



Leaving science—attrition of biologists in 38 OECD countries

 Marek Kwiek^{1,2}  and Lukasz Szymula³ 

- 1 Center for Public Policy Studies (CPPS), Adam Mickiewicz University of Poznan, Poland
 2 German Center for Higher Education Research and Science Studies (DZHW), Berlin, Germany
 3 Faculty of Mathematics and Computer Science, Adam Mickiewicz University of Poznan, Poland

Correspondence

M. Kwiek, Center for Public Policy Studies (CPPS), Adam Mickiewicz University of Poznan, ul. Szamarzewskiego 89c, 60-569 Poznan, Poland
 Tel: 0048 618292280
 E-mail: kwiekm@amu.edu.pl

(Received 9 December 2024, revised 13 February 2025, accepted 13 February 2025)

doi:10.1002/1873-3468.70028

Edited by Duncan Wright

We examine biologists leaving science in 38 OECD countries in the past two decades. In a cohort-based and longitudinal fashion, we follow individuals over time, from their first publication ($N = 86\,178$). We examine four disciplines: AGRI (agricultural, biological sciences), BIO (biochemistry, genetics, molecular biology), IMMU (immunology, microbiology), and NEURO (neuroscience). Our Kaplan–Meier survival analysis of BIO shows that 60% of women are still in science after 5 years, 40% after 10 years, and only 20% after 19 years. Women in BIO are 23.26% more likely than men to leave science after 10 years and 39.74% after 19 years. Gender differences increase consistently in later career stages. They are high, but comparing the 2000 and 2010 cohorts, have slightly decreased over time.

Keywords: academic careers; attrition and retention in science; biochemistry; biology; gender disparities; genetics; immunology; Kaplan–Meier estimator; microbiology; molecular biology; neuroscience; science of science; Scopus; survival analysis

In this study, we examine biologists in 38 OECD countries leaving science (quitting publishing in academic journals) in the past two decades. Our conceptual approach is straightforward: starting an academic career is making one's first contribution to an academic journal, and leaving science is publishing one's final research paper. We use publication metadata from a global bibliometric database (raw, structured, and curated Scopus data at the micro-level of individual scientists). We follow individual men and women scientists over time, from their first publication to their last.

Methodologically, our approach is cohort-based [1] and longitudinal [2]. We track two cohorts of scientists for up to 22 years. We examine the academic

publishing careers of 34 970 biologists who started publishing in 2000 (termed the 2000 cohort); and to see the changes from a temporal perspective, we examine the publishing careers of 51 208 biologists who started publishing in 2010 (termed the 2010 cohort).

Our focus is on the four academic disciplines closely related to biology: AGRI, agricultural and biological sciences; BIO, biochemistry, genetics, and molecular biology; IMMU, immunology and microbiology; and NEURO, neuroscience. The largest discipline studied is BIO, with ~23 000 scientists in the older cohort and 32 000 scientists in the younger cohort. In total, we analyze the publishing histories of 86 178 scientists from 38 OECD countries, tracking the entirety of their academic careers until 2022.

Abbreviations

AGRI, agricultural, biological sciences; ASJC, All Science Journal Classification; BIO, biochemistry, genetics, molecular biology; COMP, computing; ENG, engineering; IMMU, immunology, microbiology; MATH, mathematics; NEURO, neuroscience; OECD, The Organization for Economic Co-operation and Development; PHYS, physics and astronomy; STEM, Science, Technology, Engineering, and Mathematics; STEMM, Science, Technology, Engineering, Mathematics, and Medicine.

We quantify the phenomenon commonly referred to in the literature as “leaving science” [3–8]. Our original, larger study was focused on 16 STEM disciplines (science, technology, engineering, mathematics, and medicine) and 11 non-overlapping cohorts of scientists ($N = 2\,127\,803$) [9]. The present study is restricted to the four related disciplines and two cohorts of scientists.

We use micro-level bibliometric metadata regarded as the digital traces left by (publishing) academic scientists over their careers. Consequently, academic non-publishers are not included in our research – as they leave behind no metadata of a longitudinal nature to examine. In fact, “academic career” as defined by entering and leaving science here means “publishing academic career.” However, publishing research articles in STEM disciplines, following the long tradition of academic career studies, is equivalent to being an active researcher in the field [10,11].

With digital traces left behind and indexed in global publication datasets, individual scientists can be studied according to their academic age, seniority, institution type, collaboration and publishing patterns, mobility, as well as discipline and gender [12,13], with the latter two being the focus of this study. We use a unique opportunity to track individuals over their academic careers at a level of detail and at a scale previously unattainable [14]. For large-scale multi-country examinations of the science profession, only bibliometric datasets are possible sources of reliable data, despite their limitations – which are widely discussed in the scientometric literature.

Our global and longitudinal approach to academic careers represents a more general turn toward structured big data in social science research [15], with repurposing of new data [16]. In particular, we make use of remarkable progress in defining the gender of scientists with massive gender-detection tools [17,18]. Global academic career research can offer insight into the professional trajectories of men and women scientists and identify disciplinary and gender variations.

In this study, our interests go beyond biologists in any single national science system or in any given year, as in survey- and interview-based research. Our methodological approach is longitudinal (rather than cross-sectional) and global (rather than national). We view the whole traditional monolith of STEM science as divided into several segments with vastly different representations of women, with these changing at different speeds in the past 30 [19].

Our original research about 16 STEM disciplines [9], from which the data on biologists are drawn, has been discussed in 25 countries and in 15 languages, as

well as featured in *Nature* [20]. We assume that the response to this research, in both popular and professional venues, is related to its scale, scope and longitudinal dataset, which was used for new purposes – examining the academic profession at a large scale, beyond national borders, and by gender.

In the present study, we turned a large, raw, bibliometric dataset (Scopus) into a large dataset about academic careers in biology-related disciplines. The study is longitudinal in the strict sense of the term: The two cohorts are tracked over time for up to two decades (2000–2022) on a yearly basis. Scopus is particularly suitable for global analysis at the micro-level of individual scientists because it is structured around individual Scopus Author IDs, apart from its focus on publications and their metadata [21]. In terms of data flow, our sample consists of all scientists starting to publish in 2000 or 2010, publishing in one of four disciplines (AGRI, BIO, IMM, and NEURO), having at least two research articles in their lifetime publishing portfolios, and having clearly defined affiliations in an OECD country and a clearly defined gender (binary: male or female).

Previous studies

Here, our focus is on persistent gender differences in biologist attrition (leaving science/quitting academic publishing). Research conducted over the past three decades has shown that attrition is generally greater for women than for men [5,7]; women are reported to leave science earlier and much more frequently than men. Gender differences in attrition have been explained by several hypotheses and discussed using useful metaphors such as “leaky pipeline” and “chilly climate” [22,23]. In general terms, a chilly climate in STEM disciplines means an unwelcoming work environment for women. And a hostile environment can discourage women from entering and pursuing academic careers [24].

While the “pipeline” perspective is focused on individuals, the “pathways” perspective is focused on organizations [25]. The critical point in the pathways perspective is that the social structures of the science enterprise are changeable. The focus of the pathways literature is on the structural conditions of science [26]. A somehow passive flow of women between academic career stages is assumed in the first metaphor; in the second, in contrast, the focus is on possible actions to be taken to make academic workplaces more equitable [27]. The pipeline metaphor disregards “persistent exclusionary messages” directed at women [28] and the practical answer for this perspective is to block their

leakage [27] and to increase their supply (see criticism in Ref. [25]).

The “pathways” metaphor, in contrast, emphasizes progression in science and its focus is on the organizational culture [29]. The organizational culture involves shared values, beliefs, and behaviors; it defines what is valued in organizations and promotes standard “ways of doing business.” Ways of doing business also tend to influence how women collaborate internationally [30] and how they form research teams [31].

