

MAREK KWIEK\*, ŁUKASZ SZYMULA\*\*

## Znikający naukowcy. Co ustrukturyzowane Big Data mówią nam o rezygnacji z nauki w 38 krajach OECD?

### 1. Wprowadzenie

W pracy analizujemy zjawisko rezygnacji z nauki akademickiej i pokazujemy, jak odchodzenie z niej różni się między kobietami i mężczyznami, dyscyplinami akademickimi i na przestrzeni czasu. Prezentowane podejście jest kompleksowe: globalne, oparte na kohortach naukowców i podłużne – obserwujemy działalność publikacyjną indywidualnych naukowców w czasie i kwantyfikujemy zjawisko tradycyjnie określane mianem porzucania nauki (Geuna & Shibayama, 2015; Preston, 2004; White-Lewis et al., 2023; Zhou & Volkwein, 2004). Rezygnację z nauki konceptualizujemy jako trwałe zaprzestanie publikowania naukowego, ponieważ dane podłużne dotyczące odchodzenia z akademii w postaci rezygnacji z zatrudnienia (co byłoby prostsze) nie są dostępne na poziomie globalnym.

Korzystając z metadanych pochodzących z bazy Scopus – globalnej bibliometrycznej bazy danych publikacji i cytowań – analizujemy kariery publikacyjne 142 776 naukowców z 38 krajów OECD, którzy zaczęli publikować w 2000 r. oraz 232 843 naukowców, którzy zaczęli publikować w 2010 r. (zwanych dalej odpowiednio kohortą 2000 i kohortą 2010). Nasze badanie ogranicza się do 16 dyscyplin STEMM (nauki ścisłe, techniczne, inżynieryjne, matematyczne i medyczne) i umożliwia prześledzenie indywidualnego dorobku naukowego obu kohort naukowców do 2022 r.

Metadane bibliometryczne stanowią doskonały przykład śladów cyfrowych pozostawianych przez naukowców w publikacjach i wykorzystywanych do badania karier akademickich, które tradycyjnie były poddawane analizie z pomocą badań ankietowych, wywiadów pogłębionych czy rejestrowych danych administracyjnych (pochodzących z krajowych rejestrów naukowców, szeroko przez nas wykorzystywanych w ramach Laboratorium Polskiej Nauki prowadzonego na UAM).

---

\* Prof. dr hab. Marek Kwiek (marek.kwiek@amu.edu.pl), Institute for Advanced Studies in Social Sciences and Humanities (IAS), Uniwersytet im. Adama Mickiewicza w Poznaniu, ORCID: [orcid.org/0000-0001-7953-1063](https://orcid.org/0000-0001-7953-1063)

\*\* Łukasz Szymula (lukasz.szymula@amu.edu.pl), Wydział Matematyki i Informatyki, Uniwersytet im. Adama Mickiewicza w Poznaniu, ORCID: [orcid.org/0000-0001-8714-096X](https://orcid.org/0000-0001-8714-096X)

Cyfrowe naukowe bazy danych zawierające informacje bibliometryczne stanowią nowe źródło umożliwiające badanie naukowców jako populacji, chociaż wymagają odpowiedniej wstępnej modyfikacji umożliwiającej koncentrację na poszczególnych naukowcach, a nie na poszczególnych publikacjach. Ślady cyfrowe umożliwiają badanie zagadnień dotyczących nauki i naukowców na niespotykanym dotąd poziomie szczegółowości (Kashyap i in., 2022; Liu i in., 2023; Wang i Barabási, 2021). Poszczególnych naukowców można badać ze względu na wiek, długość stażu pracy, płeć, dyscyplinę naukową czy też typ instytucji – a co najważniejsze dla celów niniejszej analizy, naukowców można również obserwować w czasie (co otwiera możliwości badań podłużnych, longitudinalnych). Badania przekrojowe można zatem uzupełniać badaniami podłużnymi, w których indywidualne kariery akademickie są obserwowane przez lata, a nawet dziesięciolecia. Kariery akademickie coraz częściej bada się na poziomie globalnym i krajowym przy użyciu dużych zbiorów danych (np. Nielsen & Andersen, 2021; King et al., 2017; Boekhout et al., 2021; Nygaard et al., 2022; Spoon et al., 2023; Kwiek i Szymula 2023; Kwiek i Szymula 2024). Również nasze prace dotyczące karier polskich naukowców korzystały z metadanych miliona publikacji indeksowanych w bazie Scopus w ostatnim ćwierćwieczu (zob. Kwiek i Roszka 2023; Kwiek i Roszka 2024).

Tradycyjnie w socjologii karier akademickich uznawano, że kobiety naukowcy zazwyczaj rezygnują z pracy wcześniej niż mężczyźni. Ponadto sądzono, że kobiety rezygnują z pracy w większych odsetkach niż mężczyźni, co potwierdzają tradycyjne teorie dotyczące rezygnacji z pracy akademickiej (Alper, 1993; Blickenstaff, 2005; Deutsch & Yao, 2014; Goulden et al., 2011; Preston, 2004; Shaw & Stanton, 2012).

W naszym badaniu wychodzimy poza tradycyjne ujęcia przekrojowe (bazujące na pojedynczych lub powtarzanych punktach w czasie) i wykorzystujemy globalne dane podłużne (bazujące na wielu punktach w czasie) na mikropoziomie poszczególnych naukowców. Sformułowania „globalny” używamy w tej pracy, odnosząc się do 38 krajów OECD, najważniejszych (oprócz Chin) producentów wiedzy, w których publikuje znacząca część naukowców w świecie. Przy zaproponowanym podejściu zastosowanie naszych metod do wszystkich krajów świata nie stanowi większego problemu.

Ponadto testujemy nowe możliwości, jakie otwierają globalne zbiory danych bibliometrycznych w zakresie kwantyfikacji na dużą skalę rezygnacji z kariery naukowej. Wiemy, że naukowcy rezygnują z karier naukowych – ale skala tych rezygnacji pozostaje niemal całkowicie nieznana, zarówno w szerszym ujęciu geograficznym, jak i w szerszym ujęciu czasowym. Nie znamy skali tego zjawiska poza Stanami Zjednoczonymi – i nie wiemy, jak zmienia się ono w czasie. Nasz tekst stanowi pierwszą próbę szerokiej kwantyfikacji tej problematyki w ramach bardziej ogólnej idei, że tradycyjne naukowstwo nie docenia podstawowej roli śladów cyfrowych pozostawianych przez naukowców w ich globalnie indeksowanych publikacjach w analizie ewolucji nauki.

## 2. Kontekst teoretyczny

### Podejście do nauki i naukowców oparte na dużych zbiorach danych

Zjawisko rezygnacji z nauki można analizować w oparciu o dane jednostkowe nie tylko w odniesieniu do poszczególnych instytucji i krajów, ale także w ujęciu globalnym, wykorzystując w tym celu duże zbiory danych (ustrukturyzowane, komercyjne Big Data). Globalne i podłużne ujęcia kariery akademickiej stały się możliwe dopiero niedawno dzięki rosnącemu dostępowi do cyfrowych baz danych zawierających kompleksowe informacje o naukowcach, ich dorobku i wpływie na naukę opartym na cytowaniach (Kashyap i in., 2022; Wang i Barabási, 2021). Pojawienie się nowych cyfrowych zbiorów danych, dostęp do ogromnych komputerowych mocy obliczeniowych oraz bardziej ogólny zwrot w stronę ustrukturyzowanych dużych zbiorów danych w badaniach społecznych doprowadziły do gwałtownego wzrostu liczby publikacji dotyczących różnych aspektów kariery akademickiej. Pojawiła się też seria badań koncentrujących się na różnicach między mężczyznami i kobietami w nauce (np. King i in., 2017; Nielsen i Andersen, 2021; Sugimoto i Larivière, 2023, a w odniesieniu do polskiego systemu nauki zob. monografię Kwiek 2022 oraz Kwiek i Roszka 2021a; Kwiek i Roszka 2021b).

Duże zbiory danych zapewniają wyjątkową możliwość testowania tradycyjnych ram pojęciowych dotyczących nauki i naukowców (Liu i in., 2023). Dane cyfrowe, które pozwalają na kompleksowe monitorowanie nauki globalnej, mogą być wykorzystywane do badania jej wewnętrznego funkcjonowania na niezwykłym poziomie szczegółowości i w unikalnej skali (Wang & Barabási, 2021). Możliwe stało się systematyczne badanie historii karier zawodowych setek tysięcy indywidualnych naukowców. Ogromne komercyjne i niekomercyjne zbiory danych są dziś dostępne niemal na wyciągnięcie ręki badaczy nauki, choć nie bez istotnych ograniczeń instytucjonalnych i finansowych (np. Boekhout i in., 2021; Liu i in., 2023; Sugimoto i Larivière, 2023).

Z tego powodu nowe naukoznawstwo (zwane *science of science* czy też *quantitative science studies*) staje w obliczu niespotykanych dotąd możliwości, również w kontekście szerszego dostępu do otwartych zbiorów danych typu *OpenAlex*. Jak się jednak wydaje, tradycyjna socjologia nauki i socjologia karier akademickich stają w obliczu tych nowych możliwości nieco onieśmiałe, oddając pole niemal wyłącznie informatykom. To duża strata, którą od kilku lat staramy się wyrównywać, systematycznie prowadząc badania globalne i konfrontując je z badaniami krajowymi – jedno i drugie z wykorzystaniem Big Data o bibliometrycznym pochodzeniu.

### Rezygnacja z nauki

Młodzi naukowcy napotykają różne bariery związane z podejmowaniem i kontynuowaniem kariery akademickiej (Preston, 2004; Wohrer, 2014); jednak rezygnację z kariery

ry naukowej tradycyjnie badano jako zjawisko charakterystyczne dla kobiet: to kobiety spotykały się z „chłodną” kulturą akademickiego miejsca pracy, z trudnościami z utrzymaniem równowagi między życiem zawodowym a życiem prywatnym oraz z trudnościami związanymi z przetrwaniem okresu macierzyństwa podczas pracy w środowisku akademickim (Cornelius et al., 1988; Goulden et al., 2011; Levine et al., 2011; Maranto & Griffin, 2011; White-Lewis et al., 2023; Wolfinger et al., 2008).

Najważniejszymi wymiarami powyższych barier były awanse akademickie, produktywność badawcza, wpływ na naukę, dostęp do grantów badawczych, nagród, a także zdobywanie uznania za osiągnięcia naukowe. Badania pokazują, że kobiety są niewystarczająco reprezentowane na wyższych stanowiskach akademickich; rzadziej współpracują na arenie międzynarodowej w zakresie badań, rzadziej publikują w czasopiśmie o dużym wpływie i jednocześnie są mniej cytowane w porównaniu z mężczyznami. Ponadto kobiety częściej doświadczają dłuższych przerw w karierze naukowej, a ich wnioski grantowe są częściej odrzucane w krajowych instytucjach finansujących badania (Fochler i in., 2016; Hammarfeld, 2017; Kwiek & Roszka, 2021a, 2021b; Lindahl, 2018; Shibayama & Baba, 2015; Sugimoto & Larivière, 2023; Tang & Horta, 2023). W większości dyscyplin STEMM kobiety podejmują pracę akademicką w otoczeniu zdominowanym przez mężczyzn, w którym mogą doświadczać tradycyjnie diagnozowanego i teoretycznie badanego „chłodnego klimatu” (*chilly climate*) wobec kobiet (Santos i in., 2020; globalny rozkład mężczyzn i kobiet w nauce i jego ewolucję w ostatnim trzydziestolecu analizujemy w Kwiek i Szymula 2023).

Chociaż zarówno mężczyźni, jak i kobiety nadszpiewanie często rezygnują z nauki akademickiej – jedna trzecia wszystkich naukowców odchodzi z nauki w ciągu pierwszych pięciu lat, a połowa w ciągu dekady (jak pokażemy później) – uważa się, że wskaźnik rezygnacji jest wyższy w przypadku kobiet niż mężczyzn (Preston, 2004; Kaminski & Geisler, 2012). Hipotezy „nieszczelnego rurociągu” (*leaky pipeline*) i „chłodnego klimatu” wyjaśniają tę różnicę w przypadku dyscyplin STEM: zarówno na etapie studiów, jak i później na każdym etapie kariery akademickiej dochodzi do odpływu talentów z powodu systemowych barier utrudniających kobietom wykonywanie pracy naukowej (Blickenstaff, 2005; Goulden et al., 2011; Shaw & Stanton, 2012; Wolfinger et al., 2008), a nieprzyjazne środowisko pracy może je zniechęcać do kontynuowania kariery akademickiej (Cornelius et al., 1988; Spoon et al., 2023).

