

MAREK KWIEK\*, WOJCIECH ROSZKA\*\*

## Globalny obieg naukowy a wiek w nauce: analiza 20 000 polskich naukowców

### 1. Wprowadzenie

Wiek biologiczny jest istotnym czynnikiem demograficznym w badaniach karier akademickich (na przykład produktywności badawczej, mobilności międzynarodowej, wpływu działalności naukowej na rozwój nauki i społeczeństwa czy też wzorców współpracy w badaniach naukowych). Przyjmuje się, że w tych badaniach „wiek akademicki” – czyli liczba lat, jaka upłynęła od wydania pierwszej publikacji – jest dobrym zamiennikiem czy też przybliżeniem wieku biologicznego. W prezentowanym tekście analizujemy ograniczenia związane ze stosowaniem takiego zamiennika w przypadku kraju od niedawna uczestniczącego w szerszym zakresie w globalnym obiegu naukowym, czyli Polski. Bierzymy za przykład pełną populację polskich naukowców akademickich widocznych w ostatniej dekadzie w nauce światowej i posiadających co najmniej stopień naukowy doktora ( $N = 20\,569$ ). Okazuje się, że zamiennik ten dobrze się sprawdza w przypadku dyscyplin ścisłych, technicznych, inżynierskich, matematycznych i medycznych (określanych dalej jako dyscypliny STEMM), natomiast w przypadku dyscyplin spoza STEMM (a w szczególności w przypadku nauk humanistycznych i społecznych), wypada on radykalnie gorzej.

Ten negatywny wniosek jest szczególnie ważny dla systemów nauki, które dopiero niedawno stały się bardziej rozpoznawalne w globalnych czasopismach naukowych. Analiza polskich danych z poziomu poszczególnych naukowców – z jednej strony demograficznych (z baz OPI PIB), a z drugiej publikacyjnych (z bazy Scopus) – sugeruje opóźnione uczestnictwo naukowców z dyscyplin społecznych i humanistycznych w globalnych sieciach naukowych, co niesie z sobą praktyczne implikacje dla przewidywania wieku biologicznego na podstawie wieku akademickiego w badaniach karier akademickich.

Obliczamy w tym tekście współczynniki korelacji, przedstawiamy analizę kontyngencji etapów kariery akademickiej ze stopniami i tytułem naukowym i grupami wieku oraz tworzymy liniowy model regresji wielorakiej. Nasze badania wskazują, że w krajach roz-

---

\* Prof. dr hab. Marek Kwiek (kwiekm@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)

\*\* Dr Wojciech Roszka (wojciech.roszka@ue.poznan.pl), Uniwersytet Ekonomiczny w Poznaniu  
ORCID: [orcid.org/0000-0003-4383-3259](https://orcid.org/0000-0003-4383-3259)

wijających się naukowo wiek akademicki jako przybliżenie wieku biologicznego musi być stosowany ostrożniej niż w krajach rozwiniętych naukowo: najlepiej, aby w badaniach profesji akademickiej był stosowany tylko dla dyscyplin STEM.

W poprzedniej dekadzie badania karier akademickich prowadzone na poziomie pojedynczego naukowca, a nie na zagregowanym poziomie kraju, coraz częściej odwoływały się do kategorii „wieku akademickiego” (zob. Gingras i in. 2008; Aksnes i in. 2011; Abramo et al. 2016). Powód jest prosty: dane dotyczące wieku biologicznego naukowców są na ogół niedostępne dla badań prowadzonych na dużą skalę. A tymczasem wiek jest jedną z najważniejszych zmiennych niezależnych związanych z licznymi wymiarami kariery akademickiej (Levin & Stephan 1991; Levin & Stephan 1989). Dlatego też często, niejako z konieczności, przyjmowano, że wiek akademicki jest dobrym zamiennikiem wieku biologicznego (Robinson-Garcia i in. 2020; Milojević 2012; Nane i in. 2017). Datę pierwszej publikacji w bazach Scopus czy Web of Science, korzystając z danych surowych można wyznaczyć w sposób jednoznaczny i wykorzystać do celów badawczych dla wszystkich publikujących naukowców na poziomie instytucji lub ich wybranych grup, miast, regionów, dyscyplin, czasopism i krajów. W naszym przypadku skoncentrujemy się na różnicach między dwoma typami uczelni (10 uczelni wybranych do programu IDUB vs. uczelnie pozostałe), kobietami i mężczyznami naukowcami, etapami kariery naukowej (od doktoratu do profesury tytularnej) i dyscyplinami naukowymi.

Dostęp do informacji na temat tak zdefiniowanego wieku akademickiego, czyli do daty pierwszej publikacji naukowców w dużej skali (na przykład dla wszystkich naukowców akademickich w całym krajowym systemie nauki) wymaga wprawdzie skomplikowanych przygotowań, w szczególności uwzględnienia listy indywidualnych identyfikatorów wszystkich autorów (typu Scopus ID), ale jest technicznie możliwy. Natomiast dostęp do danych na temat wieku biologicznego naukowców – przy obecnym poziomie rozwoju dużych administracyjnych baz danych obejmujących wszystkich naukowców w danym kraju i dostępie do nich w celach naukowych – w praktyce nie istnieje (poza nielicznymi wyjątkami, do których należą Włochy i Norwegia, zob. Rørstad & Aksnes 2015; Abramo i in. 2011). Do wyjątków należy również Polska i stworzona przez nas baza danych („Obserwatorium Polskiej Nauki”). Nasza baza powstała w oparciu o integrację administracyjnego rejestru wszystkich naukowców pracujących w publicznych sektorach nauki (ok. 100 tys. osób) – prowadzonego przez powołaną w tym celu instytucję krajową (OPI-PIB) – z bazą wszystkich polskich publikacji indeksowanych w bazie Scopus w poprzedniej dekadzie (2009–2018, ok. 380 tys.). Integracja obu baz danych została przeprowadzona przy użyciu metod probabilistycznych i deterministycznych, a jej budowę opisujemy bardziej szczegółowo w innym miejscu (Kwiek i Roszka 2021a).

W niniejszym tekście analizujemy ograniczenia stosowania wieku akademickiego jako zamiennika wieku biologicznego na przykładzie krajowego systemu nauki, dla któ-

regu rozpatrujemy zarówno wiek biologiczny, jak i wiek akademicki wszystkich badaczy ze wszystkich polskich uczelni, posiadających co najmniej stopień naukowy doktora i uczestniczących w globalnej nauce akademickiej poprzez publikacje międzynarodowe. Przybliżeniem funkcjonowania naukowca w globalnej nauce jest posiadanie co najmniej jednej publikacji indeksowanej w bazie Scopus w analizowanej dekadzie 2009–2018.

Ponieważ dysponujemy jednoznacznie określonym wiekiem biologicznym (pochodzącym z rejestru naukowców) dla każdego badacza oraz jesteśmy w stanie określić wiek akademicki dla niemal wszystkich naukowców obecnych w bazie Scopus (a jedynym ograniczeniem jest działanie protokołu API w tej bazie), możemy w sposób systematyczny pokazać różnice pomiędzy tymi dwoma typami wieku w zależności od wybranych kluczowych parametrów badania. To unikalna analiza również pod względem skali – przeprowadzamy ją dla ponad 20 000 naukowców na podstawie dwóch baz danych (administracyjnej i komercyjnej) o prawie idealnej (OPI PIB) i bardzo dobrej (Scopus) jakości danych.

Jednym z ograniczeń naszego podejścia jest nieobecność części badaczy akademickich w bazie Scopus – brak przypisanych im indeksowanych publikacji związany jest z opóźnionym wejściem polskiego systemu akademickiego do światowego obiegu naukowego. Cecha ta jest jednak wspólna wielu krajom nadrabiającym naukowe zaległości w skali globalnej, co sprawia, że część empiryczna pracy jest istotna dla dużej części świata przechodzącego podobne zmiany wzorców publikacyjnych wraz z upływem czasu do zmian zachodzących w Polsce.

Innymi słowy, określamy tutaj wiek akademicki dla każdego polskiego naukowca obecnego w naszej bazie „Obserwatorium Polskiej Nauki” na podstawie daty jego lub jej pierwszej publikacji i w sposób systemowy analizujemy różnice między w taki sposób zdefiniowanym wiekiem akademickim a rzeczywistym wiekiem biologicznym. Ponadto pokazujemy, jak indywidualnie określany wiek akademicki różni się w praktyce od ściśle określonego wieku biologicznego w zależności od takich zmiennych niezależnych, jak: płeć, dyscyplina naukowa, stopień i tytuł naukowy oraz typ instytucji.

W ten sposób, na przykładzie kompleksowych danych pochodzących z całego krajowego systemu nauki szacujemy zakres ograniczeń stosowania wieku akademickiego jako zamiennika wieku biologicznego w zależności od wybranych zmiennych niezależnych oraz analizujemy zarówno praktyczne, jak i metodologiczne konsekwencje stosowania wieku akademickiego w badaniach karier akademickich, stanowiących jeden z najważniejszych komponentów ilościowych badań szkolnictwa wyższego i nauki.

Jak wynika z naszych analiz, ograniczenia stosowalności wieku akademickiego okazują się największe dla nauk humanistycznych i społecznych: dla tych dwóch dyscyplin poziom korelacji obu typów wieku okazuje się niski. W analizowanym tu polskim przypadku nie jest więc w praktyce możliwe wnioskowanie o wieku biologicznym naukow-

ców na podstawie ich wieku akademickiego dla humanistów i reprezentantów szeroko pojętych nauk społecznych, gdyż korelacja między nimi waha się między bardzo słabą a słabą (współczynnik korelacji wynosi 0,35 dla nauk humanistycznych, 0,43 dla ekonomii, ekonometrii i finansów i 0,49 dla psychologii). Z kolei w przypadku dyscyplin ścisłych poziom korelacji (i determinacji) jest zdecydowanie wyższy, a zastosowanie wieku akademickiego jako zamiennika wieku biologicznego jest obarczone mniejszą niepewnością (współczynnik korelacji waha się od 0,82 dla biochemii, genetyki i biologii molekularnej do 0,89 dla chemii i 0,90 dla immunologii i mikrobiologii). Co więcej, korelacje i determinacje przyjmują różne wartości dla różnych parametrów indywidualnych i instytucjonalnych, co szczegółowo opisujemy poniżej.

Obecność polskich naukowców w globalnym systemie nauki systematycznie rośnie w ujęciu liczbowym i procentowym, ale jest silnie zróżnicowana pod względem dyscyplinarnym. O ile globalna widoczność nauki w obszarze STEMM – w szczególności w chemii i fizyce – jest stabilna i trwała, o tyle obecność nauk humanistycznych, społecznych i ekonomicznych w globalnych zbiorach czasopism indeksowanych jest stosunkowo niewielka i obserwowana dopiero po 2010 r. Szczególnie liczba publikacji w wysoko prestiżowych czasopismach (top 10% i top 1% czasopism) rośnie silniej od stosunkowo niedawna, co z pewnością jest efektem dekady reform systemów nauki i szkolnictwa wyższego (zob. Antonowicz i in. 2020).

Wiek akademicki definiujemy tutaj jako czas, jaki upłynął między datą pierwszej publikacji indeksowanej w bazie Scopus (wszystkie typy publikacji, dowolna rola pełniona w publikacji) a rokiem 2017, czyli naszym rokiem odniesienia. Można przyjąć w dużym przybliżeniu, że w naukach ścisłych pierwsza publikacja w bazie pojawia się najczęściej w ostatnich latach pisania pracy doktorskiej – czyli w polskim przypadku w wieku około 30 lat. Zatem w dużym uproszczeniu wiek akademicki równy 0 oznacza wiek biologiczny 30 lat, a wiek akademicki równy 40 lat – wiek biologiczny 70 lat. W tym sensie kariera akademicka dla naszych potrzeb analitycznych obejmowałaby maksymalnie 40 lat życia: przedział wieku akademickiego 0–40 lat i przedział wieku biologicznego 30–70 lat. Jednak, co najważniejsze, długa kariera w sektorze akademickim nie musi oznaczać długiego okresu publikowania – zwłaszcza, jak w prezentowanym tu ujęciu, publikowania artykułów indeksowanych w bazie Scopus (szczególnie w obszarze spoza STEMM), co odróżnia polski system od systemów dłużej i intensywniej obecnych w globalnym obiegu indeksowanej nauki.

W pierwszej części sekcji poświęconej wynikom naszych badań przedstawiamy szczegółowe oszacowanie skali niespójności zachodzących między dwoma typami wieku w różnych przekrojach oraz proponujemy szersze uogólnienia. Natomiast w drugiej części tej sekcji, wykorzystując model regresji liniowej, dokonujemy oszacowania wieku biologicznego na podstawie wieku akademickiego oraz innych wybranych parametrów

dostępnych bezpośrednio (afiliacja instytucjonalna) lub pośrednio w bazie Scopus (takich jak dyscyplina ASJC dominująca w 10-letnim indywidualnym portfelu publikacyjnym naukowca; płeć określana na podstawie imion i nazwisk oraz ewentualnie zewnętrznych baz danych i specjalnych algorytmów zajmujących się określaniem płci autorów publikacji). Ponadto analizujemy użyteczność stworzonego modelu, zastanawiając się, w jakim stopniu możemy prowadzić symulacje wieku biologicznego na podstawie metadanych publikacji dostępnych w globalnych bazach danych publikacyjnych i cytowaniowych.