The conceptual explanations why faculty leave higher education include often “push” and “pull” factors (a line of research started in Ref. [32]). However, if push factors are minimal, then pull factors are expected to have less weight in departure decisions [8]. The internal push is generally more important than the external pull in most departure decisions [32–34]. Internal environmental factors influencing decisions to leave consist of both intangible and tangible benefits of the job [32].

A recent study based on surveys and interviews with ex-academics shows how the push and pull factors are interlinked: individual factors (such as lack of interest in research or unwillingness to compete for academic positions) trigger individual decisions to leave academia. Individual factors tend to mediate push and pull factors [34]. Scientists and scholars work in ever more competitive and unstable environments – which makes many of them wonder not only about changing their institutions but also changing the sector entirely, that is, leaving academia.

It is important to remember, while discussing attrition in science, that organizational cultures are always fluid rather than stable and that work expectations play fundamental roles in faculty careers [35]. Faculty before joining the academic profession have a general sense of what it is to be part of the academic community – and for many, these expectations can be unmet. The practical implication is to “leave less unsaid” to new generations of incoming scientists, to avoid “broken contracts” [35], leading to disillusionment with an institution or with the profession.

Scientists often quit science because of work–life balance [3,36]. However, what also matters in taking departure decisions is low job security and relatively low salaries in academia [37], as well as discrimination [5]. Women, at every career stage, are reported to be leaving academia at higher rates than men [8]. A recent survey study shows that work climate is more important for women than work–life balance in taking their decision to leave academic careers [38]. The authors highlight that women are more likely to report feeling pushed out of their academic institutions than

men. And men are often pulled toward attractive outside opportunities, either in the academic sector or beyond. Literature also shows that an academic culture explains better the underrepresentation of women in most STEM disciplines than family responsibilities [4].

Leaving science is a traditional scholarly topic, explored mostly through case study research (based on surveys and interviews) or through multi-year US studies of post-secondary faculty leading to single-nation results. Concepts such as “intentions to leave” and “faculty turnover” have been studied for several decades [3,8,36,37]. Most attrition studies have focused on single academic institutions (as in Refs [5,32]) and their geographical scope was restricted to the United States (exceptions include Milojevic *et al.* [39], who studied several disciplines by using data from major academic journals).

Data and methods

We analyzed the publishing careers of 34 970 biologists who started academic publishing in 2000 and 51 208 biologists who started academic publishing in 2010. We examined four biology-related disciplines: AGRI, BIO, IMMU, and NEURO. In all, we analyzed the publishing histories of 86 178 biologists from 38 OECD countries, tracking the entirety of their publishing careers until 2022. Our data come from Scopus, a global publishing and citation database. Our access to the raw Scopus dataset was granted through the International Center for the Study of Research (ICSR) Lab platform.

Leaving science was conceptualized as no longer publishing in academic journals. We used a “survival analysis” of the publishing behavior of individuals over the years – until the event of finally not publishing occurs. Using the bibliometric dataset, we were able to locate in time the event of not publishing and to analyze gender and disciplinary differences in attrition (and retention) for different cohorts of scientists, or for scientists starting their academic careers at different points in time.

We conceptualized scientific life as a sequence of scholarly publishing events. Scientists do research – and publish research findings. Leaving science (or quitting publishing) was therefore conceptualized as an event located in the academic biographies of scientists, and survival analysis was applied. In survival analysis, the timing of the occurrence of an event is explored.

In our case, some scientists did not experience the event in 2022 or earlier. They continued publishing. These were termed right-censored observations about

which we had only partial information (for instance, the event may have occurred in 2023 or 2024). Therefore, to classify an author as leaving science, the final publication had to be dated 2018 (and marked as leaving science in 2019) or earlier. We used only uncensored cases, which represent observations for which we have both the start and end year of being in science.

We were not able to distinguish between scientists stopping publishing and scientists aging out of academe: in our data, generally, both finish their (publishing) careers. However, our study is cohort-based: all scientists in our sample had their first publication in either 2000 or 2010. The tacit assumption based on age distribution of newcomers in academic science is that the chance for STEM scientists to be retiring after 15 or 25 years is small, although such an option cannot be entirely excluded (our large-scale study of the correlation between academic age and biological age in science shows that the correlation coefficient ranges from 0.82 for biochemistry, genetics, and molecular biology to 0.90 for immunology and microbiology [40]).

For the four disciplines, for each subsequent year, we had the initial number of scientists entering the time interval (which consisted of scientists who left science during the interval and those who stayed in science in the following interval – in our case, the following year). The probability of survival until a given time interval was calculated by multiplying all probabilities of survival across all time intervals preceding the time [41].

Survival analysis is presented in tables, separately for each discipline, and in graphs that highlight disciplinary differences in attrition and retention over time and by gender. We tested the power of structured and reliable big data of the bibliometric type to examine milestones in academic careers (as in Ref. [19] on the inflow of young male and female scientists in 38 OECD countries).

From a disciplinary perspective, the major dividing line is between male-dominated STEMM disciplines, with participation of women scientists usually lower than 20% and very slowly increasing over time, and the STEMM disciplines in which gender balance has been achieved. We assume that gender balance in science means the participation of women at the level of 40–60% [42].

In the four disciplines, for the two cohorts examined, gender balance has been achieved. For instance, in the largest discipline of BIO, 47.83% of the 2000 cohort and 52.75% of the 2010 cohort were women (Table 1). Our previous research showed that in gender-balanced disciplines, the attrition of women

Table 1. The sample, four disciplines, two cohorts of scientists (2000 and 2010) by gender.

	Men		Women		Total N
	N	%	N	%	
Panel 1: the 2000 cohort					
AGRI	4962	62.26	3008	37.74	7970
BIO	11 839	52.17	10 853	47.83	22 692
IMMU	805	45.35	970	54.65	1775
NEURO	1353	53.41	1180	46.59	2533
Total	18 959	54.22	16 011	45.78	34 970
Panel 2: the 2010 cohort					
AGRI	6755	52.81	6037	47.19	12 792
BIO	14 905	47.25	16 637	52.75	31 542
IMMU	939	39.77	1422	60.23	2361
NEURO	2153	47.71	2360	52.29	4513
Total	24 752	48.34	26 456	51.66	51 208

was higher than that of men. For the heavily male-dominated disciplines of MATH (mathematics), COMP (computing), ENG (engineering), and PHYS (physics and astronomy), in contrast, attrition rates for men and women were nearly equal across 38 OECD countries, with no statistically significant differences.

Membership in the four disciplines was determined by computing all cited references in all publications published during whole academic careers. All cited references for each individual scientist were linked to academic journals (using the ASJC journal classification list used in Scopus), and the most often occurring value (or the modal value) was selected. When two modal values occurred in computations, the individual scientist was removed from further analysis.

Consequently, we used a sample of scientists with clearly determined characteristics: they were all research-active in the four disciplines and they had a clearly determined country affiliation and an unambiguously determined gender. They were also assigned a year of exit from academic publishing. Details regarding determination of the various individual-level characteristics of the larger sample from which the subsample of biologists was derived were discussed in our original report and its appendices [9]. What is especially important is the validity of gender determination. We inferred gender from bibliometric metadata by using Namsor, one of the best gender-detection tools currently available [43].

