

Beata Jackowska

Katedra Statystyki, Wydział Zarządzania
Uniwersytet Gdański
beata.jackowska@ug.edu.pl
ORCID: 0000-0002-2617-0150

Zmiany tendencji rozwojowej wzorca umieralności w Polsce w latach 1958–2019

Streszczenie

W kilkudziesięcioletniej powojennej historii ewolucji wzorca umieralności w Polsce występowały okresy zaburzeń w tendencji malejącej intensywności zgonów. Zmiany wzorca umieralności przebiegały w różny sposób w grupach wieku i płci. W artykule dokonano próby identyfikacji przedziałów wieku i lat kalendarzowych, w których była utrzymana tendencja rozwojowa wzorca umieralności w Polsce. W celu wyodrębnienia tendencji rozwojowej zastosowano dwuwymiarowe wyrównywanie P-splajnami, przy założeniu, że liczba zgonów w danej grupie wieku i w określonym roku ma rozkład Poissona. Wyrównane współczynniki zgonów względem wieku i lat kalendarzowych posłużyły następnie do wyznaczenia prawdopodobieństw zgonu i przeciętnego dalszego trwania życia. W ostatnim kroku dla poszczególnych roczników wieku zidentyfikowano przedziały czasu, w których wystąpiły zaburzenia w tendencji rozwojowej prawdopodobieństw zgonu i przeciętnego dalszego trwania życia. W tym celu zaproponowano przybliżenie numeryczne pierwszej pochodnej.

Słowa kluczowe: umieralność, dwuwymiarowe wyrównywanie, P-splajny, regresja Poissona

Wprowadzenie

W latach 70. i 80. XX wieku nie tylko w Polsce, ale także w innych krajach Europy Środkowo-Wschodniej były obserwowane anomalie w tendencji rozwojowej wzorca umieralności, nazywane kryzysem zdrowotnym (Okólski, 1993). Anomalie te polegały na spowolnieniu spadku lub nawet wzroście intensywności zgonów, przy czym kierunek i tempo zmian były zależne od wieku i płci. Kryzys zdrowotny najbardziej dotknął starsze roczniki wieku produkcyjnego oraz bardziej mężczyzn niż kobiety (Kędzelski, 1991). Po 1989 r., gdy Polska weszła na drogę transformacji ustrojowej i gospodarczej, powróciła tendencja malejąca intensywności zgonów.

Celem artykułu jest przedstawienie metodyki pozwalającej na identyfikację wieku i roku, w których następuje zmiana tendencji rozwojowej wzorca umieralności, a w następnej kolejności empiryczna weryfikacja przydatności tej metodyki do wyodrębnienia przedziałów wieku i lat kalendarzowych, w których była utrzymana tendencja rozwojowa wzorca umieralności w Polsce w latach 1958–2019. Rozwiązanie problemu nie jest oczywiste, gdyż nieregularne wahania wynikające z losowej realizacji liczby zgonów nie pozwalają na jednoznaczne wskazanie punktów odciążenia na osiach wieku i czasu kalendarzowego, wyznaczających przedziały o tendencji rosnącej, malejącej bądź stałej.

Do wyodrębnienia tendencji rozwojowej wzorca umieralności wykorzystano metodę wyrównywania¹, stworzoną specjalnie do wygładzania współczynników zgonów. W ostatnich dziesięcioleciach były rozwijane zaawansowane metody wyrównywania danych (Wang, 2011), w tym P-splajny przeznaczone do wyrównywania danych populacyjnych dotyczących umieralności w układzie dwuwymiarowym wiek–okres, przy założeniu, że liczba zgonów w danej grupie wieku i w danym roku kalendarzowym ma rozkład Poissona (Camarda, 2012a). Wyrównywanie dwuwymiarowe ma tę zaletę, że nie pozwala na usunięcie „lokalnych prawidłowości” w jednym wymiarze, gdy w drugim wymiarze jest to „prawidłowość globalna”, przykładowo nie zostanie usunięty charakterystyczny garb na krzywej logarytmicznej prawdopodobieństwa zgonu osób w młodym wieku², jeżeli powtarza się w kolejnych latach. Jako kryterium rozstrzygające o monotoniczności trendu zaproponowano przybliżenie numeryczne pierwszej pochodnej, umożliwiające uchwycenie nie tylko zmian kierunku trendu, lecz także szybkości ewolucji wzorca umieralności.

¹ Wyrównywanie szeregów czasowych prowadzi do wyodrębnienia trendu poprzez częściową eliminację wahań okresowych i przypadkowych.

² Garb na krzywej obrazuje dodatkową („nadmiarową”) umieralność osób młodych spowodowaną przede wszystkim przyczynami zewnętrznymi, a w przypadku kobiet związanymi z macierzyństwem.

W części empirycznej posłużono się surowymi (niewyrównanymi) danymi przekrojowymi z bazy Human Mortality Database (HMD, 2020) dla lat 1958–2018, a dla roku 2019 z Bazy Demografia GUS (GUS, 2020a)³. Wyrównano dwuwymiarowo współczynniki zgonów i na tej podstawie skonstruowano przekrojowe tablice trwania życia dla 62 lat kalendarzowych. W analizie zmian tendencji rozwojowej wzorca umieralności wzięto pod uwagę dwie podstawowe funkcje tablicowe: logarytm prawdopodobieństwa zgonu oraz oczekiwane dalsze trwanie życia. Aproksymacja pierwszej pochodnej tych funkcji pozwoliła na wyodrębnienie przedziałów wieku i lat kalendarzowych o tendencji rosnącej, malejącej i stałej. Na wykresach przedstawiono dodatkowo zróżnicowanie szybkości zmian wzorca umieralności.

Niniejszy artykuł ma charakter metodyczno-narzędziowy i stanowi próbę nowego podejścia do analizy zmian w tendencji rozwojowej wzorca umieralności w Polsce. Celem dalszych pogłębionych badań powinna stać się próba identyfikacji źródeł wahań w danych, co pozwoliłoby na nadanie interpretacji otrzymanym wynikom.

Metodyka szacowania rozkładu czasu trwania życia

W obliczeniach wykorzystano metodykę oraz następujące dane przekrojowe z bazy Human Mortality Database (HMD, 2020) dla poszczególnych lat kalendarzowych 1958–2018:

- 1) tablice trwania życia według płci,
- 2) liczby zgonów według płci i wieku,
- 3) liczby narażonych według płci i wieku,
- 4) współczynniki zgonów według płci i wieku.

Przekrojowa tablica trwania życia jest wzorcem wymierania hipotetycznej kohorty⁴ przy założeniu, że doświadcza ona natężenia umieralności zaobserwowanej w krótkim okresie, w tym przypadku w ciągu jednego roku kalendarzowego. Dzięki takiej konstrukcji tablic można w kolejnych latach śledzić zmiany wzorca umieralności w populacji. Metodyka konstrukcji tablic trwania życia stosowana w HMD została

³ W momencie dokonywania obliczeń w bazie HMD nie było jeszcze danych dla Polski z 2019 r. W analizach zmian tendencji rozwojowej wzorca umieralności w Polsce nie można wykorzystać tablic trwania życia GUS, gdyż są to tablice modelowe, wyrównane jednowymiarowo według wieku w jednym roku kalendarzowym.

⁴ Tablica kohortowa prezentuje porządek wymierania rzeczywistej kohorty urodzeniowej i może być skonstruowana dopiero po wymarcu prawie wszystkich członków kohorty. Zatem ma znaczenie historyczne, lecz ciąg tablic kohortowych dla kolejnych generacji może stanowić podstawę do prognozowania. Dane retrospektywne dla Polski pozwalają obecnie na konstruowanie tablic kohortowych począwszy od wieku dorosłego (Jackowska, 2013).

opracowana przez międzynarodowy zespół uczonych⁵ zaangażowanych w projekt tworzenia bazy danych o otwartym dostępie i jednolitej metodyce (Wilmoth i in., 2019). Obecnie baza HMD zawiera dane dotyczące umieralności w populacjach 41 krajów.