W modelu „nieszczelnego rurociągu” w ramach kariery akademickiej naukowcy albo pokonują szereg etapów jej rozwoju, albo ostatecznie opuszczają środowisko akademickie. „Chłodny klimat”, z którym mogą spotykać się w dyscyplinach STEM („wykluczenie”, „poczucie braku przynależności” etc.), ma swoje podłoże w relatywnej demografii (czyli niskim odsetku kobiet pracujących w ramach wybranych dyscyplin). Co ważne z perspektywy polityki naukowej, koncepcja „chłodnego klimatu” przełożyła się

na długofalowe, intensywne wysiłki na rzecz promowania równości szans kobiet i mężczyzn na kampusach w Stanach Zjednoczonych i w całej Europie (Britton, 2017). Jak to zwięźle ujęto trzy dekady temu w *Science*: „kultura nauki nie przyciąga kobiet, które w innym przypadku mogłyby zostać utalentowanymi naukowcami” (Alper, 1993: 409). Pokazujemy w tej pracy, że rezygnacja z nauki z każdą kolejną badaną przez nas kohortą naukowców (a przebadaliśmy ich jedenaście) staje się problemem wszystkich naukowców, a nie tylko problemem przede wszystkim kobiet. A w centrum tego fenomenu staje kwestia atrakcyjności kariery akademickiej w szerokim sensie, na którą silnie oddziałują zarówno czynniki wewnętrzne, jak i czynniki zewnętrzne, na przykład atrakcyjność sektorów pozaakademickich jako miejsca pracy (zwane *push and pull factors*).

### **Porzucanie nauki jako motyw badawczy**

Zagadnienie rezygnacji z nauki nie było dotąd kompleksowo badane na poziomie globalnym. Tradycyjnie analizowano je albo poprzez studia przypadków na małą skalę (głównie za pomocą ankiet i wywiadów), albo poprzez wieloletnie badania wydziałów uczelni wyższych w USA (np. Rosser, 2004; Xu, 2008; Zhou & Volkwein, 2004). Ostatnio White-Lewis i współpracownicy (2023) przeanalizowali rzeczywiste decyzje dotyczące rezygnacji z pracy akademickiej 2289 amerykańskich wykładowców, którzy opuścili swoje instytucje w latach 2015–2019. Ustalono, że kobiety odchodzą ze środowiska akademickiego częściej niż mężczyźni na każdym etapie kariery zawodowej (Spoon i in., 2023). Zdaniem ankietowanych kobiet przy odchodzeniu ze środowiska akademickiego atmosfera panująca w miejscu pracy ma większe znaczenie niż równowaga między życiem zawodowym i prywatnym.

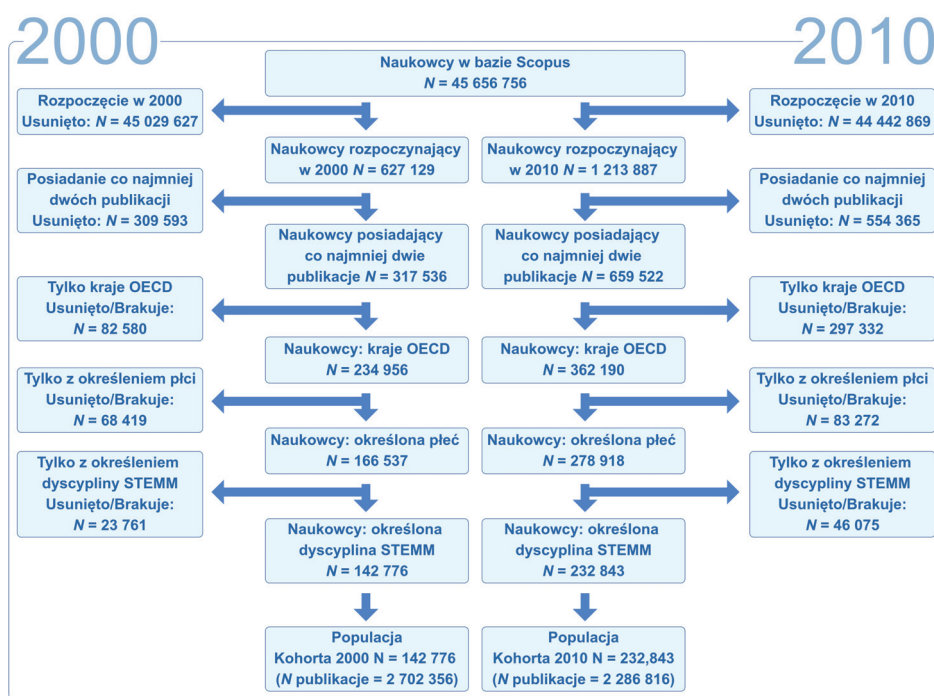
Rezygnacja z kariery akademickiej była dotąd badana przy użyciu takich pojęć, jak: „chęć odejścia z uczelni” (Zhou & Volkwein, 2004), „zamiar odejścia” (Rosser, 2004) czy „rotacja kadry akademickiej” (Ehrenberg et al., 1991; Smart, 1990; Xu, 2008). Większość badań koncentrowała się na jednej instytucji, a ich zasięg geograficzny był ograniczony do Stanów Zjednoczonych (np. Minotte i Pedersen, 2021; Levine i in., 2011).

Między odchodzeniem z instytucji i odchodzeniem ze środowiska akademickiego a naszą konceptualizacją, w ramach której „rezygnacja z nauki” jest analizowana jako zaprzestanie publikowania w czasopiśmie akademickim, istnieją istotne różnice pojęciowe. Koncentracja na publikowaniu na przestrzeni lat aż do zdarzenia w analizie przeżycia, w którym ostatecznie dochodzi do „niepublikowania”, wykracza poza badane dotąd instytucje i sektory uprawiania nauki i prowadzi do bardziej ogólnego poziomu porównań: naukowcy publikujący vs. naukowcy niepublikujący, niezależnie od sektora instytucjonalnego.

### 3. Dane i metody

#### Zbiór danych: co wiemy na temat indywidualnych naukowców na podstawie metadanych dotyczących ich publikacji?

Prezentowane badanie wykorzystuje surowe metadane bibliometryczne dotyczące publikacji i cytowań naukowców, którzy zaczęli publikować w bazie danych Scopus po raz pierwszy w 2000 i 2010 r. (ryc. 1) (a także, dodatkowo, we wszystkich latach między tymi dwoma datami, co w sumie dało nam ogląd 2,1 mln naukowców w ramach 11 kohort). Baza Scopus to największa globalna baza abstraktów i cytowań recenzowanej literatury naukowej, która szczególnie nadaje się do analiz globalnych na mikropoziomie poszczególnych naukowców, ponieważ jest zorganizowana wokół identyfikatorów autorów (*Scopus Author IDs*), oprócz organizacji wokół czasopism, publikacji i ich metadanych (Baas i in., 2020).



Ryc. 1. Schemat przygotowania danych i definiowania populacji: kolejne kroki definiujące kohorty naukowców z 2000 r. (po lewej stronie) i 2010 r. (po prawej stronie). Kroki podejmowane wobec naukowców z obydwu kohort były następujące: wybór naukowców z co najmniej dwiema publikacjami, wybór kraju przynależności jako kraju OECD, wybór płci (binarny: mężczyzna lub kobieta) oraz wybór dyscypliny naukowej jako dyscypliny STEM

Jakimi danymi użytecznymi do analizy rezygnacji z nauki dysponujemy na poziomie każdego naukowca, głównie danymi przez nas wyliczonymi?

Tabela 1. Organizacja mikro danych dla dwóch kohort: przykład danych na mikropoziomie (demograficzne, instytucjonalne i dotyczące wzorców publikowania) dla wybranych naukowców z kohorty 2000 (panel 1) i kohorty 2010 (panel 2),  $N = 375\,619$ . Dla każdego naukowca objętego badaniem dysponujemy szeregami danych i wskaźników obliczanych przy użyciu danych pochodzących ze zbioru typu bibliometrycznego (dane surowe z bazy Scopus) (\* = obserwacje prawostronnie ucięte)

Numer identyfikacyjny naukowca (ID)	Płeć	Dyscyplina	Kraj afiliacji	Typ instytucjonalny	Rok wejścia do nauki (rok pierwszej publikacji)	Rok wyjścia z nauki (rok ostatniej publikacji plus 1)	Wskaźnik współpracy międzyrodzajowej (cały dorobek)	Średnia ranga procentylowa czasopisma (skala: 1-99) (cały dorobek) (%)	Mediana wielkości zespołu (cały dorobek)	FWCI 4y – czteroletni wskaźnik wpływu w postaci cytowań znormalizowanych do dyscypliny	Liczba artykułów naukowych (cały dorobek)
<b>Panel 1: Naukowcy – kohorta 2000 (N = 142 776)</b>											
ID 1	Kobieta	MED	Hiszpania	Pozostałe	2000	2020*	60,26	31,24	6,5	0,81	78
ID 2	Mężczyzna	COMP	USA	TOP200	2000	2004	40,00	99,00	4	4,95	10
ID 3	Kobieta	AGRI	Francja	Pozostał	2000	2008	21,43	68,15	4	0,88	14
ID 4	Mężczyzna	PHYS	Japonia	TOP200	2000	2013	0,00	90,00	5	1,37	3
ID 5	Kobieta	CHEM	Dania	Pozostałe	2000	2001	75,00	1,00	3	1,19	4
...											
ID 142776	Mężczyzna	MED	Niemcy	Pozostałe	2000	2017	26,67	72,60	3	2,05	30
<b>Panel 2: Naukowcy – kohorta 2010 (N = 232 843)</b>											
ID 142777	Mężczyzna	ENER	W. Brytania	TOP200	2010	2012	33,33	98,00	5	1,15	6
ID 142778	Kobieta	IMMU	Szwajcaria	TOP200	2010	2020*	27,27	82,10	5	0,78	11
ID 142779	Kobieta	BIO	Belgia	Pozostałe	2010	2017	100,00	29,50	4	0,10	2
ID 142780	Mężczyzna	ENG	Kanada	Pozostałe	2010	2014	14,29	31,43	2,5	2,04	7
ID 142781	Mężczyzna	MED	Włochy	Pozostałe	2010	2012	100,00	14,00	10	0,13	3
...											
ID 375619	Kobieta	AGRI	Australia	TOP200	2010	2015	0,00	91,08	5	1,93	9

Tabela 2. Zmienne wykorzystane do analizy

No.	Zmienna	Opis
1.	Płeć	Dane dotyczące płci (binarne: kobieta/mężczyzna) dostarczone przez ICSR Lab firmy Elsevier. Zmienna klasyfikowana na podstawie imienia, nazwiska i kraju dominującego z pierwszego roku publikacji przy użyciu narzędzia Nam-sor. Płeć akceptowana z prawdopodobieństwem $\geq 0,85$ .
2.	Dyscyplina	Dominująca dyscyplina wyliczona na podstawie wartości modalnej uzyskanej ze wszystkich dyscyplin przypisanych do czasopism wszystkich cytowanych odniesień bibliograficznych ( <i>cited references</i> ) we wszystkich artykułach znajdujących się w portfolio publikacyjnym każdego naukowca.
3.	Kraj afiliacji	Dominujący kraj, wynik bazujący na wartości modalnej uzyskanej ze wszystkich krajów wskazanych w portfolio publikacyjnym autora.
4.	Afiliacja instytucjonalna TOP200	Wartość binarna wskazująca przynależność (prawda/fałsz) do jednej z 200 wiodących instytucji na świecie. Lista instytucji została uszeregowana na podstawie łącznej produkcji naukowej (kategoria: artykuły naukowe i prace w tomach konferencyjnych) instytucji w latach 2019–2022. Każdy badacz został przypisany do jednej instytucji jako dominującej na podstawie wartości modalnej spośród wszystkich instytucji wskazanych w jego portfolio publikacyjnym.
5.	Rok rozpoczęcia kariery naukowej	Pierwszy rok publikacji (kategoria: dowolny typ) w całym portfolio publikacyjnym naukowca.
6.	Minimum publikacyjne	Wartość binarna wskazująca posiadanie (prawda/fałsz) co najmniej dwóch publikacji (kategoria: artykuły naukowe i prace w tomach konferencyjnych) w portfolio publikacyjnym.
7.	Średnia ranga percentylowa czasopism (całego dorobku)	Średnia rang percentylowych czasopism przypisanych do wszystkich publikacji (kategoria: artykuły naukowe i prace w tomach konferencyjnych) w portfolio publikacyjnym autora. Zakres: 1–99. Wartość rangi percentylowej została uzyskana z metryki 2022 Journal CiteScore dla dyscypliny z najwyższą wartością percentylową.
8.	Wskaźnik współpracy międzynarodowej (całego dorobku)	Udział międzynarodowych publikacji zespołowych autora w jego wszystkich publikacjach zespołowych (wykluczono publikacje samodzielne). Aby publikacja została uznana za publikację zespołową, liczba wszystkich autorów w artykule musi być większa lub równa dwa. Aby publikacja została uznana za międzynarodową, liczba krajów afiliacji w artykule musi być większa lub równa dwa.
9.	4-letni wskaźnik wpływu w postaci cytowań znormalizowany do dyscypliny (FWCI 4y)	Średnia wartość metryk FWCI 4y przypisanych do każdej publikacji w portfolio publikacyjnym autora. Wartość metryki FWCI 4y publikacji oznacza stosunek liczby cytowań tej publikacji (uzyskanej w roku publikacji i trzech kolejnych latach) do średniej liczby cytowań podobnej publikacji (publikacji z tej samej grupy dyscyplin w 4-cyfrowej klasyfikacji dyscyplin ASJC) w tym samym przedziale czasowym.
10.	Mediana wielkości zespołu (z całego dorobku)	Mediana liczby autorów dla każdej publikacji (autor + liczba współautorów) w całym portfolio publikacyjnym autora. W przypadku publikacji z liczbą autorów większą niż 10, liczba autorów została określona jako 10.
11.	Rok rezygnacji z publikowania	Kolejny rok po ostatnim roku publikacji w portfolio publikacyjnym autora. Jeśli rok rezygnacji nastąpił po 2019 r., wówczas obserwacja została sklasyfikowana jako prawostronnie ucięta.
12.	Dorobek naukowy (całego życia)	Liczba publikacji w całym portfolio publikacyjnym autora.