## **2. Podstawowa literatura**

### **2.1. Wiek biologiczny w badaniach karier akademickich**

Wiek biologiczny jest ważnym czynnikiem demograficznym w socjologicznych badaniach karier akademickich od ponad pół wieku (zob. zwłaszcza Lehman, 1953; Zuckerman & Merton, 1973; Pelz & Andrews, 1976; Stern, 1978; Cole, 1979; Kyvik, 1990; Levin & Stephan, 1991; Stephan & Levin, 1992; Kyvik & Olsen, 2008). Bardziej współcześnie badania produktywności kontynuowano, wykorzystując wiek jako zmienną niezależną: zakrojone na szeroką skalę badanie 11 500 norweskich naukowców pokazało, że produktywność rośnie z wiekiem, osiągając szczyt na późnym etapie kariery, a następnie spada, wskazując zarazem, że naukowcy po 60. roku życia są znacznie mniej cytowani niż ich młodsi koledzy (Aksnes i in. 2011: 42–44). Ponadto badanie 6388 profesorów uniwersyteckich z Quebecu (Gingras i in. 2008: 6–7) wykazało, że produktywność najpierw gwałtownie rośnie z wiekiem, a następnie rośnie w wolniejszym tempie od około 40. życia, osiągając swój szczyt około 50. roku życia. Jednak średnia liczba publikacji w prestiżowych czasopismach i publikacji wysoko cytowanych stale rośnie aż do emerytury, a starsi profesorowie częściej są liderami zespołów badawczych. Inne badanie koncentrujące się na Quebecu, wykorzystujące dane dotyczące wieku, dowodzi (Larivière i in. 2011), że po przekroczeniu wieku 38 lat kobiety otrzymują średnio mniejsze środki na badania niż mężczyźni, są mniej produktywne i nieco rzadziej cytowane. Badania produktywności naukowej wskazują na silne powiązania produktywności z wiekiem – stratyfikację kadry według wieku (Kwiek 2019a: 198–2018, 2019b).

Analizując włoskich profesorów tytularnych, Abramo ze współpracownikami (2016: 318) stwierdził, że zasadniczo produktywność znacząco spada z wiekiem. Jednak profesorowie mianowani w młodym wieku mają większe szanse na utrzymanie i zwiększenie produktywności niż ich koledzy profesorowie awansowani w późniejszym wieku. Stwierdzono w tym badaniu, że ujemna monotoniczna zależność między wiekiem a produktywnością badawczą towarzyszy dodatniej zależności między stanowiskiem a produktywnością. Ponadto produktywność naukowców w zależności od wieku została również porównana z produktywnością liderów pod względem produktywności. W bibliometrycznej analizie naukowców z hiszpańskiej Krajowej Rady ds. Badań Naukowych, przeprowa-

dzonej w oparciu o podział na klasy produktywności (naukowcy podzieleni na klasy o najwyższych, średnich i niskich osiągnięciach naukowych), Costas i współpracownicy (2010) stwierdzili, że produktywność naukowców o najwyższych i średnich osiągnięciach z wiekiem rośnie lub pozostaje stabilna, a produktywność naukowców o niskich wynikach zazwyczaj z wiekiem maleje (Costas i in. 2010: 1578).

Z kolei Bonaccorsi i Daraio (2003: 75) przeprowadzili badanie wieku i produktywności naukowców z włoskiej Narodowej Rady ds. Badań Naukowych, które wykazało, że produktywność spada z wiekiem i że średnia wieku włoskich naukowców rośnie, co niesie z sobą poważne implikacje dla krajowych systemów naukowych (prowadząc do zagrożenia zazwyczaj ujmowanego poprzez hasło „starzenia się kadry akademickiej”). Z kolei Guns, Eykens i Engels (2019) przeanalizowali niedawno różne wzorce publikacyjne kolejnych kohort wiekowych na wydziale nauk społecznych i humanistycznych jednego z belgijskich uniwersytetów we Flandrii i wykazali, że najstarsza kohorta najwyraźniej utrzymuje tradycyjny wzorzec skoncentrowany na publikacjach książkowych, przy czym użyteczne okazało się rozróżnienie między zmianą wewnątrz kohort i sukcesją między kohortami. I wreszcie Rørstad i Aksnes (2015), studiując poziom produktywności w zależności od wieku, płci i stanowiska akademickiego na dużej próbie 12400 norweskich profesorów uniwersyteckich, pokazali, że stanowisko akademickie ma większy wpływ na produktywność niż wiek i płeć.

## **2.2. Wiek akademicki jako zamiennik wieku biologicznego**

Choć wiek biologiczny jest ważną zmienną demograficzną w badaniach karier akademickich, to wiek akademicki jako jego zamiennik nie był do tej pory szeroko wykorzystywany w tych badaniach. Głównym powodem takiego stanu rzeczy jest z pewnością ograniczony dostęp do kompleksowych, rzetelnych danych dotyczących pierwszej publikacji naukowca. Z wyjątkiem stosunkowo ograniczonej liczby badań podsumowanych powyżej, w których wiek biologiczny był rzeczywiście dostępny, wiek biologiczny był badany za pomocą dwóch głównych zamienników: pierwszym był wiek akademicki związany z datą pierwszej publikacji, a drugim wiek akademicki związany z datą uzyskania stopnia doktora.

Po pierwsze, w badaniach wykorzystujących wiek akademicki, data pierwszej publikacji była stosowana kilkanaście razy (zob. np. Lee & Bozeman 2005; Milojević 2012; Radicchi & Castellano 2013; Nane i in. 2017; Robinson-Garcia i in. 2020; Costas i in. 2015; Aref i in. 2019; Chan & Torgler 2020; Simoes & Crespo 2020; Petersen 2015; Wildgaard 2015). Próby w tych badaniach były różnej wielkości, co skutkowało w kilku przypadkach problemami z uogólnianiem wyników uzyskanych na małych próbach na duże krajowe populacje. Po drugie, data otrzymania stopnia doktora jako sposób określania wieku akademickiego, zwykle pochodząca z biogramów naukowców, była wyko-

rzystywana jeszcze rzadziej i odnosiła się na ogół do mniejszych prób (Perianes-Rodriguez 2014; van den Besselaar & Sandström 2016; Badar i in. 2013; Sugimoto i in. 2016; Coomes i in. 2013). Po trzecie, i bardziej szczegółowo, wiek akademicki był wykorzystywany w różnych kontekstach w celu wsparcia różnych kierunków badań. Milojević (2012) badała zachowania cytowaniowe naukowców, a Radicchi i Castellano (2013) analizowali wskaźniki bibliometryczne na poziomie poszczególnych naukowców. Aref i współpracownicy (2019) przebadali wysoce mobilnych naukowców (nazywanych *super-movers*) i przedstawili przegląd ich kluczowych cech z podziałem na wiek akademicki. Simoes i Crespo (2020: 336) wykorzystali wiek akademicki w modelu pomiaru produktywności badawczej na poziomie autora, zakładając, że każda ocena produktywności na tym poziomie musi wymagać porównania naukowców na różnych etapach kariery. Chan & Torgler (2020) zastosowali wiek akademicki w kontekście różnic między kobietami i mężczyznami dotyczących osiągnięć elity naukowej w 43 krajach, a Robinson-Garcia i współpracownicy (2020) wykorzystali wiek akademicki w badaniach dotyczących podziału zadań w ramach kariery naukowej.

Nane i współpracownicy (2017) przeanalizowali prognozowanie wieku badaczy z wykorzystaniem danych bibliometrycznych i skupili się na mocy predykcyjnej różnych modeli regresji. Ocenili, na ile wiarygodne może być szacowanie faktycznego wieku naukowców, definiowanego jako ich wiek biologiczny, na podstawie modeli, które opierają się wyłącznie na różnych wskaźnikach bibliometrycznych w kontekście empirycznym badaczy z Quebecu, dla których dysponowali wiekiem biologicznym. Autorzy uznają swoją próbę za reprezentatywną dla naukowców w ogóle (Nane i in. 2017: 714); co więcej, próba ta ma reprezentować podstawowe globalne systemy produkcji wiedzy w ten sam sposób, w jaki nasza próba może reprezentować systemy nadal doganiające liderów w nauce (lub może być przydatna do ich zrozumienia). Podczas gdy Kanada jest zaliczana do „krajów zaawansowanych” w rankingu potencjału naukowego i infrastruktury 76 krajów, Polska jest w nim zaliczana do „krajów rozwijających się” (Wagner 2008: 88). Spośród licznych wskaźników bibliometrycznych rok pierwszej publikacji okazał się najlepszym liniowym estymatorem wieku poszczególnych badaczy (Nane i in. 2017: 726). Okazało się jednak, że funkcjonuje on szczególnie dobrze, gdy pracuje się z dużymi próbami naukowców (czyli najlepiej sprawdza się z „uśrednionej” perspektywy).

### **2.3. Ograniczenia metodologiczne i dotychczasowe wnioski praktyczne**

Wykorzystanie wieku akademickiego jako zamiennika wieku biologicznego rodzi szereg typowych pytań metodologicznych. Dotyczą one między innymi nieobiektywności zbioru danych (zwykle są to bazy Scopus lub Web of Science, które mają swoje własne uwarunkowania językowe, geograficzne i dyscyplinarne). Inne zbiory danych (w polskim przypadku na przykład PBN, Polska Bibliografia Naukowa) nie są powiązane z global-

nymi źródłami bibliometrycznymi i nie mogą być swobodnie wykorzystywane w globalnych badaniach porównujących dyscypliny naukowe, wykraczających poza pojedyncze instytucje czy poszczególne kraje.

Ponadto wiek akademicki może być używany tylko w przypadku naukowców publikujących (wszyscy niepublikujący, czyli badani przez nas dosyć szczegółowo w polskim przypadku *non-publishers*, są wykluczeni z próby, zob. Kwiek 2015a: 389–395). Dodatkowo naukowcy ci muszą publikować w czasopismach indeksowanych międzynarodowo, czyli niemal wyłącznie w języku angielskim. W skali globalnej, a w szczególności wśród krajów, które są nowymi uczestnikami globalnego systemu nauki, młodsze kohorty są ogólnie silniej obecne w globalnych bazach danych niż starsze kohorty, co ma wpływ na odchylenia w strukturze wiekowej próby. Jednakże, jeśli weźmie się pod uwagę konkurencyjne źródła danych globalnych (np. bazę danych Google Scholar), skrzywienie na korzyść stosunkowo młodych naukowców może być w nich jeszcze silniejsze (Radicchi & Castellano 2013).

Co więcej, niektóre pytania badawcze mogą być analizowane w sposób znacznie bardziej rzetelny w przypadku dużych zbiorów danych, dzięki czemu wiek akademicki jest bardziej uśredniony, a mniej rzetelny w indywidualnych przypadkach lub przypadkach ograniczających się do niewielkiej liczby obserwacji (Nane i in. 2017: 726). Kraje, które dopiero niedawno dołączyły w większej skali do globalnego obiegu nauki, oraz dyscypliny, które nadal traktują priorytetowo języki narodowe oraz książki i rozdziały w książkach (a nie artykuły w czasopismach) – takie jak znaczące segmenty nauk humanistycznych – nie mogą być rzetelnie badane przy użyciu wieku akademickiego jako przybliżenia wieku biologicznego (Kwiek 2021c). W tym przypadku wiek uzyskania stopnia doktora może być lepszym wyborem metodologicznym, o ile pozwala na to dostępność danych (zob. Guns et al. 2019). Jednak dostęp do dat obrony doktoratu na skalę całego krajowego systemu nauki jest technicznie możliwy tylko w kilku krajach – tych, w których wdrażane są Current Research Information Systems (CRIS), zazwyczaj powiązane z krajowymi systemami ewaluacji osiągnięć naukowych (Włochy, Norwegia i Polska).

### 3. Dane i metody

Wykorzystaliśmy w tej pracy wspólnie stworzoną, zintegrowaną bazę danych „Obserwatorium Polskiej Nauki”. Szczegółowy opis zbioru danych i etapy jego powstawania są opisane we wcześniejszej pracy (Kwiek i Roszka 2021a).

#### 3.1. Określanie płci naukowców

Nie było w tym tekście potrzeby określania płci polskich naukowców za pomocą np. różnych metod probabilistycznych o określonym, założonym poziomie prawdopodobieństwa determinacji, za pomocą zewnętrznych zbiorów danych z obserwacjami z góry zde-



terminowanymi pod kątem płci czy też przy użyciu specjalnego oprogramowania określającego płeć np. komercyjnego programu GenderizeR (zob. Wais 2016), jak to ma miejsce w badaniach zaprezentowanych w Kwiek i Roszka 2021b, które uwzględniały płeć międzynarodowych współpracowników badawczych polskich naukowców. W tym przypadku płeć jest określona dla wszystkich naukowców ze stopniem co najmniej doktora ze wszystkich uczelni obecnych w naszej bazie danych, w tym dla 20 596 naukowców w naszej ostatecznej próbie. Próba w tej pracy jest nieco mniejsza niż we wcześniejszych pracach z uwagi na niedoskonałości protokołu API, udostępnianego przez bazę Scopus i wykorzystanego przez nas do ustalenia daty pierwszej publikacji.

### **3.2. Określanie dyscyplin naukowych naukowców**

Konieczne w ramach naszych badań było jednoznaczne ustalenie dyscyplin naukowych reprezentowanych przez polskich naukowców (tzn. każdy naukowiec może reprezentować w ramach naszych analiz tylko jedną dyscyplinę). Ponieważ dane dotyczące zatrudnienia, doktoratów, habilitacji i profesur nie pozwalały na systemowe i jednoznaczne przypisanie każdego naukowca do jednej dyscypliny (a dodatkowo ministerialna lista dyscyplin obowiązujących w szkolnictwie wyższym i nauce ewoluowała w czasie wraz z kolejnymi rozporządzeniami), przyjęliśmy inną metodę.

Każdy polski naukowiec obecny w naszej zintegrowanej bazie danych został przypisany do jednej z 334 dyscyplin ASJC na poziomie czterocyfrowym i jednej z 27 dyscyplin ASJC na poziomie dwucyfrowym reprezentowanych w bazie Scopus (zob. Abramo, Aksnes & D'Angelo (2020), którzy określili kategorię tematyczną Web of Science dla każdego włoskiego i norweskiego profesora w swojej próbie). W systemie dyscyplin ASJC stosowanym w bazie Scopus publikacja w czasopiśmie może mieć jedną lub wiele klasyfikacji dyscyplinarnych pochodzących z klasyfikacji dyscyplin czasopism. Dominująca dyscyplina dla każdego naukowca została ustalona na podstawie wszystkich publikacji (typ: artykuł naukowy) obecnych w ich indywidualnych portfelach publikacyjnych w latach 2009–2018 (i określona przez dominantę, czyli wartość modalną; wartość najczęściej występująca). W przypadku, gdy nie występowała jedna dominanta, dyscyplinę wybierano w sposób losowy.