Podstawą budowy tablic przekrojowych są obserwowane w danym roku kalendarzowym t przekrojowe współczynniki zgonów postaci (por. Wilmoth i in., 2019, s. 29; Jackowska, 2013, s. 124):

$$M(x, t) = \frac{D(x, t)}{E(x, t)}, \quad (1)$$

gdzie:

$D(x, t)$ – liczba zgonów wśród osób w wieku x w roku t ,

$E(x, t)$ – liczba osób narażonych na zgon w wieku x w roku t .

Liczba narażonych jest szacowana jako:

$$E(x, t) = \frac{1}{2}[P(x, t) + P(x, t+1)] + \frac{1}{6}[D_L(x, t) - D_U(x, t)], \quad (2)$$

gdzie:

$P(x, t)$ i $P(x, t+1)$ – liczba ludności w wieku x lat na początku lat t i $t+1$ odpowiednio, $D_L(x, t)$ – liczba zgonów w roku t wśród osób w wieku x lat, urodzonych w roku $t-x$, tzw. zgony „starsze” prezentowane na diagramie Lexisa w dolnym trójkącie (kąąt prosty na dole, indeks L – *lower*),

$D_U(x, t)$ – liczba zgonów w roku t wśród osób w wieku x lat, urodzonych w roku $x-t-1$, tzw. zgony „młodsze” prezentowane na diagramie Lexisa w górnym trójkącie (kąąt prosty u góry, indeks U – *upper*).

Występuje związek:

$$D(x, t) = D_L(x, t) + D_U(x, t). \quad (3)$$

Tablice wymieralności dla roku t buduje się przy założeniu, że obserwowane współczynniki zgonów $M(x, t)$ są równe tablicowym współczynnikom zgonów $m(x, t)$ poniżej wieku Y . Wiek Y jest najniższym wiekiem, w którym odnotowano co najwyżej 100 zgonów, pod warunkiem, że Y zawiera się w przedziale $80 \leq Y \leq 95$. Jeżeli dla każdego Y z przedziału $80 \leq Y \leq 95$ liczba zgonów jest większa niż 100, to $Y = 95$ lat⁶. W wieku Y lat i więcej wyrównuje się współczynniki zgonów za pomocą modelu analitycznego dopasowanego do wartości obserwowanych w wieku 80, 81, ...

⁵ Projekt został zapoczątkowany dzięki współpracy naukowców z Katedry Demografii Uniwersytetu Kalifornijskiego w Berkeley (USA) oraz z Instytutu Badań Demograficznych im. Maxa Plancka w Rostocku (Niemcy) (HMD, 2020).

⁶ W Polsce $Y = 95$ lat dla kobiet w całym badanym okresie i od roku 1961 dla mężczyzn, lecz warunek nałożony na wiek Y musi być spełniony dla mężczyzn i kobiet jednocześnie.

110+ lat. W tym celu przyjmuje się dla intensywności zgonów osób w wieku 80 lat i więcej funkcję logistyczną z dwoma parametrami, znaną w literaturze jako model Kannisto (Thatcher, Kannisto, Vaupel, 1998). Zatem tablice HMD poniżej wieku Y są tablicami surowymi (niewyrównanymi), a od tego wieku – modelowymi.

Współczynniki zgonów przelicza się na prawdopodobieństwa zgonu $q(x, t)$ w ciągu roku życia, pod warunkiem dożycia do wieku x lat, stosując zasadę zazwyczaj przyjmowaną przy konstrukcji tablic wymieralności (Bolesławski, 1973; Chiang, 1984; Preston, Heuveline, Guillot, 2001):

$$q(x, t) = \frac{m(x, t)}{1 + (1 - a(x, t)) \cdot m(x, t)}, \quad (4)$$

gdzie $a(x, t)$ jest przeciętną liczbą lat przeżywanych w rocznym przedziale wieku $[x, x + 1)$ przez osoby, które zmarły w tym przedziale. Przyjmowane jest założenie równomiernego rozkładu zgonów wewnątrz rocznych przedziałów wieku, tzn. $a(x, t) = \frac{1}{2}$, z dwoma wyjątkami: $x = 0$ oraz ostatniego otwartego przedziału wieku 110+, dla których stosuje się specjalne formuły (Wilmoth i in., 2019). Ciąg prawdopodobieństw zgonu $q(x, t)$ w roku t dla wieku $x = 0, 1, \dots, \omega$ lat⁷ określa rozkład czasu trwania życia i pozwala na wyznaczenie pozostałych funkcji tablicowych, w tym przeciętnego dalszego trwania życia $e(x, t)$.

W momencie wykonywania obliczeń w bazie HMD nie były dostępne dane dla Polski z 2019 r. Surowe (niewyrównane) tablice trwania życia dla 2019 r. skonstruowano według metodyki HMD na podstawie danych GUS (GUS, 2020a): liczby urodzonych w 2019 r. według płci, liczby osób na koniec lat 2018 i 2019 według płci i wieku, liczby zgonów w 2019 r. według płci, wieku i roku urodzenia.

Metodyka konstrukcji tablic GUS różni się od metodyki HMD. Wartości prawdopodobieństwa zgonu $q(x, t)$ oszacowane metodą GUS przed wyrównaniem są bardzo zbliżone do prawdopodobieństw oszacowanych metodą HMD. Przykładowo, różnice między niewyrównanymi prawdopodobieństwami zgonu w wieku 0–99 lat oszacowanymi obiema metodami dla roku 2019 zawierają się między –1% a 1% (w tym prawie wszystkie różnice zawierają się między –0,5% a 0,5%). Różnice między tablicami HMD i GUS wynikają głównie stąd, że zgodnie z metodyką GUS surowe prawdopodobieństwa zgonu od wieku 1 do 84 lat są wyrównywane mechanicznie za pomocą średnich ruchomych ważonych, a w wieku od 85 do 120 lat stosuje się ekstrapolację modelu analitycznego dopasowanego w punktach $x = 40, 45, \dots, 85$ (GUS, 2020b)⁸.

⁷ Wartość ω określa maksymalny wiek osiągnięty w populacji.

⁸ Zatem dane powyżej wieku 85 lat nie są wykorzystywane do budowy tablic trwania życia GUS. Metodyka konstrukcji polskich tablic wymieralności została opracowana przez Lecha Bolesławskiego

Tablice GUS są więc w całości tablicami modelowymi wyrównanymi jednowymiarowo według wieku, co może zakłócać badanie trendu.

W dalszej części pracy wyrównano dwuwymiarowo współczynniki zgonów dane wzorem (1) według wieku i lat kalendarzowych. Następnie zgodnie z metodyką HMD wykorzystano je do zbudowania wyrównanych tablic trwania życia dla lat 1958–2019.

Wyrównywanie dwuwymiarowe współczynników zgonów

Tablice trwania życia otrzymane bezpośrednio na podstawie współczynników zgonów (wzór (1)) charakteryzują się nieregularnymi wahaniami wynikającymi z losowej realizacji liczby zgonów⁹. Z tego powodu tablice wyrównuje się. W tym celu wyrównuje się jedną z wybranych funkcji tablicowych i na jej podstawie buduje się całą tablicę¹⁰. Istnieje wiele metod wyrównywania funkcji tablicowych. Najczęściej stosowane grupy metod to: analityczne (za pomocą z góry ustalonej postaci funkcji matematycznej), mechaniczne (różnego rodzaju formuły wielokrotnych ważonych średnich ruchomych) oraz oparte na splajnach¹¹ (wyrównywanie analityczne w podprzedziałach). Pojęcie wyrównywania (*graduation*) bywa często utożsamiane z pojęciem wygładzania (*smoothing*). W alternatywnym podejściu stopień wygładzenia jest cechą wyrównanego ciągu wartości, a ocena wyrównania opiera się na kryterium gładkości oraz kryterium dopasowania (Jackowska, 2013). Zwiększanie gładkości pociąga za sobą zmniejszanie dobroci dopasowania wartości wyrównanych do surowych. Proces wyrównywania szeregów statystycznych jest zatem poszukiwaniem kompromisu między stopniem wygładzenia a stopniem dopasowania.