Weźmy za przykład naukowca o ID 142776 (tabela 1): to mężczyzna pracujący w Niemczech, którego wszystkie odniesienia bibliograficzne wszystkich publikacji wskazują na pracę w dyscyplinie medycyna (MED); zatrudniony w instytucji nienależącej do światowej czołówki TOP 200; mężczyzna ten pierwszą publikację wydał w 2000 r. (kohorta 2000), a ostatnią w 2016 r. i był umiędzynarodowiony w badaniach naukowych na średnim poziomie (jedna czwarta jego całego dorobku powstała we współpracy z naukowcami z zagranicy); publikował średnio w dosyć dobrych, ale nie topowych czasopismach (średnia rang percentylowych CiteScore to 72 percentyl czasopism w bazie Scopus przy możliwym przedziale 0–99); pracował w stosunkowo niewielkich zespołach badawczych jak na medycynę (mediana wielkości jego zespołu to 3 osoby); jego publikacje były za to bardzo dobrze cytowane – przeciętnie dwukrotnie bardziej niż średnia dla tej dyscypliny (co pokazuje czteroletni wskaźnik wpływu w postaci cytowań znormalizowany do dyscypliny), a jego całkowity dorobek był stosunkowo duży jak na etap rozwoju naukowego i wyniósł 30 artykułów naukowych.

Danych na mikropoziomie poszczególnych naukowców mamy o wiele więcej, ale w tym badaniu nie są one przydatne (inne dane wykorzystujemy w badaniu mobilności między klasami produktywności w krajach OECD w ujęciu 25–50 lat życia naukowego w Kwiek i Szymula 2024).

Szczegółowy opis metod przypisywania zmiennych poszczególnym naukowcom przedstawiono w tabeli 2. Skupiliśmy się na 16 dyscyplinach z obszaru STEMM, zgodnie z systemem klasyfikacji czasopism stosowanym w bazie Scopus (All Science Journal Classification, ASJC): AGRI, nauki rolnicze i biologiczne; BIO, biochemia, genetyka i biologia molekularna; CHEMENG, inżynieria chemiczna; CHEM, chemia; COMP, informatyka; EARTH, nauki o Ziemi i planetach; ENER, energia; ENG, inżynieria; ENVIR, nauka o środowisku; IMMU, immunologia i mikrobiologia; MATER, materiałoznawstwo; MATH, matematyka; MED, medycyna; NEURO, neuronauka; PHARM, farmakologia, toksykologia i farmacja; oraz PHYS, fizyka i astronomia.

#### 4. Wyniki

##### **Zdarzenia publikacyjne i analiza przeżycia**

Życie naukowe konceptualizujemy jako sekwencję zdarzeń publikacyjnych, od pierwszej publikacji do kolejnych publikacji, a w wielu przypadkach do ostatniej publikacji w portfolio publikacyjnym naukowca (kiedy to naukowcy najzwyczajniej przestają publikować). W naszym badaniu naukowcy, którzy po raz pierwszy publikowali w 2000 r. tworzą kohortę 2000, z której stopniowo, każdego roku, z różnym natężeniem rezygnują z nauki. Rezygnują z niej zarówno mężczyźni, jak i kobiety, w pierwszych latach rezygnują doktoranci, a w ostatnich badanych latach rezygnują adiunkci i profesorowie.

Rezygnacja z nauki jest przez nas konceptualizowana jako zdarzenie i analizowana w ramach tzw. analizy przeżycia (Allison, 2014; Mills, 2011). Chociaż ogólny motyw „odchodzenia z nauki” był szeroko badany w literaturze (Geuna & Shibayama, 2015; Preston, 2004), rezygnacja z nauki jako zdarzenie oznaczające zaprzestanie publikowania nie była przedmiotem szeroko zakrojonych badań z wykorzystaniem analizy przeżycia (i zgodnie z naszą najlepszą wiedzą nie była również wcześniej badana z globalnej perspektywy zarazem ilościowej, kohortowej i podłużnej).

Analiza przeżycia zajmuje się czasem, jaki upływa od początku obserwacji do wystąpienia określonego zdarzenia. Zdarzenie to kończy obserwację i może być nim na przykład zachorowanie człowieka w badaniach medycznych czy awaria urządzenia w analizie niezawodności w naukach technicznych (Mills, 2011). Zdarzeniem będącym przedmiotem naszego zainteresowania jest ostatnia publikacja naukowa autora – oznaczająca ostatni rok jego pozostawania w nauce. Cały okres poprzedzający zdarzenie rezygnacji z nauki – okres od pierwszej do ostatniej publikacji naukowej – określamy jako czas przeżycia w nauce (choć nie jako czas przeżycia w szkolnictwie wyższym). Klasyczną techniką statystyczną stosowaną w analizie przeżycia jest jego szacowanie metodą Kaplana-Meiera (Mills, 2011). Wykresem estymatora funkcji przeżycia Kaplana-Meiera jest seria charakterystycznych poziomych schodków, wskazujących na prawdopodobieństwo przeżycia (na osi Y) o różnej wysokości w różnych punktach upływającego czasu (na osi X, zob. ryc. 2). Analiza za pomocą estymatora Kaplana-Meiera pozwala na szacowanie funkcji przeżycia, a wykresy tych estymacji umożliwiają graficzne pokazanie i porównywanie krzywych przeżycia dla wybranych populacji (tutaj: dla mężczyzn i kobiet).

W przypadku, gdy uczestnicy badania – nasi publikujący naukowcy posiadający publikacje indeksowane w bazie Scopus – nie doświadczają zdarzenia przed zakończeniem badania (w naszym przypadku: naukowcy kontynuują publikowanie i nie rezygnują z nauki zgodnie z naszą definicją tego terminu), są oni uznawani za obserwacje prawostronnie ucięte (*right censored*). W każdym przypadku na końcu badania pozostają naukowcy, którzy nadal publikują. W odniesieniu do takich obserwacji dysponujemy tylko częściowymi informacjami: zdarzenie mogło mieć miejsce zaraz po ostatnim roku objętym badaniem; jednak nie dysponujemy na ten temat dokładną wiedzą. W naszym przypadku autorzy, których ostatnia publikacja była datowana na 2019 rok lub później, zostali oznaczeni jako obserwacje prawostronnie ucięte (tj. jako obserwacje opuszczające naukę w 2020 roku lub później). Aby sklasyfikować autora jako rezygnującego z nauki lub pozostającego w nauce, ostatnia publikacja musi być zatem datowana na 2018 r. (rezygnacja z nauki w 2019 r.) lub wcześniej. Przypadki nieucięte (*uncensored*) reprezentują obserwacje, dla których znamy zarówno rok początkowy (2000 dla kohorty 2000, 2010 dla kohorty 2010), jak i rok końcowy pozostawania w nauce, który musi być rokiem 2019 lub wcześniejszym (określony przez datę ostatniej publikacji). Medianę długości przerw

w publikowaniu dla obu kohort podajemy w tabelach 4 i 5 w załączniku. Dla każdego roku początkowa liczba naukowców wchodzących w dany rok składa się z naukowców, którzy opuszczą naukę w tym roku, i tych, którzy pozostaną w nauce do następnego roku. Skumulowane prawdopodobieństwo przeżycia dla danego przedziału czasowego jest obliczane jako przemnożenie wszystkich prawdopodobieństw przeżycia we wszystkich wcześniejszych przedziałach czasowych (Mills, 2011). Dwie krzywe przeżycia – np. dla mężczyzn i dla kobiet – można porównywać statystycznie, między innymi sprawdzać, czy różnica między czasem przeżycia dla obu grup jest statystycznie istotna.

### **Prezentowane badanie vs dotychczasowe badania**

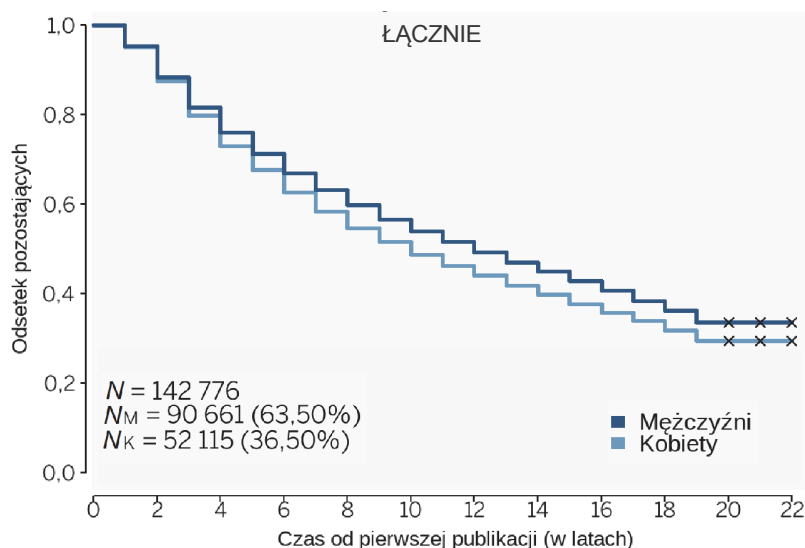
W porównaniu z dotychczasowymi badaniami nasze analizy prezentują inną skalę geograficzną (łącznie 38 krajów OECD zamiast analiz jednego kraju, niemal wyłącznie USA). Odchodzimy tym samym od studiów przypadków instytucji, skoncentrowanych na jednym kraju i opartych na ankietach i wywiadach. Nasze analizy obejmują wiele krajów (dobra jakość danych dla obszaru OECD) i koncentrują się na dyscyplinach naukowych. Ponadto nasze badanie wykorzystuje inną metodologię (tj. analizę przeżycia dla kolejnych kohort) niż dotychczasowe badania i analizuje różnice między dyscyplinami w rezygnacji z pracy na przestrzeni dwóch dekad. Stosujemy podłużne podejście do dużych, rozłącznych kohort naukowców. Testujemy moc ustrukturyzowanych, wiarygodnych i wyselekcjonowanych dużych zbiorów danych (typu bibliometrycznego: zbiór surowych danych z bazy Scopus).

Nasze podejście jest zarazem podłużne w ścisłym tego słowa znaczeniu: nie stosujemy podejścia przekrojowego (rejestracji danych w jednym okresie) ani też przekrojowego powtarzanie (rejestracji danych w kilku okresach dla różnych jednostek). Kohorty naukowców są obserwowane w ujęciu rocznym w czasie przez okres trwający maksymalnie nieco ponad dwie dekady (2000–2022). Spójne zmienne i stabilne klasyfikacje dyscyplin zostały zastosowane do wszystkich krajowych systemów nauki w 38 krajach łącznie i zostały poddane analizie czasowej w ramach kohort naukowców. Wykorzystaliśmy szereg indywidualnych danych na mikropoziomie indywidualnych naukowców, m.in. wskaźnik współpracy międzynarodowej, średnią rangę percentylową czasopism z całego życia naukowca, medianę wielkości zespołu z całego życia czy też czteroletni wskaźnik wpływu w postaci cytowań znormalizowany do dyscypliny (zob. tabela 1). Jednocześnie czas został uznany za zmienną o zasadniczym znaczeniu.

### **Rezygnacja z nauki: kohorta 2000**

Na początek prześledziliśmy losy 142776 naukowców z 16 dyscyplin z obszaru STEMM z kohorty 2000 do czasu, gdy przestali oni publikować (lub do 2022 r.). Rycina 2 przedstawia wykres estymaty Kaplana-Meiera funkcji przeżycia dla wszystkich dyscyplin łącznie, przy czym odsetek naukowców pozostających w nauce (i prawdopodobień-

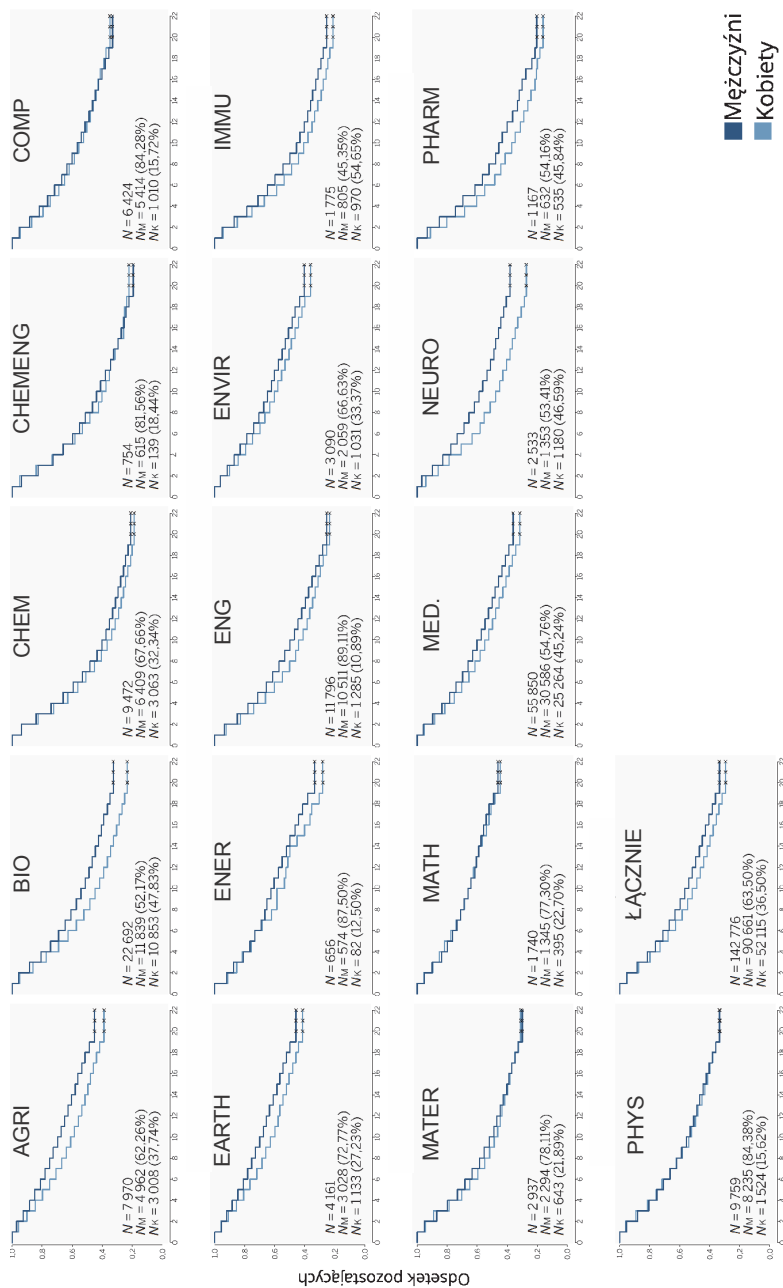
stwo pozostania w nauce) pokazano na osi Y, a lata publikowania od 2000 r. pokazano na osi X (z małymi krzyżkami oznaczającymi przypadki obserwacji prawostronnie uciętych dla trzech ostatnich lat poddanych analizie, 2020–2022).