W ten sposób otrzymaliśmy ostatecznie dane polskich naukowców jednoznacznie zdefiniowanych ze względu na płeć i dyscyplinę oraz metadane ich wszystkich publikacji zindeksowanych w bazie Scopus. Trzy dyscypliny na poziomie dwucyfrowym zostały wyłączone z dalszej analizy, ponieważ nie spełniały arbitralnie przez nas ustalonego minimalnego progu 50 naukowców przypadających na dyscyplinę.

### **3.3. Określanie wieku akademickiego naukowców**

Przy założeniu, że jednostką analizy jest dla nas indywidualny naukowiec, obliczyliśmy wiek akademicki w oparciu o rok pierwszej publikacji dla każdego naukowca objęte-

go naszą próbą. Nasza baza danych zawierała datę urodzenia oraz daty uzyskania odpowiednich stopni i tytułów naukowych (doktorat, habilitacja, profesura) w zależności od etapu kariery akademickiej naukowca.

Stosowany przez nas wiek akademicki mieści się w przedziale od 0 do 40 lat (zgodnie z ogólnym założeniem, że obejmuje on naukowców w wieku biologicznym 30–70 lat i posiadających doktorat). Daty pierwszych publikacji indeksowanych w bazie Scopus dla każdego polskiego naukowca uzyskaliśmy za pomocą protokołu API (interfejsu programowania aplikacji, czyli zestawu kodów programistycznych umożliwiających przesyłanie danych między produktami programowymi), udostępnianego przez bazę Scopus. Aby skorzystać z API, zastosowaliśmy indywidualne identyfikatory Scopus ID dla każdego naukowca znajdującego się w próbie. Naszym rokiem odniesienia był rok 2017, a różnica między rokiem 2017 a rokiem pierwszej publikacji została wykorzystana do wyznaczenia wieku akademickiego przypadającego na rok 2017.

Przykładowo, niezależnie od wieku biologicznego jednoznacznie podawanego w naszej bazie „Obserwatorium Polskiej Nauki”, jeśli naukowiec wydał swoją pierwszą publikację indeksowaną w bazie Scopus w 1997 roku, to jego wiek akademicki w 2017 roku wynosił 20 lat; z kolei jeśli wydał swoją pierwszą publikację w 2016 roku, to jego wiek akademicki w 2017 roku wynosił 1 rok, co oznacza tym samym znalezienie się na samym początku kariery akademickiej definiowanej przez globalne publikacje. Można z grubsza założyć, że w naukach przyrodniczych pierwsza publikacja w bazie danych pojawia się zwykle w ostatnich latach pisania doktoratu – czyli w polskim przypadku ok. 30. roku życia. Zatem w dużym uproszczeniu pojawienie się w nauce (wiek akademicki równy 0) oznaczałoby mniej więcej 30. rok życia, a wiek akademicki równy 20 lat oznaczałoby mniej więcej 50. rok życia. Wiek akademicki łączymy tutaj również z pojęciem wieku kariery jako liczbą lat poświęconych na naukę (dwukrotnie: na ryc. 1 i w tab. 2).

### 3.4. Próba

Zaczelśmy od zbioru 25 463 naukowców (14 886 mężczyzn i 10 577 kobiet – odpowiednio 58,5% i 41,5%). Zbiór ten obejmował wszystkich naukowców, którzy byli zatrudnieni w pełnym wymiarze czasu pracy w sektorze szkolnictwa wyższego i którzy byli autorami lub współautorami co najmniej jednego artykułu indeksowanego w bazie Scopus w latach 2009–2018 oraz którzy legitymowali się co najmniej stopniem doktora. Zasadniczo należy przyjąć, że w badanej dekadzie 2009–2018 około 95–97% naukowców w szkolnictwie wyższym było zatrudnionych w pełnym wymiarze czasu pracy (w roku 2020 niepełnozatrudnieni stanowili 5,52% nauczycieli akademickich: 4874 z 88 284 osób, GUS 2021: 126). W próbie uwzględniono wszystkich widocznych na arenie międzynarodowej (poprzez typ publikacji: artykuł w bazie danych Scopus) polskich naukowców akademickich.

Próbę do tego badania zdefiniowaliśmy w następujących krokach: zaczęliśmy od wszystkich naukowców ze wszystkich publicznych sektorów nauki obecnych w naszym zbiorze danych „Obserwatorium Polskiej Nauki” – 99 935 osób; następnie przeszliśmy do poziomu naukowców ze stopniem doktora, co ograniczyło liczbę naszych obserwacji do 70 272; a potem przeszliśmy do poziomu naukowców ze stopniem doktora i zatrudnionych w szkolnictwie wyższym, co dało nam 54 448 osób; wśród nich było 32 937 naukowców z publikacjami dowolnego typu w bazie Scopus w okresie 2009–2018. I wreszcie próbę ograniczyliśmy do 25 463 naukowców posiadających co najmniej jeden artykuł w czasopiśmie naukowym.

Do tak utworzonej bazy danych postanowiliśmy włączyć informacje o produktywności poszczególnych autorów pochodzącej z całego ich życia naukowego (nasza wyjściowa baza zawierała informacje bibliometryczne tylko dla dekady 2009–2018, a okresem referencyjnym dla danych demograficznych i zawodowych autorów był listopad 2017 r.); w szczególności postanowiliśmy uwzględnić rok pierwszej publikacji. Informacje te uzyskaliśmy na podstawie bazy Scopus przy użyciu protokołu API dla 21 285 osób. Ponadto wykluczenie z bazy danych osób starszych niż 70 lat oraz osób, dla których znaleziono nierealną datę pierwszej publikacji (wynikającą z błędów bazy Scopus), doprowadziło do zmniejszenia efektywnej liczebności próby do 20 596 autorów ( $N = 20\,596$ ). Strukturę ostatecznej próby przedstawiono w tabeli 1.

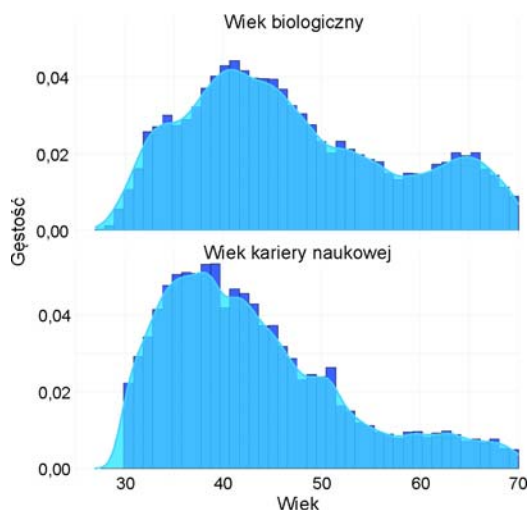
Rycina 1 przedstawia – w celach poglądowych – rozkład próby ze względu na wiek biologiczny (panel górny) i wiek kariery (panel dolny): polska profesja akademicka jest stosunkowo młoda, a większość naukowców mieści się w przedziale wiekowym 30–45 lat; jednocześnie jednak znaczna część naukowców przekroczyła 50. rok życia (oraz przekroczyła 60. rok życia). Podobnie wygląda rozkład próby pod względem wieku kariery (pochodnego od wieku akademickiego, aby porównanie było bardziej czytelne): większość polskich naukowców publikuje od co najwyżej 15 lat. Długi ogon po prawej stronie wykresu świadczy o tym, że choć odsetek starszych naukowców jest znaczny (według wieku biologicznego, jak pokazuje górny panel), publikują oni na arenie międzynarodowej znacznie krócej, niż wskazywałby na to ich wiek biologiczny, co widać na dolnym panelu. Tylko niewielki odsetek naukowców (we wszystkich dyscyplinach akademickich łącznie) publikuje od 25, 30 czy 40 lat. Różnica pokoleniowa między młodszymi i starszymi naukowcami oraz ich publikowaniem na arenie międzynarodowej jest wyraźnie widoczna jako przesunięcie próby w lewo wzdłuż osi X.

Statystyki opisowe wieku biologicznego i wieku kariery (używanego tu w celu zachowania przejrzystości i obliczanego jako wiek akademicki powiększony o 30 lat) badanej próby przedstawiono w tabeli 2 zawierającej najważniejsze parametry, takie jak płeć, stopień i tytuł naukowy, typ instytucji (typ IDUB, intensywnie prowadzący badania i pozostałe uczelnie) oraz dyscypliny naukowe według klasyfikacji ASJC.

Tabela 1. Struktura próby, polscy naukowcy akademicy widoczni globalnie w bazie publikacji Scopus w latach 2009–2018 (kategoria: artykuły naukowe), według płci, grupy wiekowej, stopnia/tytułu naukowego i dyscypliny naukowej, według typu: obszar STEMM i poza obszarem STEMM, punkt odcięcia: wiek 70 lat, przedstawiona w formie procentów kolumnowych i wierszowych

		Kobieta			Mełczyźni			Ogółem		
		<i>N</i>	Procent kolumnowy	Procent wierszowy	<i>N</i>	Procent kolumnowy	Procent wierszowy	<i>N</i>	Procent kolumnowy	Procent wierszowy
Grupa wiekowa	< 40	2828	32,3	49,5	2880	24,3	50,5	5708	27,7	100,0
	40–54	4368	49,9	44,8	5378	45,4	55,2	9746	47,3	100,0
	55–70	1564	17,9	30,4	3578	30,2	69,6	5142	25	100,0
	<b>Ogółem</b>	<b>8760</b>	<b>100,0</b>	<b>42,5</b>	<b>11836</b>	<b>100,0</b>	<b>57,5</b>	<b>20596</b>	<b>100,0</b>	<b>100,0</b>
Stopień/tytuł naukowy	Doktorat	5557	63,4	48,4	5913	50	51,6	11470	55,7	100,0
	Habilitacja	2443	27,9	39,1	3799	32,1	60,9	6242	30,3	100,0
	Profesura	760	8,7	26,4	2124	17,9	73,6	2884	14	100,0
	<b>Ogółem</b>	<b>8760</b>	<b>100,0</b>	<b>42,5</b>	<b>11836</b>	<b>100,0</b>	<b>57,5</b>	<b>20596</b>	<b>100,0</b>	<b>100,0</b>
Dyscyplina naukowa (ASJC) – obszar STEMM	AGRI	1215	13,9	54	1037	8,8	46	2252	10,9	100,0
	BIO	900	10,3	61,7	558	4,7	38,3	1458	7,1	100,0
	CHEM	612	7	52,1	562	4,7	47,9	1174	5,7	100,0
	CHEMENG	155	1,8	40,6	227	1,9	59,4	382	1,9	100,0
	COMP	141	1,6	17,1	684	5,8	82,9	825	4	100,0
	DEC	20	0,2	43,5	26	0,2	56,5	46	0,2	100,0
	EARTH	333	3,8	34,9	622	5,3	65,1	955	4,6	100,0

	ENER	60	0,7	26	171	1,4	74	231	1,1	100,0
	ENG	412	4,7	15	2336	19,7	85	2748	13,3	100,0
	ENVIR	735	8,4	51,9	680	5,7	48,1	1415	6,9	100,0
	IMMU	72	0,8	75,8	23	0,2	24,2	95	0,5	100,0
	MATER	417	4,8	34,2	801	6,8	65,8	1218	5,9	100,0
	MATH	213	2,4	26,4	595	5	73,6	808	3,9	100,0
	MED	1620	18,5	55,2	1314	11,1	44,8	2934	14,2	100,0
	PHARM	148	1,7	69,8	64	0,5	30,2	212	1	100,0
	PHYS	147	1,7	17,6	690	5,8	82,4	837	4,1	100,0
Dyscyplina naukowa (ASJC) – poza obszarem STEMM	BUS	280	3,2	53	248	2,1	47	528	2,6	100,0
	DENT	48	0,5	73,8	17	0,1	26,2	65	0,3	100,0
	ECON	143	1,6	50,4	141	1,2	49,6	284	1,4	100,0
	HEALTH	20	0,2	37,7	33	0,3	62,3	53	0,3	100,0
	HUM	405	4,6	51	389	3,3	49	794	3,9	100,0
	PSYCH	148	1,7	62,7	88	0,7	37,3	236	1,1	100,0
	SOC	381	4,3	50,8	369	3,1	49,2	750	3,6	100,0
	VET	135	1,5	45,6	161	1,4	54,4	296	1,4	100,0
	<b>Razem</b>	<b>8760</b>	<b>100,0</b>	<b>42,5</b>	<b>11836</b>	<b>100,0</b>	<b>57,5</b>	<b>20596</b>	<b>100,0</b>	<b>100,0</b>



Ryc. 1. Rozkład wieku biologicznego (górny panel) i wieku kariery (dolny panel) naukowców, wykres histogramu oraz funkcji gęstości jądrowej dla wszystkich dyscyplin akademickich łącznie

Wiek kariery został tu użyty jedynie dla celów poglądowych (wizualizacja dobrze pokazuje różnice międzypokoleniowe), a w dalszej części pracy będziemy posługiwać się wyłącznie kategorią wieku akademickiego.

W tym miejscu należy podkreślić kilka kwestii: mediana wieku biologicznego wynosząca 45 lat jest wyższa niż mediana wieku kariery zawodowej wynosząca 41 lat, a oba te wskaźniki są wyższe w przypadku mężczyzn niż kobiet (tabela 2). Różnica między tymi dwoma typami wieku rośnie wraz z kolejnymi stopniami naukowymi: o ile w przypadku osób ze stopniem doktora jest ona niewielka, o tyle w przypadku osób z tytułem profesorskim jest ona znacząca (i wynosi odpowiednio dwa i siedem lat).

