The Relationship Between Author Collaboration and Research Output Based on Age Structure Distribution

Jingran Sheng[1 a *], Lin Wang[1b], Xiaofan Wang[1,2,3c]

sheng_jingran@sjtu.edu.cn [a*], wanglin@sjtu.edu.cn [b], xfwang@sjtu.edu.cn [c]

Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, China[1], School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China[2], School of Electrical and Electronic Engineering, Shanghai Institute of Technology, Shanghai, China[3]

Abstract. Collaboration has gradually become a norm in scientific research due to the quick development of technology. The relationship between the age structure of scholars and research output has received widespread attention. How to optimize the combination of scholars of different age groups and increase research output has become a highly practical research topic. Based on the Microsoft Academic Graph (MAG) database and choosing data from the physics discipline from 1970 to 2015, we define the academic age of scholars and cluster the academic age. The scholars are divided into three types, namely young scholars, middle-aged scholars, and senior scholars. The patterns of authors' collaboration are divided using the maximum and minimum value method. We build a mathematical model to study the relationship between research output and the patterns of authors' collaboration, and the fitted parameters are statistically analyzed. A robustness test is conducted using the relationship between citations and the patterns of authors, collaboration. In terms of the number of papers published and citations, the largest number of research output is collaboration between young scholars followed by collaboration between young scholars and middle-aged scholars, while collaboration between young scholars and senior scholars is the least. At the same time, we further provide the optimal average academic age for different collaborative patterns.

Keywords: Scientific research output; Age structure distribution; Cluster analysis; Skewed distribution; Average academic age

1 Introduction

With the continuous expansion of research teams and the increase of interdisciplinary collaboration, many studies have shown that the distribution of age structure has a significant impact on research output and author collaboration [1]. For instance, some study has revealed that teams with a more balanced age structure tend to produce higher-quality scientific research outcomes [2]. Meanwhile, disparities in knowledge background, research ideas, and other features among authors of different ages in scientific research collaboration may impact scientific research output [3] [4] [5]. Therefore, an in-depth exploration of the relationship between research output, age structure distribution, and author collaboration types is significant for promoting research innovation and improving research quality.

BDEDM 2024, January 12-14, Ningbo, People's Republic of China

DOI 10.4108/eai.12-1-2024.2347252

In recent years, the collaborative relationship between authors has become increasingly close, and the collaboration network also become more complex[6]. Thesecollaborative relationships have a substantial effect on the research output of both individual scholars and teams. Similarly, the age structure distribution of authors is a significant factor in determining their research output.

Matthews et al. [7] found the average age of biomedical researchers has steadily increased. Lee and Bozeman[8] investigated the impact of collaborative relationships between authors on scientific research output. According to their research, collaborative relationships between scholars can significantly increase their research output.

Taking highly cited scholars in Physics at the American Institute of Scientific Information as a sample, Miao et al.[9]investigated the academic age characteristics of scholars using a two-dimensional perspective of academic productivity and influence. The optimal academic age range for academic influence and productivity is approximately consistent with the "golden ratio", and the fitting relationship between academic influence and productivity and academic age is a cubic polynomial distribution. From the standpoint of citation counts, Liu[10] examined the physiological age distribution of the influence of exceptional researchers in the domains of biology and genetics, mathematics, and computer science. He found that the peak citation rates of researchers in molecular biology and genetics were between 41 and 55 years old, and in mathematics and computer science were between 31 and 45 years old. Liu and Jin[11] found that the proportion of papers published during their lifetime between the ages of 55 and 70 is as high as 50%, whereas the proportion of papers published after the age of 60 is only 30%, based on their research on the publication time of highly cited scholars in the field of molecular biology and genetics who are currently around the age of 70.

However, it is necessary to analyze the academic productivity and influence of scientific researchers from the perspective of academic age. For example, Hirsch[12], the initiator of the H-index, pointed out the idea of dividing the H-index by academic age. Jian et al. [13] used variance and correlation analytic methods, with academic age as the primary factor, to examine the performance of researchers with various academic age combinations on indicators such as number of articles, total citations, average citations, H-index, G-index, A index, and so on. Falagas et al. [14] showed that the academic output of scientific experts reduces as they get older.

Bao W et al.[15]used a bidirectional fixed effect model to evaluate the association between the age structure of teaching staff in research institutes and research output. They found that the university's research output reached its optimal theoretical level when the proportions of young, middle-aged, and elderly scholars were 51.2%, 43%, and 5.8%, respectively. There are varying degrees of deviations between the age structure of the teaching staff in different types of scientific research institutions and the theoretical optimal value, and the aging problem of the teaching staff will continue to worsen [16] [17] [18].