Ryc. 2. Wykres estymaty Kaplana-Meiera funkcji przeżycia z podziałem na płeć, wszystkie dyscypliny łącznie, dla kohorty naukowców z 2000 r. ( $N = 142\,776$ , rozkład procentowy kohorty: 63,50% mężczyźni, 36,50% kobiety). Największe spadki prawdopodobieństwa pozostania w nauce zarówno dla mężczyzn, jak i kobiet pojawiają się w latach 1, 2, 3 i 4 (częściowo okres studiów doktoranckich). Począwszy od roku 4 różnica między mężczyznami i kobietami staje się coraz bardziej zauważalna: każdego roku odsetek mężczyzn, którzy pozostali (i ich prawdopodobieństwo pozostania), jest wyższy od odsetka (i prawdopodobieństwa) kobiet, które pozostały – co zasadniczo jest zgodne z prawidłowościami znanymi z wcześniejszych badań

Upraszczając wyniki, można stwierdzić, że mniej więcej jedna trzecia kohorty naukowców z 2000 r. rezygnuje z nauki po 5 latach, połowa po 10 latach, a dwie trzecie pod koniec badanego okresu (po 19 latach), przy czym odsetek rezygnujących jest konsekwentnie niższy w przypadku mężczyzn i wyższy w przypadku kobiet.

Począwszy od czwartego roku, w każdym kolejnym badanym roku wyższy odsetek kobiet niż mężczyzn rezygnuje z nauki. W związku z tym prawdopodobieństwo rezygnacji z nauki przez kobiety jest o jedną dziesiątą wyższe niż prawdopodobieństwo rezygnacji z nauki przez mężczyzn po 5 i 10 latach (odpowiednio 12,54% i 11,52%), a prawdopodobieństwo rezygnacji z nauki pod koniec badanego okresu jest o 6,33% wyższe, co wyraźnie widać na ryc. 2. Po 19 latach (tabela 3, tylko nieucięte obserwacje), prawdopodobieństwo Kaplana-Meiera pozostania w nauce dla kobiet wynosi 0,294 (29,4% kobiet z pierwotnej kohorty kontynuuje publikowanie); natomiast dla mężczyzn jest ono znacznie wyższe i osiąga poziom 0,336 (33,6%).



Ryc. 3. Wykres estymaty Kaplana-Meiera funkcji przeżycia według dyscypliny i płci dla kohorty naukowców z 2000 r. (N= 142 776)

Jednak ten zagregowany, ogólny obraz rezygnacji z publikowania w nauce zaprezentowany dla wszystkich dyscyplin łącznie przesłania różne obrazy dla różnych dyscyplin (zob. ryc. 3). Zdecydowanie najmniejsze szanse przetrwania w nauce w porównaniu z mężczyznami mają kobiety pracujące w dyscyplinach o wysokim odsetku (sięgającym 50%) kobiet, takich jak – dla kohorty 2000 – BIO, NEURO i PHARM.

Tabela 3. Prawdopodobieństwo przeżycia Kaplana-Meiera dla kohorty 2000 według płci (wszystkie dyscypliny łącznie) z całkowitą liczbą mężczyzn i kobiet, czasem (wyrażonym w latach), liczbą obserwacji naukowców rezygnujących z nauki i prawdopodobieństwem pozostania w nauce (z 95% przedziałem ufności)

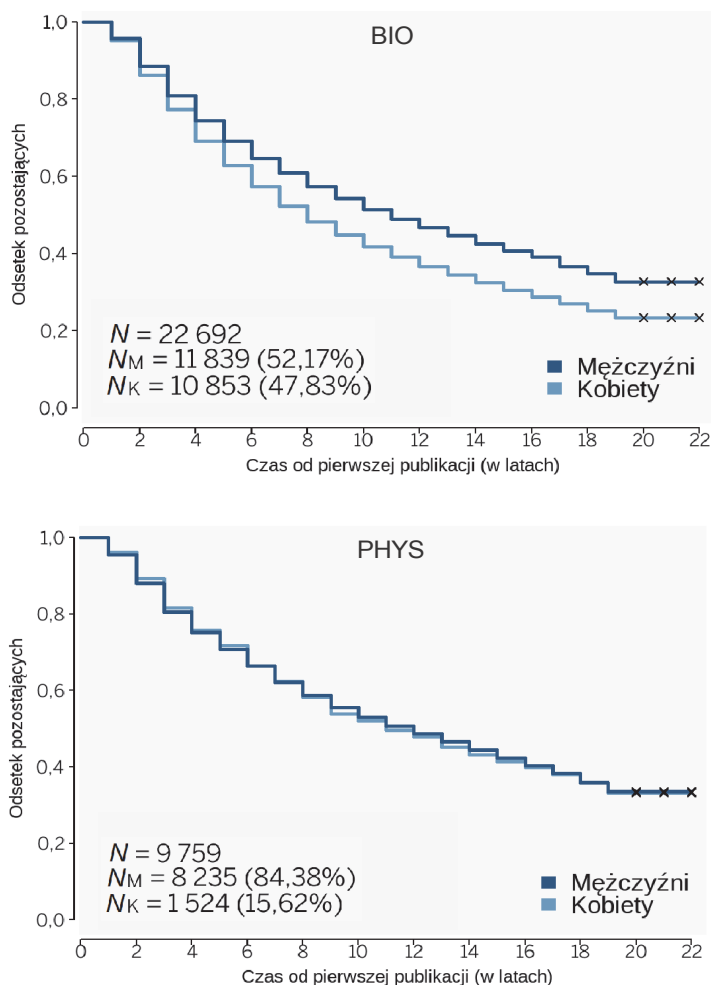
Czas (lata)	Kobiety			Mężczyźni		
	<i>N</i>	<i>N</i> wychodzących z nauki	Prawdopodobieństwo w ujęciu Kaplana-Meiera: pozostanie w nauce, 95% przedział ufności	<i>N</i>	<i>N</i> wychodzących z nauki	Prawdopodobieństwo w ujęciu Kaplana-Meiera: pozostanie w nauce, 95% przedział ufności
1	52 115	2530	0,951 (0,950–0,953) <sup>1</sup>	90 661	4151	0,954 (0,953–0,956) <sup>1</sup>
2	49 585	3985	0,875 (0,872–0,878) <sup>1</sup>	86 510	6302	0,885 (0,883–0,887) <sup>1</sup>
3	45 600	3948	0,799 (0,796–0,803) <sup>2</sup>	80 208	6114	0,817 (0,815–0,820) <sup>1</sup>
4	41 652	3553	0,731 (0,727–0,735) <sup>2</sup>	74 094	5062	0,761 (0,759–0,764) <sup>1</sup>
5	38 099	2838	0,677 (0,673–0,681) <sup>2</sup>	69 032	4356	0,713 (0,710–0,716) <sup>2</sup>
6	35 261	2602	0,627 (0,623–0,631) <sup>2</sup>	64 676	3934	0,670 (0,667–0,673) <sup>2</sup>
7	32 659	2183	0,585 (0,581–0,589) <sup>2</sup>	60 742	3458	0,632 (0,629–0,635) <sup>2</sup>
8	30 476	1961	0,547 (0,543–0,551) <sup>2</sup>	57 284	3110	0,598 (0,594–0,601) <sup>2</sup>
9	28 515	1665	0,515 (0,511–0,520) <sup>2</sup>	54 174	2774	0,567 (0,564–0,570) <sup>2</sup>
10	26 850	1472	0,487 (0,483–0,491) <sup>2</sup>	51 400	2465	0,540 (0,537–0,543) <sup>2</sup>
11	25 378	1264	0,463 (0,458–0,467) <sup>2</sup>	48 935	2225	0,515 (0,512–0,518) <sup>2</sup>
12	24 114	1158	0,440 (0,436–0,445) <sup>2</sup>	46 710	2055	0,493 (0,489–0,496) <sup>2</sup>
13	22 956	1151	0,418 (0,414–0,423) <sup>2</sup>	44 655	2032	0,470 (0,467–0,473) <sup>2</sup>
14	21 805	1089	0,398 (0,393–0,402) <sup>2</sup>	42 623	1889	0,449 (0,446–0,453) <sup>2</sup>
15	20 716	1048	0,377 (0,373–0,382) <sup>2</sup>	40 734	1884	0,429 (0,425–0,432) <sup>2</sup>
16	19 668	1033	0,358 (0,353–0,362) <sup>2</sup>	38 850	1959	0,407 (0,404–0,410) <sup>2</sup>
17	18 635	1002	0,338 (0,334–0,342) <sup>2</sup>	36 891	2020	0,385 (0,381–0,388) <sup>2</sup>
18	17 633	1064	0,318 (0,314–0,322) <sup>2</sup>	34 871	2070	0,362 (0,359–0,365) <sup>2</sup>
19	16 569	1228	0,294 (0,290–0,298) <sup>2</sup>	32 801	2350	0,336 (0,333–0,390) <sup>2</sup>

(1) Błąd standardowy 0,001; (2) Błąd standardowy 0,002

Przedmiotem naszego szczególnego zainteresowania są dyscypliny, w których występują największe i najmniejsze (lub nie występują żadne) różnice w przebiegu estymat Kaplana-Meiera funkcji przeżycia pomiędzy mężczyznami i kobietami. W dwóch najbardziej licznych dyscyplinach i dwóch dyscyplinach o największej liczbie kobiet, w medycy-

nie (MED) oraz biochemii, genetyce i biologii molekularnej (BIO), wykresy estymat funkcji przeżycia dla mężczyzn i kobiet są wyraźnie odmienne.

W przypadku dyscypliny BIO (rycina 4, górny panel, zestawionej z fizyką i astronomią PHYS, dolny panel), obejmującej 47,83% kobiet i 22 692 naukowców w badanej kohorcie, największe spadki występują w latach 2, 3 i 4; a od roku 3, pojawia się coraz większa różnica między wykresami estymat funkcji przeżycia dla kobiet i mężczyzn.



Ryc. 4. Wykres estymaty Kaplana-Meiera funkcji przeżycia dla kobiet i mężczyzn, biochemia, genetyka i biologia molekularna (BIO) ( $N = 22\,692$ , górny panel) i fizyka i astronomia (PHYS) ( $N = 9\,759$ , dolny panel), kohorta naukowców z 2000 roku

Co uderzające, w trzech dyscyplinach intensywnie zmatematyzowanych (MATH, COMP i PHYS), które charakteryzują się bardzo niską liczbą i niskim odsetkiem kobiet, wykresy

estymacji funkcji przeżycia dla mężczyzn i kobiet są niemal identyczne (patrz PHYS na ryc. 4, dolny panel), a obie krzywe niemal się pokrywają. Jak się okazuje, różnice między kobietami i mężczyznami w zakresie rezygnacji z nauki w PHYS w zasadzie nie istnieją. W przypadku połowy dyscyplin różnice są statystycznie istotne, a w przypadku drugiej połowy są statystycznie nieistotne (przeprowadziliśmy sześć typów testów statystycznych, głównie nieparametrycznych, m.in. test logarytmicznych rang, uogólnienie Gehana testu Wilcoxon'a i uogólnienie Peto i Peto; nasze badanie zawiera znaczną liczbę dodatkowych tabel i wykresów, zwłaszcza na poziomie dyscyplin, które jednak pomijamy z powodu ograniczonego miejsca).

