Journal of the Royal Society of Medicine; 2019, Vol. 112(6) 245–257
DOI: 10.1177/0141076819851666

Collaborative patterns, authorship practices and scientific success in biomedical research: a network analysis

Vanash M Patel[1], Pietro Panzarasa[2], Hutan Ashrafian[1], Tim S Evans[3], Ali Kirresh[1], Nick Sevdalis[1], Ara Darzi[1] and Thanos Athanasiou[1]
[1]Department of Surgery and Cancer, Imperial College London, London W2 1NY, UK
[2]School of Business and Management, Queen Mary University of London, London E1 4NS, UK
[3]Department of Physics, Imperial College London, London, SW7 2AZ, UK
Corresponding author: Vanash M Patel. Email: vanash.patel06@imperial.ac.uk

Summary

Objective: To investigate the relationship between biomedical researchers' collaborative and authorship practices and scientific success.
Design: Longitudinal quantitative analysis of individual researchers' careers over a nine-year period.
Setting: A leading biomedical research institution in the United Kingdom.
Participants: Five hundred and twenty-five biomedical researchers who were in employment on 31 December 2009.
Main outcome measures: We constructed the co-authorship network in which nodes are the researchers, and links are established between any two researchers if they co-authored one or more articles. For each researcher, we recorded the position held in the co-authorship network and in the bylines of all articles published in each three-year interval and calculated the number of citations these articles accrued until January 2013. We estimated maximum likelihood negative binomial panel regression models.
Results: Our analysis suggests that collaboration sustained success, yet excessive co-authorship did not. Last positions in non-alphabetised bylines were beneficial for higher academic ranks but not for junior ones. A professor could witness a 20.57% increase in the expected citation count if last-listed non-alphabetically in one additional publication; yet, a lecturer suffered from a 13.04% reduction. First positions in alphabetised bylines were positively associated with performance for junior academics only. A lecturer could experience a 8.78% increase in the expected citation count if first-listed alphabetically in one additional publication. While junior researchers amplified success when brokering among otherwise disconnected collaborators, senior researchers prospered from socially cohesive networks, rich in third-party relationships.
Conclusions: These results help biomedical scientists shape successful careers and research institutions develop effective assessment and recruitment policies that will ultimately sustain the quality of biomedical research and patient care.

Keywords

Medical careers, medical education, medical management

Received: 16th August 2018; accepted: 30th April 2019

Introduction

It is widely recognised that, over recent years, scientific research has become an increasingly collaborative enterprise.[1,2] Teamwork has also been found to dominate solitary work in the production of research of high impact across nearly all disciplinary fields.[2] Biomedical research is no exception and, for this reason, it has been suggested that collaboration should be encouraged by biomedical research institutions and centres.[1,3] Despite the growing consensus on the competitive advantage of teamwork, the distinctive collaborative practices that nurture scientific success in biomedical science still remain unknown. This may partly explain why current academic evaluation systems in healthcare are based primarily on traditional bibliometric indicators of performance, such as number of publications or citations,[4,5] and neglect to explicitly take into account structural features of the collaboration networks in which researchers are embedded.[3]

Collaborative research typically results in multiple authorship. Norms for the ordering of co-authors in a byline vary significantly across disciplinary fields but tend to be consistent within a particular field.[6–11] For biomedical publications, the sequence of co-authors' names is often determined by contribution- and supervision-related credit.[12,13] Typically, the first co-author is the researcher who has contributed the most, while the last co-author has held the most supervisory role.[14] Occasionally, researchers who appear in the middle of the sequence of co-authors are likely to be those who have failed to demonstrate adequate contribution and for this reason may have been offered authorship gratuitously.[15] Conversely, co-authors who have contributed in equal measure to the scientific work are typically listed in alphabetical order.[16–18] Despite a widespread agreement on the norms governing the ordering of authorship, how a fair allocation of credit translates into an equally fair allocation of recognition and visibility still remains largely to be investigated.[6,7,9,19–23]

In this study, drawing on a unique co-authorship network in biomedical research, we embraced a network-based perspective[1,3,24–26] to investigate the association between the scientific performance of individual biomedical researchers and various patterns of collaboration and criteria for credit allocation.

Methods

Study population

We analysed the careers of the academics at the Faculty of Medicine at Imperial College London, one of Europe's largest medical institutions (see details in the Supplementary Appendix). To this end, we used the database from the university intranet to create a list of all academics that were in employment on 31 December 2009. We included academics that were ranked in the hierarchical system in the following roles: lecturer; senior lecturer; reader; and professor. We excluded academics that were research fellows, associates, officers, assistants or staff with honorary academic status because they are not included in academic promotions. In total, the sample population included 525 academics. The composition of these academics in terms of academic rank, gender and physician status (i.e. whether the academic was a physician or not) is shown in Table 1.[27,28]

Authors' publications, positions in bylines and scientific performance

We used SciVerse Scopus Author Identifier to generate the publication list for each academic. The Author Identifier matches authors to their publications and SciVerse Scopus claims to have achieved 99% accuracy for 95% of its records.[27,29] All publication lists were divided into three time periods: 1 January 2001 to 31 December 2003; 1 January 2004 to 31 December 2006; and 1 January 2007 to 31 December 2009. For each academic and a given time period, we used the number of publications in the preceding time period to test whether the academic's past productivity was associated with the number of citations the academic received in connection with the articles published in the focal time period. For each academic and time period, we recorded the number of publications in which he or she appeared as the solo author. To assess the association between multiple authorship and performance, for each academic and across all multi-authored publications in each time period, we calculated the median number of co-authors per publication and extracted the minimum number of co-authors per publication.[30]

To examine the relationship between position in bylines and research performance, for each academic and for each time period, we extracted the number of multi-authored publications in which co-authors were listed alphabetically and non-alphabetically. For each of these two groups of publications, we recorded the number of publications in which the academic appeared as listed in each of the following five positions in bylines: first; second; penultimate; last; and 'other'. Details on how these positions were defined are reported in the Supplementary Appendix.