Sinatra et al. [19] constructed a random model that links the effects of productivity, personal ability, and luck, revealing the existence of universal patterns of scientific success based on the random distribution in scholars' publication sequences. This approach assigns each scholar a unique parameter Q that remains constant throughout their career, accurately predicting the progression of a scholar' s influence from cumulative citations to independent recognition based on the H-index.

In summary, past research has shown that author collaboration, team scientific research, and interpersonal networks all have a major impact on scientific research. Therefore, we assume that the effects of academic age structure distribution on collaboration networks and research output vary based on various disciplinary domains or research themes. Our study is based on the distribution of academic age structure and uses empirical and quantitative analysis to investigate the following issues:

1.How does the distribution of researchers' academic ages affect research output and the caliber of their research accomplishments?

2. What is the impact of the number, type, and trend of cooperation on the distribution of different age structures?

In this paper, we select data of 45 years of physics discipline from the Microsoft Academic Graph (MAG) database. We dynamically define the academic age of authors based on the published year of each paper. And we cluster scholars into three types,namely young scholars, middle-aged scholars, and senior scholars. The patterns of authors' collaboration are divided using the maximum and minimum value method. Based on existing data, we find from the scatter plot of the relationship between the number of papers published and the patterns of authors' collaboration that the results follow a skewed distribution. Therefore, we construct an exponential function model and calculate specific fitting parameter values using the curve fitting toolbox. A robustness analysis is conducted using the relationship between citations and the patterns of authors' collaboration, which verifies the rationality of the model.

2 Materials and methods

2.1 Data source and processing

Our data were derived from Microsoft Academic Graph (MAG), a database of papers from various fields and disciplines, such as biology, computer science, engineering, medicine, economics, physics, and psychology. According to the digital object identifier (DOI) and author information of the paper, the original data was integrated into a scientific dataset with a unique author/paper identifier. The dataset contained 150 million papers, including supporting information such as publication date, title, author, keywords, references, research field, and affiliated institutions. Each paper had multiple attributes that helped us comprehensively analyze from different perspectives.

We first performed a preliminary data cleaning. Papers with missing author information and unknown publication dates were removed. Then we used the method described in Sinatra' s study [19] to disambiguate the author' s name. Finally, we select papers published between 1970 and 2015 as the research subjects. At the same time, based on the field of study, papers are divided into 19 different disciplines. We select physics papers for research. The final dataset consisted of 662,065 authors and 1,674,334 papers.

Relevant data from the paper source are extracted, which include the author's name, 96 publication date, citation, and journal information. Based on the publication's metadata, we create a collaborative network to identify the collaborators and the number of collaborations for each scholar. We calculate the academic age of each scholar based on the publication date and establish distinct academic age classification requirements to divide them into distinct academic age groups. We then analyze the collaborative patterns and numbers of scholars of various academic age groups in the collaboration network. Statistics on each scholar’ s research output, including metrics like the number of papers published and the frequency of citations, and an analysis of the research output status of scholars in various age groups. Based on existing data scatter plots, we use the least squares method to fit the curve. The rationality of fitting parameters is demonstrated using regression analysis based on statistics. We examine the relationship between the age distribution, collaborative relationships and research output and confirm research theories.

2.2 Definition and calculation method of academic age structure

During the research process, it is generally possible to use the year when the authors published their first paper as the start year, and the current or final year of the study as the ending year to calculate the authors’ age during the beginning and ending periods.In order to dynamically grasp the author’s publication status, we define the author’ s academic age based on the publication time of his/her papers. The specific definition and calculation method are as follows:

Academic Age(Sa): The time interval formed by starting from the year of publication of the author’s first paper and ending with the year of publication of the current paper is defined as the academic age of the author.

$$\mathrm{S_{aj}} = PY_j - PY_1 + 1 \tag{1}$$

where PY1 represents the year of publication of the author’s first paper, PYj represents the year of publication of the author’sjth paper, j=1,2,3...,N.

