Website: http: jebhpme.ssu.ac.ir
EISSN: 2538-4716
Evid Based Health Policy Manag Econ 2026; 10: 69-85
doi: 10.18502/jebhpme.v10i7.23222

REVIEW ARTICLE

Institutional Size and Peak-Faculty Bibliometric Performance: A Theory-Driven Analysis of Research Productivity and Visibility

Kazem Barzegar

Department of Foreign Languages, School of Medicine, Yazd Shahid Sadoughi University of Medical Sciences, Yazd, Iran

Corresponding Author: Kazem Barzegar, Email: farzadbavanati1350@gmail.com, Tel: +98 935 2371469

Received: 17 August 2026 **Revised:** 11 October 2026 **Accepted:** 29 October 2026

ABSTRACT

Introduction: Research assessment increasingly relies on bibliometric indicators to represent scholarly productivity, impact, and institutional research performance. However, whether such indicators adequately represent the multidimensional nature of academic research remains contested.

Materials and Methods: This theory-driven analysis reexamined the published dataset of a previously published work, which compared 51 Iranian medical universities. The study examined the relationship between institutional size and peak-faculty bibliometric performance across the selected institutions. For each institution, one researcher was selected as the top-performing faculty member according to the study's predefined selection criteria, and all bibliometric outcomes were calculated for that individual researcher. Rather than treating publication counts, citations, h-index, and citations per paper as interchangeable indicators of institutional excellence, the present analysis conceptualized the dataset through three complementary perspectives: institutional size, peak-performance concentration, and responsible research assessment. Secondary statistical analyses included Pearson and Spearman correlations, log-transformed regression models, distributional analysis, and faculty-size-normalized indicators.

Results: Faculty size was positively associated with articles, citations, h-index, and citations per paper, but the strength of these relationships differed substantially. The strongest association was observed between faculty size and citations ($r = 0.837$), whereas the association with citations per paper was considerably weaker ($r = 0.448$). Faculty-size inequality was also substantial: the 10 largest universities accounted for approximately 55.3% of all faculty represented in the dataset. At the same time, several smaller institutions exhibited disproportionately high peak-performance indicators, demonstrating that institutional scale and scholarly distinction are not equivalent constructs.

Conclusion: The findings challenged interpretations of bibliometric rankings as direct representations of institutional quality and supported a multidimensional assessment model combining capacity, research output, impact, concentration, disciplinary context, and qualitative evidence. The analysis concluded that responsible assessment should treat bibliometrics as contextual evidence rather than definitive measures of institutional excellence.

Keywords: Bibliometrics; Medical Research; Research; Universities; Iran

Citation: Barzegar K. **Institutional Size and Peak-Faculty Bibliometric Performance: A Theory-Driven Analysis of Research Productivity and Visibility.** Evidence Based Health Policy, Management & Economics 2026; 10: 69-85.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

Copyright: ©2026 The Author(s); Published by Shahid Sadoughi University of Medical Sciences. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/ licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Universities are simultaneously educational institutions, knowledge-producing organizations, professional communities, and components of national innovation systems. Similarly, medical universities occupy an especially complex position because they combine education, clinical service, public health responsibilities, research, professional training, and increasingly intensive expectations concerning international scholarly visibility. Consequently, evaluating the performance of medical universities through a single indicator is conceptually problematic. Research productivity may be represented by publication volume, citation impact, international collaboration, methodological quality, societal influence, clinical translation, educational contribution, or the capacity to develop sustainable research communities (1). These dimensions do not necessarily converge within the same institution or researcher.

Bibliometric indicators nevertheless occupy an influential position in contemporary academic evaluation. Publication counts, citation counts, citations per paper, and the h-index are attractive because they transform complex scholarly activity into apparently objective numerical representations. Their apparent simplicity facilitates comparison between researchers and institutions and can support administrative decision-making. Yet, the increasing use of such indicators has generated substantial concern regarding construct validity, field differences, career-stage effects, cumulative advantage, and the tendency to mistake measurable attributes for the underlying construct of research quality.

The debate has intensified in recent years. The San Francisco Declaration on Research Assessment (DORA) (2) argues that quantitative indicators should not be used as substitutes for assessment of the quality of individual research outputs and recommends broader consideration of research contributions and impact. The Leiden Manifesto (3) similarly emphasizes that quantitative evaluation should support rather than replace qualitative expert judgment, that indicators should be interpreted in context, and that researchers should be evaluated through qualitative assessment of their portfolios rather than isolated numerical indicators. More recently, the Coalition for Advancing Research Assessment (CoARA) (4) has explicitly advocated recognition of diverse research contributions, the qualitative evaluation supported by responsible quantitative indicators, and avoidance of inappropriate uses of publication-based metrics such as the h-index.

These developments make older bibliometric datasets valuable for a reason that extends beyond historical description. They can be reexamined theoretically to determine what numerical patterns actually tell us, and equally importantly, what they do not tell us.

The present study therefore revisits the dataset reported by Barzegar and Barzegar (1). The original study compared 51 Iranian medical universities affiliated with the Iranian Ministry of Health and Medical Education. It examined the number of faculty members at each institution and selected bibliometric characteristics of the top faculty member, including number of articles, citations, h-index, citations per paper, scientific rank, and specialty. The original dataset represented 18,972 registered medical faculty members.

The original article concluded that substantial imbalance existed among Iranian medical universities and argued for a more balanced distribution of faculty members and resources. Moreover, the present analysis retained that observation but asked a different theoretical question:

What does the relationship between institutional size and the performance of the selected top faculty member actually reveal about research capacity, concentration, and institutional excellence?

This distinction is important. A university with 1,900 faculty members and a highly cited individual cannot straightforwardly be compared with a university with 40 faculty members and an

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

 CC BY 4.0

exceptionally productive individual as though their bibliometric values represented the same institutional phenomenon. Similarly, an h-index of 60 and an h-index of 20 cannot automatically be interpreted as a threefold difference in "research quality." Such comparisons may capture career length, field-specific citation practices, institutional resources, selection effects, cumulative advantage, or the exceptional productivity of a small number of researchers.