We used citation counts to measure scientific success, in agreement with several studies suggesting that citations correlate with research quality[2,31] and as a result play a fundamental role in promotion and funding reviews.[32] For each researcher and time period, we calculated the number of citations received by all articles published by the researcher in that time period. We collected citations in January 2013 to allow sufficient time for fairly recent publications to be cited.

The co-authorship network

For each time period, we constructed the co-authorship network in which nodes are the academics, and links are established between any two academics if they co-authored one or more articles that were published in that period (for details, see the Supplementary Appendix).[1] The network so constructed is undirected and unweighted. A graphical representation of the largest connected component[24] of the co-authorship network is shown in Figure S1 in the Supplementary Appendix.[33]

Network-based measures of authors' centrality

To capture the centrality of each academic in each time period, we calculated the following network measures: degree, eigenvector, betweenness and closeness centrality. The (normalised) degree centrality of a node is the number of links incident upon the node, divided by its maximum possible value (i.e. the

Table 1. Composition of the academics in the Faculty of Medicine.

Academic rank	Physician status n (%)		Gender n (%)		Total n (%)
	Non-physician	Physician	Male	Female	
Lecturer	85 (77)	25 (23)	66 (60)	44 (40)	110 (21)
Senior lecturer	45 (34)	89 (66)	87 (65)	47 (35)	134 (25)
Reader	42 (64)	24 (34)	44 (67)	22 (33)	66 (13)
Professor	101 (47)	114 (53)	173 (80)	42 (20)	215 (41)
Total n (%)	273 (52)	252 (48)	370 (70)	155 (30)	525 (100)

number of nodes in the network minus one) (Figure S2A in the Supplementary Appendix).[24,34] Eigenvector centrality measures the importance of a node in a network as a function of the connections the node has to other nodes that are themselves important (Figure S2A in the Supplementary Appendix).[35] Betweenness centrality measures the extent to which a node lies on the shortest paths between other nodes in the network (Figure S2A in the Supplementary Appendix).[34,36] Finally, a node's closeness centrality is defined as the inverse of the sum of the shortest distances separating the node from all other nodes and thus measures how close the node is to all other nodes in the network (Figure S2A in the Supplementary Appendix).[24,34] For details on the calculation of these network centrality measures, see Supplemental Methods in the Supplementary Appendix.

Social capital: network-based measures of closed and open structures

We investigated the role of social capital in sustaining research performance and tested the association between closed versus open network structures and academics' citation counts. To this end, we relied on two network measures: the local clustering coefficient and network constraint.

The local clustering coefficient is defined as the ratio between the number of links connecting pairs of a node's neighbours and the total number of pairs of the node's neighbours (Figure S2B in the Supplementary Appendix).[25] Clustering thus captures the extent to which a node is embedded within a cohesive social structure, rich in third-party relationships. We tested the hypothesis that academics whose local network was more socially cohesive (i.e. with a higher clustering coefficient) were associated with a better performance than academics in a less cohesive network (i.e. with lower clustering).

Network constraint measures the extent to which a node is connected to other nodes that are already connected with each other (Figure S2C in the Supplementary Appendix).[26] A low value of network constraint means that a node can broker between otherwise disconnected others and is therefore associated with network closure. We tested the hypothesis that academics with a lower value of network constraint (i.e. in an open network) were characterised by better performance than academics with a higher value (i.e. in a closed network).

Statistical analysis

To investigate the non-linear relationship between number of co-authors per publication and performance, we squared the median number of co-authors and jointly tested the main and quadratic terms in the regression model. To avoid problems of multicollinearity arising from high correlation between variables, before squaring the median, we centred it around its mean (i.e. we subtracted the mean from the variable).[37]

We estimated interaction effects between an author's academic rank and two position-related variables: the number of publications in which the author appeared as last-listed in non-alphabetical order and the number of publications in which the author was first-listed in alphabetical order. We centred the two position-related variables and multiplied each of them by each of the three rank-related indicator variables (i.e. senior lecturer, reader and professor). We thus obtained six interaction terms.

Finally, we estimated interaction effects of authors' academic rank, brokerage opportunities and position in byline. To construct the interaction terms, we multiplied each of the three rank-related indicator variables by the product between centred constraint and centred number of publications in which authors appeared as last-listed in non-alphabetical order. We thus obtained three additional interaction terms.

Modelling strategy

Because the dependent variable is a count variable, the use of linear models would have resulted in inefficient, inconsistent and biased estimates. To investigate the relationship between scientific performance and collaborative and authorship practices, we thus estimated maximum-likelihood negative binomial panel regression models, with beta-distributed random effects and bootstrapped standard errors.[38] We estimated negative binomial models instead of Poisson ones owing to the overdispersion of the dependent variable.[39] Indeed, the negative binomial estimator can explicitly handle overdispersion, and thus it accounts for unobserved heterogeneity among observations. By contrast, in case of overdispersion, the Poisson estimator would produce consistent, but inefficient estimates and standard errors that are biased downward.