Average Academic Age(ESa): The arithmetic mean of the academic age of all authors in the paper. For example, a paper has n authors, and the academic age of the ith author is Sai , (i=1,2,3,......,n), therefore

$$\mathrm{ESa} = \sum_{i=1}^{n} Sa_i / n \tag{2}$$

We cluster authors into different academic age groups. First, we find the median M of academic age in the database (if there are only two data, take the average of these two numbers), and use M * 2/3=d as the third point of the academic age group using the clustering criterion. Authors are divided into three categories: young, middle-aged, and senior scholars. The data indicators for the academic age group of the three categories of authors are as follows:

Young scholars(y): (0, d];

Middle-aged scholar(m): (d, 2d];

Senior Scholar(s): (d, —);

The papers in the database are sorted into the following seven groups based on the author's cooperative relationship, using the method of maximum and minimum values:

y0: Indicates sole authorship

y1: Indicates papers about collaborating among youth (y+y)，Max(Sa)d;

y2: Indicates papers about collaborating between young scholars and middle-aged scholars (y+m), Min(Sa)d，d<Max(Sa)2d;

y3: Indicates papers about collaborating between young scholars and senior scholars (y+s), Min(Sa)d, Max(Sa)2d;

y4: Indicates papers about collaborating between middle-aged scholars and middle-aged scholars (m+m)，d<Min(Sa)2d，d<Max(Sa)2d;

y5: Indicates papers about collaborating between middle-aged scholars and senior scholars (m+s)，d<Min(Sa)2d, Max(Sa)2d;

y6: Indicates papers about collaborating between senior scholars and senior academics (s+s), Min(Sa)2d。

To minimize potential errors in the computation of the age structure and the degree of effect of research output, genuine publication data and accurate information about academics were screened thoroughly to ensure the accuracy of the data. In this paper, he median academic age of the author M is 7, and d is 4.67.

2.3 Variables and their operational definitions

Table 1 depicts the operational definitions of variables related to this study and the symbolic representations of dependent and independent variables.

Table 1. Related variables and their operated definitions.

Variables/Indicators	operational definition	Corresponding group	Symbolic representation
Research output (dependent variable)	Total number of papers	—	Y
Sole authorship output	Number of sole authorship	y0	y0
The output of the cooperation situation	Number of papers (y+y)	y1	y1
	Number of papers (y+m)	y2	y2
	Number of papers (y+s)	y3	y3
	Number of papers (m+m)	y4	y4
	Number of papers (m+s)	y5	y5
	Number of papers (s+s)	y6	y6

This study assumes that the dependent variable is research output (the number of papers published/citations), and the specific observation indicator is the number of papers published in each database. The number of papers published is used to measure the effect of authorship collaboration on research output. Furthermore, citations are used as another observation of research output to produce additional robust checks of the analytic findings.

2.4 Model Methodology

In this study, analyzing the association between faculty age structure and research output, we split collaborations at different age groups into different variables and construct the following Model(Model I)(3):

$$Y_i = y_{i0} + y_{i1} + y_{i2} + y_{i3} + y_{i4} + y_{i5} + y_{i6} \quad (3)$$

Where Y_i, y_{i0}, y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5} and y_{i6} represent the total output at an average academic age of i years and the output of each scholar combination.

3 Results

3.1Research output statistics of scholars of different academic ages

In order to better study the relationship between scholar output and author collaboration patterns, based on the processed data in the database, we create a scatter plot of the relationship between the number of papers published and the average academic age under each type of collaboration. In Fig 1, the horizontal axis represents the average academic age of scholars, and the vertical axis is the number of published papers, which represents scientific research output.

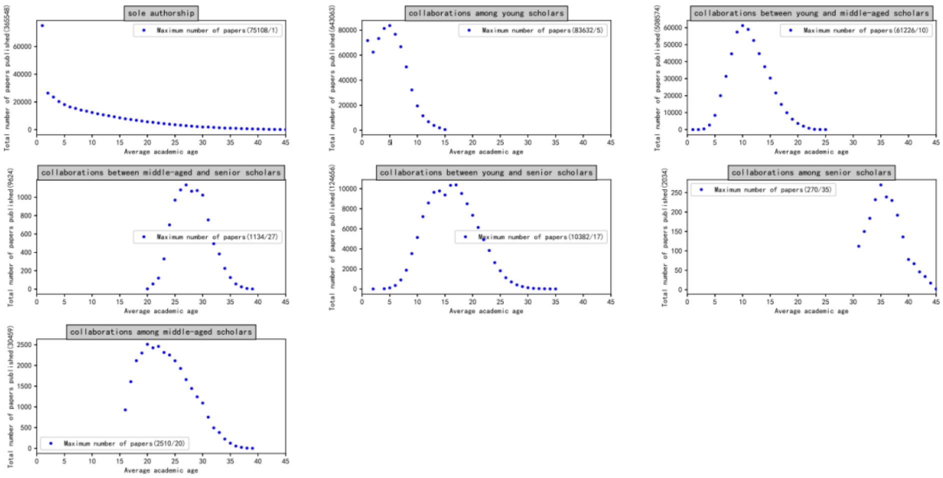

Fig 1. Scatterplot of the relationship between the number of collaborative papers and the average academic age.

