ANALIZA MIEJSCA ZAMIESZKANIA PACJENTEK KLINIKI GINEKOLOGII Z WYKORZYSTANIEM ANALIZY KORESPONDENCJI, ANALIZY KOSZYKOWEJ ORAZ METOD KLASYCZNYCH

Anna Justyna Milewska i Robert Milewski, Uniwersytet Medyczny w Białymstoku, Zakład Statystyki i Informatyki Medycznej

W roku 1999 rozpoczęto w Polsce wdrażanie reform sektora zdrowotnego. Jedną z wprowadzonych zmian było zniesienie rejonizacji. Od tego momentu pacjenci mogli wybierać miejsce leczenia i coraz popularniejsza stała się migracja pacjentów do preferowanych ośrodków. Spowodowane to było poszukiwaniem pomocy medycznej na jak najwyższym poziomie oraz chęcią rozpoczęcia leczenia po jak najkrótszym czasie oczekiwania. Celem bieżącej pracy jest analiza miejsca zamieszkania pacjentek wybranej kliniki ginekologicznej ze szczególnym uwzględnieniem przyczyn migracji pacjentek z innych województw. Zostaną przedstawione przykłady zastosowania analizy statystycznej z wykorzystaniem programu *STATISTICA Data Miner*. W podejściu klasycznym użyte zostaną testy nieparametryczne do oceny różnic pomiędzy analizowanymi grupami. Do wyszukania asocjacji między zmiennymi zastosowana zostanie analiza koszykowa, natomiast do prezentacji graficznej uzyskanych zależności – analiza korespondencji.

Wstęp

Wiele placówek medycznych dysponuje obszernymi bazami danych. Informacje te są często gromadzone przez wiele lat i mogą stanowić bogate źródło wiedzy. Umiejętne ich wykorzystanie może pozwolić na wprowadzenie usprawnień, optymalizację procesów, a także wykrycie istniejących nieprawidłowości. Prowadzenie takich analiz stało się niezbędne, gdy w roku 1999 rozpoczęto w Polsce wdrażanie reform sektora zdrowotnego i wszystkie zakłady opieki zdrowotnej zaczęły prowadzić samodzielna gospodarkę finansową. Wykorzystanie w analizach klasycznych metod statystycznych pozwala na opisanie wielu istniejących zależności, jednak przy dużych bazach danych okazuje się, że część z nich – z powodu złożoności – umyka uwadze badaczy. W takich sytuacjach coraz powszechniejsze staje się stosowanie metod data mining oraz analiz wielowymiarowych. Należy jednak pamiętać, że odkrycie zależności to jedynie pierwszy krok analizy, wymagający potwierdzenia za pomocą metod klasycznych.

Copyright © StatSoft Polska 2014

Materiał i metody

Do analizy statystycznej wykorzystano dane opisujące przebieg hospitalizacji pacjentek kliniki ginekologii leczonych w latach 1996, 2000 i 2004. Analizowane lata dotyczą odpowiednio: czasu przed wprowadzeniem reform sektora zdrowia, czasu funkcjonowania Kas Chorych i czasu funkcjonowania Narodowego Funduszu Zdrowia. Stworzona baza zawiera informacje o ponad ośmiu tysiącach pacjentek. W analizie wykorzystano dane dotyczące ich miejsca zamieszkania, przyczyn i długości hospitalizacji oraz wieku. Celem prawidłowego wykorzystania metod data mining dane zostały przekodowane z liczbowych na kategoryczne. Przedziały opisujące wiek (o szerokości 10 lat) zostały lewostronnie domknięte, natomiast czasy hospitalizacji zostały przedstawione poprzez klasy: 1 dzień, 2 dni, …, 6 dni, 7 dni, 1-2 tygodnie, 2-3 tygodnie, ponad 3 tygodnie. Jednostki chorobowe zostały zakodowane według Międzynarodowej Statystycznej Klasyfikacji Chorób i Problemów Zdrowotnych ICD-10. Pacjentki zostały również zaklasyfikowane do pięciu grup rozpoznań stanowiących główną przyczynę hospitalizacji: ginekologia, położnictwo, diagnostyka niepłodności, inseminacja (tylko rok 2004), IVF ET (procedura zapłodnienia in vitro). Miejsce zamieszkania zostało zakodowane według obecnych obszarów NFZ.