Oznacza to, że kohorta młodszych naukowców zaczęła publikować na arenie międzynarodowej wcześniej niż kohorta starszych naukowców, a w szczególności wcześniej niż obecni profesorowie tytularni. Ponadto wyraźne jest zróżnicowanie między dyscyplinami: w przypadku dyscyplin z obszaru STEMM rozbieżności między medianą wieku biologicznego a medianą wieku kariery są znacznie mniejsze niż w przypadku dyscyplin spoza obszaru STEMM. Naukowcy publikujący w obszarze STEMM zaczynają wydawać artykuły na arenie międzynarodowej znacznie wcześniej niż naukowcy spoza niego. W głównych dyscyplinach z obszaru STEMM, takich jak BIO (biochemia, genetyka i biologia molekularna) rozbieżność nie występuje w ogóle, a w CHEM, COMP, MATH, PHYS i MED rozbieżność ta wynosi zaledwie 1–2 lata. Natomiast w takich istotnych i licznie w Polsce reprezentowanych dyscyplinach niezwiązanych z naukami ścisłymi, jak BUS (biznes, zarządzanie i rachunkowość), ECON (ekonomia, ekonometria i finanse), HUM (nauki humanistyczne) oraz SOC (nauki społeczne), rozbieżność sięga 8–10 lat.

Tabela 2. Statystyki opisowe wieku biologicznego i wieku kariery według płci, stopnia i tytułu naukowego, typu instytucji i dyscypliny akademickiej

		Wiek biologiczny			Wiek kariery		
		Średnia	SD	Mediana	Średnia	SD	Mediana
<b>Płeć</b>	Razem	47	10,6	45	43,2	9,3	41
	Kobieta	45,1	9,7	43	41,2	8,2	39
	Mężczyzna	48,4	11	46	44,6	9,9	42
<b>Stopień/ tytuł</b>	Doktorat	41,3	7,8	40	39	6,2	38
	Habilitacja	50,8	8,4	49	45,6	8,4	45
	Profesura	61,3	6,8	63	54,4	10,3	56
<b>Typ uczelni</b>	IDUB	46,6	10,8	44	43,9	9,7	42
	Pozostałe	47,2	10,5	45	42,9	9,2	41
<b>Dyscyplina ASJC</b>	AGRI	47,6	10,4	46	42	7,9	40
	BIO	45,3	10,3	43	44,6	9,3	43
	BUS	46,3	9,7	44	36,3	6,6	34
	CHEM	45,8	11,1	44	46,1	9,8	45
	CHEMENG	47,3	11,7	45	43,7	10,4	41
	COMP	46,5	10,5	44	44,1	9	42
	DEC	48,1	11,2	43,5	41,6	11,8	37,5
	DENT	45	9,8	45	39,9	8,8	39
	EARTH	48,2	11,2	46	43,8	10,1	42
	ECON	44,3	9,2	42	36	5,7	35
	ENER	46,9	12	44	39,8	9	37
	ENG	47,7	11,3	45	43,4	9,4	41
	ENVIR	47,1	10,2	45	42,1	8,1	40
	HEALTH	48,8	10,4	49	38,1	4,8	37
	HUM	46,7	9,6	45	36,2	6,3	34
	IMMU	45,7	9,3	45	45,7	8,1	45
	MATER	46,6	10,9	44	45,4	9,6	43
	MATH	47,1	11,3	45	45,5	9,8	43
	MED	47,9	9,8	47	45,1	8,9	44
	PHARM	44,5	10,3	42	43,9	9,1	42
	PHYS	48,2	11,5	46	48,6	10,4	46
	PSYCH	44,6	10,5	41,5	37	6,2	36
	SOC	45,4	9,6	44	36,7	6,5	35
VET	47,3	10,4	46	44,5	7,8	43,5	

Wyraźnie widać, że w przypadku dużych dyscyplin nauk społecznych i humanistycznych rozbieżność jest zasadniczo wyższa niż w przypadku dużych dyscyplin nauk przyrodniczych, co wyraźnie wskazuje na to, że w przypadku obecnie zatrudnionych naukowców wiek biologiczny można określić za pomocą wieku akademickiego dla dyscyplin z obszaru STEMM z marginalnym błędem; jednak w przypadku dyscyplin nienależących do obszaru STEMM, błąd jest znaczny, a przypisywanie naukowców do etapów rozwoju kariery naukowej przy użyciu wieku akademickiego może być mylące.

### 3.5. Zmienne

Wykorzystujemy w tej pracy zmienne numeryczne i zmienne kategoryjne, o charakterze biologicznym, demograficznym, administracyjnym i instytucjonalnym. Lista zmiennych, ich krótkie opisy oraz źródła danych (nasze Obserwatorium Polskiej Nauki, baza Scopus lub MEiN – Ministerstwo Edukacji i Nauki) zostały przedstawione w tabeli 3.

Tabela 3. Zmienne wykorzystane w analizie

Nr	Zmienna	Opis	Źródło danych
1.	Wiek biologiczny	Zmienna ilościowa. Wiek biologiczny wykazywany zgodnie z bazą OPI ( $N = 99\ 935$ ). Wiek w pełnych latach według stanu na rok 2017.	Obserwatorium
2.	Grupa wiekowa	Zmienna jakościowa. Wykorzystano trzy główne grupy wiekowe: młodzi (39 lat i młodszy; $N = 5708$ ), w średnim wieku (40–54 lata; $N = 9746$ i starsi (55–70 lat; $N = 5142$ ) naukowcy.	Obserwatorium
3.	Płeć	Zmienna jakościowa, mężczyzna lub kobieta, zgodnie z krajowym rejestrem naukowców ( $N = 99\ 935$ ).	Obserwatorium
4.	Etap rozwoju kariery naukowej	Zmienna jakościowa. Użyto trzech stopni i tytułów naukowych: naukowcy tylko ze stopniem doktora ( $N = 11\ 470$ ), ze stopniem doktora habilitowanego ( $N = 6242$ ) oraz z tytułem profesorskim ( $N = 2884$ ). Z analizy usunięto wszystkich naukowców nieposiadających stopnia naukowego doktora oraz zatrudnionych poza sektorem szkolnictwa wyższego. Zastosowano cztery etapy kariery: początkujący etap (<5 lat doświadczenia akademickiego – od pierwszej publikacji w bazie Scopus), wczesny etap (5–14 lat), średni etap (15–29 lat) i późny etap (30 lat i więcej).	Obserwatorium
5.	Dyscyplina	Zmienna jakościowa. Wszyscy naukowcy co najmniej z doktoratem zostali przypisani do jednej z 27 dyscyplin Scopus ASJC ( <i>All Science Journal Classification</i> ). Wykorzystano dyscypliny dominujące: wyznaczone lub wylosowane z wyznaczonych. Wykorzystano ostatecznie 24 dyscypliny.	Scopus



Nr	Zmienna	Opis	Źródło danych
6.	Dyscypliny związane z naukami ścisłymi i medycyną (dyscypliny STEMM)	Zmienna jakościowa. Dyscypliny STEMM: AGRI, nauki rolnicze i biologiczne; BIO, biochemia, genetyka i biologia molekularna; CHEMENG, inżynieria chemiczna; CHEM, chemia; COMP, informatyka; DEC, nauki o podejmowaniu decyzji; EARTH, nauki o ziemi i planetach; ENER, energia; ENG, inżynieria; ENVIR, nauka o środowisku; IMMUN, immunologia i mikrobiologia; MATER, materiałoznawstwo; MATH, matematyka; PHARM, farmakologia, toksykologia i farmacja; PHYS, fizyka i astronomia; oraz MED, medycyna.	Scopus
7.	Dyscypliny niezwiązane z naukami ścisłymi i medycyną (dyscypliny spoza STEMM)	Zmienna jakościowa. Dyscypliny niezwiązane z naukami ścisłymi i medycyną (spoza STEMM): BUS, biznes, zarządzanie i rachunkowość; DENT, stomatologia; ECON, ekonomia, ekonometria i finanse; HEALTH, zawody medyczne; HUM, nauki humanistyczne; PSYCH, psychologia; SOC, nauki społeczne; oraz VET, weterynaria.	Scopus
8.	Instytucja prowadząca intensywną działalność badawczą	Zmienna jakościowa. Dziesięć instytucji (spośród 85 badanych) to instytucje z programu IDUB (czyli „Inicjatywa Doskonałości – Uczelnia Badawcza”) wybrane w ogólnopolskim konkursie przeprowadzonym w 2019 roku.	MEiN

## 4. Wyniki

### 4.1. Przykład

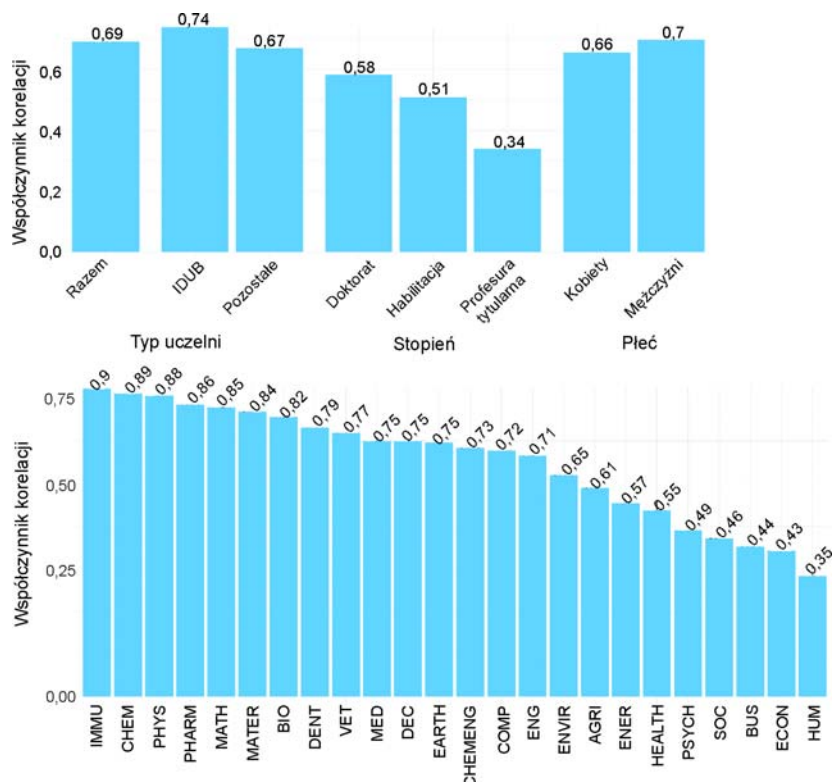
Przed przystąpieniem do dalszych rozważań posłużymy się przykładem: pierwsza publikacja badacza A (mężczyzny pracującego w dyscyplinie ECON, ekonomia, ekonometria i finanse, który otrzymał doktorat w 1995 r., habilitację w 2005 r., a profesurę tytularną w 2012 r.) ukazała się – według naszej referencyjnej bazy Scopus – w 2015 r. Przyjmujemy w związku z tym następujące założenie (na podstawie wybranych prac omówionych w części 2.2): rok 2015 jest dla tego naukowca początkiem kariery akademickiej w sensie rozpoczęcia publikowania, a jego wiek naukowy w 2015 r. wynosi 0. Zatem, jak moglibyśmy wnioskować, w roku 2017 jego wiek biologiczny musiałby wynosić 32 lata. Jednak nasza baza danych „Obserwatorium Polskiej Nauki” posługująca się datami urodzenia wskazuje 52 lata jako jego wiek biologiczny w 2017 roku.

Pojawia się zatem niespójność: gdy używamy wieku akademickiego jako zamiennika wieku biologicznego, naukowiec A okazuje się młody i dopiero rozpoczyna swoją karierę badawczą (wiek akademicki = 2, etap kariery: wczesny); natomiast gdy używamy wieku

biologicznego, stwierdzamy jednoznacznie, że naukowiec A jest starszy (ma 52 lata, etap kariery: średni). Nasze zadanie poniżej polega na szczegółowym oszacowaniu tego niedopasowania wieku akademickiego i wieku biologicznego w dużej skali całego systemu nauki, w zależności od wybranych parametrów.

#### 4.2. Korelacja wieku biologicznego i wieku akademickiego

W celu przeanalizowania wpływu rzeczywistego wieku biologicznego (w przedziale 30–70 lat) na wiek akademicki wynikający z danych publikacyjnych (w przedziale 0–40 lat) przeprowadzono analizę korelacji liniowej pomiędzy tymi zmiennymi (ryc. 2).

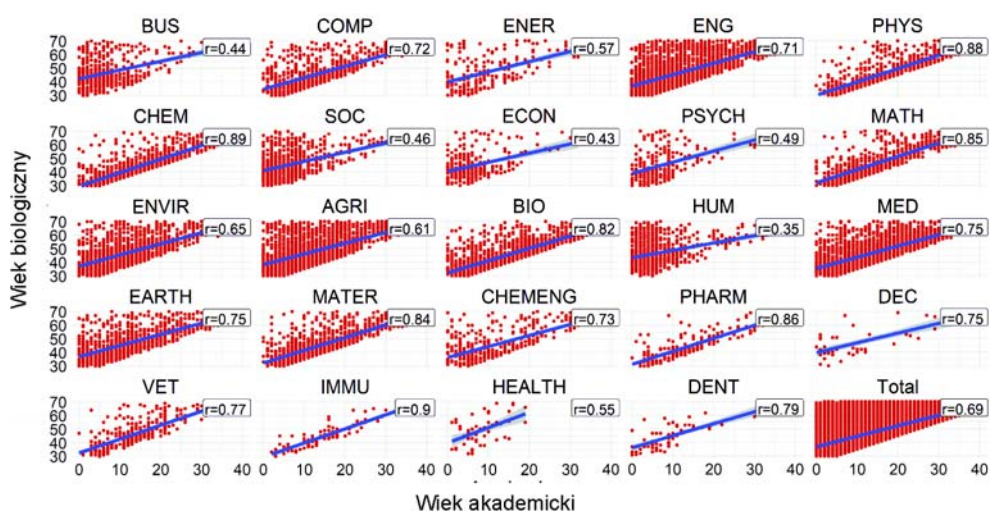


Ryc. 2. Współczynnik korelacji dla polskich naukowców pomiędzy wiekiem akademickim a wiekiem biologicznym w podziale na typ instytucji (instytucje badawcze IDUB i pozostałe), stopień naukowy, płeć oraz dyscyplinę ASJC

Niezależnie od ujęcia prezentowanych korelacji, w każdym przypadku zaobserwowano zależność o charakterze pozytywnym. Wszystkie współczynniki korelacji liniowej Pearsona są istotnie różne od zera (na poziomie istotności  $\alpha = 0,05$ ). Inną wyraźnie widoczną prawidłowością jest zazwyczaj silna lub bardzo silna korelacja obserwowana dla dys-

cyplin należących do obszaru STEMM. Dla większości z nich zaobserwowano silniejszą korelację niż dla wszystkich dyscyplin łącznie. Diagramy korelacyjne (ryc. 3) wyraźnie wskazują, że zdecydowana większość osób publikujących w dyscyplinach z obszaru STEMM wydaje swój pierwszy artykuł w stosunkowo młodym wieku (punkty przesuwają się w prawo na osi X). Co więcej, najwyższe korelacje zaobserwowano dla takich dyscyplin z obszaru STEMM, jak chemia (dla której współczynnik korelacji  $r = 0,89$ ), fizyka i astronomia ( $r = 0,88$ ), matematyka ( $r = 0,85$ ) czy medycyna ( $r = 0,75$ ).