Based on the scattering trend in Fig 1, it is consistent with a skewed distribution. We considered the exponential function to fit the curve, and thus Model 1 can be improved to the following form: (Model 2)(4)

$$Y_i = \sum_{j=0}^{6} a_{ij} e^{-\frac{(x_j - b_{ij})^2}{c_{ij}}} \quad (4)$$

Where $a_{ij}, b_{ij}, c_{ij} (j = 0,1,2 \cdots ,6)$ are the parameter to be fitted.Model 2 satisfies the trend of the scatter plot. However, there are too many parameters and no direct way to fit the form of the exponential function. Thus, we take the natural logarithm of each item in Model 2 to obtain the following improved model:(Model 3)(5)

$$\begin{cases} Y_i = y_{i0} + y_{i1} + y_{i2} + y_{i3} + y_{i4} + y_{i5} + y_{i6} \\ ln\, y_{ij} = ln\, a_{ij} - \dfrac{(b_{ij} - x_j)^2}{c_{ij}}, j = 0,1,2\cdots,6 \end{cases} \tag{5}$$

Figure 2 presents the fitted curves.

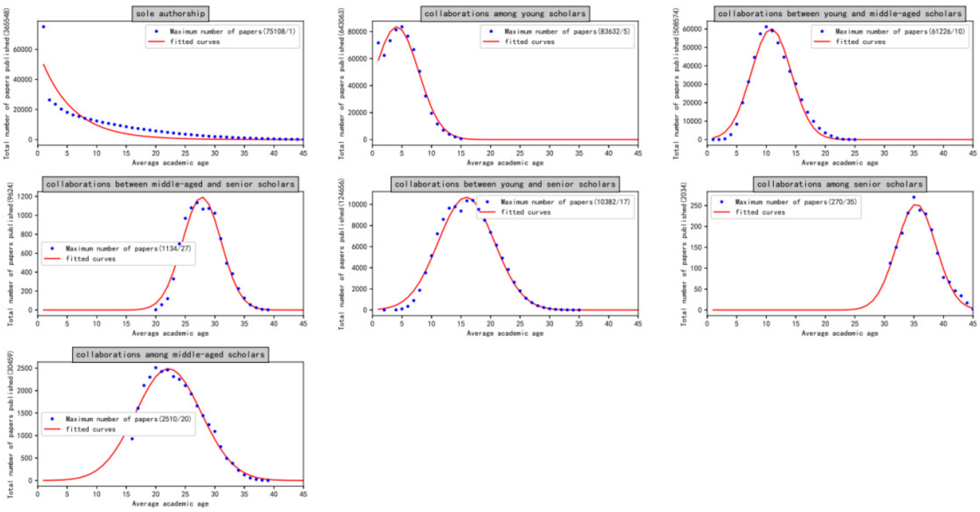

Fig 2. Fitting curve of the relationship between the number of collaborative papers and the average academic age.

3.2 Analysis of the relationship between author collaboration and research output

Model calculation results. In this paper, we counted the number of posts using the OLS regression model. Figure 1 presents the scatter plot of the number of publications, and statistical analysis is performed. The selected fitting equation is Eq(4).We simulated and analyzed all data using fitting formulas and plot the relationship between the number of published papers by different collaborative groups and average academic age. The fitting parameters are obtained during the data fitting process, as shown in Table 2.

Tab 2. Parameters of the fitted curve for the number of collaborative paper output publications.

a0	3.99507572e+163	b0	4.00997862e+003	c0	4.39695236e+004
a1	8.37028225e+04	b1	4.11127982e+00	c1	2.73810265e+01
a2	6.01849012e+04	b2	1.08513514e+01	c2	2.27268786e+01
a3	1193.92910792	b3	27.79853032	c3	21.46637797
a4	10684.86519434	b4	15.90180059	c4	44.94757693
a5	253.71699393	b5	35.39207376	c5	22.25620207
a6	2489.27791753	b6	22.14238589	c6	61.64667262

OLS regression is used in the simulation process, and least squares and F-statistics are selected to analyze the data. Table 4 depicts the results of the fitted linear regression analysis. The value of R2 is 0.956, which is very close to 1. The P is less than 0.5. The above indicates that the fitting effect of the experiment is good.

3.3 Robustness analysis

To test the robustness of the model, the observation indicator of the dependent variable is adjusted from the number of papers published to citations. We also calculate citations by different collaborative groups, and a scatter plot of citations is shown in Fig 3.

