

PRACE GEOGRAFICZNE

zeszyt 154, 2018, 93–105

doi: 10.4467/20833113PG.18.007.8759

Instytut Geografii i Gospodarki Przestrzennej UJ

Wydawnictwo Uniwersytetu Jagiellońskiego

POLSKA A I POLSKA B. ROZKŁAD PRZESTRZENNY I KLASTERYZACJA UMIERALNOŚCI NA WYBRANE NOWOTWORY ZŁOŚLIWE W POLSCE

Krzysztof Czaderny

Poland A and Poland B. Spatial distribution and clustering of mortality due to selected cancers in Poland

Abstract: This article describes the geographical distribution of mortality due to selected cancers in Poland in 2010–2014 and in 1980–1984 as reference. Using examples, three cluster identification techniques are discussed: the local Moran's index, the Kulldorff's scan statistic and the Stone's statistic. The centres of primary spatial clusters (groupings) of age-adjusted mortality of men and women due to cancers at 10 sites are situated outside the Eastern Poland macroregion. The location of spatial clusters in the 2010–2014 and 1980–1984 was different, although age-adjusted cancer mortality was lower in the Eastern Poland in both periods. Clusters of female mortality due to cancers of breast, ovary, uterus and bladder in the 2010–2014 period were detected in Pomerania and Kuyavia. The high all-site cancer mortality in the Northern and Western Poland is partly determined by a higher smoking prevalence in these macroregions and procreative attitudes. The data provide no basis for a hypothesis of an elevated risk of thyroid cancer mortality in seaside and mountainous counties, which could be hypothesised based on registry data from other countries. According to the Stone's test results, stomach cancer mortality was found to be statistically importantly lower in Warsaw and in other large agglomerations.

Keywords: malignant tumours, spatial clusters, health geography, Eastern Poland, gastric cancer

Zarys treści: W artykule zanalizowano zróżnicowanie geograficzne umieralności na wybrane nowotwory złośliwe w Polsce w okresie 2010–2014 oraz referencyjnie 1980–1984. Na przykładach omówiono trzy techniki klasteryzacji: lokalny współczynnik Morana, statystykę skanującą Kulldorffa oraz statystykę Stone’a. Środki skupień (zgrupowań) I rzędu standaryzowanej umieralności mężczyzn i kobiet na nowotwory w 10 lokalizacjach są położone poza makroregionem Polski Wschodniej. Położenie skupień przestrzennych w okresie 2010–2014 oraz 1980–1984 było odmienne, aczkolwiek w obu okresach standaryzowana umieralność na nowotwory złośliwe była niższa w Polsce Wschodniej. Na Pomorzu lub Kujawach położone są skupienia umieralności kobiet w okresie 2010–2014 na nowotwory złośliwe sutka, jajnika, szyjki macicy i pęcherza moczowego. Wysoka umieralność na nowotwory złośliwe ogółem w Polsce Północnej i Zachodniej jest po części wynikiem częstszego użytkowania tytoniu przez osoby zamieszkałe w tych makroregionach i postaw prokreacyjnych. Ponadto brak jest podstaw do wnioskowania o wyższej umieralności na nowotwory złośliwe tarczycy w powiatach nadmorskich lub górskich, co można by przypuszczać na podstawie danych rejestrowych z innych krajów. Na podstawie testu Stone’a stwierdzono również niższą umieralność na nowotwory żołądka w Warszawie i w innych dużych aglomeracjach.

Słowa kluczowe: nowotwory złośliwe, skupienia przestrzenne, geografia zdrowia, Polska Wschodnia, rak żołądka

Wprowadzenie

Celem niniejszego opracowania jest analiza rozkładu geograficznego umieralności na wybrane nowotwory złośliwe w Polsce. Na mapach zilustrowano dane o umieralności za lata 2010–2014 oraz referencyjnie 1980–1984, otrzymane z Krajowego Rejestru Nowotworów, pochodzące ze świadectw zgonów gromadzonych przez Główny Urząd Statystyczny. Czaderny (2019a, 2019b, 2018) omówił rozkład przestrzenny umieralności na nowotwory złośliwe jelita grubego, płuca oraz gruczołu krokowego w Polsce. W niniejszym artykule zanalizowano zróżnicowanie przestrzenne rejestrowanej umieralności na inne ważne nowotwory złośliwe: tarczycy (ICD10: C73; ICD9: 193), żołądka (C16; 151), pęcherza moczowego (C67; 188), jądra (C62; 186), jajnika (C56; 183), sutka u kobiet (C50; 174) oraz macicy (C53-C54; 180, 182). Większość chorób poddanych analizie w niniejszym artykule stanowią nowotwory układu moczowo-płciowego. Przedstawiono technikę identyfikacji skupień (zgrupowań przestrzennych; ang. *clusters*) najczęściej wykorzystywaną w epidemiologii przestrzennej, tj. statystykę skanującą zaproponowaną przez Kulldorffa i Nagarwallę (1995). Test ten został przeprowadzony na danych rejestrowych dotyczących umieralności na wszystkie ww. nowotwory złośliwe. Jego wyniki porównano za okresy 1980–1984 oraz 2010–2014. Dodatkowo na przykładzie nowotworów złośliwych żołądka zaprezentowano dwie inne techniki identyfikacji skupień przestrzennych: statystykę lokalną Morana (zaproponowaną przez Anselina, 1995) oraz statystykę Stone’a (1988).

Trzy wymienione techniki identyfikacji skupień przestrzennych są dostępne w przyjaznych dla użytkownika, dostępnych bezpłatnie pakietach komputerowych. Mogą one być wykorzystywane nie tylko w geografii zdrowia. Statystyki Morana i Kulldorffa mogą mieć zastosowanie w innych dziedzinach, w tym w analizach rynku nieruchomości i w kryminologii. Test Stone'a jest stosowany w ocenie oddziaływania punktowych źródeł zanieczyszczeń. Artykuł stanowi niejako uzupełnienie analizy Kretowicza (2013) nt. podejść badawczych i metodologii stosowanej w geografii zdrowia, który nie omawiał jednak metod identyfikacji skupień przestrzennych.

W sekcji następującej po wprowadzeniu metodycznym omówione zostaną wyniki analizy skupień przy zastosowaniu statystyki skanującej Kulldorffa. Na mapach zilustrowane zostaną zgrupowania przestrzenne umieralności na nowotwory złośliwe w sześciu wybranych lokalizacjach nowotworowych. W sekcji kolejnej na przykładzie umieralności na nowotwory złośliwe żołądka zastosowane zostaną algorytmy Kulldorffa, Morana i Stone'a.

Metody klasteryzacji

W niniejszej analizie zastosowano i omówiono trzy techniki identyfikacji skupień przestrzennych. Wszystkie z nich są dostępne w pakiecie *DCluster* w oprogramowaniu R CRAN. Statystyka lokalna Morana może zostać również obliczona z poziomu interfejsu przy użyciu prostszego w obsłudze oprogramowania *Geoda*, a statystyka skanująca Kulldorffa – przy użyciu oprogramowania *SaTScan*. Jednostką podstawową przeprowadzonych testów Morana i Kulldorffa jest powiat. Przedziały klas każdorazowo wyznaczono metodą naturalnej przerwy Jenksa. Należy zaznaczyć, że rozbieżna skala barwna (w tym hipsometryczna) nie jest dostatecznie dobrym wyborem dla map umieralności i zapadalności (zachorowalności) z powodu na brak punktu środkowego w tego typu danych. Tym niemniej jest ona stosowana na mapach epidemiologicznych, w tym w ramach nieraz replikowanej metodologii Pukkali i in. (1987). Wątpliwości metodologiczne może również budzić stosowanie metody izarytmicznej w celu obrazowania zróżnicowania przestrzennego umieralności i zapadalności (por. Mościbroda 1975). Ratajski (1973, s. 156) tego rodzaju izarytmy nazywał 'fałszywymi' lub pseudoizoliniami. Omawiana dalej obecność obserwacji odstających nie świadczy o ciągłości zmiennej umieralności w przestrzeni.