Do wykonania analiz wykorzystany został pakiet statystyczny *STATISTICA Data Miner*. Do oceny różnic pomiędzy grupami wykorzystano test dla dwóch wskaźników struktury, natomiast do oceny zależności test chi-kwadrat. Do wyznaczenia asocjacji występujących w analizowanej bazie danych wykorzystano jedną z metod Data Mining - analizę koszykową, natomiast do zobrazowania występujących powiązań zastosowano metodę z grupy wielowymiarowych technik eksploracyjnych - analizę korespondencji.

Cel

Celem pracy była analiza miejsca zamieszkania pacjentek hospitalizowanych w latach 1996, 2000 i 2004 z uwzględnieniem głównego rozpoznania i czasu trwania hospitalizacji. Szczególną uwagę zwrócono na różnice pomiędzy profilem pacjentek zamieszkujących województwo podlaskie oraz pacjentek z pozostałych części Polski.

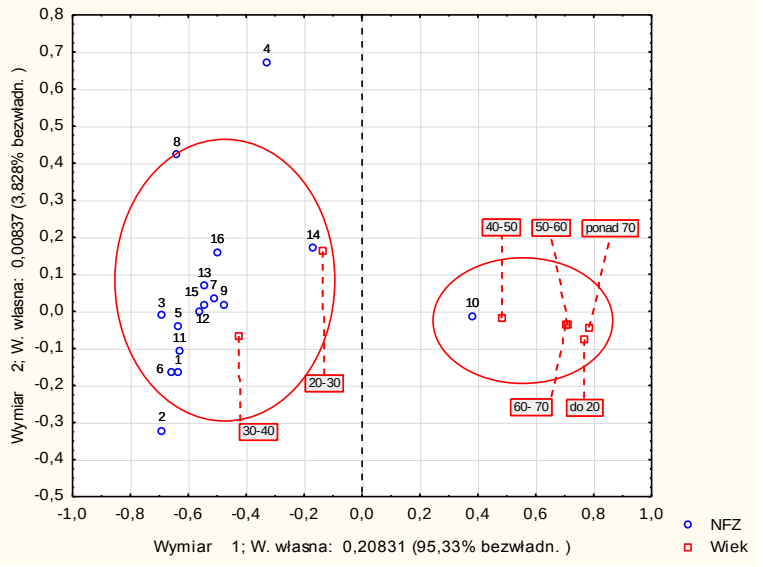
Wyniki i omówienie

Analiza korespondencji

Analiza korespondencji w pierwszej kolejności została wykorzystana do wyszukania powiązań pomiędzy miejscem zamieszkania a wiekiem pacjentek. Podczas analizy współwystępowania tych cech w poszczególnych latach dało się zaobserwować podobny schemat zależności. Prezentacja powiązań miejsca zamieszkania i wieku dla roku 1996 została przedstawiona w dwuwymiarowej przestrzeni (rys. 1), która wyjaśnia 99% ogólnej bezwładności. Pierwszy wymiar wyraźnie wyróżnia NFZ=10 (podlaski), który znajduje się po jednej stronie względem środka osi, w stosunku do pozostałych NFZ-ów znajdujących

 Copyright © StatSoft Polska 2014

się po drugiej stronie. Na rysunku widzimy zaznaczone dwa obszary skupienia punktów. Możemy zinterpretować je w następujący sposób: najmłodsze (do 20 r.ż.) oraz starsze pacjentki (po 40 r.ż.) najczęściej zamieszkują obszar NFZ-tu 10, natomiast pacjentki w wieku 20-40 lat mieszkają zazwyczaj poza woj. podlaskim. Należy pamiętać, że interpretacja skupień na mapach korespondencji mówi o współwystępowaniu cech bez podawania siły tej zależności, a uzyskane w ten sposób obserwacje wymagają potwierdzenia innymi metodami statystycznymi.