Results

A classic statistical technique for survival analysis is the Kaplan–Meier estimate of survival [41], with a plot

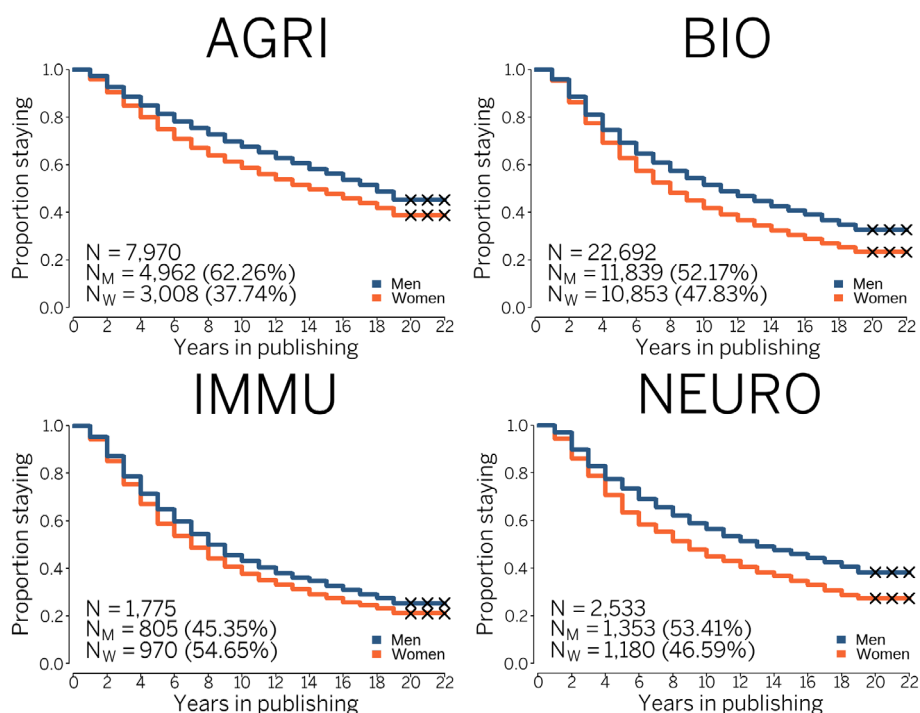


Fig. 1. 2000 cohort of scientists, Kaplan–Meier survival curve, AGRI, agricultural and biological sciences ($N = 7970$), BIO, biochemistry, genetics, and molecular biology ($N = 22\,692$), IMMU, immunology and microbiology ($N = 1775$), and NEURO, neuroscience ($N = 2533$). The four disciplines combined ($N = 34\,970$).

of the Kaplan–Meier estimator shown as a series of characteristically declining horizontal steps of various heights. Figure 1 shows gender differences in attrition (and retention) for the 2000 cohort of scientists separately for each discipline through Kaplan–Meier survival curves. Steps lead down over time, one step for 1 year, with probabilities of staying in science shown on the y -axis and the number of years spent publishing since 2000 on the x -axis. Small crosses denote right-censored observations for the three most recent years studied (i.e., 2020–2022), which were not used in computations.

For the 2000 cohort, we can observe, year after year, how the probability of staying in science is always higher for men and always lower for women. Similarly, Fig. 2 shows gender differences for the 2010 cohort. What is clearly visible at the level of survival curves is that attrition levels for the younger cohort of scientists are higher (and consequently retention levels are lower), both for men and women, than for the older cohort.

Young people, usually in the career stage of doctoral school, come to biology and disappear from it much faster in the younger cohort. At the same time, the two pictures of the two cohorts coming and

staying in science differ considerably in gender terms. The differences for the younger cohort are lower, especially in the two largest disciplines examined, namely, AGRI and BIO.

None of the disciplines was resistant to changes from a temporal perspective. Time matters in science: while there were ever more women in the disciplines (Table 1), gender differences in the attrition (and retention) rates became smaller over time.

Overall, 62.8% of women in the 2000 cohort in BIO were still in science after 5 years, 41.7% after 10 years, and only 23.4% by the end of the period examined (or after 19 years). For men in BIO, the percentages were substantially higher: 69.2%, 51.4%, and 32.7%, respectively (Table 2). These numbers highlight how differently, on average, male and female careers developed over time.

Thus, women in BIO were 10.19% more likely to drop out of science than men after 5 years, with markedly increasing chances of dropping out in later years. Women were 23.26% more likely to leave science after 10 years, and 39.74% more likely to leave science before the end of the studied period. Gender difference in attrition, only slightly visible after 5 years, consistently increased in later career stages.

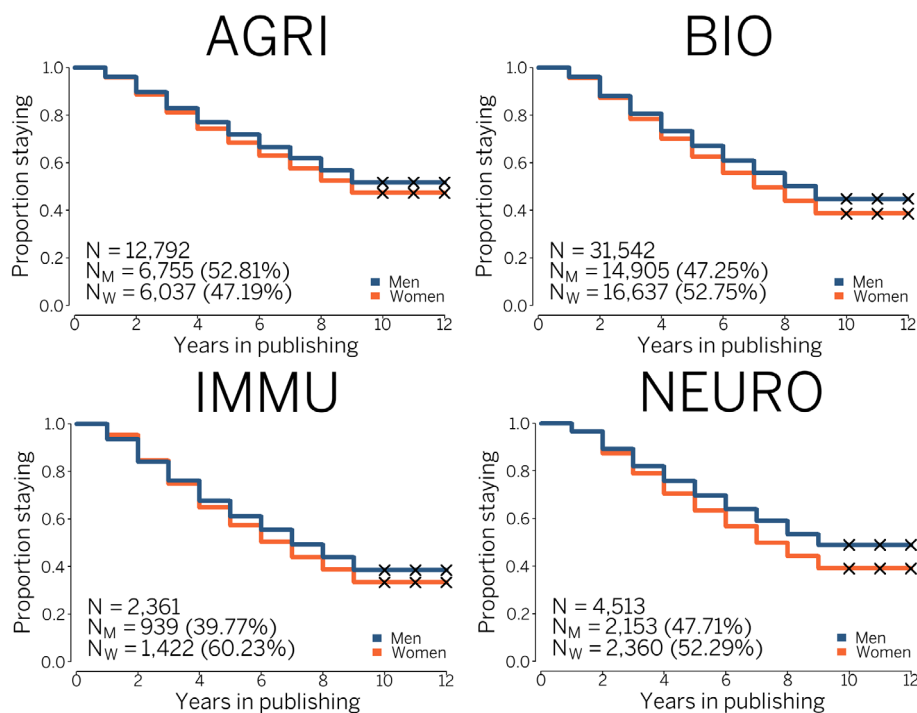


Fig. 2. 2010 cohort of scientists, Kaplan–Meier survival curve, AGRI, agricultural and biological sciences ($N = 12\,792$), BIO, biochemistry, genetics, and molecular biology ($N = 31\,542$), IMMU, immunology and microbiology ($N = 2\,361$), and NEURO, neuroscience ($N = 4\,513$). The four disciplines combined ($N = 51\,208$).

Similarly, women in NEURO were 15.77% more likely to drop out than men after 5 years, 25.61% more likely to leave science after 10 years, and 39.42% more likely to leave science by the end of the studied period. After 19 years (i.e., only uncensored observations), the chances of surviving in either BIO or NEURO were 40% higher for men than for women (Table 2 and Table S4).

In the two other disciplines, gender differences in attrition were considerably lower. For AGRI, the chances of surviving 10 years were 15.21% higher for men, and 19 years, 16.49% higher for men (Table 3). For IMMU, our data show 14.29% and 20.28%, respectively (Table S3).

The Kaplan–Meier survival curves reflect the underlying data produced for each cohort, year, discipline, and gender. Table 2 (for BIO) and Table 3 (for AGRI) highlight the changing Kaplan–Meier probability of staying in science (with a 95% confidence interval) separately for women and men in each year starting with 2000 (or 2010).