Statystyka lokalna Morana I_i jest liczona dla każdej obserwacji (jednostki przestrzennej) i . Ocena istotności statystycznej jest dokonywana przy użyciu metody Monte Carlo (tutaj: $N_{rep} = 999$ iteracji). Wartość statystyki lokalnej Morana I_i jest dodatnia, jeśli sąsiadujące jednostki przestrzenne cechują podobne wartości analizowanej zmiennej x . Ujemne wartości statystyki oznaczają negatywną autokorelację przestrzenną, tj. sąsiadowanie ze sobą jednostek przestrzennych o zróżnicowanych

wartościach zmiennej x . W przypadku losowego rozkładu przestrzennego statystyka I_i jest równa około zero

$$E(I_i) = - \frac{\sum_{j=1}^n w_{ij}}{n-1}$$

gdzie n oznacza liczbę obserwacji. Elementy macierzy wag w_{ij} definiują zależności przestrzenne między obserwacjami. Na potrzeby testu Morana w niniejszym opracowaniu przyjęto, że $w_{ij}=1$, jeśli jednostki przestrzenne i oraz j sąsiadują ze sobą, w przeciwnym przypadku $w_{ij}=0$. Test nazwano na cześć Morana, ale w wersji lokalnej został zaproponowany przez Anselina (1995):

$$I_i = \frac{(x_i - \bar{x}) \sum_{j=1}^n w_{ij} (x_j - \bar{x})}{\frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n}}$$

Podjęcie Anselina może być stosowane zarówno w celu identyfikacji skupień przestrzennych, jak i obserwacji odstających (nietypowych). Zgrupowania przestrzenne wykryte przy pomocy testu odnoszą się do rdzenia skupień przestrzennych. Hipoteza zerowa statystyki lokalnej Morana to losowy rozkład przestrzenny zmiennej x , a w tym przypadku – losowy rozkład umieralności na nowotwory w Polsce. Hipoteza alternatywna stanowi o odwrotnym – obecności czynników warunkujących nielosowy rozkład analizowanej zmiennej.

Statystyka skanująca Kulldorffa wykorzystywana jest do wykrywania najbardziej prawdopodobnych skupień (lub inaczej zgrupowań I rzędu, ang. *primary clusters*) oraz skupień II rzędu (ang. *secondary clusters*), tj. obszarów mniej wyróżniających się, o niższej istotności statystycznej. Hipotezą zerową testu jest brak klasteryzacji: wówczas liczba zgonów byłaby jednorodnym procesem Poissona z wartością oczekiwaną proporcjonalną do liczebności populacji odniesienia (populacji ryzyka). Podobnie jak w przypadku testu Morana, z powodu braku rozkładu testowego, w celu oceny istotności należy przeprowadzić symulację Monte Carlo ($N_{rep} = 9\,999$ iteracji w danym przypadku). Na mapach w dalszej części pracy zaznaczono skupienia przestrzenne, których występowanie jest istotne statystycznie na poziomie $\alpha = 0,001$. Według hipotezy alternatywnej testu liczba zgonów w granicach okna z przekracza tę oczekiwaną. Maksymalna wielkość okna w teście Kulldorffa jest określana arbitralnie, na potrzeby niniejszego opracowania określono ją na poziomie 10% populacji odniesienia. Wartość graniczna równa 10% populacji odniesienia została również zaproponowana przez Norström i in. (2000) oraz Sheehana i in. (2000). Przyjęta dopuszczalna wielkość okna skanującego z jest większa niż liczba ludności w najbardziej ludnym powiecie Polski. Statystyka skanująca dana jest wzorem (Kulldorff i Nagawalla 1995):

$$\lambda = \sup_z \left(\frac{o(z)}{e(z)} \right)^{o(z)} \left(\frac{N - o(z)}{N - e(z)} \right)^{N - o(z)} I \left(\frac{o(z)}{e(z)} > \frac{N - o(z)}{N - e(z)} \right)$$

gdzie z to wybrane okno skanujące o eliptycznym kształcie, natomiast $o(z)$ oraz $e(z)$ oznaczają obserwowaną i oczekiwaną liczbę zgonów, N oznacza liczbę zgonów na całym analizowanym obszarze, a I to funkcja indykatorowa.

Najbardziej prawdopodobne skupienie z^* (tj. zgrupowanie I rzędu) to zgrupowanie przestrzenne, dla którego wartość funkcji wiarygodności jest najwyższa. Zgrupowania II rzędu również charakteryzuje istotna statystycznie wartość funkcji wiarygodności, ale jest ona niższa niż w przypadku skupienia I rzędu. Wyznaczone skupienia nie mogą na siebie wzajemnie zachodzić. Symulacja Monte Carlo jest przeprowadzana w celu określenia rozkładu bazowego statystyki punktowej Nagarwalli λ oraz obliczenia ilorazu wiarygodności. Korzystanie z informacji z dużej liczby symulacji oznacza bardziej stabilne wyniki oraz stwarza możliwość określenia dokładniejszych wartości prawdopodobieństwa (p -wartości). Podobne podejście do tego zaproponowanego przez Kulldorffa i Nagarwallę (1995) zostało opisane przez Besaga i Newella (1991). Jednak w tym drugim przypadku z góry określona jest wielkość identyfikowanego skupienia. Na prezentowanych dalej kartogramach umieralności na nowotwory złośliwe w Polsce w latach 2010–2014 zasięgi skupień Kulldorffa zostały oznaczone czerwonymi konturami (zasięgi liniowe). Szrafem zaznaczono zasięgi skupień wykrytych przy zastosowaniu algorytmu Kulldorffa na podstawie danych z okresu 1980–1984 (zasięgi plamowe).

Na przykładzie danych o umieralności na nowotwory żołądka w obszarze metropolitalnym Warszawy przeprowadzono test Stone'a. Istotny statystycznie wynik testu oznacza klasteryzację wokół określonej lokalizacji. Test bazuje na statystyce ilorazu wiarygodności (Stone 1988):

$$T_s = 2 \sum_{i=1}^m n_i \left\{ \log \tilde{\vartheta}_i - \log \left(\frac{\sum_{j=1}^m \tilde{\vartheta}_j e_j}{\sum_{j=1}^m e_j} \right) \right\}$$

gdzie n_i oraz e_i oznaczają liczbę obserwowanych i oczekiwanych zgonów w i -tej jednostce przestrzennej (tutaj: w gminie), $\tilde{\vartheta}_i$ to estymatory największej wiarygodności. Estymatory $\tilde{\vartheta}_i$ są uzyskiwane poprzez zastosowanie algorytmu PAVA (ang. *pool-adjacent-violators algorithm*), który pozwala stwierdzić, czy między wartościami oczekiwanymi ciągu zmiennych losowych zachodzi relacja porządku. Istotność statystyczna została w danym przypadku oceniona na podstawie symulacji o $N_{rep} = 999$ iteracjach. Hipotezą zerową testu Stone'a jest stałość ryzyka relatywnego, a hipoteza alternatywna to wzrost ryzyka wraz ze zmniejszaniem się odległości od określonej lokalizacji. W niniejszym opracowaniu zanalizowano ryzyko zgonu

z powodu nowotworów złośliwych żołądka w promieniu 30 km oraz 50 km wokół Warszawy i dziewięciu innych największych miast Polski. Ograniczenie obszaru porównań do promienia kilkudziesięciu kilometrów zapewnia większą porównywalność analizowanych jednostek przestrzennych. Nałożenie tej restrikcji możliwe jest w pakiecie *Geosphere* w oprogramowaniu R CRAN.