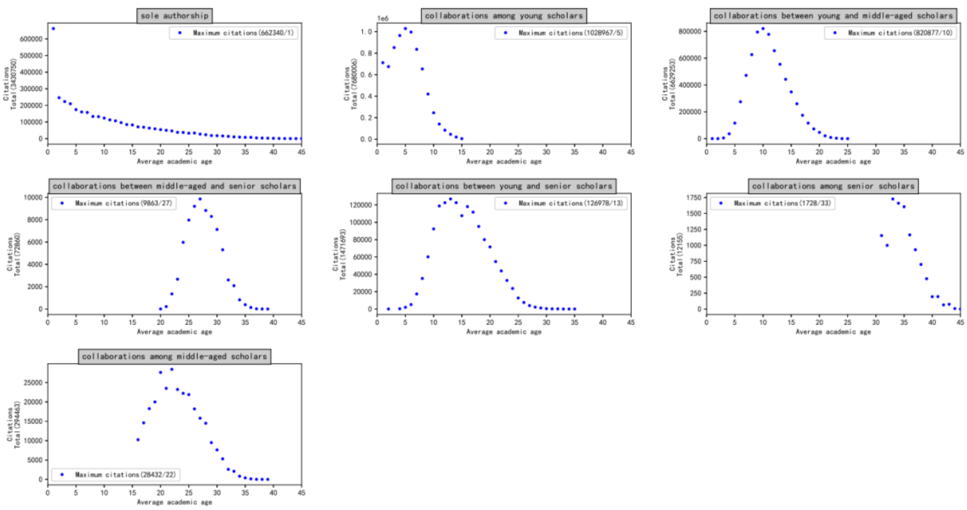

Fig 3. Scatterplot of citations of papers published by different collaborative groups.

The fitting equation chosen for the simulation process is Eq (4). Using OLS regression and F-statistics through experiments, the fitted curves of the number of times cited by others are shown in Fig 4. Table 3 depicts the parameters of citations, and the regression results are shown in Table 4. The value of R2 is 0.885, which is very close to 1, showing that the fitting effect of the experiment is good.

Tab 3. Parameters of the fitted curve of the number of citations by others for the output of collaborative papers.

a0	2.13229251e+164	b0	-4.74518884e+003	c0	6.16475824e+004
a1	1.02337529e+06	b1	4.64402301e+00	c1	2.28602809e+01
a2	8.03892663e+05	b2	1.05567044e+01	c2	2.13583414e+01
a3	9970.75971673	b3	27.37535614	c3	17.641048
a4	1.29081798e+05	b4	1.46395000e+01	c4	4.29754148e+01
a5	1621.33409931	b5	33.95181519	c5	18.68300882
a6	2.58895061e+04	b6	2.21309213e+01	c6	4.73550335e+01

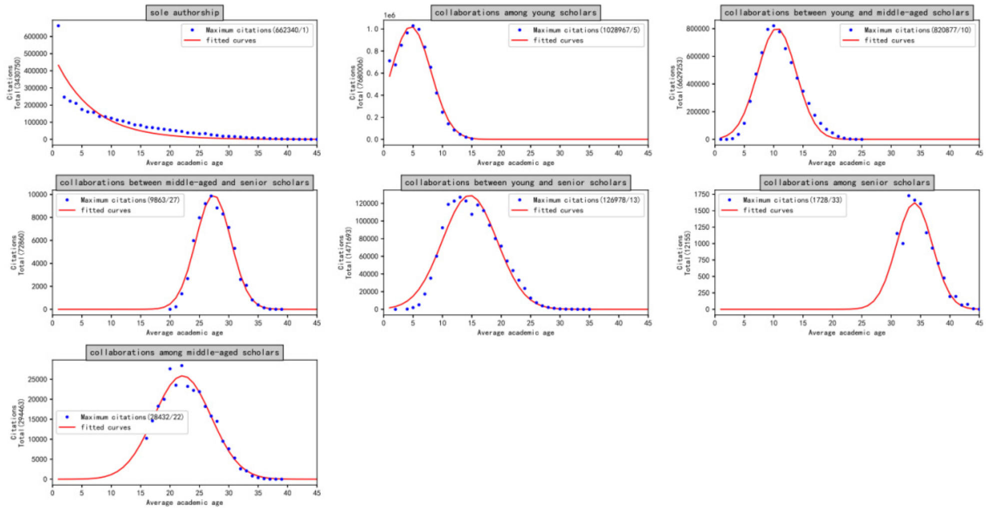

Fig 4. Fitting curve of the relationship between citations and the average academic age.

From a statistical standpoint, the simulated curves are statistically very near to the original data, but there are still varying degrees of deviation.

Tab 4. Results of the parametric regression analysis of the fitted curve of the number of published papers (citations) by different collaborative groups.