The present study consequently reconceptualized the original dataset as a case of multidimensional institutional research assessment. It distinguished institutional capacity from peak scholarly performance and absolute impact from impact relative to institutional size. In doing so, the analysis contributed to the ongoing debate about responsible research assessment.

Research Gap

Three interrelated gaps motivated this secondary analysis. First, the original study treated institutional faculty size and the bibliometric profile of a selected top faculty member as components of a common comparative framework. This was useful descriptively, but the analytical relationship between institutional capacity and individual peak performance was not fully theorized. The original analysis reported correlations among articles, scientific rank, h-index, citations per paper, and total citations, but did not examine how faculty size itself predicted these indicators.

Then, the original paper described itself as a "meta-analysis," although its methodology was primarily a descriptive-analytic comparison of institutional-level data. It did not synthesize independent effect sizes from multiple studies, which is the conventional methodological meaning of meta-analysis. Instead, the authors manually extracted institutional and faculty-level information from the Electronic Scientific Measurement System and analyzed the resulting 51 observations using descriptive statistics, rank ordering, and correlations. A theory-driven secondary analysis can therefore clarify the epistemological status of the dataset: it is best understood as a cross-institutional scientometric dataset, not a conventional meta-analytic sample.

Furthermore, developments in research-assessment theory since 2019 provide a substantially different interpretive environment. Research assessment has moved increasingly toward responsible metrics, multidimensional evaluation, and explicit recognition of contextual differences. DORA's current guidance (5) emphasizes clarity, transparency, specificity, contextualization, and fairness in the use of indicators. CoARA (4) similarly recommends that quantitative indicators support, rather than replace, qualitative assessment. Yet, relatively few secondary analyses have returned to older institutional datasets to ask whether apparently strong bibliometric relationships may partly reflect institutional scale. This evolving framework provides a useful basis for reexamining previously published institutional datasets from a contemporary perspective, particularly when apparently strong bibliometric relationships may partly reflect institutional scale.

A particularly important unresolved issue is therefore the size–performance relationship. If larger universities systematically have higher-performing "top" faculty members, this does not necessarily mean that institutional size causes better research. Larger institutions may possess more research-active faculty, more laboratories, greater access to funding, larger clinical populations, stronger postgraduate programs, greater collaboration opportunities, and longer-established reputational networks. Alternatively, the relationship may partly reflect mathematical scaling. Consequently, raw citation counts may reward institutional size without demonstrating superior research quality.

A second unresolved issue concerns the peak-performance paradox: a small institution may contain an exceptionally productive researcher whose performance is highly visible despite limited institutional capacity. Such an institution could rank poorly on aggregate volume but exceptionally on individual research intensity. Conversely, a large

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

institution may achieve impressive absolute numbers while displaying relatively ordinary performance after adjustment for size. The present study addresses these gaps by theoretically separating three constructs:

1. *Institutional capacity:* represented by faculty size.
2. *Peak scholarly performance:* represented by the published bibliometric indicators of the selected top faculty member.
3. *Relative concentration:* represented by the top faculty member's output and impact relative to the institution's faculty size.

This reframing enables a more cautious interpretation of the original data.

Purpose of the Study

The purpose of this theory-driven secondary analysis was to reinterpret the Barzegar and Barzegar (1) dataset through contemporary theories of responsible research assessment and institutional research capacity. More specifically, the study pursued five objectives:

1. To reexamine the association between institutional faculty size and the bibliometric performance of the selected top faculty member.
2. To determine whether publication quantity, citation impact, h-index, and citations per paper exhibit equivalent relationships with institutional size.
3. To identify evidence of institutional-size inequality across the 51 universities.
4. To develop size-sensitive indicators that distinguish institutional capacity from peak scholarly concentration.
5. To develop a theoretically informed interpretation of institutional research performance consistent with contemporary responsible-assessment principles.

The analysis was guided by the following research questions:

RQ1. What is the relationship between faculty size and the bibliometric indicators of the selected top faculty member?

RQ2. Does institutional size explain publication quantity, citation impact, h-index, and citations per paper to the same degree?

RQ3. What does the distribution of size-normalized peak performance reveal about the relationship between institutional capacity and scholarly concentration?

RQ4. What theoretical implications do these patterns have for the responsible assessment of Iranian medical universities?

Review of Literature

Bibliometric Indicators and the Problem of Construct Validity

Bibliometric indicators are useful because scholarly communication produces observable traces. Publications can be counted; citations can be accumulated; networks can be mapped; and indicators can be calculated. The difficulty begins when these traces are interpreted as direct measurements of latent constructs such as “quality,” “excellence,” or “institutional strength.”

The h-index is a particularly clear example. It combines productivity and citation impact by identifying the largest number h of publications that have each received at least h citations. Its popularity stems from its intuitive appeal. However, the index is influenced by career length, disciplinary citation practices, database coverage, coauthorship, and self-citation. Researchers with longer careers have more opportunities to accumulate citations, while disciplines differ substantially in citation density. Consequently, the same h-index can have different meanings in different fields or career stages (6).

The construct-validity problem is even more fundamental. If the theoretical construct is “research excellence,” no single bibliometric indicator can plausibly capture all of its dimensions. Research may be highly influential but poorly cited because it addresses a local policy issue. It may be methodologically innovative but published in a small specialized journal. It may also contribute to clinical practice without generating substantial

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

 CC BY 4.0

citation counts. DORA (2) therefore recommends that research assessment considers the content and influence of research rather than treating journal or author metrics as proxies for quality. The problem is not that bibliometric indicators are inherently useless. Rather, they are evidence of performance, not performance itself.

Institutional Scale and Research Capacity

Institutional research capacity is partly a function of human resources. A university with more faculty members has a larger potential pool of researchers, more disciplinary specializations, greater opportunities for collaboration, and potentially greater access to research infrastructure. Accordingly, one would expect faculty size to be associated with absolute publication and citation counts.