Z kolei dyscypliny takie jak (BUS) biznes, zarządzanie i rachunkowość, nauki humanistyczne (HUM) oraz nauki społeczne (SOC) – stosunkowo licznie reprezentowane przez polskich naukowców – charakteryzują się relatywnie niską siłą związku między wiekiem biologicznym a wiekiem akademickim. Wartość współczynnika korelacji w żadnym przypadku nie przekracza 0,5 i utrzymuje się w przedziale od 0,36 dla nauk humanistycznych do 0,49 dla psychologii; analiza korelacji (ryc. 3) wskazuje na wyraźne przesunięcie punktów na osi X w lewo.

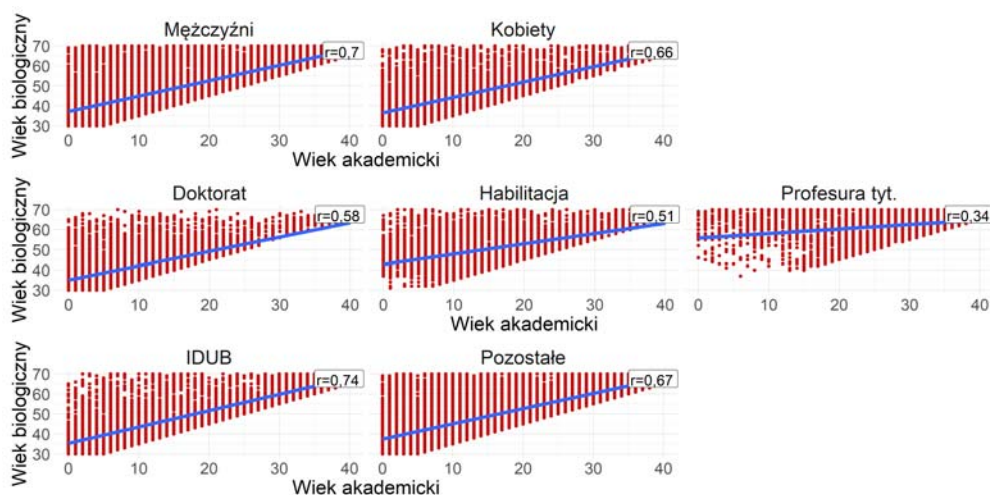


Ryc. 3. Diagramy korelacyjne dla polskich naukowców między wiekiem akademickim (zakres 0–40 lat) a wiekiem biologicznym (zakres 30-70 lat) według dyscypliny ASJC

Wyniki analiz pokazują zatem, że większość przedstawicieli dyscyplin nienależących do obszaru STEMM, mimo stosunkowo zaawansowanego wieku biologicznego, jest względnie młoda pod względem wieku akademickiego. Ich pierwsza publikacja w bazie Scopus, a więc debiut publikacyjny na globalnej arenie naukowej, rozpoczyna się wyraźnie później niż w przypadku naukowców z dyscyplin z obszaru STEMM. Obserwacja ta jest zgodna z powszechnymi intuicjami i z wcześniejszymi badaniami (zob. np. Kwiek 2015a, 2020), według których przedstawicielom nauk ścisłych w Polsce szerokie

uczestnictwo w globalnym obiegu nauki zajęło mniej czasu niż przedstawicielom nauk humanistycznych, społecznych czy ekonomicznych (m.in. ze względu na ekspansję sektora prywatnego w latach 1990–2006, na którym skupiała się duża część energii naukowców spoza obszaru STEMM, zob. o deinstytucjonalizacji misji badawczej polskich uniwersytetów w Kwiek 2015a: 175–211).

Mniejsze różnice w poziomie korelacji zaobserwowano dla pozostałych zmiennych niezależnych: typu uczelni pod względem intensywności prowadzenia badań naukowych (intensywne badawczo instytucje IDUB vs. pozostałe), stopnia naukowego oraz płci. Dla naukowców zatrudnionych na uczelniach intensywnie badawczych korelacja jest wyższa niż dla pozostałych ( $r = 0,74$  vs.  $r = 0,67$ ), ale dla obu współczynnik korelacji jest zbliżony do korelacji dla wszystkich obserwacji ( $r = 0,69$ ). Co więcej, korelacja ta jest wyższa dla mężczyzn niż dla kobiet ( $r = 0,70$  vs.  $r = 0,66$ ). Ponadto porównanie różnic dotyczących stopni naukowych wskazuje (ryc. 4), że choć współczynniki korelacji są istotnie niższe, to różnice w nachyleniu krzywej regresji wyraźnie świadczą o tym, że naukowcy na niższych etapach kariery naukowej są generalnie akademicko młodsi niż ich koledzy na wyższych etapach, czyli o wyższych stopniach naukowych (przesunięcie punktów w lewo na osi X), co zapewne wynika z korelacji między stopniem naukowym a wiekiem biologicznym. Jednocześnie wraz z kolejnymi etapami rozwoju naukowego siła tej korelacji wyraźnie maleje (od  $r = 0,58$  dla doktorów do zaledwie  $r = 0,34$  dla profesorów tytularnych).



Ryc. 4. Diagramy korelacyjne dla polskich naukowców między wiekiem akademickim (zakres 0–40 lat) a wiekiem biologicznym (zakres 30–70 lat) według płci (panel górny), typu instytucji (instytucje IDUB intensywnie prowadzące badania naukowe vs. pozostałe, panel dolny) oraz stopnia naukowego (panel środkowy)

Wyniki analiz wyraźnie pokazują znaczne zróżnicowanie naukowców pod względem momentu rozpoczęcia aktywności publikacyjnej: obecni profesorowie tytularni, najstarsi i najwyżej usytuowani w hierarchii akademickiej, zaczynają swoją międzynarodową działalność publikacyjną najpóźniej, a dla naukowców najmłodszych i dopiero rozpoczynających karierę naukową, posiadających jedynie stopień doktora, publikowanie międzynarodowe okazuje się bardziej naturalne i rozpoczyna się wcześniej.

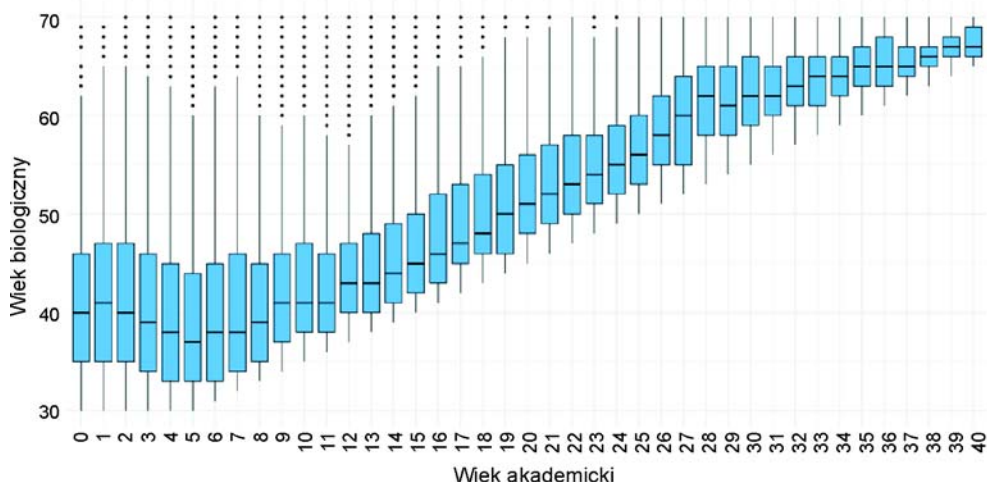
Wyniki te potwierdzają rezultaty wcześniejszych badań ankietowych i opartych na wywiadach, które pokazały radykalne różnice pokoleniowe występujące w polskiej nauce (Kwiek 2015a) oraz wyniki badań łączące „umiędzynarodowionych” i „miejscowych” w badaniach z wiekiem i pokoleniami akademickimi przedstawione w pracy (Kwiek 2020). Co znamienne, wcześniejsze badania ankietowe (przeprowadzone na podstawie 3704 zwróconych ankiet) znajdują silne potwierdzenie w prezentowanych tu szczegółowych badaniach przeprowadzonych na dużej próbie w oparciu o dane o zupełnie innym charakterze.

Zatem, jak się okazuje, niezależnie od tego, czy posługujemy się danymi administracyjnymi i publikacyjnymi, czy też danymi pochodzącymi z badań ankietowych i wywiadów, wnioski są podobne – co wskazuje na daleko idącą komplementarność metodologiczną w badaniach kadry akademickiej.

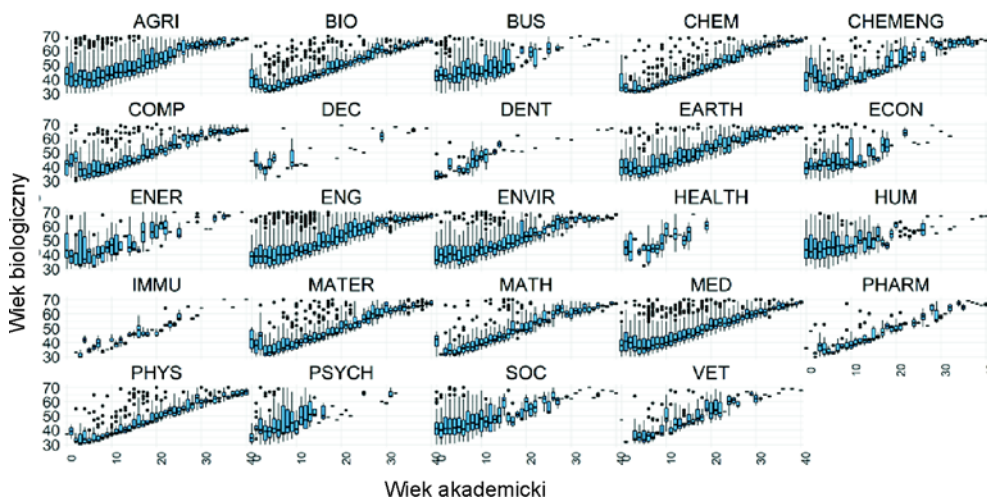
Różnice międzypokoleniowe we wzorcach publikowania międzynarodowego wynikają z wyraźnie odmiennych warunków początkowych dla różnych grup badaczy: różnych możliwości współpracy międzynarodowej i różnych wymagań instytucjonalnych na kolejnych etapach rozwoju kariery akademickiej, szczególnie rosnących po roku 2010, kiedy to zainicjowano dwie serie strukturalnych reform szkolnictwa wyższego (reformy Kudryckiej z lat 2010–2012 i reformy Gowina z lat 2016–2018). Dzisiejsza kadra naukowa z doktoratem, a więc niemal wyłącznie młoda, zaczynała swoją karierę akademicką w okresie ostatniej dekady, w radykalnie lepszych warunkach finansowych i w innych realiach politycznych, społecznych i ekonomicznych niż ich koledzy, którzy są obecnie doktorami habilitowanymi czy profesorami tytularnymi. Różnice między kohortami dobitnie wskazują na stopniową ewolucję polskiego systemu nauki – na procesy otwierania się na obieg globalny.

Analizując rozkład wieku biologicznego według wieku akademickiego dla wszystkich naukowców – bez uwzględnienia dyscypliny, typu instytucji, stopnia naukowego i płci – można zauważyć rosnący, ale gasnący co do siły, trend mediany wieku biologicznego i malejące zróżnicowanie wieku biologicznego wraz ze wzrostem wieku akademickiego (ryc. 5). W przypadku płci i typu instytucji (IDUB vs. pozostałe) nie odnotowano istotnych różnic, natomiast dyscyplina (ryc. 6) i stopień naukowy (ryc. 7) silnie wpływają na kształt rozkładu wieku biologicznego.

W przypadku dyscyplin z obszaru STEMM trend mediany w wielu przypadkach (np. BIO, CHEM, COMP, EARTH, ENG, ENVIR, MATER, MATH, MED i PHYS) nie wydaje się zanikający, lecz w dużej mierze liniowy, co po raz kolejny wskazuje na silny związek między wiekiem akademickim i wiekiem biologicznym w obszarze STEMM.



Ryc. 5. Rozkład wieku biologicznego dla polskich naukowców dla poszczególnych roczników wieku akademickiego

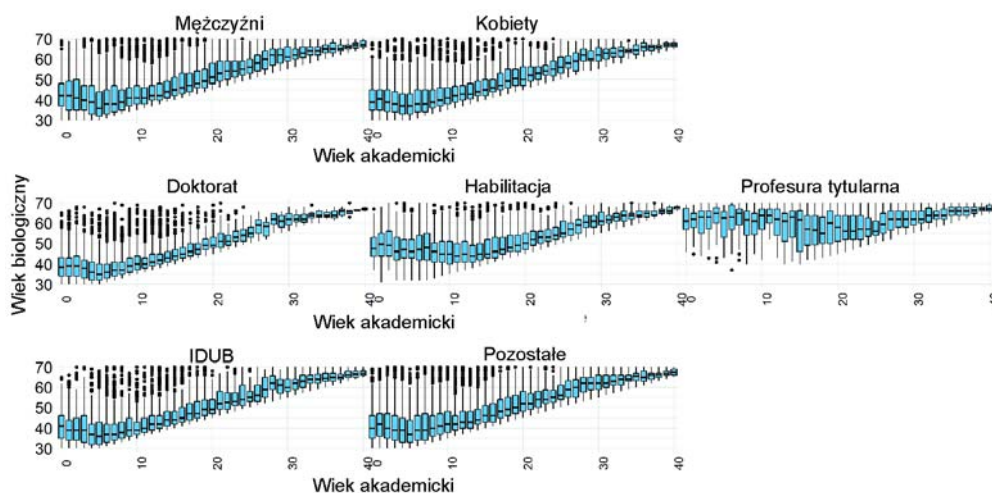


Ryc. 6. Rozkład wieku biologicznego dla polskich naukowców dla poszczególnych roczników wieku akademickiego w podziale na dyscypliny naukowe

Chociaż kariera naukowa w postaci publikacji powinna ewoluować wraz z wiekiem naukowca, w przypadku dziedzin niezwiązanych z obszarem STEMM wydaje się, że



w niektórych przypadkach trend ten w ogóle nie występuje (zwłaszcza w dużych dyscyplinach, takich jak ECON, HUM czy PSYCH) lub jest ledwo zauważalny (BUS i SOC). Również zmienność rozkładu wieku biologicznego, która maleje wraz z wiekiem akademickim dla dyscyplin z obszaru STEMM, wydaje się niezauważalna lub przebiega w sposób niewskazujący na występowanie prawidłowości.