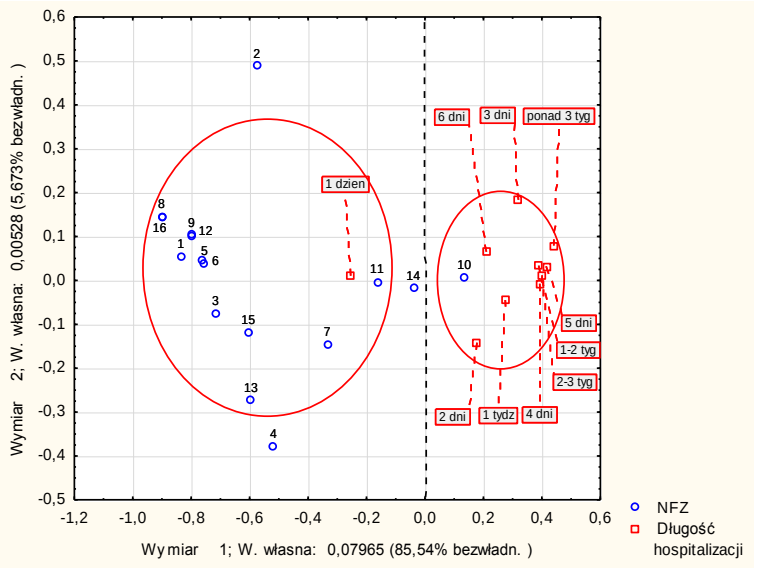


Rys. 1. Graficzna reprezentacja powiązań miejsca zamieszkania (NFZ) i wieku pacjentek hospitalizowanych w roku 1996.

W roku 2000 i 2004, w stosunku do roku 1996, nie wystąpiły istotne zmiany we wzajemnym położeniu punktów reprezentujących zarówno grupy wiekowe, jak i poszczególne NFZ-ty. Dwuwymiarowa przestrzeń dla danych z roku 2000 wyjaśnia wysoki odsetek bezwładności ogólnej - 98,5%, natomiast dla danych z roku 2004 - 96,3% bezwładności ogólnej. Zaproponowany podział na NFZ 10 i inne oraz wiek 20-40 i pozostałe grupy wiekowe został potwierdzony testem chi-kwadrat, który wykazał istotną zależność pomiędzy tak zdefiniowanymi cechami dla każdego analizowanego roku ($p<0,001$).

Kolejna analiza dotyczyła powiązań pomiędzy miejscem zamieszkania (NFZ) i długością hospitalizacji. Zależności zostały przedstawione na dwuwymiarowej mapie korespondencji, która wyjaśnia dla poszczególnych lat od 85,2% do 98,1% ogólnej bezwładności. Podobnie jak w poprzednim przypadku, również tutaj nie obserwuje się istotnych zmian w położeniu punktów reprezentujących długość hospitalizacji i poszczególne NFZ-ty na przestrzeni lat. Na rys. 2 widzimy dwa obszary skupienia. Pierwszy wymiar wyraźnie różnicuje krótkie i długie hospitalizacje. Krótkie, jednodniowe pobyty w klinice dotyczyły pacjentek mieszkających poza woj. podlaskim, a dłuższe hospitalizacje dotyczyły

pacjentek z woj. podlaskiego (NFZ=10). Test chi-kwadrat wykazał istotną zależność jednodniowych hospitalizacji od miejsca zamieszkania dla każdego roku (p<0,001).

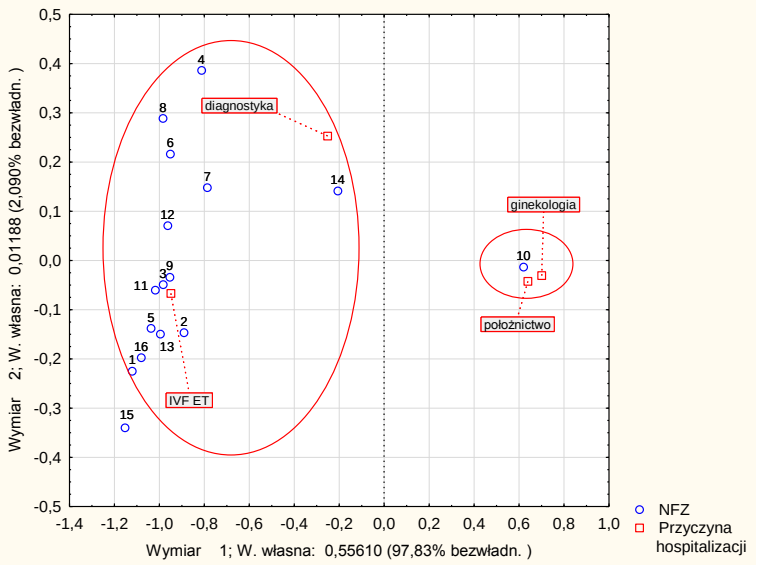


Rys. 2. Graficzna reprezentacja powiązań miejsca zamieszkania (NFZ) i długości hospitalizacji pacjentek kliniki leczonych w roku 1996.