The original Barzegar and Barzegar (1) dataset provided an unusually useful opportunity to investigate this issue because it combined a measure of institutional size with indicators of the selected top researcher at each institution. The original descriptive statistics showed substantial dispersion: faculty size ranged from 29 to 1,914, while citations of the selected top faculty members ranged from 57 to 22,460. This range immediately suggested that institutional comparison was not simply a matter of ranking. It was a problem of scale.

A large university can possess greater absolute research capacity without every researcher being highly productive. Moreover, a small university can possess limited overall capacity while producing a highly visible research leader. The two situations are theoretically distinct.

Cumulative Advantage and the Matthew Effect

The relationship between institutional resources and scholarly visibility can also be interpreted through cumulative-advantage theory. Merton's concept of the Matthew effect (7) describes how researchers who have already achieved recognition may receive disproportionately greater recognition for subsequent contributions.

In institutional research systems, cumulative advantage may operate through multiple interconnected processes. Researchers affiliated with prestigious or resource-rich institutions may have greater access to collaborators, funding opportunities, graduate students, research infrastructure, and international networks. These conditions may, in turn, be associated with greater productivity and scholarly visibility, potentially reinforcing existing differences in recognition and resources. However, this sequence should be understood here as a theoretical interpretation rather than an established causal pathway. The present cross-sectional secondary analysis does not provide evidence that institutional prestige causes greater resources, productivity, or visibility, nor can it establish the direction of these relationships. Merton's concept of the Matthew effect (7) is therefore used as a theoretical lens for interpreting how accumulated recognition may interact with institutional and individual advantages, rather than as a causal explanation of the observed associations.

In institutional research systems, cumulative advantage may operate through multiple interconnected processes. Researchers affiliated with prestigious or resource-rich institutions may have greater access to collaborators, funding opportunities, graduate students, research infrastructure, and international networks. These conditions may, in turn, be associated with greater productivity and scholarly visibility, potentially reinforcing existing differences in recognition and resources. Research on the Matthew effect in citations has specifically considered prestige as one possible component of cumulative advantage, although empirical evidence suggests that the magnitude of prestige effects may be limited after accounting for other characteristics of the research being cited (8). Thus, the present study treats cumulative advantage as a theoretical lens rather than an established causal pathway. The cross-sectional secondary data do not permit causal inference regarding whether institutional prestige, resources, productivity, or visibility produce one another. Merton's concept of the Matthew effect (7) and subsequent research on citation-based

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

cumulative advantage (8) are therefore used to contextualize the observed institutional patterns rather than to establish a causal mechanism.

Responsible Research Assessment

The modern responsible-assessment movement provides an alternative framework for interpreting bibliometric data. The Leiden Manifesto (3) emphasizes ten principles including contextualization, transparency, recognition of disciplinary differences, qualitative assessment of individuals, avoidance of false precision, and scrutiny of the systemic effects of indicators.

DORA (2) similarly argues that journal- and publication-based metrics should not be used as substitutes for assessment of individual research quality.

CoARA (4) has subsequently broadened this agenda by explicitly emphasizing diversity of research contributions, career paths, and outputs. Its agreement recommends that research assessment be primarily qualitative and peer-review based, with quantitative indicators used responsibly and contextually.

Recent scholarship suggests that implementing responsible assessment is not straightforward. Morgan-Thomas et al. (9), for example, found that the relationship between journal rankings and expert research-quality judgments remained strong even among institutions endorsing DORA principles (2). This is important because it demonstrates that changing formal assessment policies does not automatically eliminate the cultural influence of metrics.

Gärtner et al. (10) similarly argue that researchers are more likely to prioritize research quality when institutional incentives reward quality rather than simply quantity. The present analysis therefore does not advocate abandoning quantitative indicators. Instead, it asks how indicators can be interpreted without allowing them to become substitutes for the constructs they are intended to illuminate.

Iranian Medical Research Performance

The Iranian context provides particularly relevant evidence. Sadeghi-Bazargani et al. (11) reported considerable heterogeneity in the research performance of Iranian medical academics, with differences associated with institutional classification, academic position, degree, field, and other characteristics. Approximately 90% of the sampled academics had an h-index below 10, and international collaboration was relatively limited.

A later national bibliometric analysis similarly reported extensive growth in the research productivity of Iranian medical academics between 2016 and 2020 but substantial heterogeneity among researchers and universities. Importantly, university class was associated with quantitative research performance, whereas differences in certain quality-related measures such as citations per paper and high-impact publication rates were less pronounced (12).

These findings provide an important theoretical context for the present reanalysis. Institutional scale may explain part of the variation in research quantity, while other dimensions of research impact may follow different patterns.

Materials and Methods

Research Design

This study employed a theory-driven secondary quantitative analysis of the published dataset reported by Barzegar and Barzegar (SPSS19) which was cross-checked (1). Secondary analysis is appropriate when an existing dataset can be revisited using a theoretically different conceptual framework or analytical strategy.

The analytical unit was the medical university. The dataset comprised 51 Iranian medical universities affiliated with the Ministry of Health, Treatment, and Medical Education. The original study excluded independent medical schools, medical research centers, and medical schools affiliated with Islamic Azad University. Moreover, the present study did not collect new data or alter the original observations.

Source Dataset

The original database contained information on: total number of faculty members, number of articles

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

 CC BY 4.0

attributed to the selected top faculty member, total citations of that faculty member, h-index, citations per paper, scientific rank, and specialty. The original authors reported a total of 18,972 registered medical faculty members across the 51 institutions.

The present reanalysis used the five continuous variables available across all 51 universities: faculty size, articles of the selected top faculty member, citations of the selected top faculty member, h-index, and citations per paper. In addition, the top faculty member was selected based on the number of ISI papers, citations, H-index, and scientific rank.