The details on the modelling strategy, and the specification tests concerned with overdispersion, potential confounders and unobserved heterogeneity are reported in the Supplementary Appendix. Results from robustness checks based on various regression models are summarised in Tables S6a and S6b in the Supplementary Appendix. For all panel models, likelihood-ratio tests of model specification indicate that they are more appropriate than the corresponding pooled models. The computation of all models was implemented using Stata 64/MP 10.1.

Results

Descriptive statistics are shown in Table 2. Zero-order correlations and within and between variations of all variables are shown, respectively, in Tables S1 and S2 in the Supplementary Appendix. To cast more light on the variation of the dependent variable, we also calculated the transition probabilities from one period to the next (for details, see Table S3 in the Supplementary Appendix) and the first-order autocorrelations (Table S4 in the Supplementary Appendix). Findings indicate that there was considerable persistence in performance from one period to the subsequent one.

Table 3 reports the estimates for the coefficients of control and theoretical variables. Table S5 in the Supplementary Appendix shows the estimated coefficients for institutional affiliation that were not reported in the random-effects negative binomial panel models of Table 3. Models 1, 2 and 3 show the main effects of the independent variables on authors' performance. Model 4 also includes interaction effects between independent variables.

Our analysis suggests that an author's past productivity was related to future success: a good record of publications in a given period of time was likely to be associated with publications of high impact in subsequent time periods. For every additional publication in a three-year period, an author's expected citation count on the articles published in the subsequent three-year period increased by 0.20% (Model 1).

Our results do not suggest any gender-related discrimination. By contrast, we found evidence in favour of academic rank-based discrimination. Being a professor increased the expected number of citations with respect to the citations received by a lecturer by 52.20%, holding all other variables constant (Model 1). We also found that, on average, non-physicians outperformed physicians: being non-physician increased the expected citation count by 18.35% (Model 1).

Our next findings are concerned with the relationship between solo versus multiple authorship and research performance.[2] While solo authorship did not have any statistically significant association with performance, we found non-linear effects of number of co-authors per publication on citations. Figure 1 shows the inverse U-shaped relationship between median number of co-authors per publication and expected number of citations received by a female professor at the Institute of Clinical Sciences. More generally, estimates from Model 4 suggest that an author's citations increased as the size of collaborative teams expanded, but only up to a certain threshold, namely, 12 co-authors per publication. Beyond this threshold, a further increase in co-authors degraded performance. For example, for an author with a median of nine co-authors per publication (i.e. one standard deviation below 12), a one-unit increase in the median led to an increase in the expected citation count by 1.88%. However, for an author with a median of 15 co-authors per publication (i.e. one standard deviation above 12), a one-unit increase in the median led to a decrease in the expected citation count by 1.43%. Consistently, we also found a negative relationship the minimum number of co-authors per publication and performance. If the minimum number of co-authors with whom a researcher has ever published increased by one, the researcher would experience an expected decrease in citations by 3.59%, holding all other variables constant.

Next, we investigated the relation between collaborative patterns and performance. Models 1, 2 and 3 in Table 3 indicate that holding central positions in the collaboration network is associated with success.

Table 2. Means before centring, standard deviations, minimum and maximum values for all variables. Unit of analysis of our study is the individual academic.

	Mean	Standard deviation	Minimum	Maximum
Categorical variables				
1 Gender	0.705	0.456	0.000	1.000
2 Physician status	0.480	0.500	0.000	1.000
3 Institute of Clinical Sciences	0.027	0.161	0.000	1.000
4 Kennedy Institute of Rheumatology	0.046	0.209	0.000	1.000
5 Department of Medicine	0.409	0.492	0.000	1.000
6 National Heart Lung Institute	0.197	0.398	0.000	1.000
7 School of Public Health	0.109	0.311	0.000	1.000
8 Senior lecturer	0.257	0.437	0.000	1.000
9 Reader	0.126	0.332	0.000	1.000
10 Professor	0.410	0.492	0.000	1.000
Integer and continuous variables				
11 Number of citations	414.419	655.472	0.000	7,883.000
12 Number of past publications	29.270	58.252	0.000	1,013.000
13 Median number of co-authors per publication	6.600	3.313	2.000	40.000
14 Minimum number of co-authors per publication	2.900	1.620	2.000	26.000
15 Degree	0.013	0.021	0.000	0.302
16 Betweenness	0.080	0.150	0.000	2.180
17 Closeness	0.010	0.000	0.000	0.010
18 Eigenvector	0.010	0.030	0.000	0.410
19 Clustering coefficient	0.320	0.250	0.000	1.000
20 Constraint	0.160	0.328	0.000	10.583
21 Solo-authored publications	1.078	2.986	0.000	48.000
Positions in non-alphabetised bylines				
22 First	1.833	2.540	0.000	35.000
23 Last	4.380	6.900	0.000	89.615
24 Second	1.816	2.418	0.000	23.000
25 Penultimate	2.317	4.327	0.000	87.000
26 Other	4.419	7.089	0.000	104.000

(continued)

Table 2. Continued.