W celu wyeliminowania nieregularności i wyodrębnienia tendencji rozwojowej wzorca umieralności, wyrównano dwuwymiarowo współczynniki zgonów jednocześnie według wieku i lat kalendarzowych za pomocą P-splajnow. Do wyrównywania zastosowano pakiet programu R MortalitySmooth (Camarda, 2012b). Podstawy teoretyczne algorytmów użytych w pakiecie opisano w artykule Camarda (2012a).

Związek między wiekiem a umieralnością jest nieliniowy. W demografii i ubezpieczeniach na życie były rozwijane od dawna parametryczne modele umieralno-

i zastosowana po raz pierwszy do skonstruowania tablic dla roku 1995 (Bolesławski, 1996). Metodyka ta z niewielkimi korektami jest stosowana w GUS do dziś.

⁹ Wahania względne są największe w młodym wieku, gdy zgon jest zdarzeniem rzadkim oraz w zaawansowanym wieku, do którego dożywa niewiele osób (mała liczba narażonych na zgon).

¹⁰ Nie można wyrównywać każdej funkcji tablicowej z osobna, gdyż nie zostaną zachowane związki między tymi funkcjami.

¹¹ Wygładzanie splajnami cieszy się od kilkunastu lat dużym zainteresowaniem naukowców. Powstało wiele technik wygładzania splajnami adresowanych do konkretnych zastosowań (Wang, 2011).

ści. Podejście parametryczne wymaga założenia *a priori* postaci modelu, a złożona postać modelu z wieloma parametrami prowadzi do problemów numerycznych. Z tego powodu coraz częściej jest stosowane podejście nieparametryczne, które nie wymaga założenia matematycznej postaci modelu umieralności, a daje gładkie oszacowanie funkcji tablicowych. Do metod takich należy rozważana w pracy metoda wyrównywania współczynników zgonów za pomocą P-splajnow z regresją Poissona, która znalazła uznanie wśród demografów i aktuariuszy (Currie, Durban, Eilers, 2004; Djeundje, Currie, 2011; Ouellette, Bourbeau, 2011).

Splajn (*spline*), nazywany inaczej funkcją sklejaną lub funkcją giętą, składa się z fragmentów prostych funkcji, najczęściej wielomianów niskiego stopnia, gładko sklejonych na połączeniach. Wyrównywanie splajnami wartości zmiennej objaśnianej polega na podziale zakresu zmiennych objaśniających na przedziały (w tym przypadku wieku i lat kalendarzowych). Punkty odgraniczające przedziały nazywa się węzłami (*knots*). W wielowymiarowej dziedzinie, węzły tworzą siatkę (*grid*) podziału. Przy dopasowywaniu splajnow nakłada się ograniczenia: w węzłach splajny łączą się w sposób gładki, co jest osiągnięte poprzez nałożenie odpowiednich warunków na pochodne łączonych funkcji. Dla każdej ze zmiennych objaśniających określa się lokalnie splajny bazowe. Funkcja sklejana powstaje jako kombinacja liniowa splajnow bazowych (de Boor, 2001).

W zagadnieniu dwuwymiarowego wyrównywania dane dotyczące umieralności przedstawia się w następującym układzie:

$\mathbf{Y}=[Y_{ij}]$ – macierz $m \times n$ liczby zgonów, $Y_{ij} = D(x_i, t_j)$, $i = 1, 2, \dots, m, j = 1, 2, \dots, n$,
 $\mathbf{E}=[E_{ij}]$ – macierz $m \times n$ liczby narażonych, $E_{ij} = E(x_i, t_j)$, $i = 1, 2, \dots, m, j = 1, 2, \dots, n$.

Szukane jest gładkie oszacowanie intensywności zgonów $\boldsymbol{\mu}=[\mu_{ij}]$. Zakłada się, że liczba zgonów w populacji w wieku x_i oraz roku t_j ma rozkład Poissona $Y_{ij} \sim P(E_{ij} \cdot \mu_{ij})$ z wartością oczekiwaną będącą iloczynem liczby narażonych i intensywności zgonów. Jako nieparametryczne oszacowanie intensywności przyjmuje się obserwowane

współczynniki zgonów $M_{ij} = \frac{Y_{ij}}{E_{ij}}$ (por. wzór (1)), które tworzą macierz $\mathbf{M}=[M_{ij}]$.

Model regresji Poissona w ujęciu dwuwymiarowym można zapisać jako (Camarda, 2012a, s. 5):

$$\ln(E(\mathbf{Y})) = \ln(\mathbf{E}) + \ln(\mathbf{M}) = \ln(\mathbf{E}) + \mathbf{B}_a \mathbf{A} \mathbf{B}_y^t, \quad (5)$$

gdzie:

$$\ln(\mathbf{M}) = \mathbf{B}_a \mathbf{A} \mathbf{B}_y^t \quad (6)$$

oraz

\mathbf{A} – macierz współczynników regresji,

\mathbf{B}_a – macierz splajnow bazowych wyznaczonych dla zmiennej wiek (indeks a – *age*),

\mathbf{B}_y – macierz splajnow bazowych wyznaczonych dla zmiennej rok kalendarzowy (indeks y – *year*).

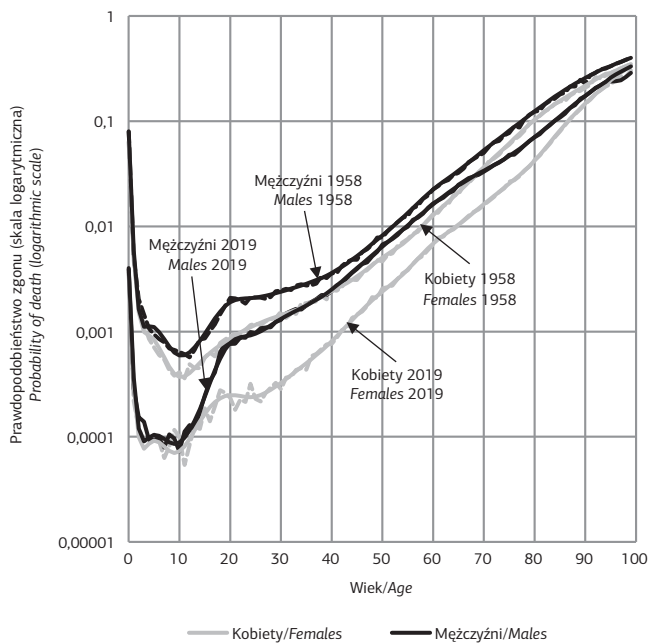
Model (5) jest uogólnionym modelem liniowym (*generalised linear model* – GLM) z logarytmem jako funkcją wiążącą. Rozwiązanie zagadnienia regresji polega na wyznaczeniu macierzy współczynników $\mathbf{A} = [a_{ij}]$. W celu łatwiejszego osiągnięcia kompromisu między dopasowaniem a wygładzeniem, do funkcji straty wprowadza się „karę” P (P – *penalty*) za różnice między sąsiednimi współczynnikami regresji a_{ij} . Dzięki zastosowaniu „kary” za brak gładkości, wybór stosunkowo dużej liczby węzłów nie doprowadzi do nadmiernego dopasowania kosztem gładkości (Eilers, Marx, 1996). W wersji dwuwymiarowej gładkość jest zapewniona dzięki nałożeniu „kar” na współczynniki w wierszach i kolumnach macierzy \mathbf{A} . Do wyznaczania współczynników a_{ij} w przypadku jednowymiarowego wygładzania proponuje się metodę IRWLS (*iteratively reweighted least squares*) (Camarda, 2012a). W przypadku dwuwymiarowym dane o umieralności dotyczące kilkudziesięciu lat życia i kilkudziesięciu lat kalendarzowych prowadzą do algorytmu obliczeniowego nieefektywnego pod względem czasu obliczeń i potrzebnej pamięci. Problem ten rozwiązali Currie, Durban, Eilers (2006), proponując wersję uogólnionego modelu liniowego GLAM (*generalized linear array model*) z algorytmem wykorzystującym specjalną strukturę danych.