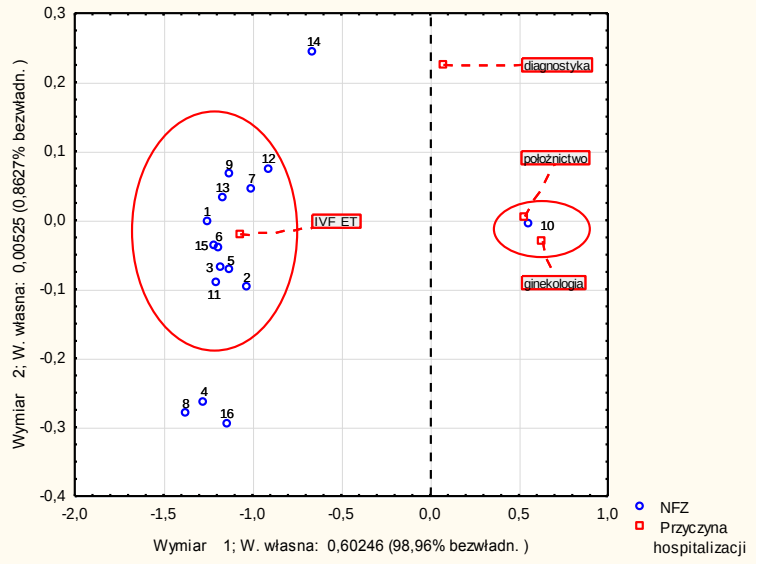
Obserwowane były również powiązania pomiędzy główną przyczyną hospitalizacji a miejscem zamieszkania (NFZ). Główna przyczyna pobytu na oddziale ginekologicznym była bardzo zróżnicowana. Z tego powodu pacjentki zostały przypisane do bardziej jednorodnych podgrup rozpoznań, których analiza ukazała wyraźną zależność przyczyny hospitalizacji od miejsca zamieszkania. Współwystępowanie cech w roku 1996 zostało przedstawione za pomocą dwuwymiarowej mapy korespondencji, wyjaśniającej 99,9% bezwładności ogólnej (rys. 3). Pierwszy wymiar wyraźnie wskazuje na skupienie obejmujące NFZ 10 i przyczyny ginekologiczno-położnicze, które znalazły się na prawo od środka osi. Natomiast z drugiej strony osi widzimy, że diagnostyka i leczenie niepłodności było przyczyną migracji pacjentek z innych województw.

Analizując rozłożenie punktów reprezentujących przyczyny hospitalizacji i poszczególne NFZ-ty dla roku 2000, widzimy istotną różnicę w stosunku do roku 1996. Na rys. 4 przedstawiającym dwuwymiarową przestrzeń wyjaśniającą 99,8% ogólnej bezwładności widzimy, że pierwszy wymiar oddziela pacjentki hospitalizowane z powodu IVF ET od wszystkich innych przyczyn. Diagnostyka niepłodności w roku 2000 znalazła się minimalnie po tej samej stronie osi co NFZ 10, czyli diagnostyka stała się przyczyną hospitalizacji nieco bardziej typową dla pacjentek z woj. podlaskiego niż przyjeżdżających z innych województw.

 Copyright © StatSoft Polska 2014



Rys. 3. Graficzna reprezentacja powiązań miejsca zamieszkania (NFZ) i przyczyny hospitalizacji pacjentek kliniki leczonych w roku 1996.



Rys. 4. Graficzna reprezentacja powiązań miejsca zamieszkania (NFZ) i przyczyny hospitalizacji pacjentek kliniki leczonych w roku 2000.

Mapa korespondencji dla roku 2004 przedstawiła analogiczny układ punktów jak w roku 1996. Dwuwymiarowa przestrzeń również wyjaśniła bardzo duży odsetek ogólnej bezwładności – 99,1%. Aby odpowiedzieć na pytanie, z jakiego powodu w roku 2000 punkt

reprezentujący diagnostykę niepłodności zmienił swoje położenie (współwystępował z NFZ-tem 10), należy wykonać kolejne analizy statystyczne.

Głównym wnioskiem z powyższych zastosowań analizy korespondencji jest taki sam podział pacjentek ze względu na miejsce zamieszkania uzyskany podczas każdej z analiz. Jedną grupę stanowią kobiety z woj. podlaskiego (NFZ 10), natomiast drugą – pacjentki mieszkające na pozostałym obszarze Polski. Ponadto możemy powiedzieć, że pacjentki z województwa podlaskiego mieściły się głównie w grupie wiekowej do 20 lat lub ponad 40 lat, były one hospitalizowane z powodów ginekologicznych i położniczych przez co najmniej 2 dni. Natomiast pacjentki migrujące z innych województw leczyły się zazwyczaj 1 dzień, były w wieku 20-40 lat i głównie poddawały się procedurze IVF ET.