Umieralność na wybrane nowotwory złośliwe. Skupienia Kulldorffa

Zarówno umieralność, jak i zapadalność na nowotwory cechują nierówności płciowe (ang. *gender gap*). Zapadalność i umieralność na nowotwory złośliwe jest w Polsce wyższa u mężczyzn. Różnice dotyczą również lokalizacji najczęściej występujących nowotworów (czyli zajętego narządu) u obu płci. Z tego powodu dane z zakresu epidemiologii nowotworów są na ogół publikowane oddzielnie dla mężczyzn i kobiet. Ponadto, z uwagi na fakt, że głównym endogenicznym czynnikiem rozwoju nowotworów złośliwych jest wiek, analizowane wartości umieralności zostały wystandaryzowane (przeważone) ze względu na strukturę wieku w całym kraju. Nowotworem złośliwym najczęściej diagnozowanym u kobiet są nowotwory sutka (17,4 tys. kobiet w 2014 r. według danych Krajowego Rejestru Nowotworów). U mężczyzn są to prawie *ex aequo* nowotwory płuca (14,7 tys.) i gruczołu krokowego (12,3 tys.). W 2014 r. 5,1 tys. mężczyzn zachorowało na nowotwory złośliwe pęcherza moczowego, 3,5 tys. – na nowotwory złośliwe żołądka, 1,1 tys. – na nowotwory złośliwe jądra, 0,5 tys. – na nowotwory złośliwe tarczycy. W 2014 r. u kobiet wykryto 5,9 tys. nowotworów złośliwych trzonu macicy, 3,7 tys. nowotworów złośliwych jajnika, 2,8 tys. nowotworów złośliwych szyjki macicy, 2,6 tys. nowotworów złośliwych tarczycy, 2,0 tys. nowotworów złośliwych żołądka oraz 1,6 tys. nowotworów złośliwych pęcherza moczowego.

Na ryc. 1, 3, 4 oraz 6 zilustrowano rozkład przestrzenny umieralności na wybrane nowotwory złośliwe w populacji mężczyzn i kobiet w Polsce. Należy podkreślić, że centra (środki) wszystkich zidentyfikowanych skupień przestrzennych według danych za lata 2010–2014 znajdowały się poza Polską Wschodnią. W niniejszym opracowaniu za Polskę Wschodnią przyjęto województwa: warmińsko-mazurskie, podlaskie, lubelskie, świętokrzyskie i podkarpackie. Takie jest również rozumienie makroregionu w polityce regionalnej (*Program Operacyjny Polska Wschodnia 2014–2020*). Polska Wschodnia z racji na niższy poziom rozwoju gospodarczego bywa nazywana „Polską B”, podczas gdy pozostała część kraju – „Polską A”. Te skrótowe, często używane określenia budzą negatywne konotacje. Naznaczenie społeczne

związane z tym podziałem jest uznawane za niesprawiedliwe i dyskryminujące¹. Przekornie warto zwrócić uwagę, że standaryzowana ze względu na strukturę wieku umieralność na nowotwory złośliwe pozostaje niższa w Polsce Wschodniej.

Niższa umieralność na nowotwory złośliwe ogółem w Polsce Wschodniej jest m.in. pochodną uwarunkowań kulturowych, w tym zróżnicowania przestrzennego w zakresie użytkowania tytoniu (ryc. 1A) i postaw prokreacyjnych. W okresie międzywojennym na obszarze obecnych województw podlaskiego, lubelskiego, świętokrzyskiego i podkarpackiego sprzedaż tytoniu była istotnie niższa niż w zachodniej części kraju (Powszechna Wystawa Krajowa 1929, s. 11). Odsetek palaczy do dzisiaj pozostaje niższy w południowo-wschodniej części kraju. Makroregion Polski Wschodniej nie jest jednolity w tym zakresie. Wysoka częstość palenia w województwie warmińsko-mazurskim – podobna do tej obserwowanej w Polsce Zachodniej – może być pochodną pozostawania historycznie tych ziem w sferze wpływów prusko-niemieckich oraz niskiego wieku (niskiej integracji) lokalnych społeczności (Czaderny 2019b). Zwyczaj użytkowania tytoniu Polacy przejęli m.in. od narodu niemieckiego (Maliniak 1933, s. 7).

Nowotwór złośliwy jądra to nowotwór najczęściej diagnozowany u mężczyzn na etapie wczesnej dorosłości. Rozkład geograficzny umieralności na nowotwory złośliwe jądra w Polsce charakteryzuje nieduża autokorelacja przestrzenna, jednak możliwe jest wyróżnienie skupienia w centralnej części kraju (ryc. 1B). Do najważniejszych czynników ryzyka nowotworów złośliwych jądra należą: wnetrostwo, obrażenia jąder, zakażenie wirusem HIV, zespół Downa, choroba nowotworowa w innej lokalizacji, przypadki nowotworów złośliwych jądra w rodzinie (Khan, Protheroe 2007). Rozkład przestrzenny ww. czynników w Polsce jest trudny do ustalenia. W literaturze wielokrotnie weryfikowano również hipotezę o związku palenia tytoniu z występowaniem nowotworów złośliwych jądra (Srivastava, Kreiger 2004). Obecnie uważa się, że palenie tytoniu nie zwiększa prawdopodobieństwa wystąpienia nowotworów złośliwych jądra, co widać również, porównując częstość palenia i umieralność mężczyzn na nowotwory złośliwe jądra w Polsce (ryc. 1).

Palenie tytoniu jest czynnikiem ryzyka nowotworów złośliwych pęcherza moczowego i żołądka (Nomura i in. 2012), umieralność w Polsce na te nowotwory zilustrowano w dalszej części pracy na ryc. 4A, 4B oraz 6. Umieralność mężczyzn na nowotwory złośliwe pęcherza moczowego (ryc. 4A) i żołądka (ryc. 6A) jest istotnie niższa w południowo-wschodniej części kraju, co po części jest pochodną wzorów przestrzennych palenia papierosów (ryc. 1A). Prezentowane dane nt. częstości palenia papierosów pochodzą z badania *Target Group Index* i dotyczą okresu 2001–2007

¹ Podział kraju na Polskę A i Polskę B za 'niesprawiedliwy' i 'pogardliwy' określał kilkakrotnie obecny prezydent RP (por. wystąpienie prezydenta RP na Krajowym Zjeździe Delegatów NSZZ „Solidarność” dn. 24.10.2016 r., wystąpienie prezydenta RP w Parczewie dn. 16.05.2017 r.).

(dane zagregowane do poziomu 380 powiatów, próba warstwowo-losowa $N=122\ 599$ mężczyzn). Mimo ograniczeń związanych z zakresem tematycznym badania, są to jedyne dane dla Polski o paleniu tytoniu pochodzące z próby tej wielkości. Badanie było realizowane techniką wywiadu wspomaganego komputerowo albo ankiety wypełnianej samodzielnie przez respondenta oraz uzupełniająco techniką ankiety internetowej. Statystyczne modele przestrzenne umieralności na nowotwory złośliwe jelita grubego oraz płuca w Polsce przedstawił Czaderny (2019a, 2019b).

Relatywnie niska umieralność na nowotwory złośliwe sutka w Polsce Wschodniej jest po części warunkowana świadomością nt. skutków doustnej antykoncepcji (ryc. 2, dane *Target Group Index*, $N=109\ 681$ kobiet). Związek między występowaniem nowotworów sutka i przyjmowaniem doustnej antykoncepcji tłumaczony jest bezpośrednim oddziaływaniem antykoncepcji doustnej na poziom estrogenu w organizmie oraz pośrednim oddziaływaniem na masę ciała (Soroush i in. 2016). Nie jest to jedyny czynnik ryzyka. Na Pomorzu lub Kujawach zlokalizowane są skupienia umieralności kobiet w okresie 2010–2014 na nowotwory złośliwe sutka, jajnika, szyjki macicy i pęcherza moczowego (ryc. 3–4). W przypadku nowotworów trzonu macicy nie zidentyfikowano skupień przestrzennych istotnych statystycznie na poziomie $\alpha = 0,001$.

Ekspozycja na promieniowanie jonizujące jest podstawowym czynnikiem ryzyka nowotworów złośliwych tarczycy. W populacjach wyspiarskich, włączając Hawaje, Japonię, Islandię, Filipiny, niektóre regiony Sycylii i Cypru, rejestrowane są jedne z najwyższych w świecie współczynników zapadalności na nowotwory złośliwe tarczycy (Poljak i in. 2011). Minelli i in. (2013) stwierdzili istotnie statystycznie wyższą umieralność na nowotwory złośliwe tarczycy na obszarach górskich we Włoszech. W Polsce w okresie 2010–2014 nie zidentyfikowano skupień wysokiej umieralności na te nowotwory w powiatach nadmorskich bądź w Karpatach i Sudetach. Obszar podwyższonej umieralności mężczyzn i kobiet na nowotwory złośliwe tarczycy w latach 1980–1984 obejmował jednak obszar obecnego województwa małopolskiego.