Ryc. 7. Rozkład wieku biologicznego polskich naukowców dla poszczególnych roczników według instytucjonalnej intensywności badań (IDUB), stopnia naukowego i płci

Jeżeli chodzi o stopnie naukowe, to rozkład wieku biologicznego dla poszczególnych lat wieku akademickiego jest zbliżony do ogólnego wzorca (tj. przypomina nieco krzywą logistyczną – gasnący wzrost) dla osób posiadających tytuł doktora i doktora habilitowanego, natomiast w przypadku profesorów tytularnych zachowuje się on odmiennie. Wśród profesorów niemal dla każdego rocznika wieku akademickiego wyraźnie dominują naukowcy starsi pod względem wieku biologicznego.

#### 4.3. Analiza kontyngencji

Wnioski z analizy korelacji potwierdzają wnioski z analizy kontyngencji etapów kariery akademickiej (cztery etapy) ze stopniami naukowymi (trzy stopnie) i grupami wiekowymi (cztery grupy wiekowe). Przydzieliliśmy wszystkich naukowców do czterech etapów kariery, czyli liczby lat aktywności publikacyjnej (przedziałów wieku akademickiego): początkujący etap (*Beginning*, mniej niż pięć lat doświadczenia liczonego od pierwszej publikacji w bazie Scopus), wczesny etap kariery (*Early*, 5–14 lat), środkowy etap kariery (*Middle*, 15–29 lat) i późny etap kariery (*Late*, 30 i więcej lat). Z dużym przybliżeniem można przyjąć, że jeśli wiek akademicki wynosi 0 lat (początek kariery publikacyjnej w sensie pierwszej publikacji = 30 lat), to naukowcy w fazie początkującej

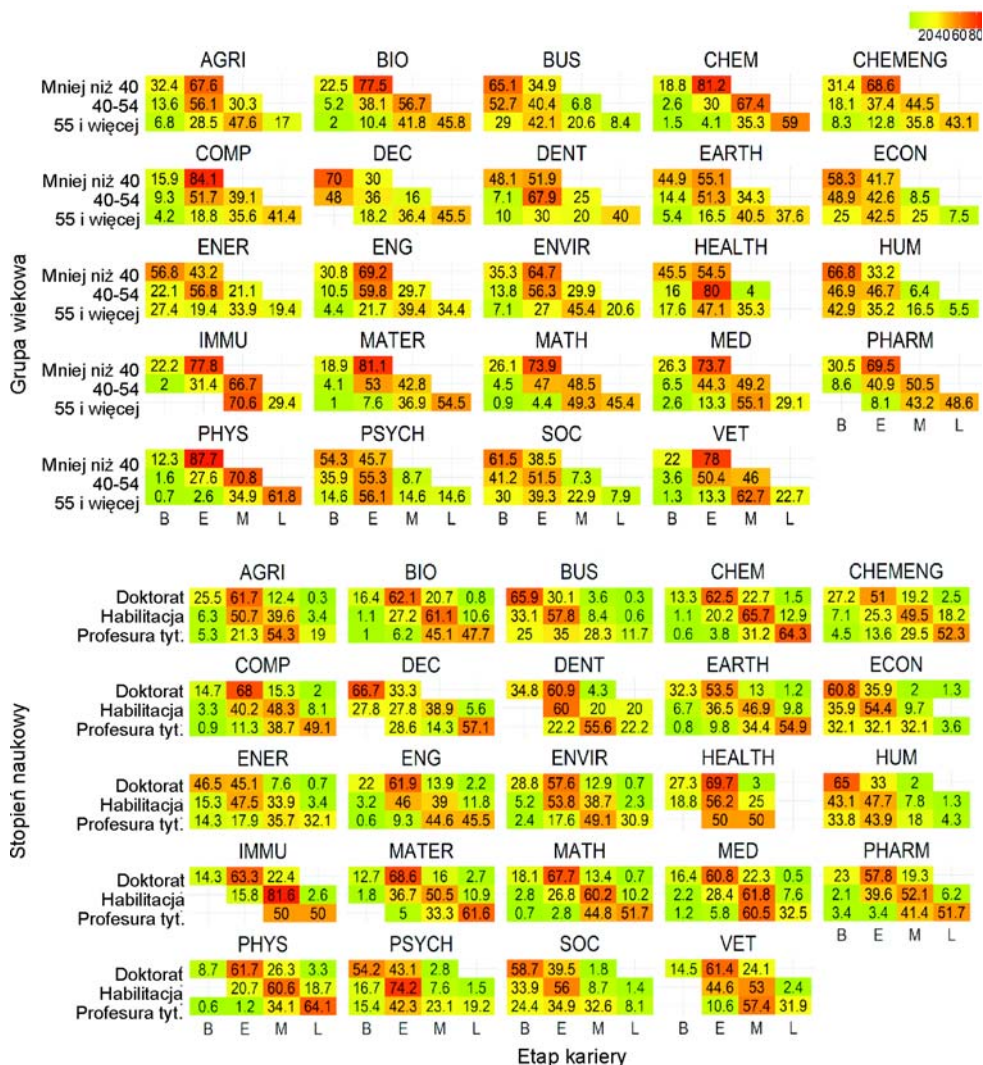
są w wieku 30–34 lat, w fazie wczesnej kariery są w wieku 35–44 lat, w fazie średniej kariery są w wieku 45–59 lat, a w fazie późnej kariery są w wieku 60 lat i więcej. We wszystkich analizach posługujemy się dobrze zdefiniowanym i ściśle określonym dla każdego badacza przedziałem wieku akademickiego 0–40 lat.

Przykład różnic między dyscyplinami w obszarze STEMM i spoza niego bardzo dobrze ilustruje porównanie dyscyplin CHEM (chemia) i HUM (nauki humanistyczne), obu stosunkowo licznych (górny panel na ryc. 8 przedstawia grupy wiekowe, a dolny – stopnie naukowe i tytuł: doktorat, habilitacja, profesura tytularna).

W modelowym przypadku chemii zdecydowana większość naukowców z doktoratem (62,5%) znajduje się w początkowej fazie kariery zawodowej. Z kolei większość osób ze stopniem doktora habilitowanego (65,7%) znajduje się w środkowej fazie kariery, a większość profesorów tytularnych (64,3%) w późnej fazie kariery. Widać wyraźnie, że korelacja kolejnych awansów i etapów kariery naukowej z zaawansowaniem własnej pracy naukowej wyrażanym poprzez międzynarodowo widoczne publikacje jest zrozumiała jako przedział wieku akademickiego. Chemia jest doskonałym przykładem dyscypliny, w której stopnie naukowe odpowiadają trajektoriom publikacyjnym naukowców: profesorowie w większości przypadków publikują znacznie dłużej niż pozostałe kategorie naukowców, co jest zgodne z oczekiwaniami. Analiza kontyngencji według stopnia naukowego (panel dolny) i według grupy wiekowej (panel górny) jest w tym przypadku podobna, przy czym na etapie wczesnej kariery jest największy odsetek młodych naukowców poniżej 40. roku życia (81,2%); jest to zgodne z przekonaniem, że młodsze pokolenia wchodzące do profesji akademickiej są już silnie umiędzynarodowione w badaniach. (Umiędzynarodowienie kadry akademickiej w 11 krajach europejskich w ramach najważniejszych klastrów dyscyplin i grup wiekowych pokazywaliśmy na podstawie rozległych badań ankietowych w Kwiek 2015b, 2018b; dysponowaliśmy ponad 17 000 zwróconych ankiet).

Inaczej jest w naukach humanistycznych: tam wzorce nie są jednoznaczne. Można zaobserwować przesunięcie etapu kariery akademickiej (definiowanego przez datę pierwszej publikacji) w kierunku młodszego wieku (ryc. 8, panel górny) i wcześniejszego etapu kariery akademickiej (ryc. 8, panel dolny). Wśród doktorów publikujących w HUM dwie trzecie (65,0%) znajduje się na początkowym etapie kariery, a nie na wczesnym etapie, jak można się było spodziewać; ale co najważniejsze, kolejne stopnie nie wskazują na wyraźne rozróżnienie między etapami kariery – to znaczy, że nie można zidentyfikować etapu wyraźnie zdominowanego przez doktorów habilitowanych czy przez profesorów tytularnych. Naukowcy ci znajdują się głównie w początkowej i w wczesnej fazie kariery, a nie w jej środkowej i późnej fazie, jak sugeruje przyjmowany przez nas model (opisywany powyżej dla chemii).





Ryc. 8. Analiza kontyngencji polskich naukowców między grupami wieku biologicznego (górny panel) i stopnia naukowego (dolny panel: doktorat, habilitacja, profesura tytułarna) oraz dyscyplinami ASJC. Analizowane są cztery etapy kariery: B (*beginning*) – początkujący w nauce, E (*early*) – wczesny etap kariery, M (*middle*) – środkowy etap kariery i L (*late*) – późny etap kariery.

Poziom kontyngencji rośnie wraz ze zmianą koloru z zielonego na czerwony

Oznacza to, że w dyscyplinach spoza nauk ścisłych i przyrodniczych, takich jak nauki humanistyczne HUM (podobnie jak w biznesie, zarządzaniu i rachunkowości BUS, ekonomii, ekonometrii i finansach ECON, psychologii PSYCH i naukach społecznych SOC), polscy naukowcy zaczęli publikować swoje artykuły w obiegu międzynarodowym radykalnie później niż w porównywanej tu chemii, podobnie jak w takich dyscyplinach

w obszarze STEMM, jak BIO biochemia, genetyka i biologia molekularna czy PHYS fizyka i astronomia. Wynika z tego również, że w naukach humanistycznych publikują głównie najmłodszy naukowcy i naukowcy, którzy są na najwcześniejszym etapie kariery naukowej (czyli używając nomenklatury stanowisk – adiunkci bez habilitacji).

Analiza zależności między stopniem naukowym a etapami kariery akademickiej (ryc. 8, dolny panel) wyraźnie wskazuje, że naukowcy w niemal wszystkich dyscyplinach spoza obszaru STEMM charakteryzują się przesunięciem w czasie o co najmniej jeden etap kariery akademickiej w stosunku do dyscyplin z obszaru STEMM. Ponadto w dyscyplinach nienależących do obszaru STEMM doktorzy habilitowani i profesorowie tytularni wcale nie dominują na późnych etapach kariery, co sugeruje, że publikują oni na arenie międzynarodowej przez krótszy okres czasu, niż można by przypuszczać, biorąc pod uwagę ich wiek biologiczny.

Analogiczny wzorec obserwujemy dla zależności grup wiekowych i karier akademickich (ryc. 8, panel górny): w dyscyplinach z obszaru STEMM wyraźnie widać awans do wyższych etapów kariery akademickiej wraz z rosnącą grupą wiekową, podczas gdy dla dyscyplin spoza obszaru STEMM zależność ta nie jest oczywista. Obserwujemy znaczące przesunięcie w czasie dla dyscyplin nienależących do STEMM – pierwsze artykuły naukowców z tego obszaru są publikowane późno, często dopiero na etapie pracy profesora tytularnego.

Wynika z tego jeden wniosek: wiek akademicki w przypadku dyscyplin spoza obszaru STEMM nie nadąża za wiekiem biologicznym, a profesorowie tytularni niejednokrotnie pozostają na tym samym etapie wieku akademickiego – określanego przez indeksowane publikacje – co doktorzy. Tym samym wnioskowanie o wieku biologicznym naukowców w dyscyplinach należących do obszaru STEMM prowadzi do adekwatnego przybliżenia do rzeczywistości, natomiast gdy to samo wnioskowanie w dyscyplinach nienależących do obszaru STEMM prowadzi do błędnych wyników.

#### 4.4. Model regresji liniowej

W celu przeprowadzenia analizy wielowymiarowej stworzono liniowy model regresji wielorakiej, w którym zmienną zależną był wiek biologiczny, a zmiennymi niezależnymi były: (1) wiek akademicki, (2) płeć, (3) typ instytucji, (4) dyscyplina akademicka oraz (5) stopień i tytuł naukowy. Uzyskany model wyjaśnia 61,5% zmienności wieku biologicznego, a błąd standardowy wynosi 6,59 lat (tzn. określając wiek biologiczny na podstawie oszacowanego modelu, odbiegamy od rzeczywistości średnio o 6,59 lat; błąd względny wynosi 14%). Wszystkie interpretacje podlegają założeniu *ceteris paribus*, a przyjęty poziom istotności to  $\alpha = 0,05$  (tab. 4).