	Value	System error	p	
y0	-3.2124(-0.7097)	2.452(0.404)	0.198(0.087)	*(**)
y1	-1.3649(-0.1695)	1.149(0.142)	0.242(0.239)	*(*)
y2	-1.2863(-0.1046)	1.123(0.148)	0.260(0.485)	*(*)
y3	316.5412(21.2979)	54.464(9.357)	0.000(0.029)	***(**)
y4	-6.8153(-1.5724)	7.948(0.939)	0.397(0.102)	*(*)
y5	3180.7149(389.9963)	223.711(52.404)	0.000(0.000)	***(***)
y6	-25.9723(-5.8007)	31.417(3.878)	0.414(0.143)	*(**)
R2 :0.956(0.885), Adjusted R2 :(0.948)0.864, *P<0.5, **P<0.1, ***P<0.01				

3.4 The Theoretical optimal value of collaboration among scholars of different academic ages

After the experimental analysis, we obtained the theoretical optimal value of cooperation among scholars of different academic ages.

From the perspective of the number of papers published: when papers have solo author, the number of papers published decreases with the increase of average academic age, while the number of papers published by co-authors shows a trend of first increasing and then decreasing, indicating that authors' collaboration contributes significantly to the number of papers published. When collaborations are among young authors, the number of papers published gradually increases with the increase of average academic age. When the average academic age is 5, the number of publications reaches its optimal level. As for collaborations between young and middle-aged authors, when the average academic age is 10, the number of publications reaches its optimal level. When young scholars collaborate with senior scholars, the optimal average academic age is 17, as shown in Fig 5(a), 6(a), and 7(a). When middle-aged scholars collaborate with senior scholars, the optimal average academic age is 27. As for collaborations among senior scholars and among middle-aged scholars, the optimal average academic ages are 35 and 20, separately.

From the perspective of citation: when papers are solo author, citations decrease with the increase in average academic age. As for collaborations among young authors, when the average academic age is 5, citations reach the optimal value. As for collaborations between young and middle-aged scholars, when the average academic age is 10, citations reach the optimal value. When young scholars collaborate with senior scholars, the optimal average academic age is 13, as shown in Fig 5(b), 6(b), and 7(b). When middle-aged scholars collaborate with senior scholars, the optimal average academic age is 27. As for collaborations among senior scholars and among middle-aged scholars, the optimal average academic ages are 33 and 22, separately.

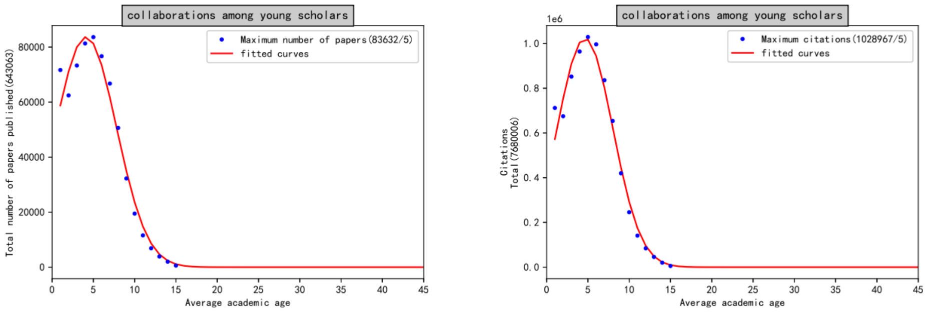

Fig 5. Comparison between the number of papers published and citations of collaborations among young scholars. A:The number of Papers published. B:Citations.

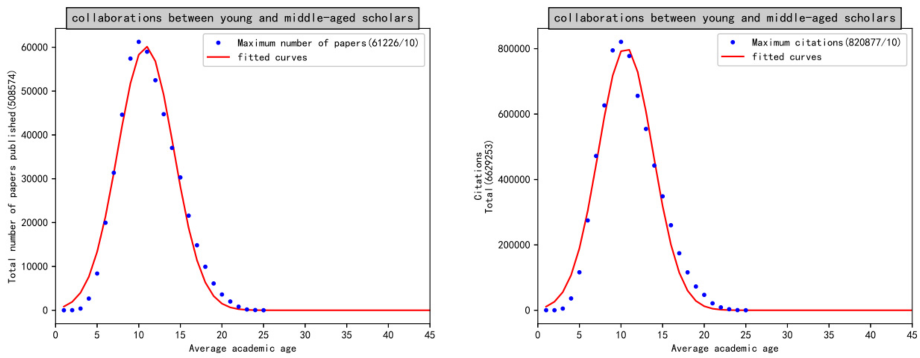

Fig 6. Comparison between the number of papers published and citations of collaborations among young and middle-aged scholars. A:The number of Papers published. B:Citations.