Umieralność na nowotwory złośliwe żołądka

Skupienia Kulldorffa i Morana

Standaryzowana umieralność na nowotwory złośliwe żołądka jest nieznacznie niższa w Polsce Wschodniej. Na ryc. 6 zilustrowano zróżnicowanie przestrzenne umieralności na nowotwory złośliwe żołądka oraz zaznaczono skupienia zidentyfikowane przy pomocy algorytmu Kulldorffa. Identyfikacja skupień przy pomocy statystyki skanującej Kulldorffa prowadzi do ujawnienia zgrupowań większych niż w przypadku

podejścia Morana (ryc. 6). W podejściu Kulldorffa identyfikowane są eliptyczne skupienia, a w algorytmie Morana kształt skupień nie jest zakładany *a priori*.

Najbardziej prawdopodobne skupienie przestrzenne umieralności mężczyzn na nowotwory złośliwe żołądka jest położone w południowo-środkowej Polsce (Górny Śląsk). Środek tego skupienia stanowi miasto Świętochłowice. Liczba zgonów mężczyzn z powodu nowotworów złośliwych żołądka w latach 2010–2014 była na tym obszarze równa 2 035, podczas gdy oczekiwana – 1 670 (logarytm ilorazu wiarygodności: $\log \lambda = 41,7$, $p = 0,001$). Empiryczny poziom istotności $p = 0,001$ stanowi, że obliczona wartość ilorazu wiarygodności λ dla skupienia jest wyższa niż w przypadku 9 999 przeprowadzonych iteracji Monte Carlo (ryc. 5). Oznacza to, że jest skrajnie mało prawdopodobne, że podwyższone ryzyko umieralności na tych obszarach jest związane z procesami losowymi. Skupienie II rzędu umieralności mężczyzn na nowotwory żołądka zostało zidentyfikowane w północno-środkowej Polsce, a środkiem zgrupowania jest powiat golubsko-dobrzyński. Obserwowana i oczekiwana liczba zgonów w skupieniu II rzędu jest równa odpowiednio 1 858 oraz 1 623 ($\log \lambda = 17,9$, $p = 0,001$). Skupienia przestrzenne I i II rzędu umieralności na nowotwory żołądka są podobnie położone w przypadku populacji mężczyzn i kobiet, co oznacza, że podobny jest również rozkład przestrzenny czynników ryzyka wśród obu płci.

Skupienie I rzędu umieralności kobiet z powodu nowotworów złośliwych żołądka jest położone w środkowo-południowej Polsce ($\log \lambda = 21,8$, $p = 0,001$), środek zgrupowania to powiat żywiecki. Obserwowana i oczekiwana liczba zgonów kobiet z powodu nowotworów żołądka na tym obszarze w latach 2010–2014 to odpowiednio 1 079 oraz 886. Skupienie II rzędu, podobnie jak w przypadku mężczyzn, położone jest w północno-środkowej Polsce wokół powiatu nowomiejskiego ($\log \lambda = 21,3$, $p = 0,001$). Na tym obszarze na nowotwory złośliwe żołądka zmarło 399 kobiet, podczas gdy oczekiwana liczba zgonów równa była 284. Zidentyfikowane skupienia II rzędu cechuje wysoki stopień heterogeniczności, co oznacza, że obejmują stosunkowo wiele obserwacji o relatywnie niskiej liczbie zgonów. Z racji na założenie o eliptycznym kształcie skupień w algorytmie Kulldorffa, są one włączone do wyróżnionych zgrupowań. Należy zarazem podkreślić, że mimo heterogeniczności skupień, ich lokalizacja była podobna w latach 1980–1984 oraz 2010–2014, zwłaszcza w przypadku populacji mężczyzn. Oznacza to, że rozkład przestrzenny czynników ryzyka na nowotwory złośliwe żołądka wykazuje wysoki stopień petryfikacji. Obserwacje odstające mogą zostać zidentyfikowane przy pomocy testu Morana, którego wyniki są zilustrowane na ryc. 7.

Według algorytmu Morana skupienia wysokiej umieralności, obejmujące przynajmniej dwie obserwacje, są położone również w północno-środkowej oraz południowo-środkowej Polsce. Każda z tych jednostek przestrzennych należy zarazem do skupienia zidentyfikowanego przy pomocy statystyki skanującej Kulldorffa. Duży obszar o niskiej umieralności na nowotwory złośliwe żołądka jest położony na

Lubelszczyźnie. Odmienność wyników klasteryzacji przedstawionych na ryc. 6 i 7 wynika z różnic metodycznych. Statystyka lokalna Morana jest liczona dla każdej obserwacji (jednostki przestrzennej), a statystyka skanująca Kulldorffa odnosi się do zbioru jednostek przestrzennych położonych w granicach okna skanującego z . Statystyka lokalna Morana dla i -tej obserwacji uwzględnia charakterystyki sąsiednich jednostek przestrzennych. Wartość statystyki skanującej Kulldorffa bazuje na rozkładzie zmiennej na całym analizowanym obszarze – w tym przypadku w całym kraju. Ponadto statystyka lokalna Morana jest też wykorzystywana do identyfikowania skupień niskich wartości.

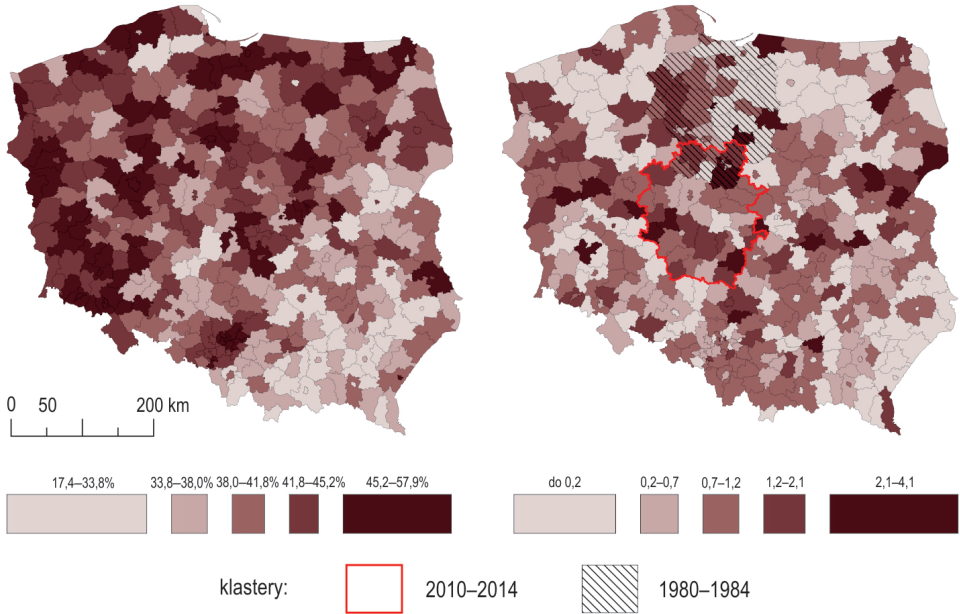
Skupienia Stone'a

Chiu i in. (2011) dowodzą, że osoby zamieszkałe na obszarach o większej liczbie stacji paliw w przeliczeniu na kilometr kwadratowy narażone są na wyższe ryzyko zgonu z powodu nowotworów złośliwych żołądka. Liczba stacji paliw była użyta przez autorów jako zmienna instrumentalna ekspozycji na benzen oraz inne węglowodory. Ryzyko zachorowania na nowotwory żołądka zostało określone jako wyższe na obszarach zurbanizowanych również przez Faivre'a i in. (1985) na podstawie danych z francuskiego departamentu *Côte-d'Or*. W Chinach umieralność na nowotwory żołądka jest wyższa na obszarach zurbanizowanych niż wiejskich (Yang 2006). Na potrzeby niniejszego opracowania przeprowadzono analizę umieralności na nowotwory złośliwe żołądka w promieniu 30 km i 50 km od Warszawy przy użyciu testu Stone'a (1988). Ryc. 8 to kartogram umieralności na nowotwory żołądka w województwie mazowieckim. Autokorelacja przestrzenna umieralności na nowotwory złośliwe żołądka na Mazowszu jest relatywnie niska. Globalna statystyka Morana (1950, 1948) umieralności na nowotwory żołądka na Mazowszu wynosi $I=0,07$ dla mężczyzn oraz $I=-0,11$ dla kobiet, co jest wartością znacznie niższą niż miara współczynnika dla całego kraju.