Conceptual Reclassification

The original variables were theoretically reorganized into three dimensions (Table 1).

Table 1. The original variables

Dimension	Operational indicator	Interpretation
Institutional capacity	Faculty members	Human-resource scale
Peak scholarly productivity	Articles	Publication volume of selected top researcher
Peak scholarly impact	Citations, h-index	Accumulated scholarly visibility
Citation intensity	Citations per paper	Average citation impact
Research concentration	Peak indicator/faculty size	Exceptional individual performance relative to institutional scale

This reclassification is central to the theory-driven nature of the analysis.

Secondary Statistical Analysis

The original Pearson correlations were retained as a baseline. Additional analyses included:

1. Pearson correlations between faculty size and each bibliometric indicator.
2. Spearman rank correlations to reduce sensitivity to extreme observations.
3. Fisher z-transformed 95% confidence intervals for Pearson correlations.
4. Log–log regression models to examine scaling relationships between faculty size and absolute bibliometric indicators.
5. Size-normalized indicators calculated as:

$$\text{Articles per 100 Faculty} = \frac{\text{Articles of Selected Top Faculty}}{\text{Faculty Size}} \times 100$$

and analogous measures for citations and h-index.

6. Distributional analysis of institutional faculty size.
7. Comparative examination of absolute versus relative performance.

Because the original data identified only the selected top faculty member rather than the complete publication portfolio of each university, the normalized measures were interpreted as peak-performance concentration indicators, not as university-wide productivity rates.

Theoretical Framework

The secondary analysis integrates four theoretical principles:

Principle 1: Scale dependence. Absolute bibliometric counts are expected to increase with the size of the underlying research population.

Principle 2: Construct multidimensionality. Publication volume, citation accumulation, h-index, and citations per paper represent related but non-identical dimensions of scholarly performance.

Principle 3: Cumulative advantage. Research visibility may be reinforced through institutional and individual accumulation of resources, reputation, collaboration, and recognition.

Principle 4: Responsible interpretation. No single numerical indicator should be treated as a direct or complete measure of institutional quality.

Results

Descriptive Profile

The original dataset contained 51 universities and 18,972 faculty members (Table 2). Faculty size also had a mean of 372.7 and a standard deviation

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

of 335.9, with a range from 29 to 1,914. The selected top faculty members had a mean of 167.0 articles, 3803.6 citations, a mean h-index of 21.0, and 14.7 citations per paper (Table 2).

The large standard deviations relative to the means indicate substantial heterogeneity. In particular, faculty size was strongly right-skewed, with a small number of very large institutions and many relatively small institutions. The 10 largest universities contained approximately 55.3% of all faculty members represented in the dataset. This is an important indicator of structural concentration.

Table 2. Descriptive statistics and theoretical interpretation

Variable	Mean	SD	Minimum	Maximum	Theoretical interpretation
Faculty members	372.70	335.90	29	1,914	Institutional capacity
Top faculty articles	167.00	146.30	6	917	Peak productivity
Top faculty citations	3803.60	2345.30	57	22,460	Peak accumulated impact
Top faculty h-index	21.00	12.70	5	60	Composite productivity/impact
Citations per paper	14.70	10.50	4.29	50.69	Citation intensity

Note. Descriptive statistics reproduce the original dataset and are reported in the original article.

Faculty Size and Peak Publication Productivity

Faculty size was positively correlated with the number of articles attributed to the selected top faculty member: Pearson r = 0.654, 95% CI [0.462, 0.788], $p < 0.001$. The corresponding Spearman rank correlation was $\rho = 0.479$, $p < 0.001$. Thus, larger universities tended to have selected top researchers with more publications. However, the relationship was not deterministic. The difference between the Pearson and Spearman coefficients suggests that extreme institutional observations contribute materially to the linear association.

Both Pearson's and Spearman's correlation coefficients indicated a positive association between the two measures. The Spearman coefficient was somewhat higher than the Pearson coefficient, suggesting that the estimated association differed according to whether the analysis was based on the original values or their ranks. However, this difference alone does not establish the influence of extreme observations or demonstrate a statistically significant difference between the two coefficients. Rather, it indicates that the observed association may not be fully characterized by a single correlation estimate and should be interpreted in light of the distinct properties of the two methods.

Because Pearson's correlation assesses the linear association and Spearman's correlation assesses the monotonic association based on ranked data, the two coefficients provide complementary information about the relationship. Accordingly, the observed association appears reasonably robust across the two approaches, while the precise magnitude of the relationship should be interpreted with caution.

A log–log regression indicated that faculty size explained approximately 26.9% of the variance in log-transformed article counts ($R^2 = .269$). The scaling coefficient was approximately .552, indicating a sublinear relationship: article counts increased with institutional size, but less than proportionally.

This is theoretically significant. If research output increased perfectly proportionally with faculty size, doubling the faculty would be associated with doubling peak publication output. The observed relationship was weaker than that expectation.

Faculty Size and Citation Impact

The strongest association was between faculty size and citations: Pearson r = .837, 95% CI [0.730, 0.904], $p < 0.001$. The Spearman correlation was also $\rho = 0.626$, $p < 0.001$. The log–log regression explained approximately 44.6% of the variance in log-transformed citations ($R^2 = 0.446$), with a scaling coefficient of approximately .857. The relationship between institutional size and citations was therefore stronger than the relationship between

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

 CC BY 4.0

institutional size and publication quantity.

This finding may reflect several mechanisms. Larger institutions may provide more opportunities for collaboration, access to clinical populations, graduate students, research infrastructure, and internationally connected research networks. However, because the dataset is cross-sectional and the citation variable refers to a selected top individual, causality cannot be inferred.

Faculty Size and h-Index

Faculty size was also strongly associated with the h-index: Pearson r = 0.752, 95% CI [0.601, .851], p < 0.001. The Spearman correlation was ρ = 0.601, p < 0.001. The log–log model explained approximately 42.0% of the variance in h-index (R^2 = 0.420).