	Mean	Standard deviation	Minimum	Maximum
Positions in alphabetised bylines				
27 First	0.400	0.928	0.000	7.597
28 Last	0.611	1.398	0.000	14.000
29 Second	0.104	0.429	0.000	6.000
30 Penultimate	0.025	0.168	0.000	2.000
31 Other	0.014	0.133	0.000	2.000

Researchers could prosper from: (i) having many collaborators (i.e. high degree); (2) having many popular collaborators (i.e. high eigenvector); (3) lying on the shortest paths between many pairs of other researchers (i.e. high betweenness); and (4) reaching many others in very few steps (i.e. high closeness).[24]

To further examine the role of social capital in biomedical research, we tested whether authors with collaborators who also collaborated with one another (closed structure) outperformed authors with collaborators who never collaborated themselves (open structure). For each author, we measured the local clustering coefficient[25] and network constraint[26] to capture these two structural positions, respectively (Supplemental Methods and Figure S2B and S2C in the Supplementary Appendix). Our findings provide evidence in favour of the positive association between brokerage (i.e. open structure) and citations. As indicated by Model 3, socially cohesive structures, rich in third-party relationships, degraded performance, while, consistently, Models 1, 2 and 4 suggest that the lack of network constraint (i.e. the availability of brokerage) had beneficial effects. Researchers could thus produce highly cited work by forging, and affiliating themselves with, non-overlapping teams and by intermediating among otherwise disconnected collaborators.

We then estimated the association between various positions in bylines and total number of citations received (Models 1, 2 and 3). We distinguished between non-alphabetised and alphabetised bylines. In the former case, first and last positions were found to have a positive association with performance. For instance, Model 1 suggests that authors with one additional publication in which they were first- or last-listed in non-alphabetical order experienced an increase in expected number of citations by 4.50% and 1.92%, respectively. These results are in qualitative agreement with the widely accepted rules for credit allocation in biomedical research, according to which first- and last-listed authors in non-alphabetised bylines are expected to have provided the two most valuable, yet distinct, contributions to the work.[12,13] However, middle positions, too, with the exception of second and penultimate ones, on articles with five or more co-authors were found to be positively associated with performance.

For publications with authors sequenced alphabetically, results show that first positions facilitated highly cited work. No other position in alphabetised bylines was found to have any statistically significant association with performance. While non-alphabetical contribution-based authorship aims chiefly to sustain proper credit assignment, alphabetical authorship is, in general, not expected to reflect the relative contributions made by the various co-authors of a publication.[16–18] Yet, even in alphabetised bylines, first-listed authors were found to accrue most credit. Alphabetical ordering may thus reduce scope for debate on contribution and credit, yet not on visibility and recognition.[23]

Combinations of position in bylines and academic rank are associated with scientific success (Model 4). In the case of non-alphabetised bylines, being last-listed remained beneficial only for higher academic ranks, especially for professors, while it hindered the scientific performance of junior academics. First positions on publications with alphabetised bylines retained their positive association with performance only for junior academics. Being first-listed alphabetically in one additional publication enabled a lecturer to witness an increase of 8.78% in the expected citation count. Thus, even when all co-authors were expected to have equally contributed to the joint publication, the first-listed author with a junior academic position could secure extra recognition within the scientific community.

Finally, we investigated the association between combination of brokerage, academic rank and position in bylines on the one hand, and performance on

Table 3. Results of random-effects negative binomial panel regressions with main and interaction effects.

	Model 1		Model 2		Model 3		Model 4	
	EC	SE	EC	SE	EC	SE	EC	SE
Control variables								
Number of past publications	1.002**	0.0006	1.002***	0.0006	1.0001	0.0005	1.002*** **	0.0005
Gender	1.077	0.059	1.072	0.059	1.070	0.081	1.076	0.063
Physician status	0.816***	0.045	0.822**	0.048	0.801***	0.045	0.815***	0.046
Senior lecturer	1.057	0.083	1.042	0.090	1.023	0.090	1.726*	0.437
Reader	1.116	0.117	1.097	0.105	1.123	0.122	1.954*	0.526
Professor	1.522***	0.129	1.509***	0.123	1.502***	0.139	2.561***	0.662
Solo versus multiple authorship								
Solo-authored articles	1.010	0.015	1.009	0.015	1.013	0.012	1.007	0.015
Median number of co-authors per publication	1.034**	0.012	1.032*	0.013	1.071***	0.014	1.030*	0.012
(Median number of co-authors per publication)^2	0.997**	0.0008	0.997***	0.0008	0.997**	0.0009	0.997***	0.0007
Minimum number of co-authors per publication	0.968	0.0008	0.968*	0.016	0.948**	0.017	0.964*	0.016
Network-based measures of centrality								
Degree	678.052*** **	1158.870					937.552***	1420.734
Eigenvector			8.986*	9.488				
Betweenness			1.433*	0.256				
Closeness					8.57e+88***	8.71e+89		
Network-based measures of social capital								
Clustering coefficient					0.455***	0.052		
Constraint	0.377**	0.137	0.342**	0.120			0.319***	0.077

(continued)

Table 3. Continued.