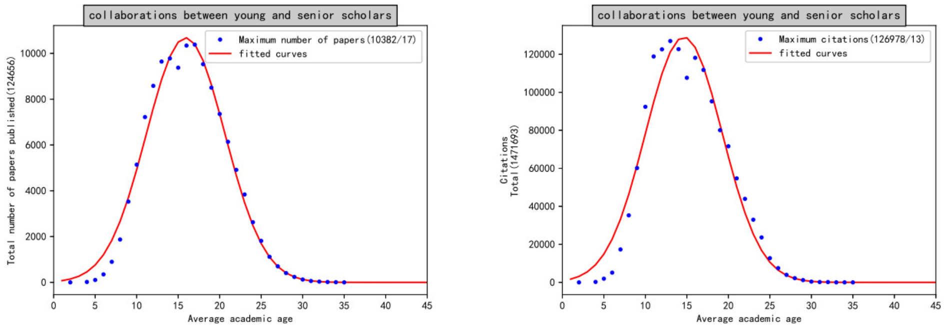

Fig 7. Comparison between the number of papers published and citations of collaborations among young and senior scholars. A:The number of Papers published. B:Citations.

From the comparison of the two data, we can find that the optimal theoretical value 235 is similar in terms of the number of papers published and the citations, and the 236 effectiveness of the data indicators and mathematical models is trustworthy.

4 Conclusion and inspiration

This paper analyzes the relationship between academic age and research output based on 45 years of paper data from the Microsoft Academic Graph (MAG) database in the field of physics and effectively groups the data using the clustering analysis method. A mathematical model and statistical analysis are developed using OLS regression and F-statistics to investigate the relationship between the number of publications and the type of author collaboration. The robustness test is performed using citations. The results of this study are found to be highly robust and reliable despite the limitations of the data sources and the possible omission of other variables. Also, the optimal average academic age for different collaborative patterns is provided.

Universities and research organizations should encourage interactions and collaboration between young, middle-aged, and senior scholars. Based on optimizing the age structure of the research team and supporting the development of young scholars, research institutions need to improve relevant systems, deepen collaboration between scholars of different age groups, encourage and support senior and middle-aged scholars to play a role in mentoring at the institutional level, and provide academic guidance and advice to young scholars while leading the construction of disciplines, thereby forming an academic community atmosphere of mutual assistance and collaboration. Scientific research institutions should also break the constraints of traditional academic power inherent models, establish a fair and just mechanism for allocating academic resources, create an atmosphere of equal communication and dialogue among scholars of different age groups, and build a value system shared by all members. In this way, the collaboration of senior, middle-aged, and young scholars can be realized to optimize the research performance of the organization. Universities and research institutions should employ scientific and technical personnel of varying ages and educational backgrounds to optimize their research teams, thereby enhancing the productivity and quality of their research output.

In our study, the number of papers published and citations in physics disciplines in the Microsoft Academic Graph database serves as a sample to study only one discipline, and the research object is from a single source, which has certain limitations. This paper only considers the effect of age structure factors on authorship collaboration and research output. However, other factors, such as different genders, different education levels, different regions, and different disciplines, may impact these relationships and need to be explored in future studies. Future research can further investigate the relationship between authorship collaboration and research output based on age structure distribution. First, a multidimensional perspective is examined, considering the age structure distribution and the influence of other factors (e.g., gender, title, education, etc.) on scientific and technological collaboration and research output. Second, horizontal expansion by broadening the scope of research to explore the differences between different countries and regions, fields, and disciplines [20] [21] [22] [23]. Third, as a methodological enhancement based on prior research, we combined technologies such as machine learning and big data to create more effective models for improved prediction and management of research collaboration and output [24] [25] [26].

References

[1] Milojevi´c S. How are academic age, productivity and collaboration related to citing behavior of researchers? *PloS one*. 2012 7(11), e49176.

[2] Liang Z, Ba Z, Mao J, Li G. Research complexity increases with scientists' academic age: Evidence from library and information science. *Journal of Informetrics*. 2023 17(1), 101375.

[3] Zhe C, Lu X, Xiong X. Analysis of influence factors on the quality of international collaboration research in the field of social sciences and humanities: The case of Chinese world class universities (2015–2019). *Sage Open*, 2021 11(4), 21582440211050381.

[4] Hou L, Pan Y, Zhu J. Impact of scientific, economic, geopolitical, and cultural factors on international research collaboration. *Journal of Informetrics*. 2021,15(3), 101194.