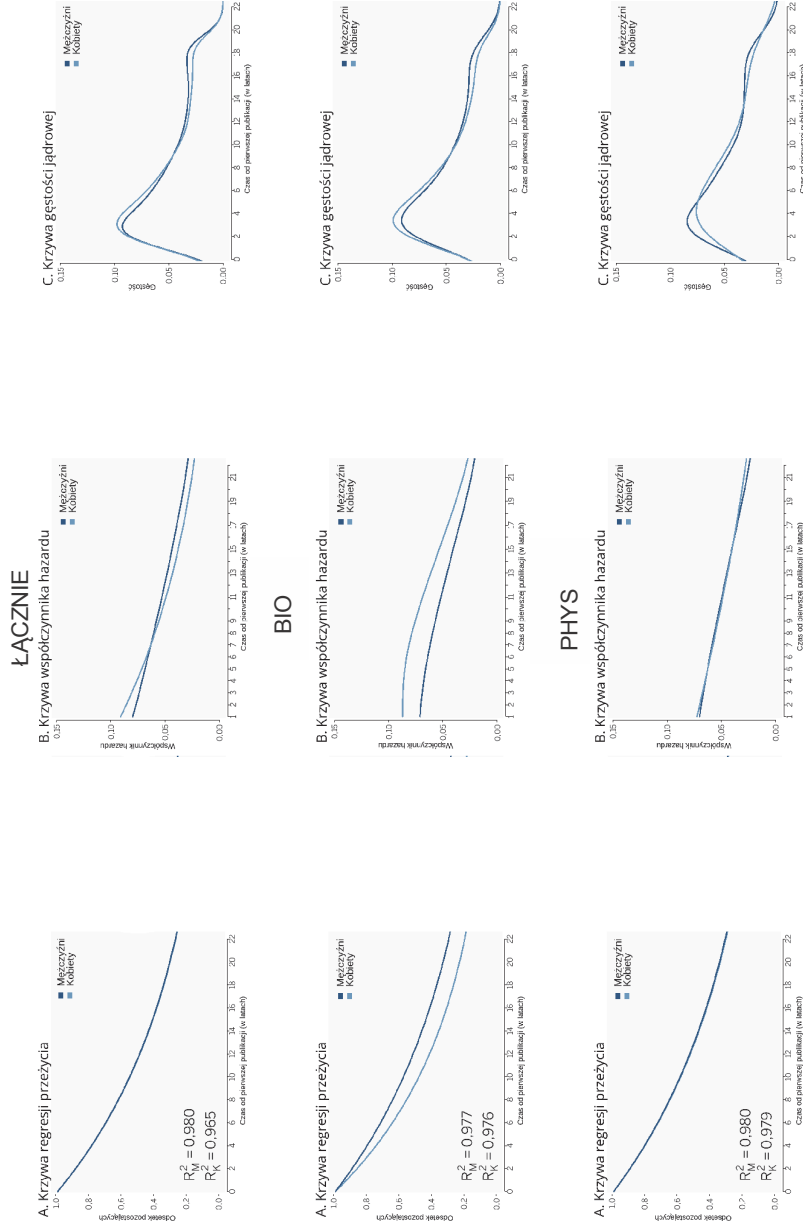
Chcąc uzyskać bardziej kompleksowy obraz rezygnacji z nauki, skorelowaliśmy dane dotyczące mężczyzn i kobiet, aby otrzymać krzywe regresji przeżycia (panele A na ryc. 6), krzywe współczynnika ryzyka/hazardu (panele B) i krzywe gęstości jądrowej (panele C). Wykresy tych krzywych pozwalają na prezentację zjawiska rezygnacji z nauki z kilku perspektyw (w każdym przypadku pomijamy obserwacje z lat 2019–2022 jako prawostronnie ucięte).

Krzywe regresji przeżycia dla wszystkich dyscyplin łącznie (ryc. 5, panel A) wskazują na znacznie bardziej stromy spadek dla mężczyzn i kobiet we wczesnych latach kariery naukowej i znacznie łagodniejszy spadek w późniejszych latach, przy rosnącej różnicy między krzywymi dla mężczyzn i kobiet. Mężczyźni pozostają w nauce częściej niż kobiety, a zarówno mężczyźni, jak i kobiety na późniejszych etapach kariery pozostają w niej częściej niż na wcześniejszych etapach kariery (funkcja regresji wykazuje większy spadek na początku kariery niż w późniejszych latach). Krzywe współczynnika hazardu/ryzyka (ryc. 5, panel B) przedstawiają podobny obraz: wskaźnik rezygnacji zarówno dla mężczyzn, jak i dla kobiet jest wyższy we wczesnych latach publikowania i znacznie niższy w późniejszych latach; a przez pierwszych 15 lat kobiety są bardziej narażone na odejście z nauki niż mężczyźni. Szczyt rezygnacji dla mężczyzn przypada na 4. rok po pierwszej publikacji, a dla kobiet na rok 6. Funkcja hazardu nie pokazuje prawdopodobieństwa, ale jest miarą ryzyka rezygnacji z nauki w czasie  $t$  (liczba lat, jakie upływają od pierwszej publikacji).

Jeszcze inny obraz przedstawiają krzywe gęstości jądrowej (ryc. 5, panel C), które pokazują, jak wygląda rozkład w czasie wszystkich mężczyzn i kobiet z kohorty 2000, którzy faktycznie zrezygnowali z nauki. Odsetek naukowców, którzy zrezygnowali z nauki, jest najwyższy w ciągu pierwszych 8 lat, przy czym lata 3–4 są krytyczne. Zarówno mężczyźni, jak i kobiety częściej rezygnują z nauki na wczesnym etapie rozwoju kariery naukowej. Po przetrwaniu pierwszych 10 lat prawdopodobieństwo rezygnacji znacznie maleje.

Ogólny obraz obejmujący wszystkie dyscypliny naukowe łącznie przesłania jednak bardziej szczegółowe obrazy charakterystyczne dla poszczególnych dyscyplin. Przeanaliz-





Ryc. 5. Wykres krzywej regresji przeżycia, krzywej współczynnika hazardu/ryzyka i krzywej gęstości jądrowej, wszystkie dyscypliny łącznie ( $N = 142\,776$ ), biochemia, genetyka i biologia molekularna (BIO) ( $N = 22\,692$ ) i fizyka i astronomia (PHYS) ( $N = 9\,750$ ), kohorta 2000

zowane powyżej wykresy estymaty Kaplana-Meiera funkcji przeżycia dla kobiet i mężczyzn dla BIO i PHYS opowiadają zasadniczo różne historie (ryc. 4). Ustalenia te znajdują potwierdzenie przy porównywaniu krzywych regresji przeżycia (BIO i PHYS, ryc. 5, panel A), z podobnym wyższym prawdopodobieństwem rezygnacji we wczesnych latach i niższym prawdopodobieństwem w późniejszych latach – ale z wyraźnymi różnicami między mężczyznami i kobietami w obu dyscyplinach. Nasze dane pokazują, że kobiety w BIO odchodzą z nauki wraz z upływem czasu w coraz większym odsetku w porównaniu z mężczyznami; a kobiety w PHYS odchodzą z niej w niemal dokładnie w takim samym odsetku, jak mężczyźni w całym badanym okresie.

Również krzywe współczynnika ryzyka (ryc. 5, panele B) opowiadają podobną historię, w której w BIO wskaźnik rezygnacji dla kobiet jest znacznie wyższy niż wskaźnik rezygnacji dla mężczyzn we wszystkich badanych latach; a w PHYS wskaźniki krzywej ryzyka dla mężczyzn i kobiet są prawie identyczne.

Krzywe gęstości jądrowej dla BIO i PHYS (ryc. 5, panel C) wyraźnie pokazują podobne wzorce wewnątrzdiscyplinarne dla mężczyzn i kobiet (wyższy wskaźnik rezygnacji we wczesnych latach) oraz różne wzorce międzydiscyplinarne dla mężczyzn i kobiet. O ile w przypadku BIO we wczesnych latach wskaźnik rezygnacji dla kobiet jest wyższy niż dla mężczyzn, o tyle w przypadku PHYS w latach 2–6 jest on wyższy dla mężczyzn. Jest to zaskakujące ustalenie, potwierdzone również w przypadku COMP i MATH.

### **Rezygnacja z nauki: kohorta 2010**

Na koniec przechodzimy do kohorty 2010 (ryc. 6). Wykresy estymaty Kaplana-Meiera funkcji przeżycia dla mężczyzn i kobiet dla obu kohort radykalnie różnią się między sobą. Co najważniejsze, wykresy estymaty dla wszystkich dyscyplin łącznie dla kohorty 2010 są niemal identyczne dla mężczyzn i kobiet. Są one również niemal identyczne dla takich dyscyplin intensywnie zmatematyzowanych jak COMP, PHYS i ENG.

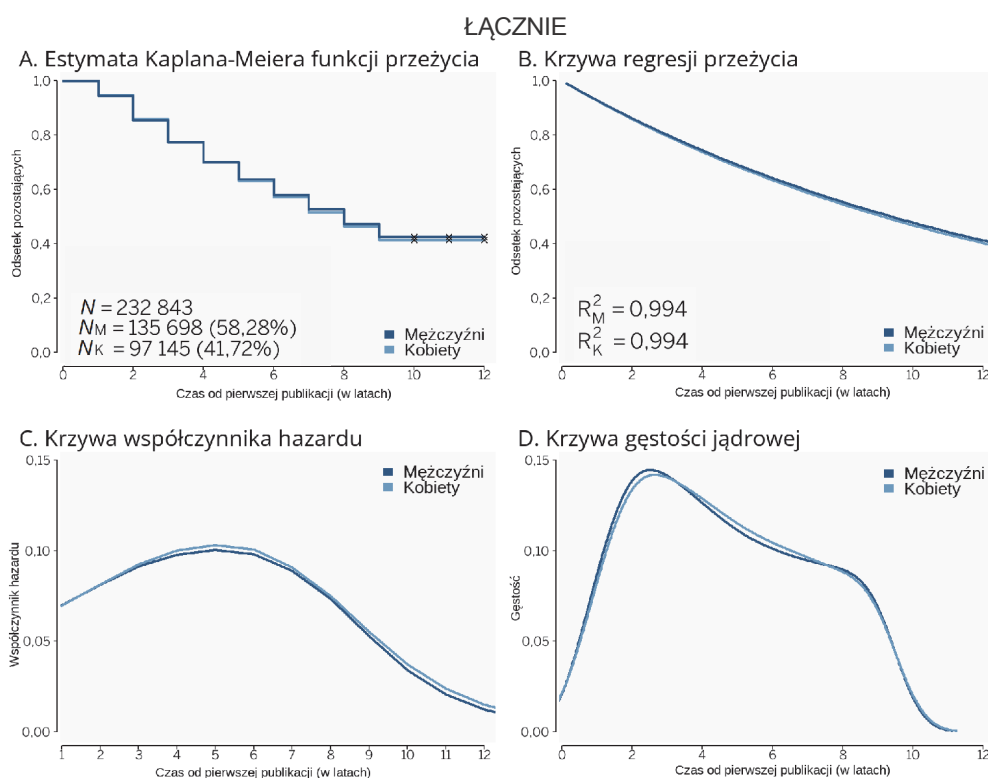
Po 9 latach od rozpoczęcia pracy naukowej prawdopodobieństwo pozostania w nauce dla kobiet z kohorty 2010 wynosi jedynie 0,414 (tylko 41,4% kobiet z pierwotnej kohorty kontynuuje publikowanie), natomiast dla mężczyzn, prawdopodobieństwo to jest tylko nieznacznie wyższe i wynosi 0,424 (42,4%). Oznacza to spektakularny brak różnic między mężczyznami i kobietami w porównaniu z kohortą 2000, dla której wyniki były w znacznym stopniu zróżnicowane ze względu na płeć. Dla dokładnie tych samych ośmiu dyscyplin (m.in. dla czwórki dyscyplin COMP, ENG, MATH i PHYS) testy statystyczne pokazują, że estymaty funkcji przeżycia dla mężczyzn i kobiet nie różnią się w sposób statystycznie istotny. Jednak dla wszystkich dyscyplin łącznie, a także dla największych dyscyplin MED, BIO i AGRI, w których kobiety stanowią około 50%, różnice są nadal statystycznie istotne.

Z perspektywy porównawczej nasze analizy wskazują na fundamentalną transformację wzorców rezygnacji z nauki w różnych dyscyplinach (i we wszystkich dyscyplinach

łącznie), jaka zaszła między kohortami 2000 i 2010. Różnice między mężczyznami i kobietami dla wszystkich dyscyplin łącznie, wyraźnie widoczne w przypadku kohorty 2000, praktycznie całkowicie zanikają w przypadku kohorty 2010 (ryc. 6).

To wnioski o ważnych potencjalnych implikacjach dla badań i polityki naukowej w świetle obszernej dotychczasowej literatury na temat rezygnacji i pozostawania w nauce akademickiej.

Sektor nauki na całym świecie przechodzi potężne transformacje, a ustalenia dotyczące starszych kohort naukowców (tutaj: kohorta 2000) mogą nie mieć zastosowania do młodszych kohort (tutaj: kohorta 2010). Czas w nauce – w tym przypadku różnica czasu rzędu dekady – ma ogromne znaczenie, ponieważ system nauki jest systemem dynamicznym, stale podlegającym zmianom strukturalnym pod wpływem różnorodnych presji zewnętrznych i wewnętrznych.