	Model 1		Model 2		Model 3		Model 4	
	EC	SE	EC	SE	EC	SE	EC	SE
Author's position in publication								
Position in non-alphabetised bylines:								
First	1.045***	0.011	1.045***	0.012	1.042***	0.010	1.047***	0.010
Last	1.019*	0.009	1.019*	0.008	1.021***	0.006	0.870*	0.055
Second	1.012	0.016	1.012	0.017	1.022	0.014	1.019	0.018
Penultimate	0.990	0.010	0.989	0.008	0.997	0.007	0.991	0.008
Other	1.020**	0.007	1.023***	0.006	1.021***	0.005	1.022**	0.008
Position in alphabetised bylines:								
First	1.094***	0.025	1.092**	0.032	1.077*	0.032	1.088*	0.042
Last	1.014	0.023	1.014	1.023	0.022	0.023	1.020	0.023
Second	1.037	0.054	1.037	0.050	1.055	0.051	0.032	0.055
Penultimate	0.938	0.108	0.899	0.097	0.932	0.085	0.935	0.094
Other	1.126	0.165	1.142	0.135	1.070	0.163	1.103	0.162
Interactions: academic rank and position								
Senior lecturer × non-alphabetical last position							1.136*	0.070
Reader × non-alphabetical last position							1.107	0.078
Professor × non-alphabetical last position							1.206**	0.077
Senior lecturer × alphabetical first position							1.090	0.086
Reader × alphabetical first position							0.926	0.076
Professor × alphabetical first position							1.004	0.050

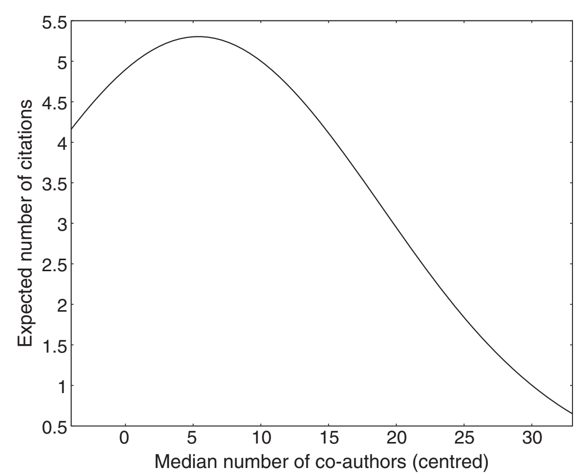
(continued)

Table 3. Continued.

	Model 1		Model 2		Model 3		Model 4	
	EC	SE	EC	SE	EC	SE	EC	SE
Interactions: academic rank, position and brokerage								
Senior lecturer × non-alphabetical last position × constraint							0.890	0.128
Reader × non-alphabetical last position × constraint							1.048	0.161
Professor × non-alphabetical last position × constraint							1.277***	0.079

Note: For ease of interpretation, the table displays incidence rate ratios. Models (1), (2) and (3) include only the main effects of the independent variables on authors' performance: Model (1) includes degree and constraint; Model (2) eigenvector centrality, betweenness centrality and constraint and Model (3) closeness centrality and the clustering coefficient (please refer to the main text and the Supplementary Appendix for the definition of these network measures). Model 4 includes interaction effects between independent variables. The reference category for academic rank is 'lecturer'. EC: estimated coefficient; SE: bootstrap standard error.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Figure 1. Association between the centred median number of co-authors per publication and the expected citation count of a professor at the Institute of Clinical Sciences, when all other independent variables are held constant at their means. Results are based on estimated coefficients from Model 4. To avoid problems of multi-collinearity, the covariate (median number of co-authors) was centred on its mean (i.e. we subtracted the mean from the variable).