Hipotezą zerową testu w danym przypadku jest zmniejszanie się ryzyka zgonu z powodu nowotworów złośliwych żołądka wraz ze wzrostem odległości od Warszawy. Na podstawie testu określono, że ryzyko zgonu z powodu nowotworów żołądka u mężczyzn nie zmniejsza się wraz ze wzrostem odległości od Warszawy, zarówno dla promienia 30 km ($T_s=1,00$, $p=0,800$), jak i 50 km ($T_s=1,00$, $p=0,962$). Analogiczne wyniki uzyskano na danych dla populacji kobiet dla promienia 30 km ($T_s=1,01$, $p=0,696$) oraz 50 km ($T_s=1,05$, $p=0,484$). Test Stone'a przeprowadzono również dla dziewięciu innych miast Polski: Krakowa, Łodzi, Wrocławia, Poznania, Gdańska, Szczecina, Bydgoszczy, Lublina i Katowic. Zarówno w przypadku mężczyzn, jak i kobiet liczba zgonów na nowotwory żołądka w okresie 2010–2014 była niższa niż oczekiwana w większości ww. miast. W połowie z nich (włączając Warszawę) umieralność na nowotwory złośliwe żołądka istotnie statystycznie zwiększała się wraz

(A) palenie papierosów przez mężczyzn

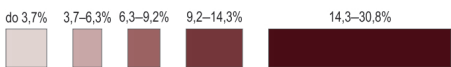
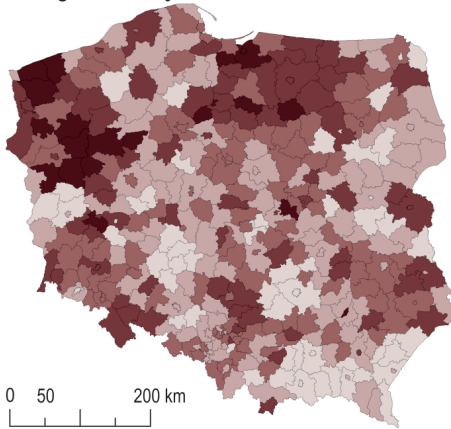
(B) umieralność na nowotwory jądra



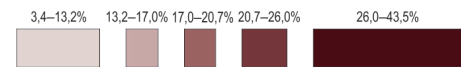
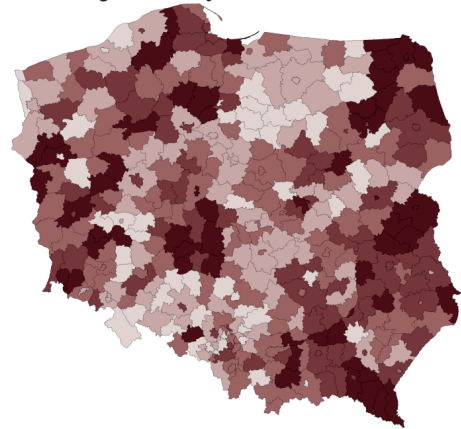
Ryc. 1. Standaryzowana umieralność (w przeliczeniu na 100 000 mężczyzn) na nowotwory złośliwe jądra w latach 2010–2014, skupienia przestrzenne Kulldorffa oraz przeciętny odsetek palaczy (włączając palaczy okazjonalnych) w latach 2001–2007

Fig. 1. Age-adjusted mortality rate (per 100,000 males) of testis cancer mortality in 2010–2014, spatial clustering by the Kulldorff's scan statistic and the average smoking prevalence rate (including light users) in Poland in 2001–2007

(A) odsetek odpowiedzi „zdecydowanie zgadzam się”

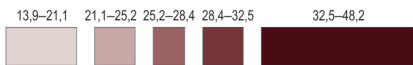
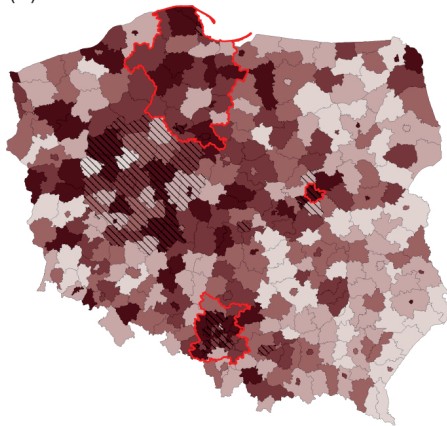


(B) odsetek odpowiedzi „zdecydowanie nie zgadzam się”

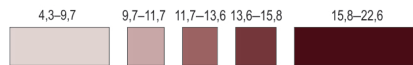
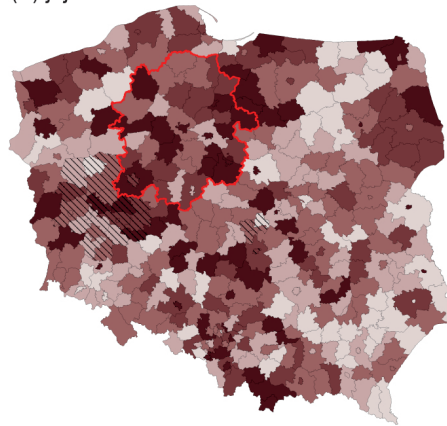


Ryc. 2. Przeciętny odsetek kobiet w latach 2004–2012, które zdecydowanie potwierdzają lub negują, że doustna antykoncepcja nie ma negatywnego wpływu na kobiece zdrowie
Fig. 2. The average percentage of women in 2004–2012 who strongly agree and disagree that oral contraception has no negative effect on female health

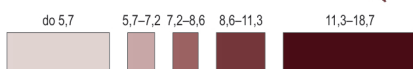
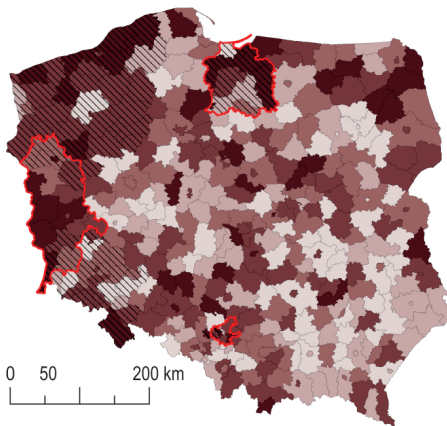
(A) sutek



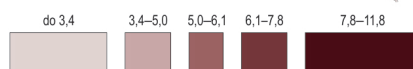
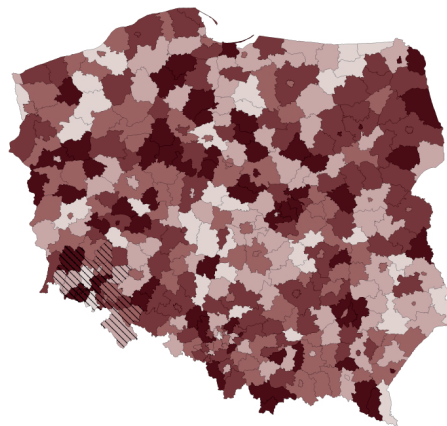
(B) jajnik



(C) szyjka macicy



(D) trzon macicy



klastery:



2010-2014

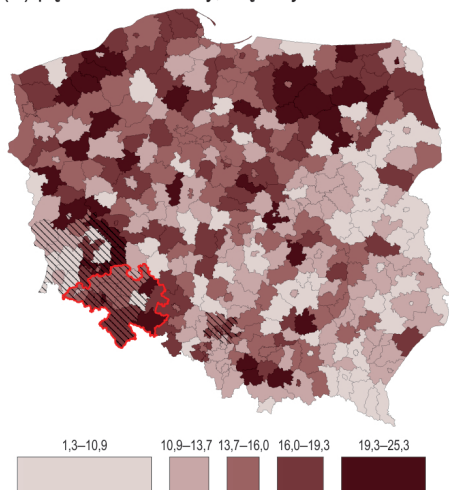


1980-1984

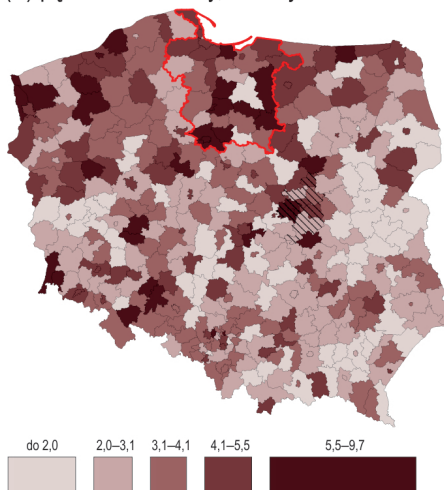
Ryc. 3. Standaryzowana umieralność (w przeliczeniu na 100 000 kobiet) na nowotwory złośliwe sutka, jajnika i macicy w Polsce w latach 2010–2014 oraz skupienia zidentyfikowane na podstawie algorytmu Kulldorffa