Taking scientists in BIO as an example, in year 2000, 10 853 women and 11 839 men had their first publication, but in year 1 (Table 2), 526 women and 494 men stopped publishing. In year 1, the difference

in retention (and, by definition, in attrition, its mirror image) between men and women was small but statistically significant. However, the gender difference in retention increased with every year, from 10.19% after 5 years to 39.74% after 19 years. Gender differences were substantial and statistically significant, with narrow confidence intervals indicating high concentrations of observations around the percentages shown.

In the case of the 2010 BIO cohort, the gender difference in retention was lower, and after 9 years, that is, in the last year examined in our study, 38.7% of women stayed in science, as opposed to 44.8% of men. For this younger cohort, the probability of staying in science after 9 years was 15.76% higher for men, compared with 20.94% in the older cohort.

Importantly, while gender differences in attrition diminished, attrition became an ever more acute problem: after 9 years (with no data beyond this period), retention rates decreased by more than 15 percentage points (from 61.2% to 47.3% for women and from 69.7% to 51.8% for men). The chances of surviving in science became smaller for both men and women. This indicates an issue requiring reflection not only on women but on the science workforce in general.

Table 2. 2000 and 2010 cohorts of scientists, BIO, biochemistry, genetics, and molecular biology. Kaplan–Meier estimate by gender with total counts for men and women, time (in years), number of observations of scientists leaving science. Kaplan–Meier probability of staying in science with a 95% confidence interval.

Time (years)	Women (cohort 2000)			Men (cohort 2000)			Probability of staying is higher for men by ... (in %)
	<i>n</i>	No. leaving science	KM probability (staying) with 95% CI and SE	<i>n</i>	No. leaving science	KM probability (staying) with 95% CI and SE	
1	10 853	526	0.952 (0.948–0.956) ^a	11 839	494	0.958 (0.958–0.958) ^a	0.63
2	10 327	971	0.862 (0.856–0.869) ^a	11 345	865	0.885 (0.885–0.885) ^a	2.67
3	9356	956	0.774 (0.766–0.782) ^a	10 480	900	0.809 (0.809–0.809) ^a	4.52
4	8400	901	0.691 (0.682–0.700) ^a	9580	754	0.746 (0.746–0.746) ^a	7.96
5	7499	678	0.628 (0.619–0.638) ^a	8826	637	0.692 (0.692–0.692) ^a	10.19
6	6821	607	0.573 (0.563–0.582) ^a	8189	530	0.647 (0.647–0.647) ^a	12.91
7	6214	526	0.524 (0.515–0.534) ^a	7659	445	0.609 (0.609–0.609) ^a	16.22
8	5688	465	0.481 (0.472–0.491) ^a	7214	435	0.573 (0.573–0.573) ^a	19.13
9	5223	353	0.449 (0.439–0.458) ^a	6779	349	0.543 (0.543–0.543) ^a	20.94
10	4870	345	0.417 (0.408–0.426) ^a	6430	341	0.514 (0.514–0.514) ^a	23.26
11	4525	290	0.390 (0.381–0.400) ^a	6089	292	0.490 (0.490–0.490) ^a	25.64
12	4235	255	0.367 (0.358–0.376) ^a	5797	254	0.468 (0.468–0.468) ^a	27.52
13	3980	249	0.344 (0.335–0.353) ^a	5543	255	0.447 (0.447–0.447) ^a	29.94
14	3731	210	0.324 (0.316–0.333) ^a	5288	251	0.425 (0.425–0.425) ^a	31.17
15	3521	212	0.305 (0.296–0.314) ^a	5037	212	0.408 (0.408–0.408) ^a	33.77
16	3309	189	0.287 (0.279–0.296) ^a	4825	205	0.390 (0.390–0.390) ^a	35.89
17	3120	198	0.269 (0.261–0.278) ^a	4620	279	0.367 (0.367–0.367) ^a	36.43
18	2922	184	0.252 (0.244–0.261) ^a	4341	221	0.348 (0.348–0.348) ^a	38.10
19	2738	203	0.234 (0.226–0.242) ^a	4120	243	0.327 (0.327–0.327) ^a	39.74

Time (years)	Women (cohort 2010)			Men (cohort 2010)			Probability of staying is higher for men by ... (in %)
	<i>n</i>	No. leaving science	KM probability (staying) with 95% CI and SE	<i>n</i>	No. leaving science	KM probability (staying) with 95% CI and SE	
1	16 637	715	0.957 (0.954–0.960) ^a	14 905	600	0.960 (0.960–0.960) ^a	0.31
2	15 922	1423	0.871 (0.866–0.877) ^a	14 305	1199	0.879 (0.879–0.879) ^a	0.92
3	14 499	1464	0.783 (0.777–0.790) ^a	13 106	1085	0.807 (0.807–0.807) ^a	3.07
4	13 035	1389	0.700 (0.693–0.707) ^a	12 021	1092	0.733 (0.733–0.733) ^a	4.71
5	11 646	1246	0.625 (0.618–0.633) ^a	10 929	940	0.670 (0.670–0.670) ^a	7.20
6	10 400	1158	0.556 (0.548–0.563) ^a	9989	901	0.610 (0.610–0.610) ^a	9.71
7	9242	980	0.497 (0.489–0.504) ^a	9088	805	0.556 (0.556–0.556) ^a	11.87
8	8262	931	0.441 (0.433–0.448) ^a	8283	806	0.502 (0.502–0.502) ^a	13.83
9	7331	895	0.387 (0.380–0.394) ^a	7477	797	0.448 (0.448–0.448) ^a	15.76

^aStandard error of 0.01 (total in Table S1).

Comparing the four disciplines, for the older cohort, gender differences in retention were highest for NEURO after 5 years, after 10 years, and at end of the period examined, reaching 10% in each case. Cross-disciplinary differences were marked, with only about 20% of women in IMMU contrasted with about 40% in AGRI remaining in science at the end of the period.

Comparing the four disciplines for the younger cohort after 9 years, the highest probability of staying in science was observed for AGRI (47.3% for women, 51.8% for men), with women in the other disciplines having probabilities below 40%. The gender difference was highest for NEURO, reaching nearly 10 percentage points (39.1% vs. 48.9%).

The estimated probability that a woman would survive in science 15 years differed considerably between the four disciplines: it varied from as much as 47.8% in AGRI (95% confidence interval: 0.461–0.496) to as little as 27.6%, or 20 p.p. less, in IMMU (0.250–0.306). For men, the probabilities were much higher at 56.2% and 32.7% (higher by 17.57% and 18.48%, respectively; Table 2 and Table S2).

Our data show that women in BIO (as well as those AGRI and NEURO) disappeared from science, with the passage of time, in ever-larger proportions compared with men. In contrast, women in male-dominated disciplines – with shares of women in the range of 15–20%, namely PHYS, COMP, and MATH, disappeared from science in almost exactly the same

Table 3. 2000 and 2010 cohorts of scientists, AGRI, agricultural and biological sciences. Kaplan–Meier estimate by gender with total counts for men and women, time (in years), number of observations of scientists leaving science. Kaplan–Meier probability of staying in science with a 95% confidence interval.