The stronger relationship between faculty size and h-index compared with faculty size and articles suggests that institutional scale may be associated not only with publication volume but also with accumulated visibility.

However, h-index is itself a cumulative indicator, and therefore, potentially sensitive to career duration, disciplinary norms, and citation opportunity. Consequently, it should not be interpreted as an unqualified measure of research quality.

Faculty Size and Citations per Paper

The relationship with citations per paper was considerably weaker: Pearson r = 0.448, 95% CI [0.197, 0.644], p = 0.001. Moreover, the Spearman correlation was ρ = 0.534, p < 0.001 (Table 3).

This result is theoretically important because citations per paper is less directly dependent on the absolute number of publications than total citations. The weaker relationship suggests that institutional scale explains substantially less of the variation in citation intensity than it does in total citation accumulation. The result, therefore, supports a distinction between quantity-dependent impact and impact intensity.

Table 3. Secondary correlations between institutional size and peak-performance indicators

Outcome	Pearson *r*	95% CI	Spearman ρ	Interpretation
Articles	0.65***	[0.46, 0.78]	0.47***	Strong positive association
Citations	0.83***	[0.73, 0.90]	0.62***	Very strong positive association
H-index	0.75***	[0.60, 0.85]	0.60***	Strong positive association
Citations per paper	0.44**	[0.19, 0.64]	0.53***	Moderate association

Note. **p < .01; ***p < .001. Pearson confidence intervals were calculated using Fisher's z transformation.

Institutional Concentration

A central finding of the secondary analysis is the degree of institutional-size concentration (Figure 1). It displays the rank-size distribution of faculty members across the 51 universities. The figure demonstrates the concentration of human-resource capacity in a relatively small group of institutions.

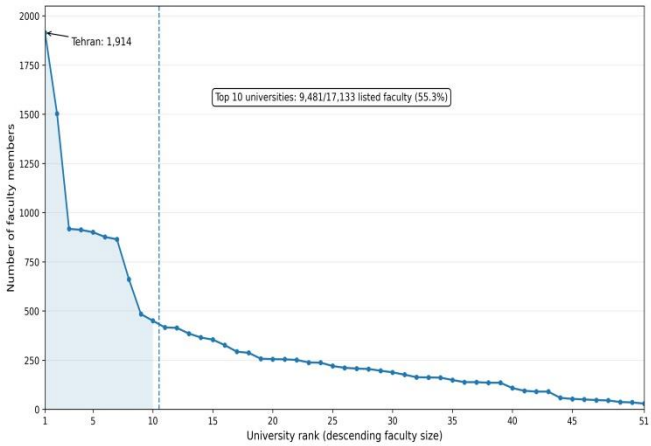


Figure 1. Rank-Size Distribution of Faculty Members Across 51 Iranian Universities

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

The 10 largest universities were accounted for approximately 55.3% of all the faculty members in the dataset. In contrast, many institutions had fewer than 100 faculty members. The faculty distribution therefore resembles a highly concentrated research system rather than an evenly distributed national network.

Peak-Performance Concentration

The most revealing result emerges when peak performance is considered relative to institutional size (Table 4). For example, Tarbiat Modarres University had 94 faculty members, while the selected top faculty member had 429 articles. This yields approximately 456 articles per 100 faculty members as a peak-performance concentration indicator.

Babol University of Medical Sciences had 37 faculty members and a selected top faculty member with 123 articles, yielding approximately 332 articles per 100 faculty members.

Golestan University of Medical Sciences had only 29 faculty members but a selected top faculty member with 93 articles and 1,926 citations. Its citations-per-100-faculty concentration was exceptionally high. These values should not be interpreted as institutional productivity rates. Moreover, the selected researcher demonstrated notably high research productivity relative to the size of the faculty. However, because institution-wide publication data were unavailable, this finding should not be interpreted as indicating a disproportionately large share of the institution's total research output. Rather, it reflects peak individual performance relative to faculty size.

Table 4. Selected examples of peak-performance concentration

University	Faculty	Top faculty articles	Articles/100 faculty	Top faculty Citations	Citations/100 faculty
Tarbiat Modarres	94	429	456.4	4,902	5,214.9
Babol	37	123	332.4	835	2,256.8
Golestan	29	93	320.7	1,926	6,641.4
Baqiatallah	257	515	200.4	4,659	1,813.2
Gilan	45	85	188.9	814	1,808.9
Shahre Kord	238	344	144.5	7,142	3,000.8

Note. These indicators are secondary analytical constructs and should not be interpreted as university-wide publication or citation rates.

Quantity–Impact Discordance

Another important pattern is the imperfect alignment between publication quantity and citation intensity. The original dataset showed that the highest article count was 917, whereas the highest citations per paper was 50.69. These maxima occurred at different universities. The selected top faculty member at Shahid Beheshti University had the largest number of articles, whereas the highest citations-per-paper value was associated with Urmia University of Medical Sciences.

Similarly, Tehran's selected top faculty member had the highest citation count (22,460) and h-index (60), whereas other institutions showed high citation intensity despite substantially lower publication counts.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

 CC BY 4.0

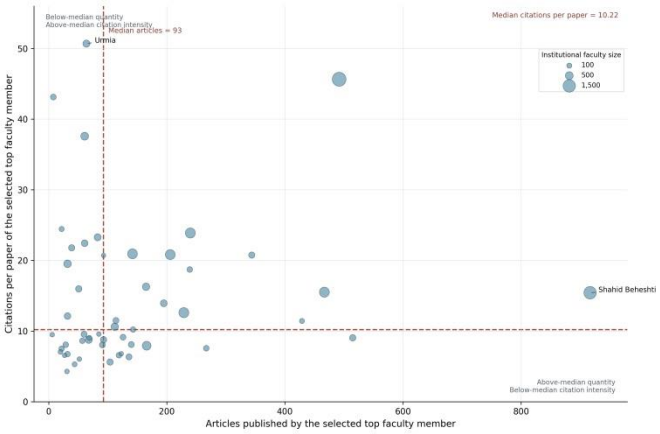


Figure 2. Publication Quantity and Citation Impact of the Most Highly Published Faculty Members across 51 Iranian Universities

Figure 2 (a bubble scatter plot) presents the publication quantity and citation intensity of the selected top faculty members. Bubble size represents institutional faculty size. Moreover, the vertical and horizontal reference lines represent the dataset medians. The figure illustrates why institutional performance cannot be reduced to a single ranking. Some universities combine high publication quantity with high citation intensity, while others demonstrate relatively high citation intensity despite moderate or low publication counts.