Fig. 3. Age-adjusted mortality rate (per 100,000 females) of breast, ovarian and uterine cancers in Poland in 2010–2014 and the results of the Kulldorff's spatial scan statistic

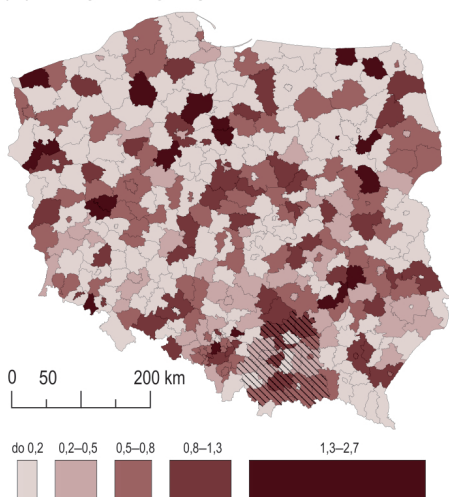
(A) pęcherz moczowy, mężczyźni



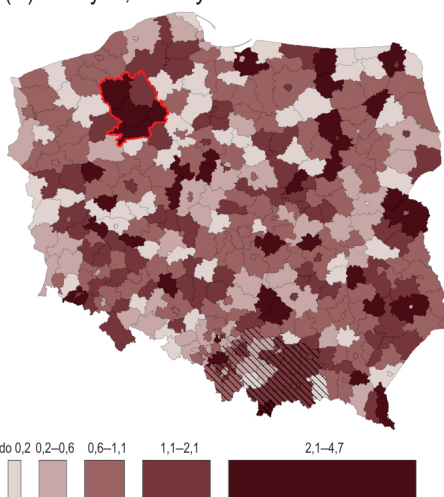
(B) pęcherz moczowy, kobiety



(C) tarczycy, mężczyźni



(D) tarczycy, kobiety



klastery:



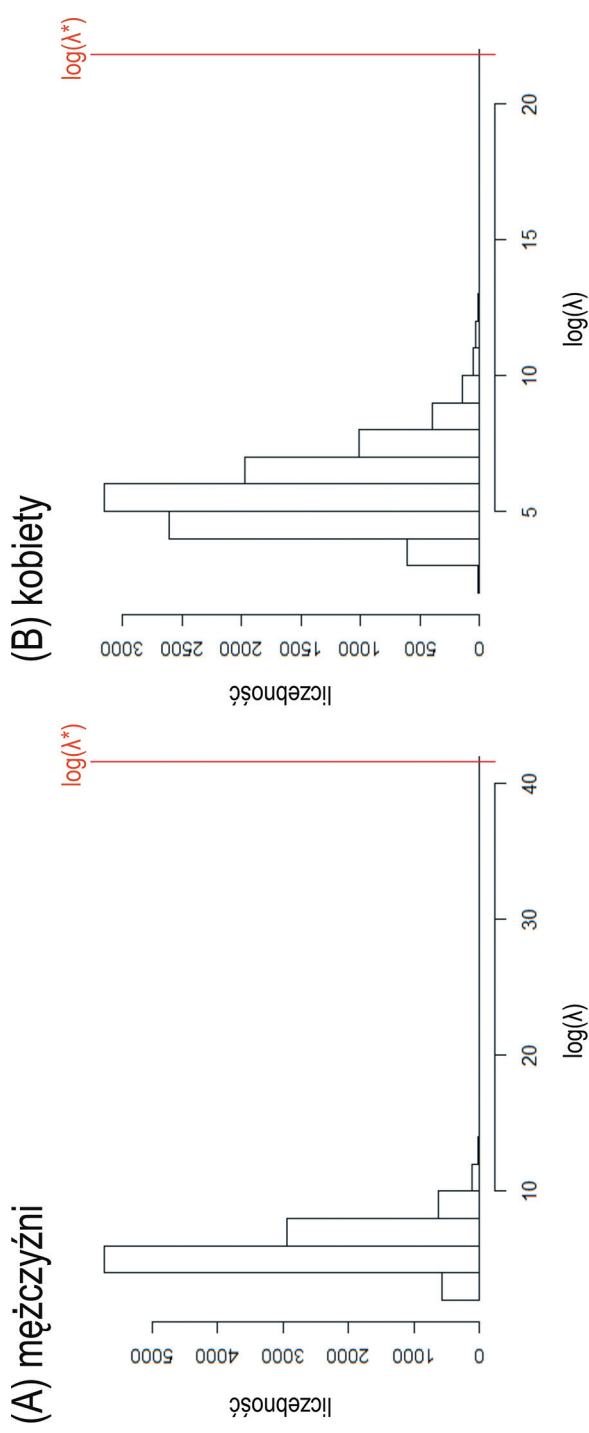
2010-2014



1980-1984

Ryc. 4. Standaryzowana umieralność (w przeliczeniu na 100 000 osób) na nowotwory złośliwe pęcherza moczowego i tarczycy w Polsce w latach 2010–2014 oraz skupienia zidentyfikowane na podstawie algorytmu Kulldorffa

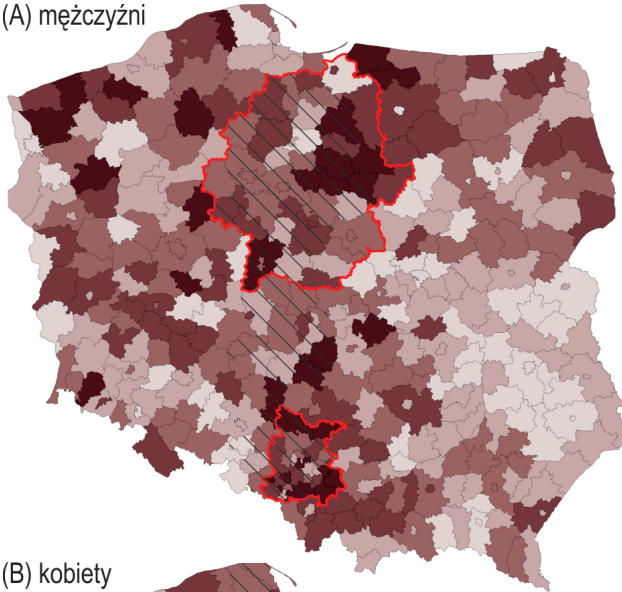
Fig. 4. Age-adjusted mortality rate (per 100,000 individuals) of bladder and thyroid cancers in Poland in 2010–2014 and the results of the Kulldorff's spatial scan statistic



Ryc. 5. Histogram statystyk Nagawalli uzyskanych na podstawie symulacji Monte Carlo oraz obserwowana wartość dla skupień I rzędu umieralności na nowotwory złośliwe żołądka

Fig. 5. Histogram of the Nagawalla's scan statistic based on the Monte Carlo testing with repetitions and the location of the observed Nagawalla's scan statistic for primary clusters of gastric cancer mortality

(A) mężczyźni



klastery:

2010–2014

1980–1984

standaryzowane współczynniki umieralności (2010–2014):