Time (years)	Women (cohort 2000)			Men (cohort 2000)			Probability of staying is higher for men by ... (in %)
	<i>n</i>	No. leaving science	KM probability (staying) with 95% CI and SE	<i>n</i>	No. leaving science	KM probability (staying) with 95% CI and SE	
1	3008	125	0.958 (0.951–0.966) ^a	4962	138	0.972 (0.972–0.972) ^a	1.46
2	2883	167	0.903 (0.892–0.914) ^a	4824	226	0.927 (0.927–0.927) ^a	2.66
3	2716	170	0.846 (0.834–0.859) ^a	4598	212	0.884 (0.884–0.884) ^a	4.49
4	2546	145	0.798 (0.784–0.813) ^a	4386	169	0.850 (0.850–0.850) ^a	6.52
5	2401	144	0.750 (0.735–0.766) ^a	4217	182	0.813 (0.813–0.813) ^a	8.40
6	2257	125	0.709 (0.693–0.725) ^a	4035	158	0.781 (0.781–0.781) ^a	10.16
7	2132	115	0.671 (0.654–0.688) ^a	3877	142	0.753 (0.753–0.753) ^a	12.22
8	2017	95	0.639 (0.622–0.656) ^a	3735	131	0.726 (0.726–0.726) ^a	13.62
9	1922	82	0.612 (0.595–0.629) ^a	3604	146	0.697 (0.697–0.697) ^a	13.89
10	1840	79	0.585 (0.568–0.603) ^a	3458	113	0.674 (0.674–0.674) ^a	15.21
11	1761	76	0.560 (0.543–0.578) ^a	3345	113	0.651 (0.651–0.651) ^a	16.25
12	1685	70	0.537 (0.519–0.555) ^a	3232	115	0.628 (0.628–0.628) ^a	16.95
13	1615	65	0.515 (0.498–0.533) ^a	3117	109	0.606 (0.606–0.606) ^a	17.67
14	1550	59	0.496 (0.478–0.514) ^a	3008	124	0.581 (0.581–0.581) ^a	17.14
15	1491	53	0.478 (0.461–0.496) ^a	2884	97	0.562 (0.562–0.562) ^a	17.57
16	1438	59	0.458 (0.441–0.477) ^a	2787	126	0.536 (0.536–0.536) ^a	17.03
17	1379	60	0.438 (0.421–0.457) ^a	2661	104	0.515 (0.515–0.515) ^a	17.58
18	1319	61	0.418 (0.401–0.436) ^a	2557	144	0.486 (0.486–0.486) ^a	16.27
19	1258	90	0.388 (0.371–0.406) ^a	2413	169	0.452 (0.452–0.452) ^a	16.49

Time (years)	Women (cohort 2010)			Men (cohort 2010)			Probability of staying is higher for men by ... (in %)
	<i>n</i>	No. leaving science	KM probability (staying) with 95% CI and SE	<i>n</i>	No. leaving science	KM probability (staying) with 95% CI and SE	
1	6037	252	0.958 (0.953–0.963) ^a	6755	265	0.961 (0.961–0.961) ^a	0.31
2	5785	431	0.887 (0.879–0.895) ^a	6490	442	0.895 (0.895–0.895) ^a	0.90
3	5354	456	0.811 (0.802–0.821) ^a	6048	452	0.828 (0.828–0.828) ^a	2.10
4	4898	415	0.743 (0.732–0.754) ^a	5596	397	0.770 (0.770–0.770) ^a	3.63
5	4483	351	0.684 (0.673–0.696) ^a	5199	351	0.718 (0.718–0.718) ^a	4.97
6	4132	330	0.630 (0.618–0.642) ^a	4848	354	0.665 (0.665–0.665) ^a	5.56
7	3802	331	0.575 (0.563–0.588) ^a	4494	317	0.618 (0.618–0.618) ^a	7.48
8	3471	305	0.524 (0.512–0.537) ^a	4177	342	0.568 (0.568–0.568) ^a	8.40
9	3166	311	0.473 (0.460–0.486) ^a	3835	338	0.518 (0.518–0.518) ^a	9.51

^aStandard error of 0.01 (total in Table S2).

proportions as men – with the two Kaplan–Meier curves overlapping and no statistically significant differences – in the entire period examined [9]. For a more comprehensive view of gender differences in attrition and retention in the four disciplines, we correlated the data on both cohorts to discuss three other approaches – survival regression curves, hazard rate curves, and kernel density curves – separately for each discipline.

Survival analysis models the time until an event of interest (here, ceasing publishing) occurs. For some proportion of the 2000 and 2010 cohorts, the event occurred. Hazard is the probability of an event occurring during any given time. We calculated the probability that any given scientist would cease publishing each year and plotted the results. The survival

regression curve is actually a smoothed Kaplan–Meier curve, with the share of authors gradually declining over time.

The survival regression curves for the four disciplines for both cohorts (Figs 3 and 4, Panels A; as well as Figs S1 and S2, Panels A) show a steeper decline in early publishing careers. The decline is smoother later on, and divergence between the curves for men and women from the 2000 cohort increases. Top panels refer to cohort 2000 and bottom panels refer to cohort 2010.

The hazard rate curves for both cohorts (Panels B) show that attrition was greater in the early publishing years. Especially, in BIO and NEURO for the 2000 cohort and for the first 15 years, women had greater chances of leaving science than men. The short-term

AGRI

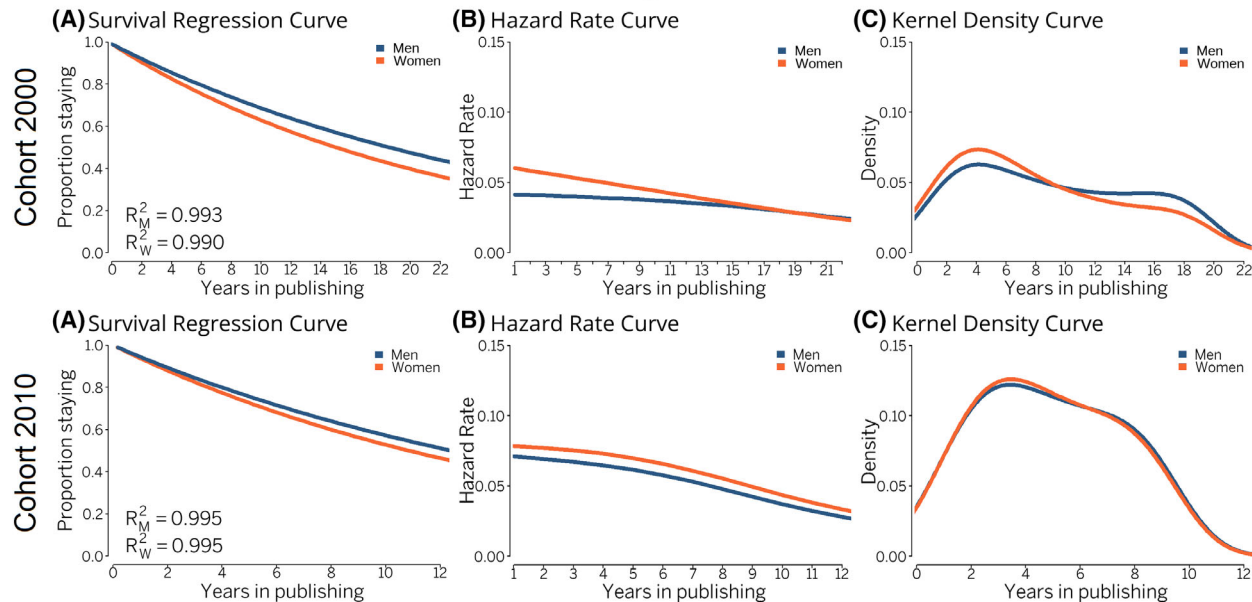


Fig. 3. 2000 (top panel, $N = 7970$) and 2010 (bottom panel, $N = 12\,792$) cohorts of scientists, AGRI, agricultural and biological sciences. Panels (A) Survival regression curve (exponential distribution fitting of Kaplan–Meier curve). Panels (B) Hazard rate curve (B-splines smoothing method, smoothing parameter 10 k). Panels (C) Kernel density curve (B-splines smoothing method, bandwidth 2, component per point based on Gaussian curve).