Chociaż ostateczny akceptowalny efekt wyrównywania danych jest kwestią subiektywną, to przy wyborze można kierować się obiektywnymi kryteriami. W rozważanym zagadnieniu regresji z P -splajnami takim wyznacznikiem mogą być kryteria informacyjne Akaike (AIC) lub kryterium bayesowskie Schwarza (BIC), które znajdują kompromis między dopasowaniem a złożonością modelu. Parametry wygładzenia dla obu wymiarów są dobierane tak, aby zminimalizować kryterium informacyjne. W modelowaniu umieralności przy optymalizacji wygładzenia preferowane jest kryterium BIC (Currie, Durban, Eilers, 2004).

Do dwuwymiarowego wyrównywania współczynników zgonów za pomocą pakietu programu R *MortalitySmooth* (Camarda, 2012b) wykorzystano P -splajny z węzłami równoodległymi (co 5 lat, zarówno dla wieku, jak i lat kalendarzowych), wielomiany stopnia 3, różnice rzędu 2 dla współczynników regresji oraz BIC jako kryterium kontroli wygładzenia. Wyrównywanie przeprowadzono dla rocznych przedziałów wieku 0–99 lat oraz lat 1958–2019. Efekty wyrównywania według wieku dla wybranych lat kalendarzowych (1958 i 2019) przedstawiono na rysunku 1, natomiast według czasu kalendarzowego dla wybranego wieku (0 i 40 lat) na rysunkach 2 i 3.

Rysunek 1. Prawdopodobieństwo zgonu surowe (skala logarytmiczna) (linie przerywane) i wyrównane (linie ciągłe) według wieku i płci w Polsce w 1958 r. i 2019 r.

Figure 1. Probability of death (logarithmic scale), non-smoothed values (dashed lines) and smoothed values (solid lines) by age and sex in Poland in 1958 and 2019



Źródło: obliczenia własne na podstawie danych HMD dla 1958 r. i GUS dla 2019 r.

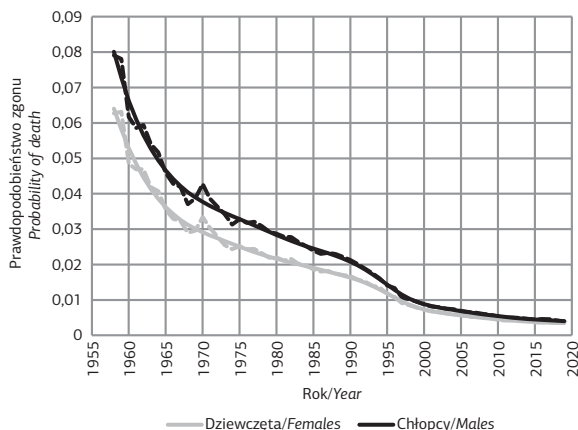
Source: own calculations based on data from HMD for 1958 and GUS for 2019.

W badanym okresie 62 lat prawdopodobieństwa zgonu kobiet i mężczyzn zmniejszyły się (rysunek 1), lecz w tendencji spadkowej wystąpiły pewne zaburzenia, które łatwiej jest wyodrębnić, analizując wartości wyrównane. Przykładowo dla 40-letnich mężczyzn (rysunek 3) jest widoczny wyraźny wzrost prawdopodobieństw zgonu w latach ok. 1965–1992, podczas gdy w tym samym czasie dla 40-letnich kobiet najpierw nastąpiło zahamowanie spadku, a dopiero ok. 1980–1991 wystąpił lekki wzrost. W przypadku niemowląt (rysunek 2) można zauważyć zmianę szybkości spadku prawdopodobieństw zgonu w latach ok. 1965, 1990, 1997.

W konsekwencji zahamowania lub odwrócenia tendencji malejącej prawdopodobieństwa zgonu w pewnych okresach dla pewnych grup wieku, pojawiły się zaburzenia wzrostu przeciętnego dalszego trwania życia. Przykładowo na rysunkach 4 i 5 przedstawiono przeciętne trwanie życia od momentu urodzenia oraz przeciętne dalsze trwanie życia 40-latków, obliczone dla tablic surowych i wyrównanych.

Rysunek 2. Prawdopodobieństwo zgonu surowe (linie przerywane) i wyrównane (linie ciągłe) dla niemowląt w Polsce w latach 1958–2019

Figure 2. Probability of death, non-smoothed values (dashed lines) and smoothed values (solid lines) for infants in Poland in 1958–2019

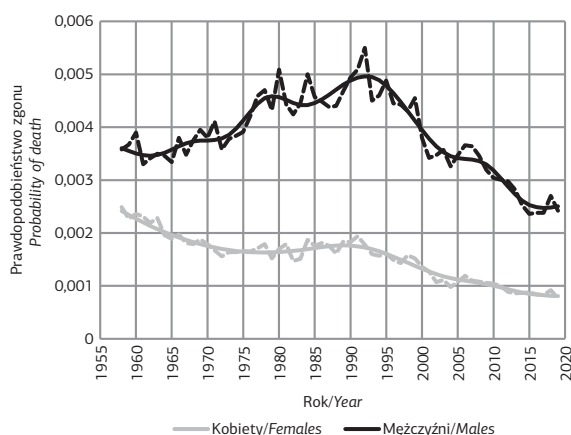


Źródło: wartości surowe – HMD, wartości wyrównane – obliczenia własne na podstawie danych HMD dla lat 1958–2018 i GUS dla 2019 r.

Source: non-smoothed values – HMD, smoothed values – own calculations based on data from HMD for 1958–2018 and GUS for 2019.

Rysunek 3. Prawdopodobieństwo zgonu surowe (linie przerywane) i wyrównane (linie ciągłe) kobiet i mężczyzn w wieku 40 lat w Polsce w latach 1958–2019

Figure 3. Probability of death, non-smoothed values (dashed lines) and smoothed values (solid lines) at age 40 in Poland in 1958–2019

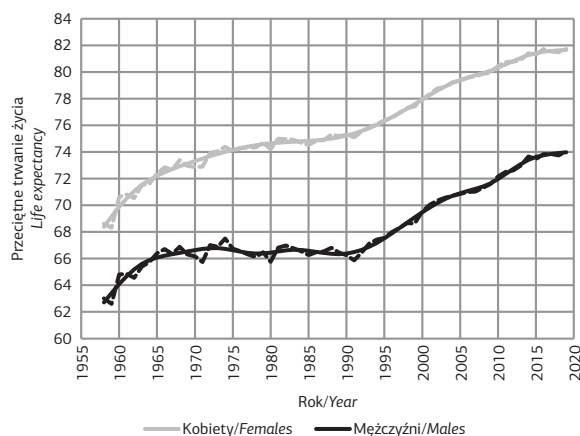


Źródło: wartości surowe – HMD, wartości wyrównane – obliczenia własne na podstawie danych HMD dla lat 1958–2018 i GUS dla 2019 r.

Source: non-smoothed values – HMD, smoothed values – own calculations based on data from HMD for 1958–2018 and GUS for 2019.