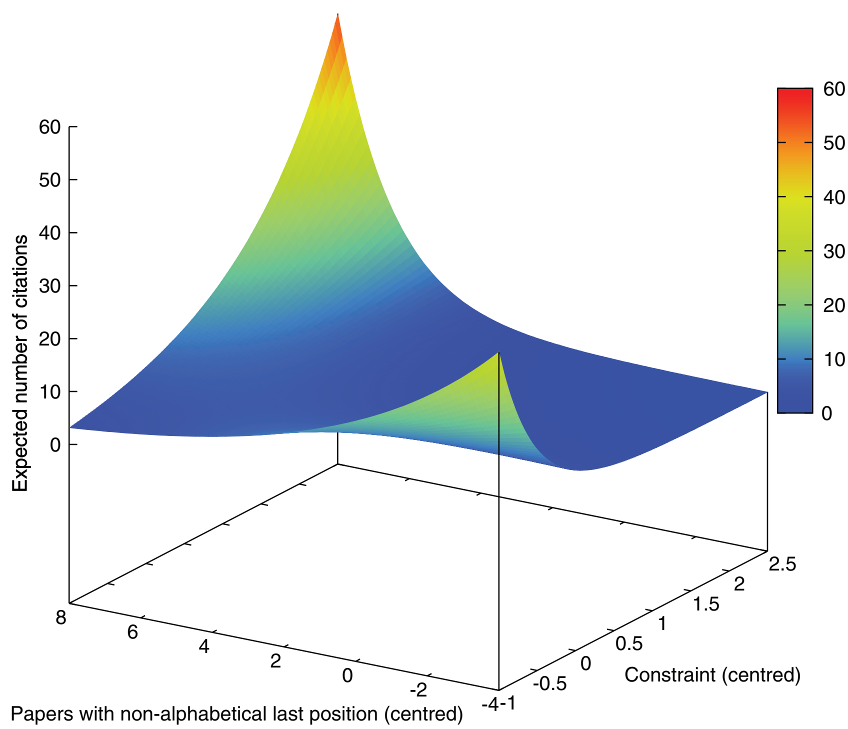


the other. Results suggest that, for professors, the positive association between brokerage opportunities and performance was intensified by a decrease in number of publications on which they were last-listed non-alphabetically. Alternatively, the positive association between being last-listed and a professor's performance was mitigated by an increase in brokerage opportunities. Unlike other academics in junior positions, researchers in senior academic roles acting as coordinators of many collaborative groups could therefore gain a competitive advantage by enhancing social cohesion and facilitating third-party relationships among collaborators from different groups. Figure 2 shows the association between combinations of brokerage and last positions in non-alphabetised bylines on the one hand, and the expected performance of a female professor at the Institute of Clinical Sciences on the other, when all other variables are held constant at their means.

Discussion

Our study has shown that, overall, collaboration is positively associated with research performance[2] and, more importantly, that researchers should embrace different collaborative strategies as their academic career progresses. For example, junior scientists, in

Figure 2. Association between combinations of brokerage and last position in non-alphabetised bylines and the expected citation count of a female professor at the Institute of Clinical Sciences, when all other independent variables are held constant at their means. Results are based on estimated coefficients from Model 4. To avoid problems of multi-collinearity, both covariates were centred on their respective means.



addition to securing first positions in bylines, should aim for positioning themselves in open collaborative structures rich in brokerage opportunities.[26] They should affiliate themselves with diverse research teams and collaborate with otherwise disconnected others from whom they will be able to acquire novel and non-overlapping knowledge, expertise, ideas and insights. However, different network structures become beneficial when researchers are promoted to more senior academic roles. For example, senior scientists, such as professors, who typically act as coordinators of large groups of researchers, prosper from some degree of network closure that enables them to develop an established scientific vision and research strategy, promote the transfer and sharing of complex knowledge, consolidate their groups' identity and minimise the risk of missing important opportunities of synergies and cross-fertilisation of ideas.

Strengths and weaknesses of the study

Previous studies have shown that diverse and open networks, rich in brokerage opportunities, may thrive in sustaining innovation.[26,40] Our results have extended these studies by suggesting that the benefits of brokerage are contingent upon the researcher's career stage.

Previous research has also pointed to the 'gender gap' and the obstacles that women were likely to face in academic medicine over the past decades, which hindered peer-review publications and delayed advancement in their careers.[41] Our analysis provides more encouraging results, for it suggests that gender equality has recently become more widespread in biomedical research.

Our study is not without limitations that, in turn, open up new avenues for future investigation. It seems reasonable to argue that co-authorship represents only one of the main forms of scientific collaboration.[1] Indeed, there are other peripheral or indirect forms of collaboration that are not reflected in formal co-authorship, and yet represent genuine instances of intellectual co-operation. Researchers may motivate, inspire and contribute to each other's scientific work without always being listed as co-authors, for example, by mentoring and supervising junior colleagues or by providing commentary at conferences, workshops and professional meetings.[3] Moreover, although the results from this study were derived from a large-scale collaboration network, the data were obtained from a single institution and a single disciplinary field. Therefore, the results of this study may not be generalisable to other institutions or disciplinary fields. We used a single citation database to extract bibliometric data; however, alternative databases, such as Scopus, Web of Science and Google Scholar, may produce different citation counts for biomedical researchers, although recent studies suggest that large changes in the results would be unlikely.[42,43] Even though special care has been taken to disambiguate authors' names, multiple databases could be used to further validate the bibliometric data concerning authors who have a common name, have worked in numerous institutions or have different research interests.

Implications for practice

It is widely recognised that biomedical studies depend heavily on non-clinical scientific research, and consequently basic medicine is cited three to five times more than clinical medicine.[44] Equally, academic medicine requires excellent physicians, who greatly facilitate translation of laboratory research into clinical practice.[45] Unlike non-physicians, academic physicians tend to split their time between clinical and research activities. However, current systems for assessing academic performance and guiding decisions on academic promotions are largely based on scholarly productivity and often do not take these differences into account.[46] As a result, academic physicians may become increasingly liable to isolation from academic medicine, which may ultimately stymie translational research.[45] The discrepancy in citations found in our study between academic physicians and non-physicians thus highlights the underlying need for greater incentives and support for collaborative relationships between these two groups of academics, which in turn can circumvent any obstacles to translational research.

Over the years, biomedical research has witnessed increasingly blurring boundaries between distinct specialties.[47] The lone researcher struggles to answer prominent questions without collaborating with other scientists, often from other disciplines or institutions.[48] Evidence suggests that biomedical studies are moving towards a team-based approach to research.[47,48] Our study has indeed provided support in favour of an association between collaboration and scientific performance in biomedical research. While junior researchers (lecturers) and senior researchers (professors) may benefit from distinctive collaborative strategies, there are advantages of intellectual cooperation and joint publications over solo research and authorship, across all academic ranks. However, fostering a collaborative culture in academic medicine is challenging because promotion committees and tenure systems often discourage collaboration by focusing on researchers' independent contributions. Conversely, our study suggests that academic institutions should encourage joint research projects and also aim to engage in a systematic measurement of scientific collaboration with a view to helping researchers to improve their academic productivity and quality.