Discussion

From Institutional Ranking to Institutional Ecology

The principal contribution of this secondary analysis is conceptual rather than merely statistical. Moreover, the original dataset was constructed to compare universities. The present analysis suggests that comparison should be replaced by institutional profiling.

The distinction is important because ranking assumes that the underlying construct is unidimensional. If university A has a higher h-index than university B, ranking implies that A is "better." But the present findings demonstrate that different dimensions of performance can diverge.

Faculty size is strongly associated with total citations, strongly associated with h-index, moderately-to-strongly associated with articles, and only moderately associated with citations per paper. Therefore, these indicators cannot be treated as interchangeable manifestations of a single latent construct.

This finding is consistent with contemporary responsible-assessment principles. DORA (2) emphasizes that indicators should not substitute for qualitative judgment, while Leiden Manifesto (3) recommends that quantitative evaluation be interpreted in context. Finally, the present data provide empirical support for that conceptual argument.

The Size–Impact Relationship

The strongest new finding is the very strong relationship between faculty size and citations ($r = 0.837$). At first glance, this could be interpreted as evidence that larger universities perform better. Such an interpretation, however, would be theoretically premature.

Several explanations are possible. First, larger universities may genuinely possess stronger research ecosystems. Greater numbers of faculty members create more opportunities for interdisciplinary collaboration, co-authorship, postgraduate supervision, and research specialization.

Second, institutional size may act as a proxy for age and accumulated reputation. Large universities

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

are often established institutions with extensive alumni networks, established laboratories, clinical infrastructure, and international partnerships.

Third, larger institutions may contain more researchers, making it statistically more likely that one individual will emerge as an exceptionally productive or highly cited scholar.

Fourth, cumulative advantage may amplify the visibility of researchers located within larger or more prestigious research environments.

And finally, the original selection of the "top faculty member" creates a statistical selection mechanism. The most visible individual in a large institution is selected from a larger population of researchers than the most visible individual in a small institution. Consequently, larger institutions may have a greater probability of producing extreme observations even without a proportional difference in average researcher quality.

This last possibility is particularly important and has received insufficient attention in simplistic institutional rankings. The observed association therefore should be interpreted as compatibility with a capacity effect, not proof of a causal capacity effect.

Why Citations Are More Strongly Associated with Size than Articles

The correlation between faculty size and citations (0.837) was stronger than the correlation between faculty size and articles (0.654).

One interpretation is that larger institutional ecosystems may provide access not merely to publication opportunities but to research networks that enhance visibility and citation accumulation.

However, citations are also cumulative. Unlike publication counts, citations depend on the age of publications, field citation density, collaboration patterns, journal visibility, and network structure. Consequently, institutional size may interact with cumulative advantage.

This is precisely why responsible assessment discourages interpreting citation counts without context. The h-index and citation counts have known limitations related to career length, discipline, database coverage, and other factors. The finding should therefore be interpreted as indicating that institutional capacity is more strongly associated with accumulated scholarly visibility than with citation intensity.

The Importance of Citations per Paper

The weaker relationship between faculty size and citations per paper is arguably the most theoretically informative finding.

In addition, citations per paper can be viewed as an indicator of citation intensity rather than cumulative output. It is still imperfect, but it partially reduces the mechanical advantage associated with producing a large number of publications.

The implication is not that citations per paper is a superior measure of research quality. Rather, it represents a somewhat different dimension of scholarly visibility.

This supports a multidimensional approach in which:

-publication count represents productivity,
-total citations represent accumulated visibility,
-h-index represents a combined productivity-impact profile,
-citations per paper represent citation intensity, and
-faculty size represents institutional capacity.

These indicators should therefore not be collapsed into a single "research excellence" score without explicit theoretical justification.

The Peak-Performance Paradox

The size-normalized analysis reveals a second major phenomenon: small institutions can display exceptionally concentrated peak performance. The examples of Tarbiat Modarres, Babol, Golestan, and Guilan are especially instructive. Their selected top faculty members produced bibliometric profiles that were very large relative to the size of their institutions.

This phenomenon can be described as the peak-

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

 CC BY 4.0

performance paradox: An institution can have low aggregate capacity but high exceptional performance. This is important for policy because conventional rankings can obscure such institutions. If policymakers allocate resources solely according to aggregate publication or citation totals, institutions with small research populations may be perpetually disadvantaged.

The reverse problem also exists. A large institution may have impressive absolute performance simply because it has a large number of researchers. Without size-sensitive analysis, this can be mistaken for uniformly superior research quality. Therefore, a responsible national research policy should distinguish: *capacity* from *productivity*, *productivity* from *impact*, and *institutional impact* from *individual peak performance*.

Cumulative Advantage and Institutional Inequality

The strong size–citation relationship can also be interpreted through cumulative advantage. Research systems may develop positive feedback loops:

Resources → researchers → publications → visibility → citations → reputation → additional resources.

This process can generate increasing differentiation between institutions. The Matthew-effect literature provides a theoretical basis for understanding how existing recognition can amplify subsequent recognition (6).

However, the present analysis also provides evidence against a simple deterministic cumulative-advantage model. Some smaller institutions exhibit highly visible research peaks. Thus, institutional disadvantage is not equivalent to scholarly incapacity.