Rysunek 4. Przeciętne trwanie życia od momentu urodzenia surowe (linie przerywane) i wyrównane (linie ciągłe) w Polsce w latach 1958–2019

Figure 4. Life expectancy at birth, non-smoothed values (dashed lines) and smoothed values (solid lines) in Poland in 1958–2019

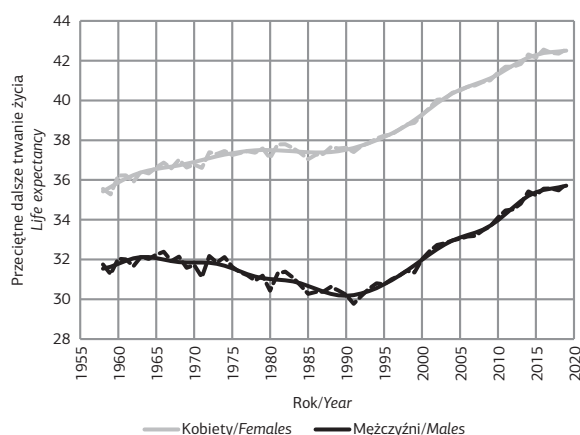


Źródło: wartości surowe – HMD, wartości wyrównane – obliczenia własne na podstawie danych HMD dla lat 1958–2018 i GUS dla 2019 r.

Source: non-smoothed values – HMD, smoothed values – own calculations based on data from HMD for 1958–2018 and GUS for 2019.

Rysunek 5. Przeciętne dalsze trwanie życia kobiet i mężczyzn w wieku 40 lat surowe (linie przerywane) i wyrównane (linie ciągłe) w Polsce w latach 1958–2019

Figure 5. Life expectancy at age 40, non-smoothed values (dashed lines) and smoothed values (solid lines) in Poland in 1958–2019



Źródło: wartości surowe – HMD, wartości wyrównane – obliczenia własne na podstawie danych HMD dla lat 1958–2018 i GUS dla 2019 r.

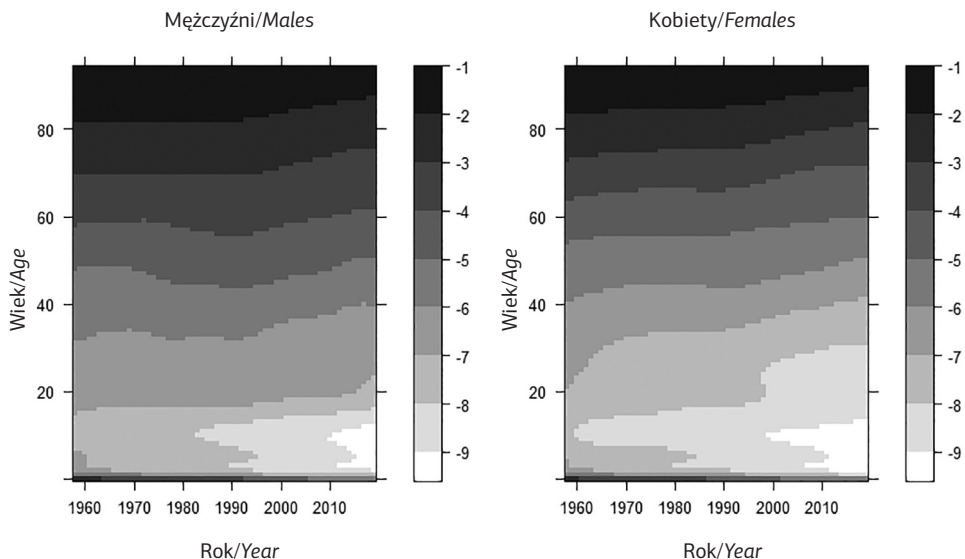
Source: non-smoothed values – HMD, smoothed values – own calculations based on data from HMD for 1958–2018 and GUS for 2019.

W latach ok. 1965–1992 przeciętne trwanie życia mężczyzn oscylowało wokół stałego poziomu, a w tym czasie wzrost przeciętnego trwania życia kobiet spowolnił (rysunek 4). Jednak w pewnych grupach wieku nastąpił w tym czasie wyraźny spadek przeciętnego dalszego trwania życia, np. dla mężczyzn w wieku 40 lat (rysunek 5).

Przedstawione wykresy prezentują poziom zjawiska w wybranym wieku lub dla wybranych lat. Szczegółową analizę zmian tendencji rozwojowej wzorca umieralności należałoby przeprowadzić jednocześnie ze względu na wszystkie przedziały wieku i lata kalendarzowe. Poziom prawdopodobieństwa zgonu, jednocześnie ze względu na wszystkie grupy wieku i wszystkie lata, można przedstawić na wykresie, gdzie w układzie współrzędnych X – rok kalendarzowy, Y – wiek, trzeci wymiar Z – logarytm prawdopodobieństwa zgonu jest reprezentowany przez natężenie odcieni szarości (rysunek 6). Rysunek 6 daje jedynie ogólny pogląd na dynamikę zjawiska. W dalszej części artykułu został przedstawiony sposób wyodrębnienia okresów zaburzeń w tendencji rozwojowej wzorca umieralności.

Rysunek 6. Wyrównane prawdopodobieństwo zgonu (skala logarytmiczna) według wieku i okresu w Polsce w latach 1958–2019

Figure 6. Smoothed probability of death (logarithmic scale) by age and period in Poland in 1958–2019



Źródło: obliczenia własne na podstawie danych HMD dla lat 1958–2018 i GUS dla 2019 r.

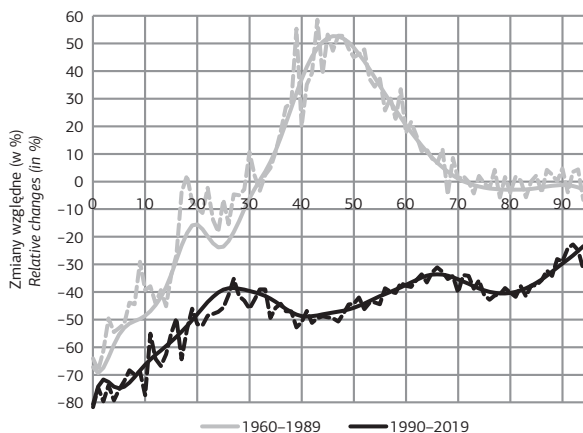
Source: own calculations based on data from HMD for 1958–2018 and GUS for 2019.

Identyfikacja okresów zaburzeń w tendencji rozwojowej wzorca umieralności

W Polsce w latach 50. XX wieku prawdopodobieństwo zgonu malało w szybkim tempie wraz z poprawą warunków życia po wojnie. Na początku lat 60. spadek wyhamował, a w następnych latach odnotowywano różną szybkość i różne kierunki zmian w zależności od wieku i płci. Po 1990 r. wraz z transformacją ustrojową i gospodarczą powrócił trend spadkowy umieralności we wszystkich grupach wieku i płci. Na rysunkach 7 i 8 przedstawiono względne zmiany prawdopodobieństw zgonu mężczyzn i kobiet w dwóch 30-letnich okresach: 1960–1989 oraz 1990–2019 (punkt podziału przebiega w roku, w którym Polska weszła na drogę transformacji). W przypadku mężczyzn (rysunek 7) w okresie 1960–1989 wyrównane prawdopodobieństwa zgonu w najmłodszych grupach wieku spadły, natomiast w wieku 32–71 lat wzrosły (nawet o 35–53% w wieku 40–55 lat), a w wieku 70 lat i więcej umieralność pozostała na zbliżonym poziomie.

Rysunek 7. Zmiany względne (w %) surowych i wyrównanych prawdopodobieństw zgonu mężczyzn według wieku w latach 1960–1989 oraz 1990–2019

Figure 7. Relative changes (in %) of non-smoothed and smoothed probabilities of death for males by age in 1960–1989 and 1990–2019



Źródło: obliczenia własne na podstawie danych HMD dla lat 1958–2018 i GUS dla 2019 r.