24,9–33,7

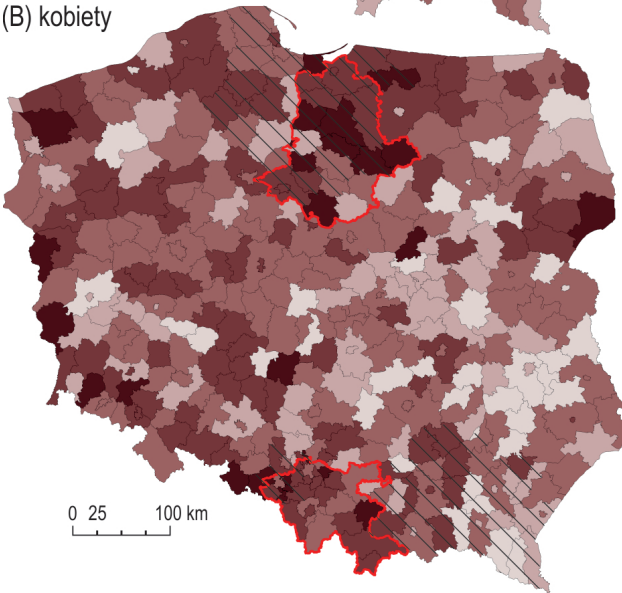
21,2–24,9

18,2–21,2

14,7–18,2

7,0–14,7

(B) kobiety



klastery:

2010–2014

1980–1984

standaryzowane współczynniki umieralności (2010–2014):

13,7–22,2

10,4–13,7

8,0–10,4

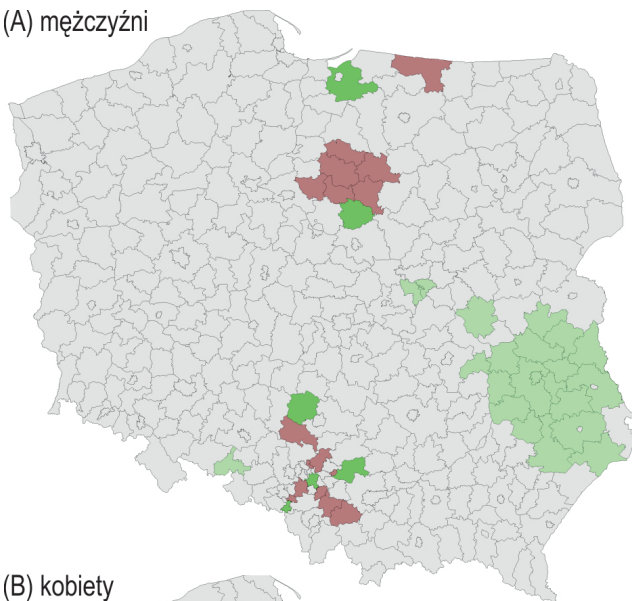
6,1–8,0

3,2–6,1

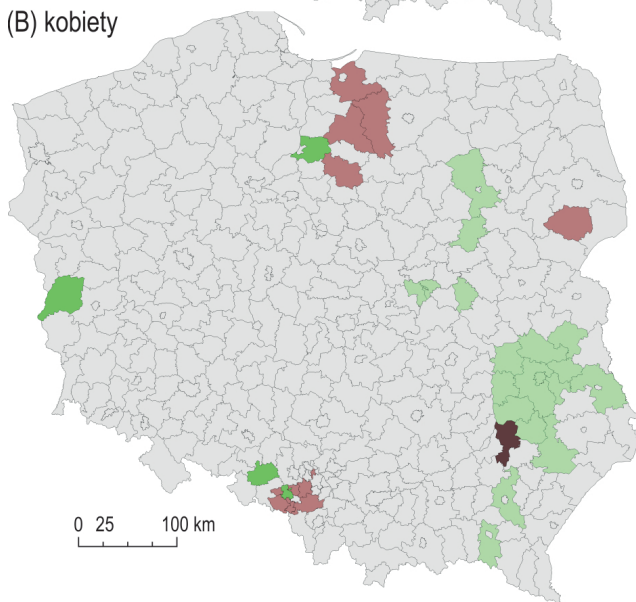
Ryc. 6. Standaryzowana umieralność (w przeliczeniu na 100 000 osób) na nowotwory złośliwe żołądka w Polsce oraz skupienia Kulldorffa

Fig. 6. Age-adjusted mortality rate (per 100,000 individuals) of gastric cancer in Poland and the results of the Kulldorff's spatial scan statistic

(A) mężczyźni



(B) kobiety

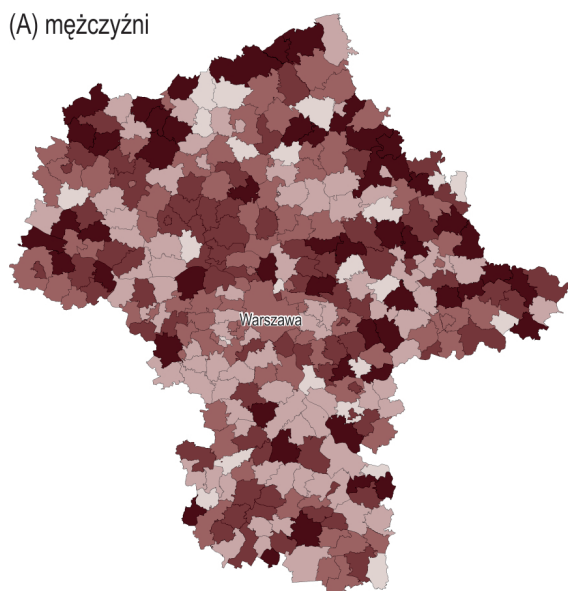


-  klasterzy wysokich wartości
-  klasterzy niskich wartości
-  obserwacje odstające wysokich wartości
-  obserwacje odstające niskich wartości

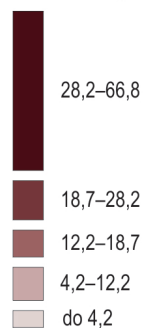
0 25 100 km

Ryc. 7. Skupienia i obserwacje odstające standaryzowanej umieralności na nowotwory złośliwe żołądka w Polsce w latach 2010–2014 zidentyfikowane przy pomocy algorytmu Morana
Fig. 7. Clusters and outliers of age-adjusted mortality of gastric cancer in Poland in 2010–2014 identified by the local Moran's index

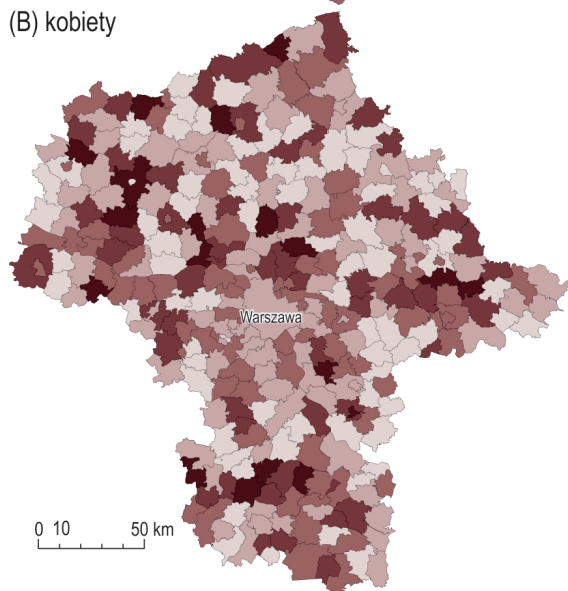
(A) mężczyźni



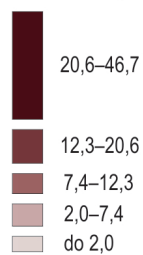
Standaryzowane współczynniki umieralności (2010–2014):



(B) kobiety



Standaryzowane współczynniki umieralności (2010–2014):



Ryc. 8. Standaryzowana umieralność (w przeliczeniu na 100 000 osób) na nowotwory złośliwe żołądka na Mazowszu w latach 2010–2014. Dane wejściowe testu Stone'a

Fig. 8. Age-adjusted mortality rate (per 100,000 individuals) of gastric cancer in Mazovia in 2010–2014. Input data of the Stone's test

z odległością od Warszawy, zarówno w przypadku mężczyzn, jak i kobiet. Wyniki te świadczą o polaryzacji zależności przestrzennych umieralności na nowotwory złośliwe żołądka na świecie. W Polsce różnice w poziomie umieralności na nowotwory złośliwe żołądka między obszarami wiejskimi i zurbanizowanymi są odwrotne niż te prezentowane przez Faivre'a i in. (1985), Yanga (2006) oraz Chiu i in. (2011).