A more appropriate conceptualization is probabilistic cumulative advantage. Institutional capacity changes the probability of producing highly visible research but does not determine the outcome.

Toward a Multidimensional Institutional Research Assessment Framework

The results support the development of a conceptual framework that can be termed the Institutional Capacity–Peak Performance–Impact Framework (ICPPIM).

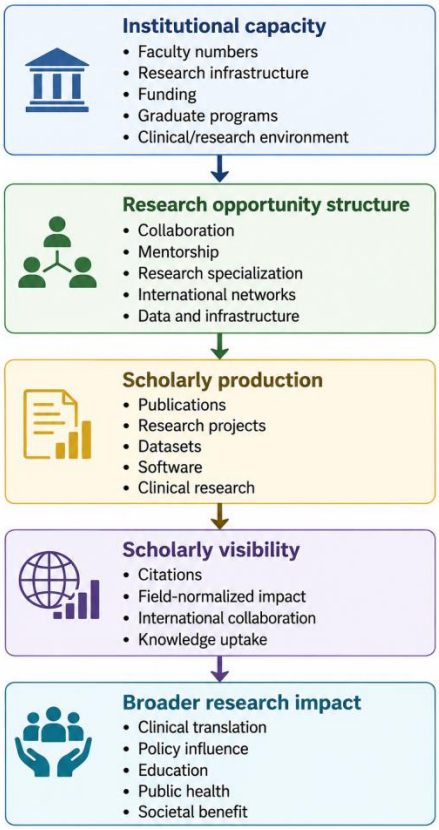


Figure 3. Proposed Institutional Capacity–Peak Performance–Impact Framework

The model contains a critical feedback loop:

Recognition → reputation → resources → future research opportunity.

The model also includes a contextual moderation layer. Discipline, career stage, institutional mission, geographic location, language, funding structure, and research field can modify the relationship between resources and observable metrics.

This model is more compatible with the current direction of responsible research assessment than a simple ranking system. CoARA (4) explicitly recommends recognition of diverse research contributions and cautions against using publication-based indicators as proxies for quality.

Implications for Medical University Policy

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

The findings have several implications for policymakers.

1. *Resource allocation should not simply follow rankings:* If rankings reproduce existing advantages, allocating resources exclusively according to bibliometric performance may intensify inequality. Moreover, the Leiden Manifesto (3) explicitly warns that indicators can create systemic effects that alter the behavior of the research system itself.
2. *Small institutions require capacity-sensitive evaluation:* Smaller universities should not automatically be interpreted as low-performing because of lower absolute publication or citation counts. Size-normalized indicators can also identify institutions with unusually high peak performance and may help policymakers identify emerging research centers.
3. *High-impact individuals should not substitute for institutional capacity:* The reverse problem is equally important. An institution whose research visibility depends heavily on one exceptional researcher may require broader investment in research teams, infrastructure, doctoral training, and succession planning. Thus, high peak performance is not necessarily evidence of sustainable institutional research capacity.
4. *Citation intensity deserves separate consideration:* The weaker association between faculty size and citations per paper suggests that citation intensity may provide information not captured by institutional scale. However, it should be interpreted with field and career context.
5. *Assessment should incorporate qualitative evidence:* Quantitative indicators should be supplemented by evidence concerning: research originality, methodological rigor, mentoring, open science, collaboration, clinical translation, policy influence, teaching and supervision, datasets and software, and societal impact. This direction is consistent with DORA (2), Leiden Manifesto (3), and CoARA (4).

Implications for Responsible Research Assessment in Iran

The Iranian medical-university system provides a particularly interesting context for responsible assessment because substantial differences exist among institutions.

The original study already identified substantial differences in faculty numbers and bibliometric performance and concluded that manpower and resources were not evenly distributed. Later national analyses similarly documented substantial heterogeneity in Iranian medical academics' research performance.

The present analysis suggests that Iranian research assessment could benefit from a portfolio approach rather than a ranking-only approach.

A university's research profile could be represented through five domains:

1. Capacity: researchers, infrastructure, doctoral programs.
2. Productivity: publications and research outputs.
3. Impact: citations and field-normalized indicators.
4. Concentration and sustainability: distribution of performance across researchers.
5. Societal relevance: clinical, educational, policy, and public-health influence.

Such an approach would reduce the risk of rewarding institutions simply for size.

Limitations of the Study

For this study, several limitations should be recognized. First, this was a secondary analysis of a historical dataset collected in 2018 and published in 2019. It does not represent the current research profile of Iranian medical universities.

Second, the dataset contains only 51 institutions within the Ministry of Health system. Islamic Azad University medical schools, independent medical schools, and medical research centers were excluded. Therefore, the results should not be generalized to all Iranian higher-education institutions.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

 CC BY 4.0

Another limitation is that the analysis uses the selected "top faculty member" from each institution rather than institution-wide publication and citation distributions. Consequently, the size-normalized indicators are measures of peak-performance concentration, not average institutional productivity.

Furthermore, citation counts and h-index values are cumulative and influenced by career duration, field, database coverage, and citation practices.

In addition, faculty size is an imperfect proxy for institutional capacity. It does not directly measure research funding, infrastructure, doctoral students, research time, laboratory resources, or clinical research opportunities.

Also, the secondary analysis is observational, and the correlations cannot establish that increasing faculty size causes higher research performance.

A further limitation is that the analyses were based on published aggregate data, preventing multilevel modeling of individual researchers within institutions.

Finally, responsible research assessment cannot be reduced to a new numerical index. Creating another composite metric without qualitative contextualization would reproduce the very problem that this study seeks to critique.

Conclusion

This theory-driven secondary analysis revisited a 2019 dataset comparing 51 Iranian medical universities and demonstrated that institutional size, publication productivity, citation impact, h-index, and citation intensity represent related but distinct dimensions of research performance.

The strongest relationship was between faculty size and citations, whereas citations per paper showed a substantially weaker association with institutional size. This pattern indicates that absolute indicators of scholarly visibility are strongly influenced by institutional scale, while citation intensity follows a somewhat different logic.