Ryc. 6. Wykres estymaty Kaplana-Meiera funkcji przeżycia (panel A), krzywej regresji przeżycia (panel B), krzywej współczynnika hazardu/ryzyka (panel C) i krzywej gęstości jądrowej (panel D), kohorta 2010 ( $N = 232\ 843$ )

Wykresy estymaty Kaplana-Meiera funkcji przeżycia dla wszystkich dyscyplin łącznie przyjmują zdecydowanie odmienne kształty dla obu kohort. Dla kohorty z 2010 r.

(ryc. 6), po pierwsze, wskaźniki rezygnacji są znacznie wyższe, a spadki znacznie bardziej dramatyczne niż dla kohorty 2000 (50% lub więcej zarówno mężczyzn, jak i kobiet przestaje publikować w ósmym roku). Po drugie, nie ma żadnych istotnych różnic między kobietami i mężczyznami co do prawdopodobieństwa rezygnacji z nauki. Wniosek ten potwierdzają kształty krzywych regresji przeżycia dla dwóch kohort, ponownie bez różnicy między kobietami i mężczyznami. Potwierdzenie pochodzi również z obrazu krzywych współczynnika ryzyka, które są prawie identyczne dla mężczyzn i kobiet, oraz z obrazu krzywych gęstości jądrowej, które świadczą o zasadniczo zbliżonym rozkładzie naukowców, którzy istotnie zrezygnowali z nauki (w ujęciu czasowym).

Zwłaszcza rozkład gęstości jądrowej ujawnia dramatycznie wysokie wskaźniki rezygnacji zarówno dla mężczyzn, jak i kobiet w ciągu pierwszych 4 lat pozostawania w nauce – co może odnosić się wysokiego poziomu rezygnacji z nauki ze strony doktorantów lub młodszych postdoków (ryc. 6, panel D). Ponownie, podobnie jak w przypadku kohorty z 2000 r., za zagregowanym obrazem wszystkich dyscyplin łącznie kryją się mniejsze i bardziej złożone obrazy poszczególnych dyscyplin, czego jednak w tej pracy nie pokazujemy z powodu braku miejsca.

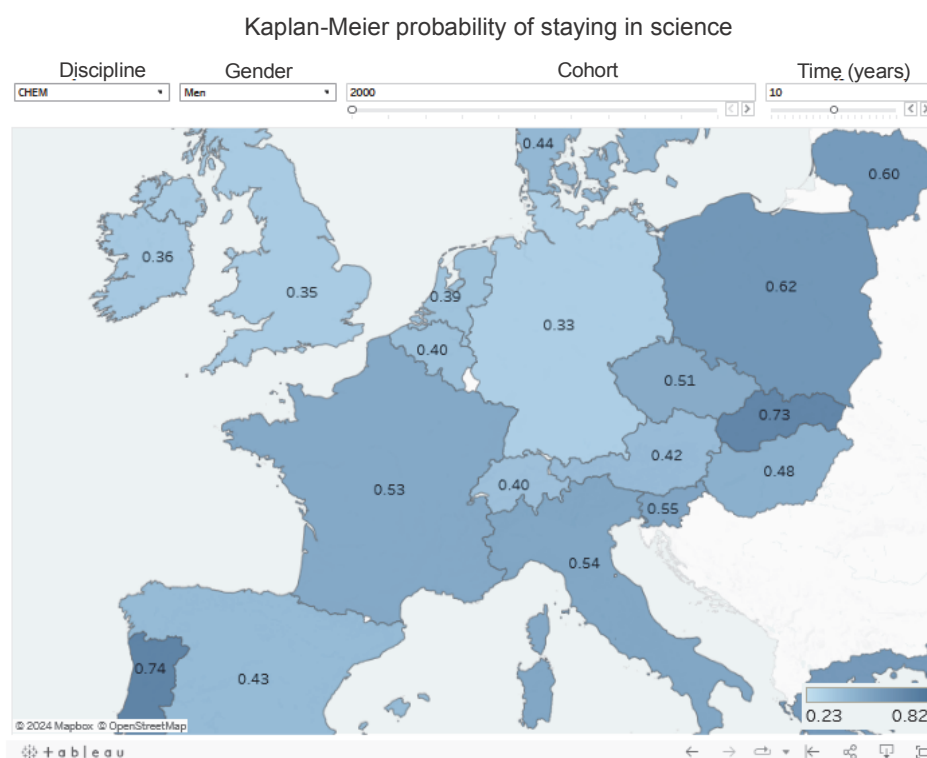
Nasze badanie ma charakter globalny: analizujemy globalną profesję naukową za pomocą globalnych kohort naukowców (łącznie w 38 krajach OECD) w różnych okresach i we wszystkich dyscyplinach STEMM. Możliwe jest jednak również, co pozostaje poza zakresem naszego zainteresowania, porównywanie różnic między krajami w panelu interaktywnym, w którym przedstawiamy wyniki prawdopodobieństwa według Kaplana-Meiera dla poszczególnych dyscyplin dla wszystkich 11 kohort (kohorty 2000–2010, 2,1 mln naukowców).

Nasze badania pokazują ogromną skalę różnic dotyczących szans na przetrwanie w nauce między krajami zarówno dla mężczyzn, jak i dla kobiet (na przykład bardzo wysoką stabilność kariery naukowej – związaną z dużą stabilnością zatrudnienia – w Polsce w porównaniu z większością krajów Europy Zachodniej).

Polska w proponowanym ujęciu pozostawania i rezygnacji z nauki okazuje się niezwykle atrakcyjnym miejscem uprawiania nauki pod jednym kątem: stabilności zatrudnienia (w sensie: kontynuacji publikowania). Znikamy z nauki rzadziej – i znikamy z niej w radykalnie niższych odsetkach niż nasi koledzy z najbogatszych gospodarek świata. Polska notuje niezwykle niskie poziomy rezygnacji z nauki niezależnie od wybranej kohorty i we wszystkich dyscyplinach; co zadziwiające w globalnym kontekście, różnice między mężczyznami i kobietami nawet dla najstarszych kohort nie są duże.

Przykładowo w Polsce (zob. panel interaktywny dostępny online): dla wszystkich dyscyplin łącznie dla kohorty 2000 prawdopodobieństwo pozostania w nauce po 10 latach wynosi 67% dla kobiet i 75% dla mężczyzn – i w obu przypadkach należy do najwyższych wśród badanych 38 krajów OECD. W medycynie (MED), największej dyscyplinie w na-

szym ujęciu, prawdopodobieństwo to wynosi odpowiednio 69% i 73%, a w matematyce (MATH) odpowiednio 75% i 76%. Wysoki poziom pozostawania w nauce świadczyć może zarówno o dużej atrakcyjności kariery akademickiej (słabe *push factors*, czynniki wypychające z nauki), jak i o braku dobrych możliwości zatrudnienia dla naukowców poza sektorem akademickim (słabe czynniki *pull factors*, odciągające od nauki).



Ryc. 7. Panel interaktywny: prawdopodobieństwo pozostania w nauce w ujęciu kraju, dyscypliny, płci, kohorty naukowców i czasu, jaki upłynął od pierwszej publikacji. Przykład dla mężczyzn z kohorty naukowców 2000 publikujących w dyscyplinie chemia (CHEM), po 10 latach: szanse wynoszą 62% w Polsce, ale tylko 33% w Niemczech, 35% w Wielkiej Brytanii i 36% w Irlandii. <https://public.tableau.com/app/profile/marek.kwiek/viz/Attrition-in-science-OECD/Dashboard>

Dla kohorty 2010 w przypadku Polski szanse mężczyzn na pozostanie w nauce po 10 latach dla wszystkich dyscyplin łącznie są niemal największe spośród wszystkich badanych krajów (i wynoszą 58%, ustępując tylko niewielkiemu systemowi naukowemu Łotwy z poziomem 59%). W oparciu o nasze dane można przeprowadzić szczegółowe oszacowania dla polskiej nauki we wszystkich wymiarach zaproponowanych powyżej dla krajów OECD: nasze badanie objęło 2014 polskich naukowców z kohorty 2000 i 3900 naukowców z kohorty 2010 (oraz wszystkie kohorty pośrednie). Jednak nie Polska jest tematem tej pracy, ponieważ interesuje nas tu kadra akademicka ujmowana globalnie.

## 5. Dyskusja, wnioski i ograniczenia

Odchodzenie z nauki i pozostawanie w niej to procesy o charakterze długofalowym i do ich badania w dużej skali potrzebne są kompleksowe, podłużne i globalne zbiory danych – jeśli tylko chcemy wyjść poza ramy jednego kraju i analizować je w podziale na dyscypliny i na przestrzeni czasu. Nasze analizy pokazują, że zjawisko rezygnacji z nauki przechodzi istotne transformacje, ponieważ nowe kohorty naukowców wchodzą do nauki w nowych warunkach (zawodowych, ekonomicznych i innych) (Milojevic i in., 2018).

Nasze badanie pokazuje, że za zagregowanymi zmianami na poziomie wszystkich dyscyplin STEMM łącznie kryją się wysoce zróżnicowane przeobrażenia, które zachodzą z różną intensywnością na poziomie dyscyplin i w ujęciu czasowym. Zjawisko rezygnacji z nauki ma odmienne znaczenie dla mężczyzn i kobiet w różnych dyscyplinach, a także dla naukowców z różnych kohort wchodzących po raz pierwszy do nauki.

W ujęciu wszystkich dyscyplin STEMM łącznie jedna trzecia kohorty naukowców z 2000 r. rezygnuje z nauki po 5 latach, połowa po 10 latach, a dwie trzecie przed końcem badanego okresu (po 19 latach), przy czym odsetek odchodzących jest niezmiennie niższy w przypadku mężczyzn i wyższy w przypadku kobiet. Kobiety są o jedną dziesiątą bardziej narażone na rezygnację z nauki niż mężczyźni po 5 i 10 latach (odpowiednio o 12,54 i 11,52%) i są o 6,33% bardziej narażone na rezygnację pod koniec badanego okresu.

Różnice między kobietami i mężczyznami pod względem prawdopodobieństwa rezygnacji z nauki maleją wraz z każdą kolejną kohortą naukowców rozpoczynających karierę naukową w latach 2000–2010: wraz z rosnącą liczbą kobiet w nauce i rosnącym odsetkiem kobiet w kohortach naukowców rezygnacja z nauki staje się z czasem coraz słabiej zależna od płci. Dla 38 krajów OECD różnice między kobietami i mężczyznami, wyraźnie widoczne dla kohorty 2000 dla wszystkich dyscyplin łącznie, zanikają dla kohorty 2010. Jednak zagregowane dane kryją za sobą szczegółowe, zdezagregowane dane dla obu kohort, przedstawiające różne procesy zachodzące w różnych dyscyplinach.

Wykresy estymat Kaplana-Meiera funkcji przeżycia dla dwóch szczegółowo porównywanych dyscyplin: biochemii, genetyki i biologii molekularnej (BIO, duża dyscyplina, 47,83% kobiet) oraz fizyki i astronomii (PHYS, niewielka dyscyplina, tylko 15,62% kobiet) dla kohorty 2000 opowiadają zasadniczo odmienne historie. W BIO kobiety charakteryzują się znacznie niższym prawdopodobieństwem przetrwania niż mężczyźni, przy czym różnica rośnie wraz z każdym kolejnym rokiem pracy, a w PHYS wskaźniki przetrwania dla mężczyzn i kobiet przez dwie dekady pracy są niemal identyczne. Kobiety w BIO rezygnują z nauki w coraz większym odsetku w ciągu dwóch dekad w porównaniu z mężczyznami; natomiast kobiety w PHYS odchodzą z nauki w niemal dokładnie takim samym odsetku co mężczyźni w całym badanym okresie. Różnice między tymi dyscy-

plinami są zadziwiające i z pewnością wymagają dalszych badań, zwłaszcza na poziomie krajowym.

Jednak w przypadku kohorty 2010 dramatyczny brak różnic w rezygnacji z nauki mężczyzn i kobiet dla wszystkich dyscyplin łącznie idzie w parze ze znacznie mniej wyraźnymi, ale wciąż istniejącymi różnicami w poszczególnych dyscyplinach. W obu przypadkach w dyscyplinach intensywnie zmatematyzowanych (takich jak MATH, COMP i PHYS) różnice między kobietami i mężczyznami nie występują. W tych dyscyplinach odsetek kobiet rozpoczynających karierę naukową jest niewielki, ale kobiety pozostają w nauce w dokładnie takich samych proporcjach co mężczyźni – co może sugerować, że odnoszą one duże sukcesy zawodowe pomimo ewentualnego nieprzyjawnego klimatu (*chilly climate*) w miejscu pracy. W dyscyplinach o bardzo niskiej reprezentacji kobiet, w których są one widocznymi mniejszościami (poniżej 20% publikujących naukowców), często pełniącymi rolę symboli czy wzorcowych postaci na wydziałach (zob. Kanter, 1977), kobiety pozostają w systemie nauki z takim samym powodzeniem co mężczyźni.