Podsumowanie

Analiza skupień jest jedną z podstawowych metod badawczych w epidemiologii przestrzennej. Identyfikacja obszarów o ponad przeciętnej liczbie przypadków nowotworów złośliwych powinna mieć przełożenie na prewencję nowotworów. Wyjaśnianie przyczyn nierównomiernego rozkładu geograficznego umieralności i zapadalności na nowotwory wymaga posiadania wysokiej jakości danych nt. zachowań konsumenckich ludności. Reprezentatywność wyników *Badania Budżetów Gospodarstw Domowych* na poziomie powiatów jest wątpliwa z racji na stosowany w tym badaniu schemat losowania gospodarstw domowych. Wykonywanie analiz ekologicznych wiąże się ponadto z ryzykiem błędu na skutek korelacji pozornej.

Na podstawie przeprowadzonej analizy skupień przestrzennych można wnioskować o wyraźnym reżimie przestrzennym umieralności na nowotwory złośliwe w Polsce. Poza makroregionem Polski Wschodniej są zlokalizowane centra skupień przestrzennych I rzędu umieralności na nowotwory złośliwe w okresie 2010–2014 w 10 lokalizacjach, włączając płuco, jelito grube, pęcherz moczowy, żołądek, tarczycę, gruczoł krokowy, jądro, sutek, jajnik, macicę. Standaryzowana umieralność na nowotwory złośliwe ogółem w okresie 2010–2014 była w Polsce Wschodniej niższa niż w pozostałej części kraju o ponad 15% w populacji kobiet oraz o ponad 7% w populacji mężczyzn. Obie różnice były nieznacznie wyższe w okresie 1980–1984. Umieralność na tytoniozależne nowotwory złośliwe cechuje wyższe zróżnicowanie przestrzenne niż umieralność na nowotwory złośliwe ogółem.

Wyższa pogładowość charakteryzuje wyniki klasteryzacji według algorytmu Kulldorffa niż Morana. Identyfikacja skupień na podstawie statystyki lokalnej Morana nie wiąże się jakkolwiek z założeniem odnośnie do kształtu wyróżnionych skupień (np. eliptycznego, jak w przypadku algorytmu Kulldorffa). Możliwości aplikacyjne statystyki skanującej Kulldorffa ograniczają się do zmiennych, których rozkład w przestrzeni jest zbliżony do rozkładem Poissona.

Adnotacje

Artykuł to fragment wyników projektu realizowanego w ramach grantu wewnętrznego Centrum Onkologii – Instytutu im. Marii Skłodowskiej-Curie (numer grantu GW35KC).

Literatura

- Anselin L., 1995, *Local indicators of spatial association – LISA*, Geographical Analysis, 27 (2), 93–115.
- Besag J., Newell J., 1991, *The detection of clusters in rare diseases*, Journal of the Royal Statistical Society: Series A, 154 (1), 143–155.
- Chiu H.F., Tsai S.S., Chen P.S., Liao Y.H., Liou S.H., Wu T.N., Yang C.Y., 2011, *Traffic air pollution and risk of death from gastric cancer in Taiwan: petrol station density as an indicator of air pollutant exposure*, Journal of Toxicology and Environmental Health, Part A, 74 (18), 1215–1224.
- Czaderny K., 2019a, *Increasing deaths from colorectal cancer in Poland. Insights for optimising colorectal cancer screening in society and space*, Annals of Agricultural and Environmental Medicine, w druku.
- Czaderny K., 2019b, *Lung cancer mortality in historical context. How stable are spatial patterns of smoking over time?*, Demographic Research, w druku.
- Czaderny K., 2018, *Wysoka umieralność na nowotwory gruczołu krokowego w Polsce. Analiza przestrzenna, czasowa i strukturalna*, Przegląd Epidemiologiczny – Epidemiological Review, 72 (2).
- Faivre J., Boutron M.G., Arveux P., Bedenne L., Meny B., Klepping C., 1985, *Gastric carcinoma in Côte d’Or (France): A population-based study*, Gastroenterology, 88 (6), 1874–1879.
- Khan O., Protheroe A., 2007, *Testis cancer*, Postgraduate Medical Journal, 83 (984), 624–632.
- Kretowicz P., 2013, *Nowe podejścia badawcze w geografii zdrowia w literaturze anglosaskiej*, Przegląd Geograficzny, 85 (4), 549–571.
- Kulldorff M., Nagarwalla N., 1995, *Spatial disease clusters: detection and inference*, Statistics in Medicine, 14 (8), 799–810.
- Maliniak J., 1933, *Tytoń w dawnej Polsce*, Wydawnictwo M. Arcta, Warszawa.
- Minelli G., Conti S., Manno V., Olivieri A., Ascoli V., 2013, *The geographical pattern of thyroid cancer mortality between 1980 and 2009 in Italy*, Thyroid, 23 (12), 1609–1618.
- Moran P.A.P., 1948, *The interpretation of statistical maps*, Journal of the Royal Statistics Society, Series B (Methodological), 10 (2), 243–251.
- Moran P.A.P., 1950, *Notes on continuous stochastic phenomena*, Biometrika, 37 (1–2), 17–23.
- Mościbroda J., 1975, *Rozwój poglądów na metodę izarytmiczną oraz jej zastosowań w kartografii ludnościowej i gospodarczej*, 7 (2), 55–65.
- Nomura A.M.Y., Wilkens L.R., Henderson B. E., Epplein M., Kolonel L.N., 2012, *The association of cigarette smoking with gastric cancer: the multiethnic cohort study*, Cancer Causes & Control: CCC, 23(1), 51–58.
- Norström M., Pfeiffer D.U., Jarp J., 2000, *A space-time cluster investigation of an outbreak of acute respiratory disease in Norwegian cattle herds*, Preventive Veterinary Medicine, 47 (1–2), 107–119.

- Poljak N.K., Kontić M., Colović Z., Jerončić I., Lukšić B., Mulić R., 2011, *Iodine intake and epidemiological characteristics of thyroid cancer: comparison between inland and littoral Croatia*, Acta Clinica Croatica, 50 (3), 329–339.
- Powszechna Wystawa Krajowa, 1929, *Polski Monopol Tytoniowy: sprzedaż wyrobów tytoniowych w Polsce*, Polski Monopol Tytoniowy, Poznań.
- Program Operacyjny Polska Wschodnia 2014–2020, 2018, Ministerstwo Rozwoju Regionalnego.
- Pukkala E.I., Gustavsson N., Teppo, L., 1987, *Suomen syöpäkartasto 1953–82 – Atlas of cancer incidence in Finland, 1953–82*, Suomen syöpärekisteri, Helsinki.
- Ratajski L., 1973, *Metodyka kartografii społeczno-gospodarczej*, Państwowe Przedsiębiorstwo Wydawnictw Kartograficznych, Warszawa.
- Sheehan J.T., Gershman S.T., MacDougall L.A., Danley R.A., Mroszczyk M., Sorenson A.M., Kulldorff M., 2000, *Geographic assessment of breast cancer screening by towns, zip codes, and census tracts*, Journal of Public Health Management Practice, 6 (6), 48–57.
- Soroush A., Farshchian N., Komasi S., Izadi N., Amirifard N., & Shahmohammadi A. (2016), *The Role of Oral Contraceptive Pills on Increased Risk of Breast Cancer in Iranian Populations: A Meta-analysis*, Journal of Cancer Prevention, 21 (4), 294–301.
- Srivastava A., Kreiger N., 2004, *Cigarette smoking and testicular cancer*, Cancer Epidemiology, Biomarkers & Prevention, 13 (1), 49–54.
- Stone R.A., 1988, *Investigating of excess environmental risks around putative sources: Statistical problems and a proposed test*, Statistics in Medicine, 7 (6), 649–660.
- Yang L., 2006, *Incidence and mortality of gastric cancer in China*, World Journal of Gastroenterology, 12 (1), 17–20.

Krzysztof Czaderny
Centrum Onkologii – Instytut im. Marii Skłodowskiej-Curie,
Zakład Epidemiologii i Prewencji Nowotworów,
ul. Wawelska 15C, 02-034 Warszawa:
krzysztof.czaderny@coi.pl