Jak wynika z modelu, w przypadku polskiego systemu nauki akademickiej, jeśli wiek akademicki rośnie o jeden rok, to wiek biologiczny rośnie średnio o 0,6 roku. Jest on

zarazem zmienną o najsilniejszym wpływie na zmienną zależną (współczynnik standaryzowany wynosi 0,523 i jest najwyższy spośród wszystkich zmiennych niezależnych). Cechą o drugim najsilniejszym wpływie jest stopień i tytuł naukowy. Tytuł profesorski ma dodatni wpływ w porównaniu ze stopniem doktora (jako kategorią odniesienia; przyrost wieku jest średnio o 11 lat większy); jednocześnie stopień doktora habilitowanego również charakteryzuje się wysokim wpływem (przyrost wieku średnio o 5,4 lata). Zależności te dotyczą jednak wyłącznie badanej populacji naukowców publikujących w bazie Scopus, a nie wszystkich polskich naukowców. Ponadto predyktor związany z pracą na uczelniach innych niż 10 uczelni badawczych wybranych do programu IDUB (kategoria odniesienia) pozytywnie wpływa na wiek biologiczny (przyrost wieku średnio o 1,3 roku). Co ważne, płeć nie wpływa w istotny sposób na przewidywanie wieku biologicznego.

Kilka interesujących wniosków płynących z modelu pochodzi z analizy wpływu poszczególnych dyscyplin naukowych na przewidywany wiek biologiczny. Naukowcy publikujący w dyscyplinach BUS, DEC i ENER charakteryzują się podobnym wiekiem biologicznym jak naukowcy publikujący w dyscyplinie HUM (która jest kategorią odniesienia). Jedynie przypisanie do dyscypliny HEALTH ma istotnie dodatni wpływ na wiek biologiczny (średnio zwiększa go o dwa lata). Publikowanie w zdecydowanej większości dyscyplin wywiera ujemny wpływ na wiek biologiczny w porównaniu z HUM (z wyjątkiem małej dyscypliny HEALTH, małej dyscypliny DEC oraz dużych dyscyplin BUS i ENER). Przypisanie do dyscyplin PHYS, CHEM, IMMU, BIO, PHARM, MATH, MATER, COMP i MED wywiera najsilniejszy ujemny wpływ na wiek biologiczny (średnio zmniejszając go o 4–6 lat). Dyscypliny te należą do tradycyjnego obszaru STEMM, co wyraźnie wskazuje na wcześniejsze rozpoczynanie kariery akademickiej (mierzonej według wieku akademickiego). Sześć pozostałych dyscyplin z obszaru STEMM (tj. AGRI, ENVIR, CHEMENG, DENT, ENG i EARTH) również ujemnie wpływa na wiek biologiczny w porównaniu z HUM, zmniejszając go średnio o 2–2,5 roku. Natomiast dyscypliny niezwiązane z obszarem STEMM (zwłaszcza największe, takie jak ECON, PSYCH i SOC) również wykazują ujemny wpływ, ale wyraźnie mniejszy niż pozostałe (1–1,5 roku). Analiza ta wskazuje na dominację dyscyplin należących do obszaru STEMM w widocznej na arenie międzynarodowej produkcji naukowej autorstwa młodych naukowców.

Analiza współczynników standaryzowanych pokazuje, że najważniejszym predyktorem wieku biologicznego jest wiek akademicki, dla którego wartość współczynnika wyniosła 0,523 i była znacznie wyższa niż dla drugiego najważniejszego czynnika – stopnia i tytułu naukowego (dla habilitacji wartość współczynnika wyniosła 0,234, a dla profesury tytularnej 0,357). Zatem te dwie cechy stanowią zdecydowanie najsilniejsze determinanty wieku biologicznego. Dyscypliny CHEM, BIO, MED, PHYS, MATER i MATH

(ze wskaźnikami wynoszącymi odpowiednio: 0,13, 0,13,0,12, 0,11, 0,11, 0,09) charakteryzowały się stosunkowo wysokimi, choć znacznie niższymi niż powyżej, wartościami współczynników standaryzowanych, co wskazuje na silny wpływ dyscyplin należących do obszaru STEMM na wiek biologiczny. Pozostałe dyscypliny charakteryzowały się wartością współczynnika standaryzowanego istotnie poniżej 0,1. Ponadto, co ciekawe, silnym predyktorem nie okazała się cecha przynależności do grupy uczelni badawczych (IDUB), dla której wartość współczynnika wyniosła zaledwie 0,06. Zjawisko współliniowości (istotnego skorelowania wektora zmiennych niezależnych) nie wystąpiło w modelu – wartości współczynników VIF w niemal każdym przypadku były niższe od 2 (przy czym przyjmuje się, że dopiero wartość 4 pozwala stwierdzić występowanie istotnej współliniowości).

Tabela 4. Współczynniki modelu liniowego (zmienna zależna: wiek biologiczny; kategorie odniesienia: płeć – kobieta, typ instytucji: uczelnia badawcza (IDUB), dyscyplina ASJC – nauki humanistyczne (HUM) oraz stopień lub tytuł naukowy – doktorat)

$R^2 = 0,615$ SE = 6,59	Współ- czynnik niestanda- ryzowany	Współ- czynnik standa- ryzowany	Błąd std.	Statystyka t-Studenta	Wartość $p$	Współcz. VIF
(Stała)	38,269	–	0,25	152,998	< 0,001	–
Wiek akademicki	0,595	0,523	0,01	90,076	< 0,001	1,867
Płeć: mężczyzna	0,094	0,004	0,101	0,934	0,35	1,179
Instyt. intens. badawcza: pozostałe	1,253	0,055	0,103	12,206	< 0,001	1,089
Dyscyplina: AGRI	–2,46	–0,072	0,278	–8,846	< 0,001	1,585
Dyscyplina: BIO	–5,352	–0,129	0,299	–17,876	< 0,001	1,226
Dyscyplina: BUS	0,29	0,004	0,371	0,78	0,436	1,512
Dyscyplina: CHEM	–5,983	–0,131	0,312	–19,182	< 0,001	1,159
Dyscyplina: CHEMENG	–2,709	–0,034	0,416	–6,517	0	1,383
Dyscyplina: COMP	–3,768	–0,07	0,336	–11,231	< 0,001	1,018
Dyscyplina: DEC	–1,91	–0,008	1,001	–1,909	0,056	1,027
Dyscyplina: DENT	–2,616	–0,014	0,853	–3,068	0,002	1,417

Dyscyplina: EARTH	-2,118	-0,042	0,322	-6,576	0	1,123
Dyscyplina: ECON	-1,724	-0,019	0,457	-3,775	0	1,102
Dyscyplina: ENER	-0,818	-0,008	0,495	-1,653	0,098	2,086
Dyscyplina: ENG	-2,162	-0,069	0,275	-7,875	0	1,521
Dyscyplina: ENVIR	-2,515	-0,06	0,297	-8,463	< 0,001	1,022
Dyscyplina: HEALTH	2,039	0,01	0,937	2,177	0,029	1,367
Dyscyplina: IMMU	-5,787	-0,037	0,72	-8,04	0	1,044
Dyscyplina: MATER	-4,705	-0,105	0,309	-15,243	< 0,001	1,535
Dyscyplina: MATH	-4,912	-0,09	0,336	-14,606	< 0,001	1,356
Dyscyplina: MED,	-3,698	-0,122	0,273	-13,568	< 0,001	2,052
Dyscyplina: PHARM	-5,534	-0,053	0,514	-10,765	< 0,001	1,094
Dyscyplina: PHYS	-5,663	-0,105	0,338	-16,744	< 0,001	1,441
Dyscyplina: PSYCH	-1,533	-0,015	0,49	-3,13	0,002	1,098
Dyscyplina: SOC	-0,878	-0,015	0,336	-2,614	0,009	1,332
Dyscyplina: VET	-4,115	-0,046	0,456	-9,031	< 0,001	1,119
Stopień naukowy: doktor habilit.	5,407	0,234	0,114	47,567	< 0,001	1,279
Tytuł naukowy: profesor tytularny	10,923	0,357	0,171	63,708	< 0,001	1,764

## 5. Dyskusja i wnioski

Nasze badania pokazują, że wykorzystanie wieku akademickiego jako zamiennika wieku biologicznego w badaniach karier akademickich dobrze sprawdza się w przypadku

dyscyplin z obszaru STEMM (przypomnijmy: nauki ścisłe, techniczne, inżynierskie, matematyczne i medyczne).

Natomiast w przypadku dyscyplin spoza obszaru STEMM (w szczególności takich jak HUM, nauki humanistyczne, SOC, nauki społeczne, ECON, ekonomia, ekonometria i finanse czy BUS, biznes, zarządzanie i rachunkowość), wykorzystanie wieku akademickiego wypada zdecydowanie gorzej. Ten negatywny wniosek jest szczególnie ważny dla systemów nauki, które dopiero od niedawna (w ciągu ostatnich 10–15 lat) są szerzej widoczne w globalnych czasopismach naukowych – a więc dla krajów klasyfikowanych jako naukowo „rozwijające się” (do których, oprócz Polski, zalicza się m.in. Portugalie, Słowację, Bułgarię, Rumunię, Iran czy Turcję) (Wagner 2008: 88).

Okazuje się zatem, że różnice w poziomie przydatności stosowania wieku akademickiego do badania karier akademickich na poziomie indywidualnego naukowca dotyczą w różnym stopniu różnych grup krajów: z jednej strony mamy kraje, które tworzą rdzeń globalnej nauki, z drugiej zaś kraje, które tworzą jej peryferia. Zatem chociaż globalna nauka nie powstaje w równym stopniu wszędzie, to w coraz większym stopniu powstaje w miejscach w poprzednich dekadach w nauce globalnej nieobecnych lub słabo obecnych (o globalizacji nauki, zob. Kwiek 2021b). Przykładem procesów rosnącej widzialności kraju w globalnej nauce jest Polska.

W przypadku grupy krajów zaawansowanych naukowo, stanowiących od dawna centrum globalnego obiegu wiedzy, wiek akademicki stosowany jako zamiennik wieku biologicznego doskonale sprawdza się dla wszystkich dyscyplin (Nane i in., 2017; Costas i in., 2015). Natomiast w przypadku grupy krajów „rozwijających się” naukowo, reprezentowanych w tym tekście przez Polskę, zamiennik ten dobrze funkcjonuje tylko dla naukowców z dyscyplin należących do obszaru STEMM. Między wiekiem biologicznym i akademickim stwierdzono dla nich bardzo silną współzależność – co oznacza, że naukowcy pracujący w ramach obszaru STEMM zazwyczaj działają według globalnie obowiązujących wzorców, w ramach których relacja między wiekiem akademickim i wiekiem biologicznym jest wysoce przewidywalna. I tak na przykład dla naukowców zajmujących się chemią korelacja między wiekiem akademickim a wiekiem biologicznym jest bardzo wysoka (współczynnik korelacji  $r = 0,89$ ), podobnie jak w przypadku fizyki i astronomii ( $r = 0,88$ ) oraz matematyki ( $r = 0,85$ ). Korelacja ta jest radykalnie niższa dla naukowców spoza obszaru STEMM, dla których ten wskaźnik sprawdza się jedynie w przedziale 35–50%: jest najniższa dla naukowców z takich dyscyplin jak HUM ( $r = 0,35$ ), BUS ( $r = 0,42$ ), ECON ( $r = 0,43$ ) czy PSYCH ( $r = 0,49$ ).

Polski przypadek potwierdza istnienie znaczących różnic w globalnych wzorcach publikacyjnych między krajami „rozwijającymi się” naukowo a krajami naukowo „zaawansowanymi” (Wagner 2008: 88). Polska stanowi dobry przykład kraju pozostającego poza tradycyjnymi akademickimi centrami nauki, w którym jednak systematycznie roś-

nie liczba publikacji w globalnych czasopismach naukowych. O ile kraje zachodnie funkcjonują we wspólnych anglojęzycznych globalnych sieciach naukowych od kilkadziesiąt lat, a od 1989 r. są w nich obecne z radykalnie rosnącą intensywnością, o tyle Polska i inne kraje rozwijające się naukowo uczestniczą w tych sieciach – na większą skalę – dopiero od 10–15 lat, co dokładnie pokazują coroczne dane pochodzące z globalnych baz indeksujących publikacje. W 2000 r. polscy naukowcy opublikowali 12 861 artykułów naukowych, w 2010 r. – 22 377, w 2015 r. – 32 408, a w 2021 r. już 46 967, co obrazuje dynamikę zmian (SciVal 2022).

Różnice geopolityczne spowodowały, że postkomunistyczne kraje członkowskie Unii Europejskiej, w tym Polska (oraz inne biedniejsze kraje świata), stosunkowo niedawno stały się szerzej widoczne w czasopismach indeksowanych w bazach Scopus czy Web of Science. Z kolei kraje zachodnie od samego początku funkcjonowały na dużą skalę (i za pośrednictwem znacznego odsetka swoich naukowców) w systemie czasopism indeksowanych. Przypomnijmy, że w dekadzie 2009–2018 tylko 25 463 naukowców co najmniej z doktoratem i zatrudnionych w szkolnictwie wyższym funkcjonujących w bazie OPI („Nauka Polska”) było autorami lub współautorami co najmniej jednego artykułu indeksowanego w bazie Scopus. Różnice geopolityczne przekładają się na praktyczne konsekwencje związane z wykorzystaniem wieku akademickiego w analizach na poziomie indywidualnym poszczególnych naukowców.

Nasze badania potwierdzają, że globalizacja nauki wiąże się z istotnym zróżnicowaniem między dyscyplinami z obszaru STEMM i spoza tego obszaru (na temat szczegółowych różnic wzorców publikacyjnych i cytowaniowych 25 największych producentów publikacji naukowych, zob. Kwiek 2021a). Dane na poziomie indywidualnym odzwierciedlają opóźniony udział naukowców społecznych i humanistów w globalnych sieciach naukowych, w przeciwieństwie do nieprzerwanej w nich obecności naukowców specjalizujących się w naukach przyrodniczych. Tak więc podczas gdy nauka globalizuje się szybko, globalizuje się ona zarazem nierównomiernie: szybciej w przypadku nauk z obszaru STEMM niż w przypadku nauk spoza obszaru STEMM.

Nasze analizy wzmocniają tezę, że warto gromadzić kompleksowe dane na poziomie indywidualnym w krajowych bazach danych: do szczegółowych analiz kariery akademickiej w jej różnych wymiarach (produktywność badawcza, międzynarodowa współpraca badawcza, mobilność międzynarodowa) przydatne jest bowiem dysponowanie datami urodzenia wszystkich naukowców ze wszystkich sektorów nauki, tak jak w rozbudowanym obecnie polskim systemie POL-on. W takich systemach analizy krajowe nie wymagają stosowania zamienników wieku biologicznego – czyli nie ma potrzeby stosowania wieku akademickiego, a wyniki analiz są jednoznaczne co do wieku czy grupy wiekowej naukowca.