Tradycyjnie uważa się (Preston, 2004; 2014; Goulden et al., 2011; Shaw & Stanton, 2012; Wolfinger et al., 2008), że kobiety odchodzą z nauki wcześniej i w większym odsetku niż mężczyźni. Zasadniczo potwierdzają to nasze analizy kohorty 2000 w oparciu o estymaty Kaplana-Meiera funkcji przeżycia, z wyjątkiem czterech dyscyplin wysoce zmatematyzowanych: MATH, COMP, PHYS i ENG. Jednak w przypadku nowszych kohort naukowców różnice między kobietami i mężczyznami dotyczące prawdopodobieństwa rezygnacji z nauki mogą już nie występować bądź występować w małej skali, szczególnie w przypadku dyscyplin o niskiej liczbie i małym odsetku kobiet.

W przypadku nowych pokoleń naukowców prawdopodobieństwo rezygnacji z nauki wyraźnie rośnie i jest bardzo wysokie (58,6% kobiet i 57,6% mężczyzn z kohorty 2010 zrezygnowało z nauki w ciągu 9 lat) – ale poziom rezygnacji jest znacznie mniej uwarunkowany przez płeć naukowców niż tradycyjnie zakładano.

Nasze analizy przynoszą zatem dobre i złe wieści: pierwsze wskazują na to, że kobiety i mężczyźni (zwłaszcza w dyscyplinach silnie zmatematyzowanych, do których jednak mają dosyć ograniczony dostęp, zob. Kwiek i Szymula 2023) w nauce radzą sobie niemal tak samo dobrze; jednak drugie wskazują na to, że z każdą kolejną kohortą rośnie poziom rezygnacji z nauki – i jest on dzisiaj, po 10 latach publikowania, nadszpodziewanie wysoki zarówno dla kobiet, jak i dla mężczyzn.

W tym kontekście Polska wypada bardzo dobrze: polski system nauki opiera się na rzadko spotykanej w najbogatszych gospodarkach świata stabilności zatrudnienia, co stanowi o atrakcyjności akademickiego miejsca pracy. Warto w tym miejscu wspomnieć również o niezwyklej stabilności klas produktywności w polskiej nauce z perspektywy długich karier akademickich (na przykładzie 2326 polskich profesorów tytularnych publikujących 25–40 lat, Kwiek i Roszka 2024).

Zmiany dotyczące uczestnictwa kobiet w nauce w ostatnich trzech dekadach w dyscyplinach z obszaru STEMM miały charakter tektoniczny (Kwiek i Szymula, 2023). Badania podłużne wspierane przez ustrukturyzowane, komercyjne Big Data o charakterze bibliometrycznym otwierają nowe możliwości globalnego ujmowania nauki i naukowców. W szybko zmieniającym się środowisku naukowym (Fochler et al., 2016; Wang & Barabási, 2021; Stephan, 2012), z setkami tysięcy nowych osób rozpoczynających karierę naukową każdego roku, nasze tradycyjne założenia dotyczące tego, jak mężczyźni i kobiety odchodzą z nauki, mogą wymagać ostrożnych korekt. Naszym zamiarem było nakreślenie ogólnych odpowiedzi na dawne pytania w oparciu o do niedawna jeszcze niedostępne Big Data i wskazanie możliwych kierunków dalszych, bardziej szczegółowych analiz.

Prezentowane studium nie jest rzecz jasna pozbawione ograniczeń. Nasze badania jednoznacznie skupiają się na publikujących naukowcach reprezentujących wszystkie sektory, a nie na kadrze akademickiej, dominującej w tradycyjnych opisach zjawiska rezygnacji z kariery naukowej ze względu na płeć (White-Lewis et al., 2023; Zhou & Volkwein, 2004; Deutsch & Yao, 2014; Goulden et al., 2011; Kaminski & Geisler, 2012).

Aktywne uczestnictwo w nauce jest definiowane w jednoznaczny sposób: jako publikowanie wyników badań w globalnym obiegu czasopism i idei. Nauka globalna to zatem w naszym ujęciu nauka indeksowana w komercyjnej bazie Scopus (zob. monografię Kwiek 2022). Tym samym niepublikowanie jest definiowane jako rezygnacja z nauki, zgodnie z Mertonowską tradycją w socjologii nauki. Chociaż niepublikujący naukowcy w dziedzinach STEMM mogą kontynuować pracę w sektorze akademickim na innych stanowiskach lub pracować w innych sektorach, to nie ma aktualnie możliwości weryfikacji ich zatrudnienia na poziomie globalnym przy użyciu dostępnych zbiorów danych.

Konceptualizacja rezygnacji z nauki jako zaprzestanie publikowania rzecz jasna pomija inne role akademickie (prowadzenie dydaktyki czy zaangażowanie w pracę administracyjną) i pozaakademickie (praca w startupach, korporacjach czy instytucjach publicznych). Rezygnacja z nauki jest trudnym do zdefiniowania pojęciem, ponieważ mamy do czynienia z porołatymi granicami oddzielającymi uprawianie nauki od miejsc pracy z nim niezwiązanych (Preston, 2004).

W tekście omawiamy cały cykl życia akademickiego od wejścia do nauki do jej opuszczenia, poprzez wskaźniki pomocnicze określające opublikowanie pierwszej i opublikowanie ostatniej pracy naukowej (indeksowanej w bazie danych Scopus). Tym samym sekwencja następujących po sobie indeksowanych publikacji zastępuje sekwencję znacznie szerszych procesów poznawczych i społecznych obejmujących różne wymiary uprawiania nauki (Sugimoto & Larivière, 2023). Zakładamy zatem na potrzeby tego badania, że brak publikacji w czasopismach naukowych jest równoznaczny z zaprzestaniem uprawiania nauki. Dlatego też nasze modele kariery naukowej są z konieczności uprosz-



czone, a nasza interpretacja dorobku naukowego – z konieczności zredukowana do globalnie indeksowanych publikacji (globalnej nauki). W indywidualnych portfelach publikacyjnych nieindeksowane publikacje (i większość publikacji w językach innych niż angielski) z konieczności nie są brane pod uwagę. Stąd nasza koncentracja na dziedzinach STEMM i pominięcie dziedzin humanistycznych i społecznych, generalnie słabo reprezentowanych w tak rozumianej nauce globalnej (w wyjątkami).

Pomijamy zatem szeroki zakres działań naukowców w środowisku akademickim (takich jak opieka nad studentami, recenzowanie artykułów naukowych i wniosków grantowych czy redagowanie czasopism) (Liu i in., 2023). Ponadto zasadniczo za każdym razem, gdy dane są wykorzystywane w innym celu niż ich pierwotne przeznaczenie – a tak jest w naszym przypadku wykorzystywania surowych danych bibliometrycznych – należy uwzględniać również kwestie epistemiczne wyższego rzędu (Kashyap i in., 2022): komercyjne i urzędowe zbiory danych typu rejestry naukowców nie zostały pierwotnie stworzone przez naukowców w celach badawczych. Jednak lepszym zbiorem danych do badania rezygnacji z nauki w ujęciu globalnym aktualnie nie dysponujemy. Dysponujemy takimi danymi (OPI, RADON, Scopus, a wkrótce nowy portal Ludzie Nauki) w przypadku Polski – ale nie w interesującym nas ujęciu globalnym.

W ramach nowego naukoznawstwa (znanego w świecie jako *science of science*) znajdujemy się zatem w sytuacji permanentnego szacowania zysków i strat w prowadzonych badaniach: albo warto ze śledzenia śladów cyfrowych zostawianych przez naukowców korzystać, albo nie warto. Mamy unikalną okazję testować dawne teorie, używając Big Data, i proponować nowe teorie, korzystając z jeszcze do niedawna niedostępnych danych empirycznych. Osobiście największe nadzieje w badaniach karier akademickich pokładamy w komplementarności różnych metod badawczych (globalne badania ankietowe wspierane przez duże dane, globalne wywiady pogłębione, analiza ustrukturyzowanych Big Data). Naszym zdaniem zdecydowanie warto wykorzystywać możliwości badań longitudinalnych prowadzonych z wykorzystaniem danych bibliometrycznych – jednak ze świadomością istnienia ich na bieżąco diagnozowanych ograniczeń o różnym charakterze.

### **Podziękowania**

Marek Kwiek dziękuje gospodarzom i słuchaczom swoich międzynarodowych seminariów poświęconych globalnej rezygnacji z nauki: DZHW w Berlinie (Torger Möller, listopad 2023), University of Hong Kong (Hugo Horta, styczeń 2024), University of Oxford (Simon Marginson, styczeń 2024), Leiden University (Ludo Waltman, luty 2024) i Stanford University (John Ioannidis, marzec 2024). Marek Kwiek dziękuje za wsparcie udzielone w ramach grantu Nds/529032/2021/2021. Łukasz Szymula jest wdzięczny za grant NCN 2019/35/0/HS6/02591. Dziękujemy również za pomoc dr. Wojciechowi Roszce z Uniwersytetu Ekonomicznego w Poznaniu oraz Kristy James i Aliczowi Birdowi z ośrodka badawczego International Center for the Studies of Research (ICSR) Lab.

## Bibliografia

- Allison, P.D. (2014). *Event History and Survival Analysis*. Sage.
- Alper, J. (1993). The Pipeline Is Leaking Women All the Way Along. *Science*, Vol. 260, 16 April.
- Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a Curated, High-Quality Bibliometric Data Source for Academic Research in Quantitative Science Studies. *Quantitative Science Studies*, 1(1), 377–386. 10.1162/qss\_a\_00019
- Blickenstaff, J.C. (2005). Women and science careers: leaky pipeline or gender filter? *Gender and Education*, 17(4), 369–386.
- Boekhout, H., van der Weijden, I., & Waltman, L. (2021). Gender Differences in Scientific Careers: A Large-Scale Bibliometric Analysis. Preprint: <https://arxiv.org/abs/2106.12624>
- Britton, D. M. (2017). Beyond the Chilly Climate: The Salience of Gender in Women's Academic Careers. *Gender & Society*, 31(1), 5–27.
- Cornelius, R., Constantinople, A., & Gray, J. (1988). The Chilly Climate: Fact Or Artifact? *The Journal of Higher Education*, 59(5), 527–55.
- Deutsch, F.M., & Yao, B. (2014). Gender Differences in Faculty Attrition in the USA. *Community, Work & Family*, 17, 392–408.
- Ehrenberg, R.G., Kasper, H., & Rees, D.I. (1991). Faculty Turnover in American Colleges and Universities. *Economics of Education Review*, 10(2), 99–110.
- Fochler, M., Felt, U., & Müller, R. (2016). Unsustainable Growth, Hyper-Competition, and Worth in Life Science Research: Narrowing Evaluative Repertoires in Doctoral and Postdoctoral Scientists' Work and Lives. *Minerva*, 54(2), 175–200.
- Geuna, A., & Shibayama, S. (2015). Moving out of Academic Research: Why Do Scientists Stop Doing Research? [w:] A. Geuna (Ed.), *Global Mobility of Research Scientists* (s. 271–297). Amsterdam: Elsevier.
- Goulden, M., Mason, M.A., & Frasch, K. (2011). Keeping women in the science pipeline. *The ANNALS of the American Academy of Political and Social Science*, 638, 141–162 (2011).
- Hamermesh, D.S., & Pfann, G.A. (2011). Reputation and Earnings: The Roles of Quality and Quantity in Academe. *Economic Inquiry*, 50(1), 1–16.
- Hammarfelt, B. (2017). Recognition and Reward in the Academy: Valuing Publication Oeuvres in Biomedicine, Economics and History. *Aslib Journal of Information Management*, 69(5), 607–623.
- Heckman, J.J., & Moktan, S. (2018). Publishing and Promotion in Economics. The Tyranny of the Top Five. *NBER Working Paper* 25093.
- Ioannidis, J.P.A., Boyack, K.W., & Klavans, R. (2014). Estimates of the Continuously Publishing Core in the Scientific Workforce. *PLOS One*, 9(7), e101698.
- Kaminski, D., & Geisler, C. (2012). Survival Analysis of Faculty Retention in Science and Engineering by Gender. *Science*, 335, 864–866.
- Kanter, R.M. (1977). Some Effects of Proportions on Group Life: Skewed Sex Ratios and Responses to Token Women. *American Journal of Sociology*, 82(5), 965–990.
- Kashyap, R., Rinderknecht, R.G., Akbaritabar, A., Alburez-Gutierrez, D., Gil-Clavel, S., Grow, A., Zhao, X., et al. (2022). *Digital and Computational Demography*. <https://doi.org/10.31235/osf.io/7bvypt>
- King, M.M., Bergstrom, C.T., Correll, S.J., Jacquet, J., & West, J.D. (2017). Men Set their Own Cites High: Gender and Self-Citation across Fields and over Time. *Socius*, 3.
- Kwiek, M. (2016). The European Research Elite: A Cross-National Study of Highly Productive

