data and analysis tools; Performed the analysis; Wrote the paper.

Saeed-Ul Hassan: Conceived and designed the analysis; Collected the data; Contributed data and analysis tools; Performed the analysis; Wrote the paper.

Craig W. Abbey: Conceived and designed the analysis; Performed the analysis.

Lee Inn Beng: Conceived and designed the analysis.

Acknowledgements

We would like to thank the GRBS international advisory board for their significant contributions to the design and creation of GRBS: Elizabeth Capaldi Phillips, John V. Lombardi, Hyun-Ku Rhee, Terry Nolan, Henk Moed, Andrea Bonaccorsi, Wu Yishan, Yuan Sun, Rujhan Mustafa, RF Shangraw, H.A. Ranganath, and Kazuo Okamoto. This work was partially supported through a fellowship from the Hanse-Wissenschaftskolleg Institute for Advanced Study, Delmenhorst, Germany to Haddawy.

References

ARWU. (2017). *Academic Ranking of World Universities*.. http://www.shanghairanking.com/, [Accessed 21 January 2017]

Abbey, C., Capaldi, E., & Haddawy, P. (2011). *Defining the Concept of University for Research Benchmarking*. http://www.researchbenchmarking.org/files/Defining%20a%20University.pdf, [Accessed 2 February 2017]

Academic Analytics. (2017). *Benchmarking for Academic Excellence*.. http://www.academicanalytics.com/, [Accessed 22 January 2017]

Ahlgren, P., & Waltman, L. (2014). The correlation between citation-based and expert-based assessments of publication channels: SNIP and SJR vs Norwegian quality assessments. *Journal of Informetrics*, *8*(4), 985–996.

Altbach, P. G. (2015). The dilemmas of ranking. *International Higher Education*, *42*.

Baty, P., & Hong, B. C. (2012). Rankings without reason. *Inside Higher Education*, 31.

Billaut, J. C., Bouyssou, D., & Vincke, P. (2010). Should you believe in the Shanghai ranking? *Scientometrics*, *84*(1), 237–263.

Bowman, N. A., & Bastedo, M. N. (2011). Anchoring effects in world university rankings: exploring biases in reputation scores. *Higher Education*, *61*(4), 431–444.

CWTS Leiden. (2017). *Leiden Ranking*. https://www.leidenranking.com/, [Accessed 18 January 2017]

ERA. (2017). *Excellence in Research for Australia*.. http://www.arc.gov.au/era/, [Accessed 20 January 2017]

GRUP. (2017). *Global Research University Profiles*.. http://www.shanghairanking.com/grup/index.html, [Accessed 16 January 2017]

Group of Eight Executive Dashboard (2017), https://go8.edu.au, [Accessed 04 January 2017].

HEFCE. (2009). *Report on the pilot exercise to develop bibliometric indicators for the Research Excellence Framework, Sept 2009*.. http://www.hefce.ac.uk/pubs/year/2009/200939/, [Accessed 14 August 2014]

Haddawy, P., Hassan, S. U., Asghar, A., & Amin, S. (2016). A Comprehensive Examination of the Relation of Three Citation-Based Journal Metrics to Expert Judgment of Journal Quality. *Journal of Informetrics*, *10*(1), 162–173.

Hazelkorn, E. (2014). Reflections on a decade of global rankings: what we've learned and outstanding issues. *European Journal of Education*, *49*(1), 12–28.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proc. National Academy of Sciences*, *102*(46), 16569–16572.

InCites. (2017). *Supporting objective analysis of people, programs and peers by connecting research to impact*.. https://incites.thomsonreuters.com, [Accessed 21 January 2017]

Kehm, B. M. (2014). Global University Rankings —Impacts and Unintended Side Effects. *European Journal of Education*, *49*(1), 102–112.

Marginson, S. (2014). University Ranking and Social Science. *European Journal of Education*, *49*(1), 45–59.

Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, *4*(3), 265–277.

Moed, H. F. (2016). A critical comparative analysis of five world university rankings. *Scientometrics*, 1–24.

QS (2017), http://www.topuniversities.com/university-rankings, [Accessed 21 January 2017].

REF. (2017). *Research Excellence Framework*.. http://www.ref.ac.uk/, [Accessed 20 January 2017]

Rauhvargers, A. (2013). *Global university rankings and their impact, report II*. European University Association. http://www.eua.be/Libraries/publications-homepage-list/Global_University_Rankings_and_Their_Impact.pdf, [Accessed 20 January 2017]

Salmi, J. (2009). *The Challenge of Establishing World-Class Universities*. The World Bank. ISBN: 978-0-8213-7865-6.

SciVal (2017), https://www.elsevier.com/solutions/scival, [Accessed 14 January 2017].

Scimago. (2017). *Scimago Institutes Ranking*.. http://www.scimagoir.com/, [Accessed 17 January 2014]

THE. (2017). *Times Higher Education*. http://www.timeshighereducation.co.uk/world-university-rankings/, [Accessed 02 January 2017]

U-Multirank (2017), http://www.u-multirank.eu/, [Assessed 03 January 2017].

Van Raan, A. F. (2005). Challenges in ranking of universities. In *Invited paper for the first international conference on world class universities, shanghai jaio tong university* (pp. 133–143).

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., van Eck, N. J., et al. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, *63*(12), 2419–2432.