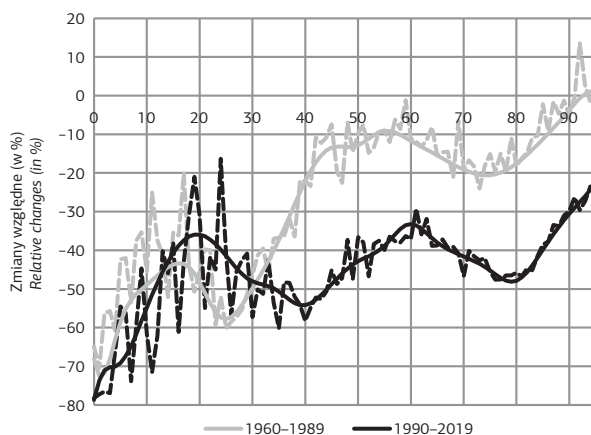
Source: own calculations based on data from HMD for 1958–2018 and GUS for 2019.

W tym samym okresie wyrównane prawdopodobieństwa zgonu kobiet spadły (rysunek 8), przy czym widoczna jest granica wieku ok. 40 lat, po której spadki były niewielkie jak na okres 30-letni. W latach 1990–2019 wyrównane prawdopodobieństwa

zgonu kobiet i mężczyzn spadły we wszystkich grupach wieku. Spadek ten dla osób w wieku 20–90 lat zawierał się w granicach 30–55% (por. Jackowska, 2017).

Rysunek 8. Zmiany względne (w %) surowych i wyrównanych prawdopodobieństw zgonu kobiet według wieku w latach 1960–1989 oraz 1990–2019

Figure 8. Relative changes (in %) of non-smoothed and smoothed probabilities of death for females by age in 1960–1989 and 1990–2019



Źródło: obliczenia własne na podstawie danych HMD dla lat 1958–2018 i GUS dla 2019 r.

Source: own calculations based on data from HMD for 1958–2018 and GUS for 2019.

Przy analizie zmian tendencji rozwojowej wzorca umieralności, dowolny podział badanego okresu na podokresy (np. równej długości 10-letnie), może spowodować, że wewnątrz podokresów znajdują się przeciwstawne tendencje rozwojowe. Dodatkowo dla różnych grup wieku i płci zmiany tendencji mogą występować w różnych momentach. Zatem wielkość i kierunek zmian w czasie t prawdopodobieństwa zgonu $q(x, t)$ czy przeciętnego dalszego trwania życia $e(x, t)$ w poszczególnych grupach wieku x zależy od wyboru „okna czasu”, tzn. nie tylko od długości okresu branego pod uwagę, ale także od umiejscowienia tego okresu na osi czasu. Z tego powodu zostanie zaproponowana metoda identyfikacji przedziałów wieku i lat kalendarzowych, w których była utrzymana tendencja rozwojowa.

Poszukiwane jest rozwiązanie zagadnienia polegającego na ustaleniu przedziałów monotoniczności wygładzonych funkcji prawdopodobieństwa zgonu $q(x, t)$ i przeciętnego dalszego trwania życia $e(x, t)$ względem czasu t dla ustalonego wieku x . Badane funkcje nie są określone wzorem, lecz są dane w postaci tablic wartości funkcji w równo oddalonych punktach. W celu ustalenia monotoniczności można wyznaczyć przybliżoną wartość pierwszej pochodnej funkcji, stosując różniczkowanie numeryczne, polegające na tym, że pochodna w danym punkcie jest szacowana

za pomocą wartości funkcji w sąsiednich punktach. Powstało wiele wzorów przybliżających pochodne funkcji na podstawie ilorazów różnicowych, rozwinięcia funkcji w szereg Taylora lub za pomocą pochodnej funkcji aproksymującej, najczęściej wielomianu (Dahlquist, Björck, 1983). Przybliżony wzór pięciopunktowy na pierwszej pochodną funkcji $f(x)$ o wartościach określonych w równoodległych punktach x_1, x_2, \dots, x_n , otrzymany na podstawie różnic centralnych z rozwinięcia funkcji w szereg Taylora, przyjmuje następującą postać:

$$f'(x_i) \approx \frac{f(x_{i-2}) - 8f(x_{i-1}) + 8f(x_{i+1}) - f(x_{i+2}))}{12h}, \quad (7)$$

gdzie h jest odległością między sąsiednimi punktami.

Na kolejnych wykresach przedstawiono aproksymację pierwszej pochodnej po czasie kalendarzowym z logarytmu wyrównanego prawdopodobieństwa zgonu mężczyzn (rysunek 9) i kobiet (rysunek 10) w Polsce według wieku w latach 1960–2017. Zastosowanie różnic centralnych we wzorze pięciopunktowym (7) spowodowało skrócenie szeregu czasowego o dwa lata na początku i dwa lata na końcu badanego okresu. Na wykresach ujemne wartości oznaczone kolorem zielonym odpowiadają korzystnej tendencji – spadkowi prawdopodobieństwa zgonu, żółty kolor oznacza stabilizację poziomu umieralności (pochodna bliska zera), a czerwony – wzrost umieralności (pochodna dodatnia). Ciemniejszy odcień oznacza większą szybkość zmian.

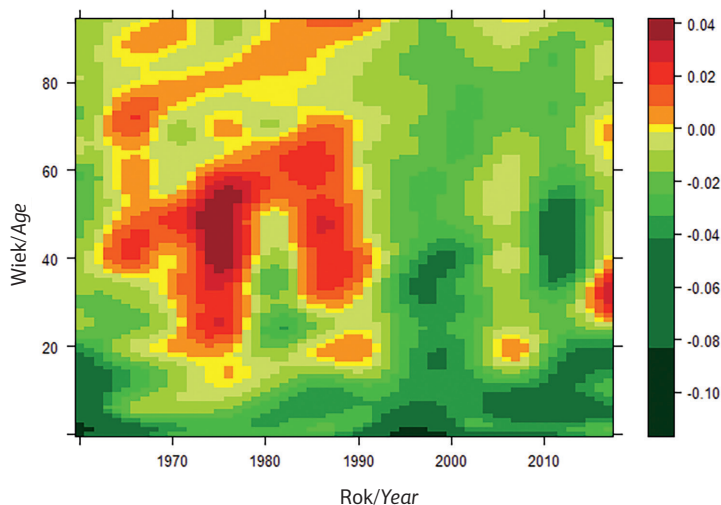
W odniesieniu do tendencji rozwojowej umieralności mężczyzn (rysunek 9) należy zwrócić uwagę na to, że w analizowanym okresie prawdopodobieństwo zgonu chłopców do wieku 10 lat spadało, a w wieku 10–19 lat pojawiły się okresy stabilizacji i lekkich wzrostów. Dla mężczyzn w wieku 20–35 lat wyraźne wzrosty umieralności wystąpiły w latach 1970–1978, a lata 1984–1991 były okresem stabilizacji lub wzrostów. Natomiast w wieku 36–65 lat okres wzrostów zaczął się już ok. 1965 r. i trwał do ok. 1990 r., przedzielany okresami stabilizacji lub lekkich spadków. Najsilniejsze wzrosty prawdopodobieństwa zgonu w przypadku mężczyzn wystąpiły w latach 1972–1978 w wieku 25–60 lat oraz latach 1985–1990 w wieku 30–65 lat. Dla mężczyzn w wieku 65 lat i więcej po 1965 r. wystąpiła stabilizacja lub wzrost umieralności, który w kolejnych latach przesunął się do coraz starszych grup wieku. Po 1990 r. we wszystkich grupach wieku dominowały spadki umieralności, z okresami spowolnienia w latach 2005–2008 (z wyjątkiem niewielkiego wzrostu w wieku 17–22 lat) oraz 2015–2017 (z wyjątkiem wzrostu w wieku 26–37 lat).