At the same time, our results also suggest that researchers should strike a balance between solo authorship and excessive co-authorship so as to optimise the size of collaborative teams and, ultimately, sustain performance over time. Concerns have been raised on the increasingly widespread abuse of inappropriate authorship and the threats that the subsequent lack of transparency and accountability poses to the integrity of scientific research.[15] Honorary authorship granted to those who made no significant contribution to the work has been partly responsible for the large increase in the number of co-authors per publication over the last 50 years.[49] In the face of the recent upsurge in such practices, our study uncovered non-linear effects of number of co-authors per publication on citations and indicated that publications with very large numbers of co-authors were disadvantageous to the scientific performance of biomedical researchers. Our findings encourage scientific journals to reinforce authorship policies and ensure that each author has made a genuine contribution to the published work.[12,15]

As a result of the increasingly large number of multi-authored publications, it has become imperative that a proper assessment of research performance through such publications be based on the identification of the most important co-authors that played fundamental roles in the scientific work.[9,12,19,21] Our study has indicated that junior researchers can enhance scientific performance by producing publications in which they appear as the first-listed authors. Subsequently, when they are promoted to more senior academic roles, they should publish papers in which they are the last-listed authors, traditionally associated with mentoring and supervisory roles. Our findings thus suggest that, if promotion committees use bibliometric indicators for measuring researchers' scientific performance, they can produce accurate and fair assessments only by explicitly taking into account the number of co-authors and the positions in the bylines of the researchers' publications.

Conclusions

In summary, our study has four fundamental implications for biomedical research and policy: (1) intellectual collaboration has a competitive advantage over solo authorship across all academic ranks; (2) while researchers gain from expanding their network of collaborators, they should redistribute collaborators across multiple teams and refrain from excessive co-authorship; (3) there are differential benefits of collaborative and authorship practices depending on the stage of the researcher's career; and (4) research institutions and funding agencies should encourage and reward a team-based approach to scientific production and engage in a systematic and transparent measurement of collaboration and multiple authorship, in order to develop effective assessment and recruitment policies that will ultimately translate into the highest-quality research and patient care.

Declarations

Competing Interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that all authors have no relationships with any company that might have an interest in the submitted work in the previous three years; their spouses, partners or children have no financial relationships that may be relevant to the submitted work and all authors have no non-financial interests that may be relevant to the submitted work.

Funding: None declared.

Ethics approval: Not applicable.

Guarantor: VMP.

Contributorship: VMP – literature search, study design, data collection, data analysis and data interpretation; PP – study design, data collection, data analysis, data interpretation and figures; HA – study design, data collection, data interpretation and figures; TSE – study design, data analysis, data interpretation and figures; AK – study design and data collection; NS – data interpretation; AD study design and data interpretation; TA – study design, data analysis and data interpretation. All authors contributed to drafting the article or revising it critically for important intellectual content. All authors have approved the final version of the manuscript. All authors had full access to all of the data (including statistical reports and tables) in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Acknowledgements: None.

Provenance: Not commissioned; peer-reviewed by Syed Raza and Julie Morris.

ORCID iD

Vanash M Patel https://orcid.org/0000-0001-9579-8227
Pietro Panzarasa https://orcid.org/0000-0001-7596-4806
Hutan Ashrafian https://orcid.org/0000-0003-1668-0672

References

1. Newman MEJ. The structure of scientific collaboration networks. *Proc Natl Acad Sci USA* 2001; 98: 404–409.
2. Wuchty S, Jones BF and Uzzi B. The increasing dominance of teams in production of knowledge. *Science* 2007; 316: 1036–1039.
3. Yousefi-Nooraie R, Akbari-Kamrani M, Hanneman RA and Etemadi A. Association between co-authorship network and scientific productivity and impact indicators in academic medical research centers: a case study in Iran. *Health Res Policy Syst* 2008; 6: 9.
4. Ashrafian H, Rao C, Darzi A and Athanasiou T. Benchmarking in surgical research. *Lancet* 2009; 374: 1045–1047.
5. Patel VM, Ashrafian H, Ahmed K, Arora S, Jiwan S, Nicholson J, et al. How has healthcare research performance been assessed? – a systematic review. *J R Soc Med* 2011; 104: 251–261.
6. Editorial. Games people play with authors' names. *Nature* 1997; 387: 831.
7. Greene M. The demise of the lone author. *Nature* 2007; 450: 1165.
8. Stallings J, Vance E, Yang J, Vannier MW, Liang J, Pang L, et al. Determining scientific impact using a collaboration index. *Proc Natl Acad Sci USA* 2013; 110: 9680–9685.
9. Stubbs C. The serious business of listing authors. *Nature* 1997; 388: 320.
10. Frandsen TF and Nicolaisen J. What is in a name? Credit assignment practices in different disciplines. *J Inf* 2010; 4: 608–617.
11. Kosmulski M. The order in the lists of authors in multi-author papers revisited. *J Inf* 2012; 6: 639–644.
12. Rennie D, Yank V and Emanuel L. When authorship fails. A proposal to make contributors accountable. *JAMA* 1997; 278: 579–585.
13. Shapiro DW, Wenger NS and Shapiro MF. The contributions of authors to multiauthored biomedical research papers. *JAMA* 1994; 271: 438–442.
14. Baerlocher MO, Newton M, Gautam T, Tomlinson G and Detsky AS. The meaning of author order in medical research. *J Investig Med* 2010; 55: 174–180.
15. Wislar JS, Flanagin A, Fontanarosa PB and Deangelis CD. Honorary and ghost authorship in high impact biomedical journals: a cross sectional survey. *BMJ* 2011; 343: d6128.
16. Akhabue E and Lautenbach E. "Equal" contributions and credit: an emerging trend in the characterization of authorship. *Ann Epidemiol* 2010; 20: 868–871.
17. Levitt JM and Thelwall M. Alphabetization and the skewing of first authorship towards last names early in the alphabet. *J Inf* 2013; 7: 575–582.
18. Waltman L. An empirical analysis of the use of alphabetical authorship in scientific publishing. *J Inf* 2012; 6: 700–711.