BIO

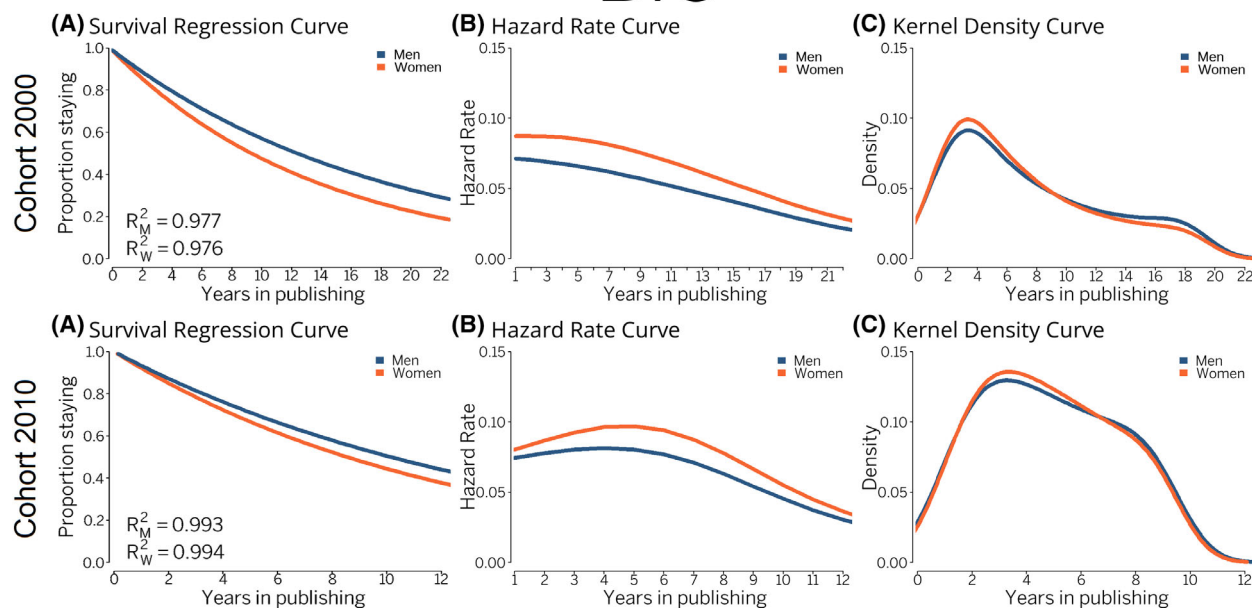


Fig. 4. 2000 (top panel, $N = 22\,692$) and 2010 (bottom panel, $N = 31\,542$) cohorts of scientists, BIO, biochemistry, genetics, and molecular biology. Panels (A) Survival regression curve (exponential distribution fitting of Kaplan–Meier curve). Panels (B) Hazard rate curve (B-splines smoothing method, smoothing parameter 10 k). Panels (C) Kernel density curve (B-splines smoothing method, bandwidth 2, component per point based on Gaussian curve).

risk of ceasing publishing is greater in the early years of publishing careers, after which the risk decreases over time. The riskiest years are always the early years in science. Most importantly, gender differences are visible for all disciplines examined, with the curves for men and women never overlapping (except for AGRI, cohort 2000).

Kernel density estimations for both cohorts (Figs 3 and 4, Panels C) display the distribution of values in a dataset using one continuous curve identifying the shape of the distribution. Kernel density curves show how all men and women who actually left science are distributed over time. The curves are easily interpretable because the area under the curve always reaches 100%. For both cohorts, the shares of scientists who left science were highest for the first 2–4 years (representing the doctoral stage).

The survival regression curves mirror the Kaplan–Meier curves. They highlight consistent and large gender differences in attrition, especially in BIO and NEURO. The differences visible for the older cohort are still visible for the younger cohort. The hazard rate curves show that in two disciplines (NEURO and IMMU), the chances of leaving science were as gendered for the younger cohort as they were for the older one. Finally, kernel density curves show largely the same stories about the four disciplines: for both cohorts, disappearing from science was more intensive in the first 4 years, and much less intensive after 10 years. The distribution patterns of the leavers were largely similar for men and women.

Discussion

The conceptualization of leaving science as quitting academic publishing does not entail any other academic or non-academic roles (teaching and administration, as well as working for governments, NGOs, industry, etc.). In our simplified methodological approach, a sequence of globally indexed publications replaces a sequence of much wider cognitive and social processes encompassing the various dimensions of 'doing science' [44,45]. In it, not publishing in scholarly journals means not doing science anymore – which is a fairly simple approach driven by data availability for a global study like ours.

The metadata about publications at an individual level, with considerable cleaning and computing, are globally (and longitudinally) available for comparative research. In contrast, micro-data about employment in the science sector – first employment, changing sectors – including academia, are not available. Our research would not be possible without access to micro-data

drawn from the raw Scopus dataset and computed using commercial, cloud-based computing infrastructure.

Our study has an OECD focus: all data refer to 38 OECD countries combined. However, it is also possible to compare differences by country. We have published an interactive tool that provides the results of the Kaplan–Meier probability of staying in science by country, discipline, and gender for 11 cohorts (the 2000–2010 cohorts, $N = 2\,127\,803$). The snapshot in Fig. 5 shows the probability of staying in science after 10 years for women from the 2000 cohort for major European countries in biochemistry, genetics, and molecular biology (BIO). It reaches only 30% in Germany, as opposed to the much higher 83% in Portugal and 64% in Poland (see also an interactive presentation in *Nature* news covering this research [46]).

The chances of surviving in biochemistry, genetics, and molecular biology greatly differ across Europe. The various retention rates reflect different national conditions in the science sector changing over time. In some countries, push factors are important in leaving decisions (e.g., the science sector being an unattractive, perhaps unwelcoming workplace for women), while in others, pull factors matter (e.g., outside opportunities are attractive, pulling scientists to industry). Examples include computing in the USA, chemistry and engineering in Germany, with attractive non-academic workplaces – as opposed to almost all disciplines in Poland and Portugal. There, pull factors have traditionally been weak, with outside, non-academic opportunities for scientists being limited.

Exploring attrition and retention in science at a multi-country scale has limitations – but it also creates new opportunities. Disciplines can be examined across many countries, and conclusions can be drawn about cross-disciplinary differences within the STEMM disciplines. The traditional “leaky pipeline” and “chilly climate” metaphors can be tested in detail, using large populations of scientists from different countries and different disciplines.

While regarded as a monolith, there are large variations among STEMM disciplines, especially in their gender composition over time. With global datasets, attrition and retention – coming to science and leaving it – can be explored from a longitudinal perspective in which the same individual scientists are tracked for years and decades, rather than from the perspective of annually changing averages in the science profession by gender in countries and institutions. These are known from traditional science and engineering statistics at an aggregated level, which hide the different career trajectories of men and women.