The analysis also revealed substantial concentration of human resources: more than half of the faculty represented in the dataset were located in the 10 largest institutions. At the same time, several smaller institutions exhibited exceptionally high peak-performance indicators. This demonstrates that institutional capacity and scholarly excellence are not synonymous.

The central theoretical implication is therefore that institutional research performance should be conceptualized as multidimensional and scale-sensitive.

A university is not simply a number of publications, citations, or h-index points. It is a research ecosystem in which human resources, infrastructure, institutional history, collaboration, disciplinary culture, career structure, and cumulative advantage interact.

The original study's observation that Iranian medical universities were unevenly distributed in terms of faculty and research resources remains important. However, the present reanalysis suggests that the policy response should move beyond simply attempting to equalize numerical indicators. The objective should be to develop sustainable research capacity while preserving and recognizing exceptional scholarly performance.

Contemporary responsible-assessment frameworks provide an appropriate foundation for such a transition. DORA (2), the Leiden Manifesto (3), and CoARA (4) all converge on a central principle: quantitative indicators can be useful, but they should be transparent, contextualized, and subordinate to broader judgments of research quality and impact.

The strongest conclusion from the present analysis is therefore not that bibliometrics should be abandoned. Rather, bibliometrics should be demoted from verdicts to evidence.

That conceptual shift is particularly important for medical universities, where research performance must ultimately be understood in relation to education, clinical practice, public health, scientific capacity, and societal benefit.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

CC BY 4.0

Authors' Contributions

K.B designed research, conducted research, analyzed data and wrote the paper. K.B had primary responsibility for final content. The author read and approved the final manuscript.

Acknowledgements

The author acknowledges the authors of the original dataset, Kazem Barzegar and Somayeh Barzegar, and the Iranian Ministry of Health, Treatment, and Medical Education's Scientific Measurement System for making the underlying institutional information available.

Competing Interests

The authors declare no competing interests.

Data Availability Statement

The present study is based on data published in Barzegar and Barzegar (1). The secondary analytical dataset was reconstructed from the published Table 1 and accompanying statistical information. The original article describes the Scientific Measurement System as an open-access database and provides the source of the institutional information.

Ethical Considerations

No new human participants were recruited and no new human-subject data were collected. The present study is a theoretical analysis of a previously published paper of the author. Accordingly, no new institutional ethics approval was sought for the present analysis. The present study involved no new human participants, identifiable individuals, interventions, interviews, surveys, or experimental procedures. It constitutes a secondary analysis of previously published, institution-level information.

The original study explicitly reported that the dataset was derived from the open-access Electronic Scientific Measurement System of the Iranian Ministry of Health and that the study involved no human or animal subjects. The original article also stated that the database could be used freely with appropriate citation.

Nevertheless, secondary analysis of bibliometric information requires ethical caution. Numerical indicators can influence perceptions of researchers and institutions, and rankings can unintentionally stigmatize institutions with lower absolute performance. The present analysis therefore avoided interpreting individual researchers as "better" or "worse".

Funding

The present study did not receive any financial support for secondary data analysis.

Artificial Intelligence Use Disclosure Statement

AI (ChatGpt) was used to create figures in this paper.

References

1. Barzegar K, Barzegar S. Comparative meta-analysis of fifty-one Iranian medical universities regarding faculty members information. Evidence Based Health Policy, Management & Economics 2019; 3(1): 12-22. doi:10.18502/jebhpme.v3i1.577.
2. San Francisco Declaration on Research Assessment. San Francisco Declaration on Research Assessment [Internet]. Available from URL: https://sfdora.org/read/. Last Access: Aug 16, 2026
3. Hicks D, Wouters P, Waltman L, de Rijcke S, Rafols I. Bibliometrics: The leiden manifesto for research metrics. Nature. 2015;520:429-31. doi: 10.1038/520429a
4. Coalition for Advancing Research Assessment. Agreement on reforming research assessment [Internet]. Available from URL: https://coara.eu/agreement/the-agreement-full-text/. Last Access: Aug 16, 2026.
5. San Francisco Declaration on Research Assessment. Guidance on the responsible use of quantitative indicators in research assessment. San Francisco: San Francisco Declaration on Research Assessment; 2024.
6. Mondal H, Deepak KK, Gupta M, Kumar R. The h-index: Understanding its predictors, significance, and criticism. Journal of Family Medicine and Primary Care. 2023;12(11): 2531–7. Doi: 10.4103/jfmpc.jfmpc_1613_23
7. Merton RK. The Matthew effect in science. Science. 1968;159(3810):56-63. doi:10.1126/science.159.3810.56
8. Wang J. Unpacking the Matthew effect in citations.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

 CC BY 4.0

J Informetr. 2014;8(2):329-339. doi:10.1016/j.joi.2014.01.006
9. Morgan-Thomas A, Tsoukas S, Dudau A, Gąska P. Beyond declarations: Metrics, rankings and responsible assessment. Res Policy. 2024; 53(10):105093. doi:10.1016/j.respol.2024.105093
10. Gärtner A, Leising D, Schönbrodt FD. Towards responsible research assessment: How to reward research quality. PLoS Biol. 2024;22(2):e3002553. doi:10.1371/journal.pbio.3002553
11. Sadeghi-Bazargani H, Bakhtiary F, Golestani M, Sadeghi-Bazargani Y, Jalilzadeh N, Saadati M. The research performance of Iranian medical academics: A national analysis. BMC Med Educ. 2019;19:449. doi:10.1186/s12909-019-1892-4
12. Khanali J, Malekpour M-R, Kolahi A-A. Assessing the research performance of the Iranian medical academics and universities: A bibliometric analysis. Med J Islam Repub Iran. 2023;37:31. doi:10.47176/mjiri.37.31

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-16]

[DOI: 10.18502/jebhpme.v10i7.23222]

CC BY 4.0