- Academics Across 11 European Systems. *Higher Education*, 71(3), 379–397. <https://doi.org/10.1007/s10734-015-9910-x>
- Kwiek, M. (2022). *Globalna nauka, globalni naukowcy*. Warszawa: PWN.
- Kwiek, M., & Roszka, W. (2021a). Gender Disparities in International Research Collaboration: A Large-Scale Bibliometric Study of 25,000 University Professors. *Journal of Economic Surveys*, 35(5), 1344–1388. <https://doi.org/10.1111/joes.12395>
- Kwiek, M., & Roszka, W. (2021b). Gender-Based Homophily in Research: a Large-Scale Study of Man-Woman Collaboration. *Journal of Informetrics*, 15(3), 1–38. <https://doi.org/10.1016/j.joi.2021.101171>
- Kwiek, M., & Roszka, W. (2022). Are Female Scientists Less Inclined to Publish Alone? The Gender Solo Research Gap. *Scientometrics*, 127, 1697–1735. <https://doi.org/10.1007/s11192-022-04308-7>
- Kwiek, M., Roszka, W. (2023). The Young and the Old, the Fast and the Slow: A Large-Scale Study of Productivity Classes and Rank Advancement. *Studies in Higher Education*. Online first: <https://doi.org/10.1080/03075079.2023.2288172>
- Kwiek, M., & Roszka, W. (2024). Once Highly Productive, Forever Highly Productive? Full Professors' Research Productivity from a Longitudinal Perspective. *Higher Education*. Online first: <https://doi.org/10.1007/s10734-023-01022-y>
- Kwiek, M., & Szymula, L. (2023). Young Male and Female Scientists: A Quantitative Exploratory Study of the Changing Demographics of the Global Scientific Workforce. *Quantitative Science Studies*, 4 (4): 902–937. [https://doi.org/10.1162/qss\\_a\\_00276](https://doi.org/10.1162/qss_a_00276)
- Kwiek, M., & Szymula, L. (2024). Quantifying Lifetime Productivity Changes: A Longitudinal Study of 325,000 Late-Career Scientists. Preprint SocArXiv: <https://osf.io/preprints/socarxiv/y2a5e>
- Levine, R.B., Lin, F., Kern, D.E., Wright, S.M., & Carrese, J. (2011). Stories from Early-Career Women Physicians Who Have Left Academic Medicine: a Qualitative Study at a Single Institution. *Academic Medicine*, 86, 752–758.
- Lindahl, J. (2018). Predicting Research Excellence at the Individual Level: The Importance of Publication Rate, Top Journal Publications, and Top 10% Publications in the Case of Early Career Mathematicians. *Journal of Informetrics*, 12(2), 518–533.
- Liu, L., Jones, B.F., Uzzi, B., et al. (2023). Data, Measurement and Empirical Methods in the Science of Science. *Nature Human Behaviour*, 7, 1046–1058.
- Maranto, C. L., & Griffin, A. E. (2010). The antecedents of a “chilly climate” for women faculty in higher education. *Human Relations*, 64(2), 139–159.
- Mihaljević, H., & Santamaría, L. (2020). Authorship in Top-Ranked Mathematical and Physical Journals: Role of Gender on Self-Perceptions and Bibliographic Evidence. *Quantitative Science Studies*, 1(4), 1468–1492.
- Mills, M. (2011). *Introducing Survival and Event History Analysis*. Sage.
- Milojevic, S., Radicchi, F., & Walsh, J.P. (2018). Changing Demographics of Scientific Careers: the Rise of the Temporary Workforce. *Proceedings of the National Academy of Sciences*, 115. 12616–12623.
- Nielsen, M.W., & Andersen, J.P. (2021). Global Citation Inequality is on the Rise. *Proceedings of the National Academy of Sciences*, 118(7), e2012208118.
- Nygaard, L.P., Piro, F., & Aksnes, D. (2022). Gendering Excellence Through Research Productivity Indicators. *Gender and Education*, 34(6), 690–704.

- O'Meara, K., Bennett, J.C., & Niehaus, E. (2016). Left Unsaid: The Role of Work Expectations and Psychological Contracts in Faculty Careers and Departure. *The Review of Higher Education*, 39(2), 269–297
- Preston, A.E. (2004). *Leaving Science. Occupational Exit from Scientific Careers*. Russell Sage Foundation.
- Rosser, V.J. (2004). Faculty Members' Intentions to Leave: A National Study on their Worklife and Satisfaction. *Research in Higher Education*, 45(3), 285–309.
- Santos, J. M., Horta, H., & Amâncio, L. (2020). Research agendas of female and male academics: a new perspective on gender disparities in academia. *Gender and Education*, 1–19.
- Shaw, A.K., & Stanton, D.E. (2012). Leaks in the Pipeline: Separating Demographic Inertia from Ongoing Gender Differences in Academia. *Proceedings of the Royal Society B. Biological Sciences*, 279(1743), 3736–41.
- Shibayama, S., & Baba, Y. (2015). Impact-Oriented Science Policies and Scientific Publication Practices: The Case of Life Sciences in Japan. *Research Policy*, 44(4), 936–950.
- Smart, J.C. (1990). A Casual Model of Faculty Turnover Intentions. *Research in Higher Education*, 31(5), 405–424.
- Spoon, K. et al. (2023). Gender and Retention Patterns Among U.S. Faculty. *Science Advances*, 9, eadi2205. DOI:10.1126/sciadv.adi2205
- Stephan, P.E. (2012). *How Economics Shapes Science*. Harvard University Press.
- Sugimoto, C., & Larivière, V. (2023). *Equity for Women in Science. Dismantling Systemic Barriers to Advancement*. Harvard University Press.
- Tang, L. & Horta, H. (2023). Supporting Academic Women's Careers: Male and Female Academics' Perspectives at a Chinese Research University. *Minerva*. <https://doi.org/10.1007/s11024-023-09506-y>
- Wang, D., & Barabási, A.-L. (2021). *The Science of Science*. Cambridge University Press.
- White-Lewis, D.K., O'Meara, K., Mathews, K., et al. (2023). Leaving the Institution or Leaving the Academy? Analyzing the Factors that Faculty Weigh in Actual Departure Decisions. *Research in Higher Education*, 64, 473–494.
- Wohrer, V. (2014). To Stay or to Go? Narratives of Early-Stage Sociologists About Persisting in Academia. *Higher Education Policy*, 27, 469–487.
- Wolfinger, N.H., Mason, M.A., & Goulden, M. (2008). Problems in the Pipeline: Gender, Marriage, and Fertility in the Ivory Tower. *Journal of Higher Education*, 79(4), 388–405.
- Xu, Y.J. (2008). Gender Disparity in STEM Disciplines: A Study of Faculty Attrition and Turnover Intentions. *Research in Higher Education* 49, 607–624.
- Zhou, Y., & Volkwein, J.F. (2004). Examining the influence on faculty departure intentions: A comparison of tenured versus nontenured faculty at research universities using NSOPF-99. *Research in Higher Education*, 45(2), 139–176.

## Załącznik

Tabela 4. Mediana długości przerw w publikowaniu dla kohorty 2000 dla subpopulacji naukowców publikujących w ciągu ostatnich 12 badanych lat (2010–2021), odsetek naukowców mężczyzn (lewy panel) i kobiet (prawy panel). Zerowa przerwa oznacza publikowanie co najmniej jednego artykułu rocznie; roczna przerwa oznacza publikowanie co najmniej jednego artykułu co drugi rok; dwuletnia przerwa oznacza publikowanie co najmniej jednego artykułu co trzeci rok itd. (w procentach)

	Mężczyźni					Kobiety				
	0 lat	1 rok	2 lata	3 lata	4 lata i więcej	0 lat	1 rok	2 lata	3 lata	4 lata i więcej
AGRI	85,75	8,34	2,65	0,90	2,36	84,29	9,75	2,29	1,26	2,41
BIO	87,52	8,23	2,27	0,74	1,24	81,78	11,28	3,07	1,94	1,93
CHEM	86,85	7,92	2,00	1,08	2,15	85,09	8,80	2,68	0,96	2,47
CHEMENG	85,05	7,48	3,74	1,87	1,86	82,14	14,29	–	–	3,57
COMP	88,51	7,20	1,94	0,80	1,55	89,58	6,55	2,38	0,30	1,19
EARTH	86,25	7,15	2,29	1,92	2,39	86,51	8,73	1,59	1,06	2,11
ENER	83,23	11,61	2,58	1,94	0,64	81,25	12,50	–	6,25	0,00
ENG	83,22	9,41	3,40	1,52	2,45	86,33	7,03	3,12	1,17	2,35
ENVIR	84,71	7,48	4,62	1,11	2,08	82,98	11,70	2,13	1,42	1,77
IMMU	86,44	8,47	3,39	1,13	0,57	79,26	10,11	3,72	5,32	1,59
MATER	89,09	7,44	1,65	0,83	0,99	88,16	7,89	2,63	0,66	0,66
MATH	82,43	13,18	2,72	1,05	0,62	75,00	15,32	6,45	1,61	1,62
MED	86,18	8,14	2,21	1,26	2,21	83,12	10,00	2,84	1,44	2,60
NEURO	89,58	7,71	1,25	0,83	0,63	85,53	9,54	2,30	0,66	1,97
PHARM	80,99	12,40	1,65	1,65	3,31	82,43	9,46	5,41	2,70	0,00
PHYS	90,67	5,28	2,01	0,63	1,41	90,62	4,79	2,40	0,80	1,39
<b>Łącznie</b>	<b>86,70</b>	<b>7,96</b>	<b>2,36</b>	<b>1,10</b>	<b>1,88</b>	<b>83,66</b>	<b>9,82</b>	<b>2,79</b>	<b>1,47</b>	<b>2,26</b>

Tabela 5. Mediana długości przerw w publikowaniu dla kohorty 2010 dla subpopulacji naukowców publikujących w ciągu 12 badanych lat (2010–2021), odsetek naukowców mężczyzn (lewy panel) i kobiet (prawy panel) (w procentach)

	Mężczyźni					Kobiety				
	0 lat	1 rok	2 lata	3 lata	4 lata i więcej	0 lat	1 rok	2 lata	3 lata	4 lata i więcej
AGRI	62,30	18,97	7,44	3,59	7,70	62,45	19,69	7,68	3,58	6,60
BIO	68,78	17,94	6,23	2,85	4,20	64,36	20,96	6,92	3,24	4,52
CHEM	72,49	15,27	5,50	2,57	4,17	70,11	16,92	5,96	2,89	4,12
CHEMENG	71,67	15,80	5,35	2,35	4,83	65,30	21,27	5,60	1,12	6,71
COMP	74,07	15,80	4,83	2,01	3,29	74,20	15,01	5,77	2,03	2,99
EARTH	65,17	17,76	6,66	3,82	6,59	64,54	19,19	6,93	3,52	5,82
ENER	64,95	19,20	7,44	2,43	5,98	68,04	19,30	5,06	2,85	4,75

ENG	70,92	16,45	5,51	2,74	4,38	70,17	16,39	6,00	2,53	4,91
ENVIR	62,88	18,68	7,93	3,29	7,22	61,38	20,18	7,87	3,91	6,66
IMMU	65,95	17,88	7,97	2,85	5,35	66,77	19,65	6,94	3,25	3,39
MATER	74,41	15,26	5,20	2,27	2,86	74,24	16,18	4,67	2,04	2,87
MATH	66,87	18,39	7,26	3,38	4,10	61,59	23,34	7,78	3,48	3,81
MED	66,53	17,15	6,49	3,31	6,52	63,44	18,46	7,15	3,71	7,24
NEURO	68,91	18,86	5,29	2,94	4,00	65,77	20,23	6,27	2,98	4,75
PHARM	61,90	20,30	6,39	3,88	7,53	59,58	20,56	7,78	3,47	8,61
PHYS	80,65	12,28	3,25	1,53	2,29	81,83	11,68	3,25	1,19	2,05
<b>Łącznie</b>	<b>69,55</b>	<b>16,66</b>	<b>5,89</b>	<b>2,85</b>	<b>5,05</b>	<b>65,10</b>	<b>18,70</b>	<b>6,85</b>	<b>3,36</b>	<b>5,99</b>

### Znikający naukowcy. Co ustrukturyzowane Big Data mówią nam o rezygnacji z nauki w 38 krajów OECD?

W pracy analizujemy zjawisko rezygnacji z nauki akademickiej i pokazujemy, jak odchodzenie z nauki różni się między kobietami i mężczyznami, dyscyplinami akademickimi i na przestrzeni czasu. Prezentowane podejście jest kompleksowe: globalne, oparte na kohortach naukowców, i podłużne – obserwujemy działalność publikacyjną indywidualnych naukowców w czasie. Korzystając z metadanych pochodzących z bazy Scopus – globalnej bibliometrycznej bazy danych publikacji i cytowań – analizujemy kariery publikacyjne naukowców z 38 krajów OECD, którzy rozpoczęli publikowanie w 2000 r. ( $N = 142\,776$ ) i w 2010 r. ( $N = 232\,843$ ). W pracy przetestowano przydatność dużych zbiorów danych bibliometrycznych do globalnych analiz karier naukowych.

**Słowa kluczowe:** kariera akademicka, rezygnacja z nauki, ustrukturyzowane Big Data, estymator Kaplana-Meiera, analiza przeżycia, nauka o nauce, naukoznawstwo, kraje OECD

### Disappearing Scientists. What Structured Big Data Tell Us about Leaving Science in 38 OECD Countries?

In this paper we analyze the phenomenon of quitting academic science and show how quitting differs between men and women, academic disciplines and over time. The approach presented is comprehensive: global, based on cohorts of scientists, and longitudinal – we observe the publication activity of individual scientists over time. Using metadata from Scopus, a global bibliometric database of publications and citations, we analyze the publication careers of scientists from 38 OECD countries who began publishing in 2000 ( $N = 142\,776$ ) and 2010 ( $N = 232\,843$ ). The paper tests the usefulness of large bibliometric datasets for a global analysis of academic careers.

**Key words:** academic career, attrition in science, structured Big Data, Kaplan-Meier estimator, survival analysis, science of science, OECD countries