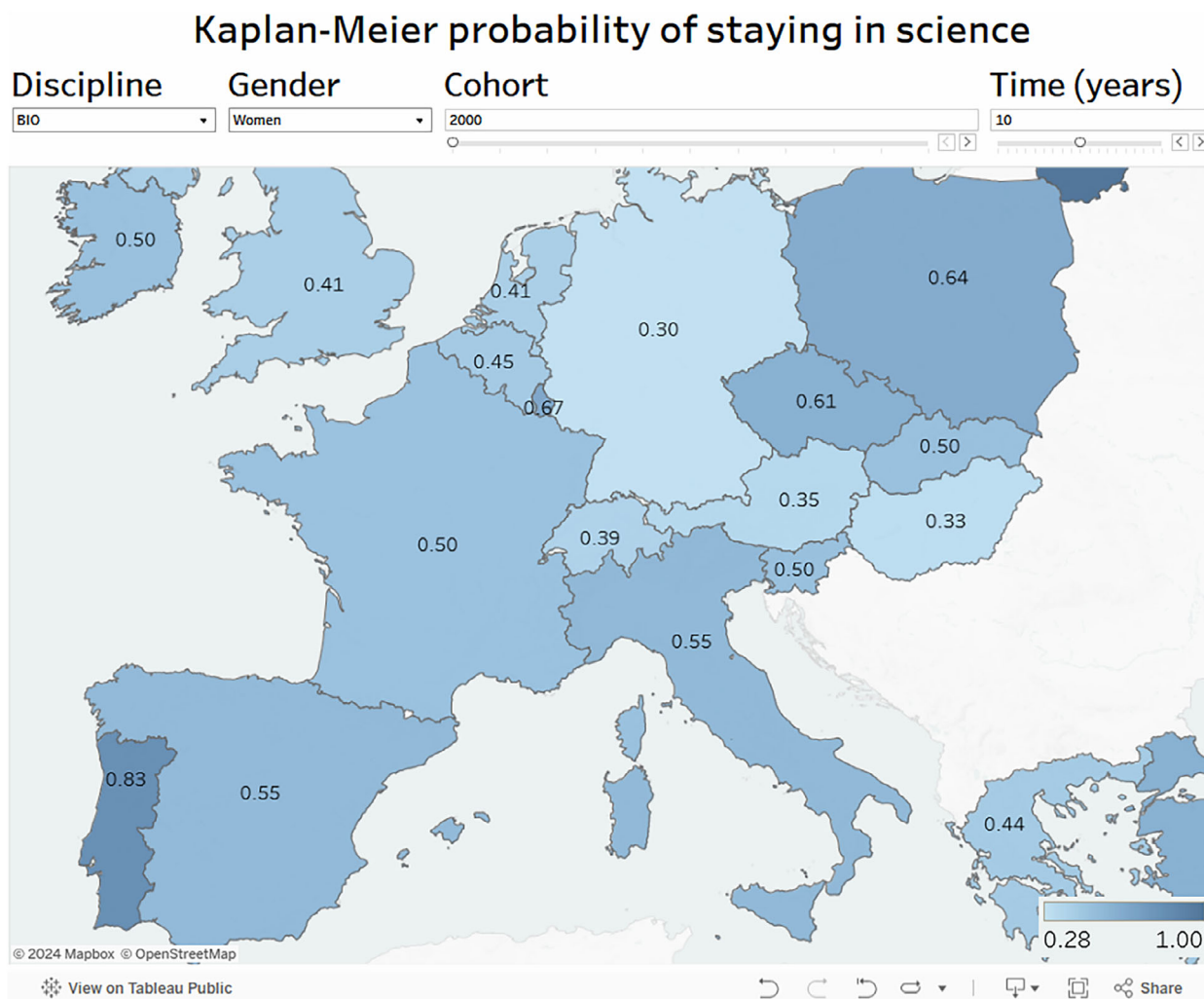


Fig. 5. BIO, biochemistry, genetics, and molecular biology. Women scientists from cohort 2000, probability of staying in science after 10 years. A snapshot of an interactive map (available at <https://public.tableau.com/app/profile/marek.kwiek/viz/Attrition-in-science-OECD/Dashboard>) in which Kaplan–Meier probabilities of remaining in science (i.e., continuing publishing) are provided by country (38 OECD countries), discipline, and gender and for all 11 cohorts from 2000 to 2010 ($N = 2\,127\,803$ scientists in 16 STEMM disciplines, 1 289 756 men and 838 047 women).

Attrition, as it emerged from our results for scientists in AGRI, BIO, IMMU, and NEURO, is a permanent challenge in science. Attrition affects men and women alike – but in these disciplines, with gender balance already achieved (according to the 40/60 formula), attrition affects women much more powerfully than men.

About 60% of women from the 2000 cohort in BIO were still in science after 5 years, 40% after 10 years, and only 20% at the end of the period examined (i.e., after 19 years). In stark contrast, the percentages were substantially higher for men: ~70%, 50%, and 30%, respectively (Table 2). These reflect enormous differences in how, on average, male and female careers

develop over time. Kaplan–Meier estimations show that women in BIO were one quarter more likely to leave science after 10 years (23.26%), and as much as 40% more likely to leave science by the end of the study period (39.74%). In the four disciplines examined here, gender differences in attrition and retention have been and continue to be high and are only slightly decreasing over time. The data throw new light on careers in BIO (and related fields) in the 38 OECD countries combined.

Our online interactive tool (Fig. 5) reports data about retention in BIO for the USA (the largest OECD science system): 60% vs. 70% after 5 years, 40% vs. 51% after 10 years, and 20% vs. 30% at the

end of the study period for women vs. men, respectively. The Kaplan–Meier probability of surviving 19 years in BIO in the USA was 50% higher for men (61.11% higher for men in IMMUNO, 52% in NEURO, but only 20.59% in AGRI – for which the highest retention rates were also observable for each year).

It is intriguing that in the disciplines where gender balance has been achieved (as in the four studied here), for both older and younger cohorts of scientists, the chances of women surviving were consistently lower than for men. However, in disciplines where there is a long way to go to achieve equal representation of men and women (as in mathematics, computing, or engineering), for both cohorts, women's chances of surviving were equal to those of men, with no statistically significant differences at all.

Consequently, increasing the representation of women in science does not automatically guarantee increasingly equal attrition rates for men and women. In short, the problem for women in male-dominated disciplines is how to get there (conceptualized with our dataset as how to start publishing). The problem for women in the biology-related disciplines examined here, with equal gender representation, is not so much getting there—it is actually staying there (i.e., continuing to publish) for as long as men do.

We are currently studying the why question in attrition and retention: what makes men and women leave – implementing a global survey of 100 000 carefully selected “survivors in science” or scientists with 25–35 years of publishing careers across STEM and social sciences.

Acknowledgements

MK gratefully acknowledges the support provided by NDS grant no. Nds-II/SP/0010/2023/01. LS is grateful for the support of his doctoral studies provided by NCN grant 2019/35/0/HS6/02591. We gratefully acknowledge the assistance of the International Center for the Studies of Research (ICSR) Lab with Kristy James and Alick Bird. We also want to thank Dr Wojciech Roszka from the CPPS Poznan Team for fruitful discussions.

Conflict of interest

The authors declare no competing interests.

Author contributions

MK contributed to conceptualization, data curation, formal analysis, investigation, methodology, resources,

software, validation, writing—original draft, writing—review and editing. LS contributed to conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, writing—original draft, writing—review and editing.

Data accessibility

We used data from Scopus, a proprietary scientometric database. For legal reasons, data from Scopus received through collaboration with the ICSR Lab owned by Elsevier cannot be made openly available.

References

- 1 Glenn ND (2005) Cohort Analysis. Sage, London.
- 2 Singer JD and Willett JB (2003) Applied Longitudinal Data Analysis. Modeling Change and Event Occurrence. Oxford University Press, Oxford.
- 3 Rosser VJ (2004) Faculty members' intentions to leave: a national study on their work-life and satisfaction. *Res High Educ* **45**, 285–309.
- 4 Xu YJ (2008) Gender disparity in STEM disciplines: a study of faculty attrition and turnover intentions. *Res High Educ* **49**, 607–624.
- 5 Preston AE (2004) Leaving Science: Occupational Exit from Scientific Careers. Russell Sage Foundation, New York, NY.
- 6 Geuna A and Shibayama S (2015) Moving out of academic research: why do scientists stop doing research? In Global Mobility of Research Scientists (Geuna A, ed.), pp. 271–297. Elsevier, Amsterdam.
- 7 Kaminski D and Geisler C (2012) Survival analysis of faculty retention in science and engineering by gender. *Science* **335**, 864–866.
- 8 White-Lewis DK, O'Meara K, Mathews K and Havey N (2023) Leaving the institution or leaving the academy? Analyzing the factors that faculty weigh in actual departure decisions. *Res High Educ* **64**, 473–494.
- 9 Kwiek M and Szymula L (2024) Quantifying attrition in science: a cohort-based, longitudinal study of scientists in 38 OECD countries. *Higher Educ* 1–29. doi: [10.1007/s10734-024-01284-0](https://doi.org/10.1007/s10734-024-01284-0)
- 10 Hermanowicz J (2012) The sociology of academic careers: problems and prospects. In Higher Education: Handbook of Theory and Research (Smart JC and Paulsen MB, eds), Vol. 27, pp. 207–248. Springer, Dordrecht.
- 11 Leisyte L and Dee JR (2012) Understanding academic work in a changing institutional environment. In Higher Education: Handbook of Theory and Research (Smart JC and Paulsen MB, eds), Vol. 27, pp. 123–206. Springer Netherlands, Dordrecht.