W tendencji rozwojowej umieralności kobiet można zauważyć, że okresy i wiek, w których wystąpiło zahamowanie spadku lub wzrost umieralności, zawierają się w analogicznych obszarach okres–wiek dla mężczyzn, przy czym w przypadku kobiet można zauważyć zdecydowanie mniej grup wieku i lat kalendarzowych z ten-

dencją rosnącą (por. rysunki 9 i 10). Wzrosty prawdopodobieństwa zgonu kobiet były głównie odnotowane w wieku 36–70 lat w latach 1974–1990, przedzielanych okresami stabilizacji poziomu umieralności. Do 1992 r. występowała także tendencja wzrostowa przedzielana okresami stabilizacji wśród osób w sędziwym wieku 84–94 lat. Po 1990 r. przeważały spadki we wszystkich grupach wieku. Wyjątkiem są lata 2005–2008, kiedy to nastąpił wzrost prawdopodobieństwa zgonu w przedziale wieku 59–65 lat, który przesunął się na lata 2015–2017 do przedziału wieku 68–76 lat. W latach 2015–2017 odnotowano także wzrost umieralności kobiet w młodszej grupie wieku 19–36 lat.

Rysunek 9. Aproxymacja pierwszej pochodnej logarytmu wyrównanego prawdopodobieństwa zgonu mężczyzn w Polsce w latach 1960–2017

Figure 9. Approximation of the first derivative of the logarithm of the smoothed probability of death for males in Poland in 1960–2017



Źródło: obliczenia własne na podstawie danych HMD.

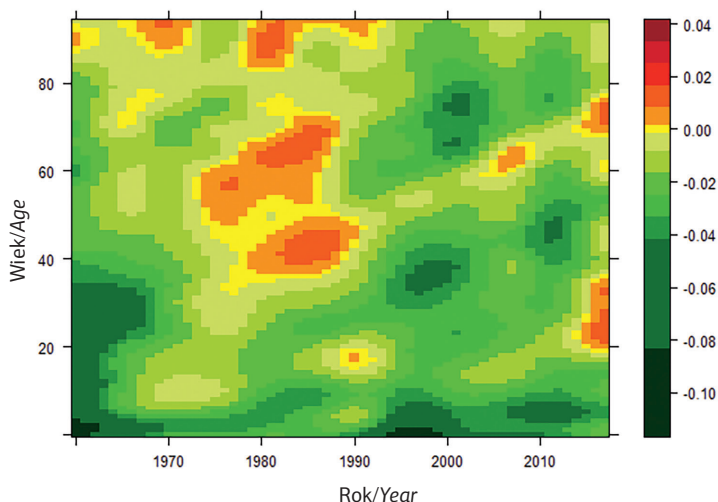
Source: own calculations based on data from HMD.

Zaburzenia w tendencji malejącej natężenia zgonów występujące przed 1990 r. przełożyły się na zmiany w tendencji rozwojowej przeciętnego dalszego trwania życia. Na wykresach przedstawiono aproksymację pierwszej pochodnej po czasie kalendarzowym z wyrównanego przeciętnego dalszego trwania życia mężczyzn (rysunek 11) i kobiet (rysunek 12) w Polsce według wieku, w latach 1960–2017. Na wykresach dodatnie wartości oznaczone kolorem zielonym odpowiadają korzystnej tendencji – wzrostowi przeciętnego dalszego trwania życia, żółty kolor oznacza stabilizację (pochodna bliska zera), a czerwony – spadek przeciętnego dalszego trwa-

nia życia (pochodna ujemna). Tak jak wcześniej, ciemniejszy odcień jest związany z większą szybkością zmian.

Rysunek 10. Aproksymacja pierwszej pochodnej logarytmu wyrównanego prawdopodobieństwa zgonu kobiet w Polsce w latach 1960–2017

Figure 10. Approximation of the first derivative of the logarithm of the smoothed probability of death for females in Poland in 1960–2017



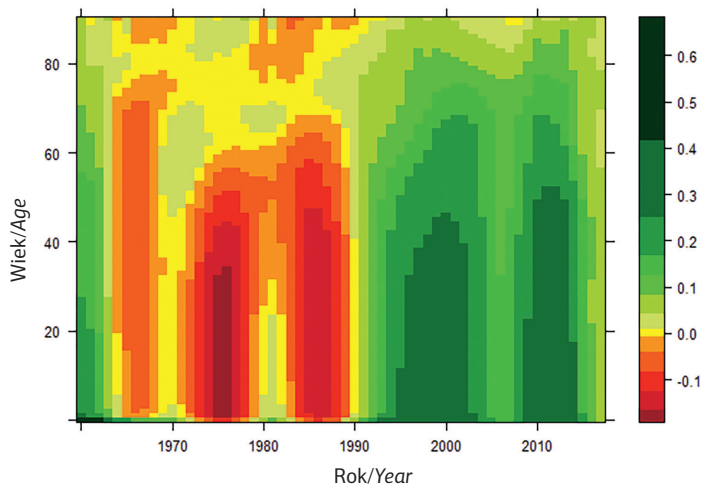
Źródło: obliczenia własne na podstawie danych HMD.

Source: own calculations based on data from HMD.

W kierunkach zmian wyrównanego przeciętnego dalszego trwania życia mężczyzn (rysunek 11) jest widoczny wyraźny podział badanego okresu w roku 1990. Anomalie wzorca umieralności przejawiały się w tym, że przeciętne trwanie życia (od urodzenia) mężczyzn w latach 1970–1989 oscylowało wokół stałego poziomu, a w grupach wieku 1–62 lat w okresie 1964–1990 dominowały spadki przeciętnego dalszego trwania życia. Natomiast w wieku 63–70 lat spadki występowały w krótszym okresie 1964–1969, a do roku 1990 analizowana funkcja tablicowa oscylowała wokół stałego poziomu. Wśród osób najstarszych, w wieku 70 lat i więcej, przed rokiem 1990 przeważały okresy stabilizacji wartości $e(x, t)$. W przypadku kobiet anomalie nie były tak duże, jak u mężczyzn (rysunek 12). Przyrosty przeciętnego trwania życia (od urodzenia) kobiet zmniejszyły się w latach 1965–1990. Największe anomalie wystąpiły w latach 1980–1985, kiedy to odnotowano stabilizację lub bardzo lekki spadek przeciętnego dalszego trwania życia kobiet. Po 1990 r. przeciętne dalsze trwanie życia rosło dla obu płci we wszystkich grupach wieku, przy czym w latach 1995–2015 wzrost był szybszy.

Rysunek 11. Aproksymacja pierwszej pochodnej wyrównanego przeciętnego dalszego trwania życia mężczyzn w Polsce w latach 1960–2017

Figure 11. Approximation of the first derivative of the smoothed life expectancy for males in Poland in 1960–2017

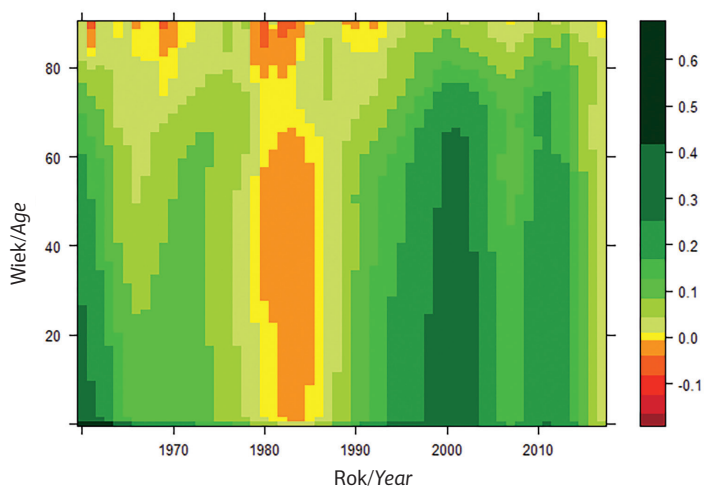


Źródło: obliczenia własne na podstawie danych HMD.