Nasze badania pokazują – po raz pierwszy na tak dużą skalę – że w krajach rozwijających się naukowo wiek akademicki jako zamiennik wieku biologicznego musi być stosowany z większą ostrożnością niż w krajach rozwiniętych. Byłoby najlepiej, gdyby był on stosowany tylko w przypadku dyscyplin należących do obszaru STEMM. Nie-spójności między tymi dwoma typami wieku w dyscyplinach spoza STEMM są bowiem większe, a korelacje radykalnie niższe, co czyni ten wskaźnik niedoskonałym narzędziem analitycznym.

Jednocześnie analizowany tutaj polski przypadek nawiązuje do kilku równoległych procesów widocznych w nauce globalnej: (1) obecności na peryferiach globalnej nauki z powodów geopolitycznych; (2) niskiej, ale rosnącej indywidualnej produktywności badawczej (Kwiek 2016, 2018a); (3) opóźnionego wejścia do globalnej nauki i do obiegu globalnie indeksowanych czasopism akademickich; oraz (4) permanentnego niedofinansowania akademickiego systemu nauki (zwłaszcza w pierwszych dwóch dekadach po 1989 r.). Procesy te są charakterystyczne dla znacznej części świata i dotyczą, w różnym stopniu, zarówno kraje postkomunistyczne w Europie i Azji Środkowej, jak i kraje Ameryki Łacińskiej oraz Afryki. Transformacje polskiego systemu nauki akademickiej były do niedawna bliższe transformacjom zachodzącym na globalnym Południu niż na globalnej Północy, a najważniejszym czynnikiem politycznym i ekonomicznym, decydującym o trwałej przynależności do krajów globalnej Północy, było pełne członkostwo Polski w Unii Europejskiej uzyskane w 2004 roku i płynące w związku z tym, chociaż z czasem malejące, dodatkowe środki na naukę.

Instytucje akademickie na całym świecie dysponują pełnymi danymi na temat wieku biologicznego swojej kadry. Natomiast dostępność tak szczegółowych danych na makropoziomie całych krajów pozostaje niezwykle ograniczona, co utrudnia możliwość prowadzenia rzetelnych analiz profesji akademickiej w ujęciu demograficznym. Przeprowadzone przez nas badania mogą być powielane – można je prowadzić wszędzie tam, gdzie istnieją jednoznaczne dane dotyczące wieku biologicznego naukowców i wiarygodne dane historyczne dotyczące publikacji. Globalne badania porównawcze profesji akademickiej wymagają dysponowania solidnymi danymi demograficznymi na skalę krajową, dlatego niezwykle istotne jest tworzenie nowych i doskonalenie istniejących baz danych.

### **Podziękowania**

Dziękujemy za wsparcie w postaci grantu Dialog 0022/DLG/2019/10 (RESEARCH UNIVERSITIES) oraz za pomoc ze strony International Center for the Studies of Research Lab (Elsevier). Jesteśmy szczególnie wdzięczni Kristy James, analityczce danych w tym laboratorium, za stałe wsparcie. Jesteśmy również wdzięczni Panu Łukaszowi Szymuli, doktorantowi w Centrum Studiów nad Polityką Publiczną i na Wydziale Informatyki UAM.



## Literatura

- Abramo G., Aksnes D.W., D'Angelo C.A. (2020). *Comparison of research productivity of Italian and Norwegian professors and universities*. *Journal of Informetrics*, 14(2), 101023.
- Abramo G., C.A. D'Angelo, Murgia, G. (2016). *The combined effect of age and seniority on research performance of full professors*. *Science and Public Policy*, 43(3), 301–319.
- Abramo G., D'Angelo C.A., Solazzi M. (2011). *The relationship between scientists' research performance and the degree of internationalization of their research*. *Scientometrics*, 86, 629–643.
- Aksnes D.W., Rørstad K., Piro F.N., Sivertsen G. (2011). *Age and Scientific Performance. A Large-Scale Study of Norwegian Scientists*. [w:] Ed. Noyons P. Ngulube, J. Leta, *Proceedings of ISSI 2011 – the 13th International Conference of the International Society for Scientometrics and Informetrics*, t. 1, Durban, South Africa, 4–7 July 2011, 34–45.
- Antonowicz D., Kulczycki E., Budzanowska A. (2020). *Breaking the deadlock of mistrust? A participative model of the structural reforms in higher education in Poland*. *Higher Education Quarterly*, On-line first February 14, 2020, <https://doi.org/10.1111/hequ.12254>.
- Aref S., Zagheni E., West J. (2019) *The demography of the peripatetic researcher: evidence on highly mobile scholars from the web of science*. [w:] I. Weber et al. (red.) *Social Informatics. SocInfo 2019*. Lecture Notes in Computer Science (t. 11864). Springer.
- Badar K., M. Hite, J., F. Badir Y. (2014). *The moderating roles of academic age and institutional sector on the relationship between co-authorship network centrality and academic research performance*. *Aslib Journal of Information Management*, 66(1), 38–53. doi: 10.1108/ajim-05-2013-0040
- Bonaccorsi A., Daraio C. (2003). *Age effects in scientific productivity. The case of the Italian National Research Council (CNR)*. *Scientometrics*, 58(1), 49–90.
- Chan H.F., Torgler B. (2020). *Gender differences in performance of top cited scientists by field and country*. *Scientometrics*, 125, 2421–2447.
- Cole S. (1979). *Age and scientific performance*. *American Journal of Sociology*, 84(4), 958–977.
- Coomes O.T., Moore T., Paterson J., Breau S., Ross N.A., Roulet N. (2013). *Academic performance indicators for departments of geography in the United States and Canada*. *The Professional Geographer*, 65 (3), 433–450.
- Costas R., Nane G.F., Larivière V. (2015). *Is the year of first publication a good proxy of scholars academic age?* [w:] A.A. Salah, Y. Tonta, A.A. Akdag Salah (red.), *Proceedings of the 15th international conference on scientometrics and informetrics* (ss. 988–998). Istanbul: Boga-zici University Printhouse.
- Costas R., van Leeuwen T.N, Bordons M. (2010). *A bibliometric classificatory approach for the study and assessment of research performance at the individual level: the effects of age on productivity and impact*. *Journal of the American Society for Information Science and Technology*, 61 (8), 1564–1581
- Gingras Y., Larivière V., Macaluso B., Robitaille J.P. (2008). *The effects of aging on researchers' publication and citation patterns*. *PLoS One*. 3(12), e4048. doi: 10.1371/journal.pone.0004048. Epub 2008 Dec 29. PMID: 19112502; PMCID: PMC2603321.
- Guns R., Eykens J., Engels T.C.E. (2019). *To what extent do successive cohorts adopt different publication patterns? Peer review, language use, and publication types in the social sciences and humanities*. *Frontiers in Research Metrics and Analytics* 3. doi: 10.3389/frma.2018.0003
- GUS (2021). *Szkoły Wyższe i ich finanse w 2020 r.* Warszawa: GUS.

- Kwiek M. (2015a). *Uniwersytet w dobie przemian. Instytucje i kadra akademicka w warunkach rosnącej konkurencji*. Warszawa: PWN.
- Kwiek M. (2015b). *The internationalization of research in Europe. A quantitative study of 11 national systems from a micro-level perspective*. *Journal of Studies in International Education*, 19(2), 341–359.
- Kwiek M. (2015c). *Academic generations and academic work: Patterns of attitudes, behaviors and research productivity of Polish academics after 1989*. *Studies in Higher Education*, 40(8), 1354–1376.
- 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.
- Kwiek M. (2018a). *High research productivity in vertically undifferentiated higher education systems: Who are the top performers?* *Scientometrics*, 115(1), 415–462.
- Kwiek M. (2018b). *International research collaboration and international research orientation: Comparative findings about European academics*. *Journal of Studies in International Education*, 22(1), 1–25.
- Kwiek M. (2019a). *Changing European academics. A comparative study of social stratification, work patterns and research productivity*. London and New York: Routledge.
- Kwiek M. (2019b). *Social Stratification in Higher Education: What It Means at the Micro-Level of the Individual Academic Scientist*. *Higher Education Quarterly*, 73(3), 419–444.
- Kwiek M. (2020). *Internationalists and locals: International research collaboration in a resource-poor system*. *Scientometrics*, 124, 57–105.
- Kwiek M. (2021a). *What large-scale publication and citation data tell us about international research collaboration in Europe: Changing national patterns in global contexts*. *Studies in Higher Education*, 46(12), 2629–2649.
- Kwiek M. (2021b). *Globalizacja nauki: rosnąca siła indywidualnych naukowców*. *Nauka* 4/2021, 37–66.
- Kwiek M. (2021c). *Globalny system akademicki i stratyfikująca rola badań naukowych*. *Człowiek i Społeczeństwo*. Tom LII, 72–90.
- 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), 1344–1380.
- Kwiek M., Roszka, W. (2021b). *Gender-based homophily in research: A large-scale study of man-woman collaboration*. *Journal of Informetrics*, 15(3), article 101171. 1–38.
- Kwiek, M., Roszka, W. (2022). *Are female scientists less inclined to publish alone? The gender solo research gap*. *Scientometrics*. Online first March 5, 2022. <https://doi.org/10.1007/s11192-022-04308-7>
- Kyvik S. (1990). *Age and scientific productivity. Differences between fields of learning*. *Higher Education*, 19, 37–55.
- Kyvik, S., Aksnes, D.W. (2015). *Explaining the increase in publication productivity among academic staff: A generational perspective*. *Studies in Higher Education*, 40, 1438–1453.
- Kyvik S., Olsen T.B. (2008). *Does the aging of tenured academic staff affect the research performance of universities?* *Scientometrics*, 76(3), 439–455.
- Larivière V., Vignola-Gagné E., Villeneuve C. et al. (2011). *Sex differences in research funding, productivity and impact: an analysis of Québec university professors*. *Scientometrics*, 87, 483–498.
- Lehman H.C. (1953). *Age and achievement*. Princeton: Princeton University Press.

- Levin S., Stephan P.E. (1991). *Research productivity over the life cycle: Evidence for academic scientists*. The American Economic Review, March 1991, 114–132.
- Milojević S. (2012). *How are academic age, productivity, and collaboration related to citing behavior of researchers?* PLoS One, 7(11), e49176. doi: 10.1371/journal.pone.0049176. Epub 2012 Nov 7. PMID: 23145111; PMCID: PMC3492318.
- Olechnicka A., Ploszaj A., Celinska-Janowicz D. (2019). *The geography of scientific collaboration*. London and New York: Routledge.
- Pelz D.C., Andrews F.W. (1976). *Scientists in organizations*. New York: Wiley.
- Perianes-Rodriguez A., Ruiz-Castillo, J. (2014). *Within- and between-department variability in individual productivity: the case of economics*. Scientometrics, 102(2), 1497–1520.
- Petersen A.M. (2015). *On the impact of super ties in scientific careers*. Proceedings of the National Academy of Sciences Aug 2015, 112 (34), E4671-E4680. doi: 10.1073/pnas.1501444112
- Radicchi F., Castellano, C. (2013). *Analysis of bibliometric indicators for individual scholars in a large data set*. Scientometrics, 97, 627–637.
- Robinson-Garcia N., Costas R., Sugimoto C.R., Larivière V., Nane G.F. (2020). *Task specialization across research careers*. eLife, 9, e60586 doi: 10.7554/eLife.60586
- Rørstad K., Aksnes D.W. (2015). *Publication rate expressed by age, gender and academic position – A large-scale analysis of Norwegian academic staff*. Journal of Informetrics, 9, 317–333.
- SciVal (2022). *Funkcjonalność SciVal bazy danych Scopus*, strona www.scival.com (dostęp ograniczony do uprawnionych użytkowników).
- Simoes N., Crespo, N. (2020). *A flexible approach for measuring author-level publishing performance*. Scientometrics, 122, 331–355.
- Stephan P. (2012). *How economics shapes science*. Cambridge, MA: Harvard University Press.
- Stephan P.E., Levin S.G. (1992). *Striking the mother lode in science: the importance of age, place, and time*. New York: Oxford University Press.
- Stern N. (1978). *Age and achievement in mathematics: A case-study in the sociology of science*. Social Studies of Science, 8(1), 127–140.
- Sugimoto C.R., Sugimoto T.J., Tsou A., Milojević S., Larivière, V. (2016). *Age stratification and cohort effects in scholarly communication: A study of social sciences*. Scientometrics, 109(2), 997–1016.
- Van den Besselaar, P., Sandström, U. (2016) *Gender differences in research performance and its impact on careers: A longitudinal case study*. Scientometrics, 106(1), 143–162.
- Wagner C.S. (2008). *The new invisible college. Science for development*. Brookings Institution Press.
- Wais K. (2016). *Gender Prediction Methods Based on First Names with genderize R*. The R Journal, 8(1), 17–37.
- Zuckerman H., Merton R.K. (1973). *Age, aging, and age structure in science*. [w:] M.W. Riley, M. Johnson, A. Foner (red.), *Aging and Society*, t. 3. *A Sociology of Age Stratification*. Russell Sage Foundation.

### Global science and academic age of Polish researchers

In this text, we analyze the limitations of using academic age as a proxy for biological age in the whole national science system, for which we consider both bio-

logical age and academic age of all researchers from all Polish universities, holding at least a PhD degree and participating in global academic science through international publications ( $N= 20\ 569$ ). An approximation of a researcher's functioning in global science is having at least one publication indexed in the Scopus database in the analyzed decade 2009–2018. Thus, using the example of comprehensive data from the entire national system of science, we estimate the extent of limitations of using academic age as a proxy for biological age depending on selected independent variables and analyze both the practical and methodological implications of using academic age in academic career research, which is one of the most important components of quantitative studies of science.

**Key words:** global science, academic age, sociology of science, quantitative science studies, science of science, publishing patterns, academic profession