19. Allen L, Scott J, Brand A, Hlava M and Altman M. Publishing: credit where credit is due. *Nature* 2014; 508: 312–313.
20. Hodge SE and Greenberg DA. Publication credit. *Science* 1981; 213: 950–952.
21. Kennedy D. Multiple authors, multiple problems. *Science* 2003; 301: 733.
22. Sekercioglu CH. Quantifying coauthor contributions. *Science* 2008; 322: 371.
23. Shen HW and Barabási AL. Collective credit allocation in science. *Proc Natl Acad Sci USA* 2014; 111: 12325–12330.
24. Newman MEJ. *Networks: An Introduction*. Oxford: Oxford University Press, 2010.
25. Watts DJ and Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1988; 393: 440–442.
26. Burt RS. *Structural Holes. The Social Structure of Competition*. Cambridge, MA: Harvard University Press, 1992.
27. Patel VM, Ashrafian H, Bornmann L, Mutz R, Makanjuola J, Skapinakis P, et al. Enhancing the h index for the objective assessment of healthcare researcher performance and impact. *J R Soc Med* 2013; 106: 19–29.
28. National Academy of Sciences, National Academy of Engineering, Institute of Medicine and Committee on Maximizing the Potential of Women in Academic Science and Engineering. *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*. Washington, DC: National Academies Press, 2006.
29. Elsevier. *Scopus Website*. See http://www.scopus.com (last checked 2 May 2019).
30. Levitt JM. What is the optimal number of researchers for social science research? *Scientometrics* 2014; 102: 213–225.
31. Radicchi F, Fortunato S and Castellano C. Universality of citation distributions: toward an objective measure of scientific impact. *Proc Natl Acad Sci USA* 2008; 105: 17268–17272.
32. Ali SN, Young HC and Ali NM. Determining the quality of publications and research for tenure or promotion decisions: a preliminary checklist to assist. *Libr Review* 1996; 45: 39–53.
33. Brandes U and Wagner D. Visone—analysis and visualization of social networks. In: Jünger M, Mutzel P, eds. *Graph Drawing Software*. New York: Springer-Verlag, 2004:321–340.
34. Wasserman S and Faust K. *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press, 1994.
35. Bonacich PF. Power and centrality: a family of measures. *Am J Sociol* 1987; 92: 1170–1182.
36. Freeman LC. A set of measures of centrality based upon betweenness. *Sociometry* 1977; 40: 35–41.
37. Jaccard J, Turrisi R and Wan C. Interaction effects in multiple regression. Newbury Park, CA: SAGE, 1990.
38. Cameron AC and Trivedi PK. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press, 2005.
39. Cameron AC and Trivedi PK. *Regression Analysis of Count Data*. Cambridge: Cambridge University Press; 1998.
40. Guerrero-Medina G, Feliú-Mójer M, González-Espada W, Díaz-Muñoz G, López M, Díaz-Muñoz S, et al. Supporting diversity in science through social networking. *PLoS Biol* 2013; 11: e1001740.
41. Jagsi R, Guancial EA, Worobey CC, Henault LE, Chang Y, Starr R, et al. The "gender gap" in authorship of academic medical literature – a 35-year perspective. *N Engl J Med* 2006; 355: 281–287.
42. Patel VM, Ashrafian H, Almoudaris A, Makanjuola J, Bucciarelli-Ducci C, Darzi A, et al. Measuring academic performance for healthcare researchers with the H index: which search tool should be used? *Med Princ Pract* 2013; 22: 178–183.
43. Martin-Martin A, Orduna-Malea E, Thewall M and Lopez-Cozar E. Google Scholar, Web of Science, and Scopus: a systematic comparison of citations in 252 subject categories. *J Inf* 2018; 12: 1160–1177.
44. Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ* 1997; 314: 498–502.
45. Sheridan DJ. Reversing the decline of academic medicine in Europe. *Lancet* 2006; 367: 1698–1701.
46. Atasoylu AA, Wright SM, Beasley BW, Cofrancesco J Jr, Macpherson DS, Partridge T, et al. Promotion criteria for clinician-educators. *J Gen Intern Med* 2003; 18: 711–716.
47. Lee S and Bozeman B. The impact of research collaboration on scientific productivity. *Soc Stud Sci* 2005; 35: 673–702.
48. Cohen J and Siegel E. Academic medical centers and medical research: the challenges ahead. *JAMA* 2005; 294: 1367–1372.
49. Al-Herz W, Haider H, Al-Bahhar M and Sadeq A. Honorary authorship in biomedical journals: how common is it and why does it exist? *J Med Ethics* 2014; 40: 346–348.