Source: own calculations based on data from HMD.

Rysunek 12. Aproksymacja pierwszej pochodnej wyrównanego przeciętnego dalszego trwania życia kobiet w Polsce w latach 1960–2017

Figure 12. Approximation of the first derivative of the smoothed life expectancy for females in Poland in 1960–2017



Źródło: obliczenia własne na podstawie danych HMD.

Source: own calculations based on data from HMD.

Zakończenie

W Polsce przed transformacją ustrojową i gospodarczą były obserwowane anomalie w tendencji rozwojowej wzorca umieralności. Obecnie, po długim okresie (od 1991 r.) wydłużania przeciętnego dalszego trwania życia, wobec pandemii koronawirusa pojawiło się zagrożenie, że okresy anomalii zdrowotnych mogą się powtarzać. Konstruowane co rok przekrojowe tablice trwania życia pozwalają na bieżąco monitorować zmianę wzorca umieralności, a przeciętne dalsze trwanie życia, parametr tablic, można traktować jako syntetyczną miarę stanu zdrowia populacji. Analiza zmiany trendu prawdopodobieństwa lub intensywności zgonów umożliwia wyodrębnienie okresów nietypowych w tendencji rozwojowej wzorca umieralności oraz identyfikację grup wieku odpowiedzialnych za wzrost czy spadek przeciętnego dalszego trwania życia.

Celem artykułu nie było zmierzenie wielkości zmian poziomu umieralności, lecz identyfikacja momentów zmian trendu i grup wieku, których te zmiany dotyczą. Zamiast nieelastycznych modeli analitycznych, wykorzystano metody wyodrębniania tendencji rozwojowej za pomocą regresji Poissona i P-splajnów oraz zaproponowano numeryczne oszacowanie pierwszej pochodnej dla ustalenia monotoniczności badanych funkcji tablicowych. Podejście takie pozwala na identyfikację kierunków i szybkości zmian tendencji rozwojowej wzorca umieralności, a tym samym zdiagnozowanie zmian w sytuacji zdrowotnej populacji, co może być przydatne szczególnie w okresach powracającej pandemii.

Przedstawiona metodyka wymaga rozwinęcia w zakresie ustalenia umownego kryterium, określającego dla oszacowanej pochodnej szerokość otoczenia liczby 0, dla którego można by uznać, że trend jest stały. Wówczas pozostałe wartości ujemne pochodnej świadczyłyby o tendencji malejącej, a pozostałe dodatnie o tendencji rosnącej. Wybór szerokości przedziału wokół liczby 0 jest uzależniony od dynamiki badanego zjawiska i wielkości lokalnych zaburzeń (związanych m.in. z liczebnością grup wieku i płci). Problem należałoby zatem rozwiązać w sposób doświadczalny. Innym otwartym polem do dalszych badań jest identyfikacja źródeł wahań w celu nadania interpretacji otrzymanym wynikom.

Bibliografia

- [1] Bolesławski, L. (1973). *Budowa tablic trwania życia. Teoria i praktyka*. Warszawa: GUS.
- [2] Bolesławski, L. (1996). *Trwanie życia i umieralność według przyczyn w 1995 r.* Warszawa: GUS.
- [3] Camarda, C.G. (2012a). MortalitySmooth: An R Package for Smoothing Poisson Counts with P-Splines, *Journal of Statistical Software*, 50(1), s. 1–24, www.jstatsoft.org/v50/i01/.
- [4] Camarda, C.G. (2012b). MortalitySmooth: Smoothing and Forecasting Poisson Counts with P-splines, *R package version 2.3.4*, CRAN.R-project.org/package=MortalitySmooth.
- [5] Chiang, C.L. (1984). *On Constructing Current Life Tables*. Florida: Robert E. Krieger Company Malabar.
- [6] Currie, I.D., Durban, M., Eilers, P.H.C. (2004). Smoothing and Forecasting Mortality Rates, *Statistical Modelling*, 4, s. 279–298.
- [7] Currie, I.D., Durban, M., Eilers, P.H.C. (2006). Generalized Linear Array Models with Applications to Multidimensional Smoothing, *Journal of the Royal Statistical Society, Series B*, 68, s. 259–280.
- [8] Dahlquist, G., Bjorck, A. (1983). *Metody numeryczne*. Warszawa: PWN.
- [9] De Boor, C. (2001). *A Practical Guide to Splines* (Revised edition). New York: Springer-Verlag.
- [10] Djeundje, V.A.B., Currie, I.D. (2011). Smoothing Dispersed Counts with Applications to Mortality Data, *Annals of Actuarial Science*, 5, s. 33–52.
- [11] Eilers, P.H.C., Marx, B.D. (1996). Flexible Smoothing with B-splines and Penalties (with Discussion), *Statistical Science*, Vol. 11, No. 2, s. 89–121.
- [12] GUS. (2020a). *Baza Demografia*, Główny Urząd Statystyczny, demografia.stat.gov.pl/bazademografia.
- [13] GUS. (2020b). *Trwanie życia w 2019 r.* Warszawa: GUS.
- [14] HMD. (2020). *Human Mortality Database*, University of California (Berkeley, USA), Max Planck Institute for Demographic Research (Rostock, Germany), www.mortality.org.
- [15] Jackowska, B. (2013). *Modele dalszego trwania życia oraz ich zastosowania w przypadku osób starszych*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.
- [16] Jackowska, B. (2017). Analiza przekrojowa i kohortowa w badaniu umieralności a ubezpieczenia na życie, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 500, s. 48–61.
- [17] Kędelski, M. (1991). Regres potencjału życiowego dorosłych mężczyzn w krajach Europy Wschodniej (1956–1985), *Studia Demograficzne*, 2(104), s. 47–68.
- [18] Okólski, M. (1993). Kryzys zdrowotny w Polsce, *Polityka Społeczna*, 1, s. 8–15.

- [19] Ouellette, N., Bourbeau, R. (2011). Changes in the age-at-death distribution in four low mortality countries: A nonparametric approach, *Demographic Research*, 25, Article 19, s. 595–628.
- [20] Preston, S.H., Heuveline, P., Guillot, M. (2001). *Demography. Measuring and Modeling Population Processes*. Oxford: Blackwell Publishers.
- [21] Thatcher, A.R., Kannisto, V., Vaupel, J.W. (1998). The force of mortality at ages 80 to 120, *Monographs on Population Aging*, No. 5, Odense University Press.
- [22] Wang, Y. (2011). *Smoothing Splines. Methods and Applications*. New York: Chapman and Hall/CRC.
- [23] Wilmoth, J.R., Andreev, K., Jdanov, D., Gleijeses, D.A., Riffe, T., Boe, C., Bubenheim, M., Philipov, D., Shkolnikov, V., Vachon, P., Winant, C., Barbieri, M. (2019). *Methods Protocol for the Human Mortality Database*, Last Revised: October 5, 2019 (Version 6), www.mortality.org/Public/Docs/MethodsProtocol.pdf.

Changes in the trend of the mortality pattern in Poland in 1958–2019

Abstract

In the several decades of post-war history of the evolution of the mortality pattern in Poland, there were periods of disturbances in the trend of decreasing force of mortality. The trend changes were different in age and gender groups. The aim of the article was to identify the age intervals and calendar years in which the mortality trend in Poland was maintained. In order to identify the trend, two-dimensional smoothing with P-splines were used. The number of deaths over a specified age and year was assumed to have a Poisson distribution. Then, smoothed death rates were employed to determine the probabilities of death and the life expectancy. Finally, to identify the time intervals in which disturbances occurred in the trend of the probabilities of death and the life expectancy, the numerical approximation of the first derivative was proposed.

Keywords: mortality, two-dimensional smoothing, P-splines, Poisson regression