- 12 Kashyap R, Rinderknecht RG, Akbaritabar A, Albrez-Gutierrez D, Gil-Clavel S, Grow A and Zhao X (2023) Digital and computational demography. In *Research Handbook on Digital Sociology*, pp. 48–86. Edward Elgar Publishing, Cheltenham.
- 13 Kwiek M and Roszka W (2024) Are scientists changing their research productivity classes when they move up the academic ladder?. *Innovative Higher Education*, 1–39. doi: [10.1007/s10755-024-09735-3](https://doi.org/10.1007/s10755-024-09735-3)
- 14 Wang D and Barabási AL (2021) *The Science of Science*. Cambridge University Press, Cambridge.
- 15 Liu L, Jones BF, Uzzi B and Wang D (2023) Data, measurement, and empirical methods in the science of science. *Nat Hum Behav* **7**, 1046–1058.
- 16 Salganik MJ (2018) *Bit by Bit: Social Research in the Digital Age*. Princeton University Press, Princeton, NJ.
- 17 Karimi F, Wagner C, Lemmerich F, Jadidi M and Strohmaier M (2016) Inferring gender from names on the web: a comparative evaluation of gender detection methods. In *Proceedings of the 25th International Conference Companion on World Wide Web*, pp. 53–54.
- 18 Sebo P (2021) Performance of gender detection tools: a comparative study of name-to-gender inference services. *J Med Libr Assoc* **109**, 414–421.
- 19 Kwiek M and Szymula L (2023) Young male and female scientists: a quantitative exploratory study of the changing demographics of the global scientific workforce. *Quant Sci Stud* **4**, 902–937.
- 20 Naddaf M (2024, October 3) Nearly 50% of researchers quit science within a decade, huge study reveals. *Nature News*.
- 21 Baas J, Schotten M, Plume A, Côté G and Karimi R (2020) Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quant Sci Stud* **1**, 377–386. doi: [10.1162/qss_a_00019](https://doi.org/10.1162/qss_a_00019)
- 22 Blickenstaff JC (2005) Women and science careers: leaky pipeline or gender filter? *Gend Educ* **17**, 369–386.
- 23 Shaw AK and Stanton DE (2012) Leaks in the pipeline: separating demographic inertia from ongoing gender differences in academia. *Proc Biol Sci* **279**, 3736–3741.
- 24 Cornelius R, Constantinople A and Gray J (1988) The chilly climate: fact or artifact? *J Higher Educ* **59**, 527–555.
- 25 Fox MF and Kline K (2016) Women faculty in computing. A key case of women in science. In *Pathways, Potholes, and the Persistence of Women in Science: Reconsidering the Pipeline* (Branch EH, ed.), pp. 54–69. Lexington Books, Lanham, MD.
- 26 Fox MF (2020) Gender, science, and academic rank: key issues and approaches. *Quant Sci Stud* **1**, 1001–1006.
- 27 Xie Y and Shauman KA (2003) *Women in Science: Career Processes and Outcomes*. Harvard University Press, Cambridge, MA.
- 28 Branch EH and Alegria S (2016) Gendered responses to failure in undergraduate computing. Evidence, contradictions, and new directions. In *Pathways, Potholes, and the Persistence of Women in Science. Reconsidering the Pipeline* (Branch EH, ed.), pp. 17–31. Lexington Books, Lanham, MD.
- 29 Branch EH, ed. (2016) *Pathways, Potholes, and the Persistence of Women in Science: Reconsidering the Pipeline*. Lexington Books, Lanham, MD.
- 30 Kwiek M and Roszka W (2021) Gender disparities in international research collaboration: a large-scale bibliometric study of 25,000 university professors. *J Econ Surv* **35**, 1344–1388.
- 31 Kwiek M and Roszka W (2021) Gender-based homophily in research: a large-scale study of man-woman collaboration. *J Informetr* **15**, 1–38.
- 32 Matier MW (1990) Retaining faculty: a tale of two campuses. *Res High Educ* **31**, 39–61.
- 33 Nair S, Mee LE and Cheik AN (2016) Internal push factors and external pull factors and their relationship with lecturer's turnover intention. *Int J Bus Manag* **11**, 110–126.
- 34 Kallio T, Siekkinen T, Pekkola E, Kivistö J, Nokkala T and Kuoppakangas P (2024) Push and pull factors affecting in leaving academia. *Tertiary Educ Manag* **30**, 65–80.
- 35 O'Meara K, Bennett JC and Neihaus E (2016) Left unsaid: the role of work expectations and psychological contracts in faculty careers and departure. *Rev Higher Educ* **39**, 269–297.
- 36 Smart JC (1990) A causal model of faculty turnover intentions. *Res High Educ* **31**, 405–424.
- 37 Zhou Y and Volkwein JF (2004) Examining the influence on faculty departure intentions: a comparison of tenured versus nontenured faculty at research universities using NSOPF-99. *Res High Educ* **45**, 139–176.
- 38 Spoon K, LaBerge N, Wapman KH, Zhang S, Morgan AC, Galesic M, Fosdick BK, Larremore DB and Clauset A (2023) Gender and retention patterns among U.S. faculty. *Sci Adv* **9**, eadi2205.
- 39 Milojevic S, Radicchi F and Walsh JP (2018) Changing demographics of scientific careers: the rise of the temporary workforce. *Proc Natl Acad Sci USA* **115**, 12616–12623.
- 40 Kwiek M and Roszka W (2022) Academic vs. biological age in research on academic careers: a large-scale study with implications for scientifically developing systems. *Scientometrics* **127**, 3543–3575.
- 41 Mills M (2011) *Introducing Survival and Event History Analysis*. Sage, London.
- 42 Kwiek M and Roszka W (2022) Are female scientists less inclined to publish alone? The gender solo research gap. *Scientometrics* **127**, 1697–1735.
- 43 NamSor (2024) *NamSor API documentation*. <https://namsor.app/api-documentation/>

- 44 Sugimoto C and Larivière V (2023) *Equity for Women in Science: Dismantling Systemic Barriers to Advancement*. Harvard University Press, Cambridge, MA.
- 45 Kwiek M (2019) *Changing European Academics: A Comparative Study of Social Stratification, Work Patterns and Research Productivity*. Routledge, London.
- 46 Naddaf M (2025, January 9) Who's quitting academia? Data reveal gender gaps in surprising fields. *Nature News*.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. 2000 ($N = 1775$) and 2010 ($N = 2361$) cohorts of scientists, IMM, immunology and microbiology.

Fig. S2. 2000 ($N = 2533$) and 2010 ($N = 4513$) cohorts of scientists, NEURO, neuroscience.

Table S1. 2000 and 2010 cohorts of scientists, BIO, biochemistry, genetics, and molecular biology.

Table S2. 2000 and 2010 cohorts of scientists, AGRI, agricultural and biological sciences.

Table S3. 2000 and 2010 cohorts of scientists, IMM, immunology and microbiology.

Table S4. 2010 and 2010 cohorts of scientists, NEURO, neuroscience.