[5] Jung I.International Collaboration in Educational Technology Research: A Personal Reflection on Research Process, Experience, and Outcomes. *Information and Technology in Education and Learning.*Volume 2 , Issue 1 . 2022. PP Inv-p003-Inv-p003

[6] Zhang J,Yang X, Hu X,Li T. Author Cooperation Network in Biology and Chemistry Literature during 2014-2018: Construction and Structural Characteristics.*Information*. Volume 10,Issue 7.2019. PP 236-236.

[7] Matthews KR, Calhoun KM, Lo N, Ho V. The aging of biomedical research in the United States. *PloS one*. 2011 6(12), e29738.

[8] Lee S, Bozeman B. The impact of research collaboration on scientific productivity. *Social Studies of Science*. 2005 35(5), 673-702.

[9] YaJun M, Wei QI, Qi Z. Study on the academic age characteristics of scientists:based on the two dimensional of academic productivity and academic influence. *Studies in Science of Science*. 2013 31(2), 177-183.

[10] Junwan L. An Age Distribution of Outstanding Scientists' Scientific Influence. *Journal of the China Society for Scientific and Technical Information*. 2010 29(1), 121-127.

[11] Junwan L, Bihui J. An age distribution of productivity of scientific papers for highly cited scientists. *Science Research Management*. 2009 30(3), 96-103.

[12] Hirsch JE. An index to quantify an individual's scientific research output.*Proceedings of the National academy of Sciences*. 2005 102(46), 16569-16572.
[13] Jian D, Bin Z, Yang L, Xiaoli T, Peiyang X. Optimization of the evaluation indicators of scholars' research impact and comparative analysis between national and international academic behaviors of researchers. *Library and Information Service*. 2011 55(10),98-102
[14] Falagas ME, Ierodiakonou V, Alexiou VG. At what age do biomedical scientists do their best work? *The FASEB Journal*. 2008 22(12),4067-4070.
[15] Bao W, Hongbao J, Mingzhou T. The relationship between the age structure of faculty and the research performance of research universities in China. *Journal of Higher Education*. 2020 41(5),54-62.
[16] Gingras Y, Lariviere V, Macaluso B, Robitaille JP. The effects of aging on researchers' publication and citation patterns. *PloS one*. 2008 3(12), e4048.
[17] Zuckerman H, Merton RK. Age, aging, and age structure in science. *Higher Education*. 1972 4(2), 1-4.
[18] Liu M, Ajay J,Bu Y,Min C, Yang S,Liu Z, Daniel A, Ding Y,Team formation and team impact: The balance between team freshness and repeat collaboration. *Journal of Informetrics*,Volume 16, Issue 4,2022,101337,ISSN 1751-1577.
[19] Sinatra R, Wang D, Deville P, Song C, Barab´asi AL. Quantifying the evolution of individual scientific impact. *Science*. 2016 354(6312), aaf5239.
[20] Zhang M, Zhang G, Liu Y, Zhai X,Han X.Y. "Scientists' genders and international academic collaboration: An empirical study of Chinese universities and research institutes." *J. Informetrics* 14 (2020): 101068.
[21] Ma R, Li Z.High Citation or Zero Citation: Exploring the Optimal Scale of Research Cooperation Based on the Citation of Scientific Publication-Evidence from the Financial Times TOP 45 Journals.*Journal of the China Society for Scientific and Technical Information*, Nov. 2020, 39(11): 1182-1190
[22] Esva B, Jccd, E, Aurora A, Teixeira A. Which distance dimensions matter in international research collaboration? A cross-country analysis by scientific domain. *J. Informetr*. 2022, 16, 101259
[23] Yang R, Li X. The Correlation Study of Scientific Collaboration and the Influence of Paper.*Journal of Modern Information*,2019,39(4),125-133.
[24] Wang W, Xia F, Wu J, Gong Z, Tong H, Davison, BD. Scholar2vec: vector representation of scholars for lifetime collaborator prediction. *ACM Transactions on Knowledge Discovery from Data (TKDD)*. 2021 15(3), 1-19.
[25] Zhang X, Zhang Z, Chen X. Collaborative Features of Authors Based on Academic Journal Papers and Their Influence on Scientific Research Output-Taking Highly Published Authors of International Medical Informatics as an Example.*Journal of the China Society for Scientific and Technical Information*, Jan. 2019, 38(1): 29-37
[26] Rørstad K, Aksnes DW, Piro FN. Generational differences in international research collaboration: A bibliometric study of Norwegian University staff. *PLoS One*. 2021 16(11), e0260239.