Analiza koszykowa

Analiza koszykowa zastosowana na tych samych danych dotyczących przebiegu hospitalizacji ukazuje zależności pomiędzy analizowanymi cechami za pomocą ich procentowego współwystępowania. Reguły asocjacji były tworzone osobno dla dwóch grup pacjentek, które zostały wydzielone podczas analizy korespondencji. W ten sposób analiza koszykowa pozwoliła na dogłębną obserwację asocjacji występujących wśród kobiet z województwa podlaskiego oraz wśród kobiet z innych województw, połączonych w jedną grupę.

Obserwacja asocjacji dotyczących najmłodszych pacjentek z woj. podlaskiego dla lat 1996, 2000 i 2004 została przedstawiona w tabeli 1. Reguła asocjacji dla roku 1996 mówi o tym, że jeżeli pacjentka miała do 20 lat, to w 82,6% przypadków była hospitalizowana z powodów ginekologicznych ("grupa = g"). W roku 2000 odsetek przyczyn ginekologicznych w grupie najmłodszych pacjentek zwiększył się do 87,5%, a w roku 2004 wzrósł dalej do 88,7%. Patrząc z drugiej strony na uzyskane odsetki, widzimy, że zmniejsza się udział innych przyczyn hospitalizacji u pacjentek najmłodszych - 17,4%, 12,5%, 11,3% (jednak jest on nieistotny statystycznie).

Tabela 1. Reguły asocjacji dotyczące pacjentek w wieku do 20 lat pochodzących z woj. podlaskiego hospitalizowanych w klinice w latach 1996, 2000, 2004.

Rok 1996						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
1	wiek_kod == do 20	==>	grupa == g	2,32416	82,6087	16,5074
Rok 2000						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
1	wiek_kod == do 20	==>	grupa == g	2,37020	87,5000	17,3347
Rok 2004						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
1	wiek_kod == do 20	==>	grupa == g	2,90725	88,7324	18,5703

Następnie wśród uzyskanych reguł asocjacji zostały wybrane te, które dotyczyły najstarszych pacjentek (tabela 2). W każdym analizowanym roku współczynnik zaufania dla tych asocjacji wyniósł 100%, co oznacza, że wszystkie pacjentki w wieku ponad 70 lat były

 Copyright © StatSoft Polska 2014

hospitalizowane z przyczyn ginekologicznych. Jest to informacja oczywista, ponieważ najstarsze kobiety we względu na wiek nie mogły być leczone z przyczyn związanych z rozrodem. Jednak zwróćmy uwagę na wysokość parametru wsparcia w tych regułach, który mówi nam o tym, jaki odsetek wszystkich pacjentek stanowiły kobiety po 70 r.ż. Zauważamy, że w kolejnych latach odsetek najstarszych kobiet wzrasta, w roku 1996 wynosił 4,1%, w roku 2000 wynosił podobnie 4,2%, natomiast w 2004 wzrósł do 6,9%. Ta tendencja wzrostowa sugeruje nam, że na przestrzeni analizowanych lat możemy dostrzec efekt starzenia się społeczeństwa. Wzrost odsetka (obserwowany pomiędzy latami 1996 i 2004) osób najstarszych jest istotny statystycznie (p=0,0002).

Tabela 2. Reguły asocjacji dotyczące pacjentek w wieku ponad 70 lat pochodzących z woj. podlaskiego hospitalizowanych w klinice w latach 1996, 2000, 2004.

Rok 1996						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
5	wiek_kod == ponad 70	==>	grupa == g	4,09786	100,0000	24,1163
Rok 2000						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
5	wiek_kod == ponad 70	==>	grupa == g	4,17607	100,0000	24,5982
Rok 2004						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
5	wiek_kod == ponad 70	==>	grupa == g	6,87587	100,0000	30,3181

Najliczniejszą grupę wiekową hospitalizowaną w klinice stanowiły kobiety w wieku 40-50 lat. Z tego powodu ważne było dokładniejsze przyjrzenie się przyczynom leczenia, jak też długościom hospitalizacji. Asocjacje zebrane w tabeli 3 mówią, że pacjentki w wieku 40-50 lat były hospitalizowane w ponad 90% przypadków z przyczyn ginekologicznych. W roku 1996 było to 92,3% przyczyn ginekologicznych, w roku 2000 – 94,4%, natomiast w 2004 roku 95,5%. Widzimy, że na przestrzeni lat odsetek ten zwiększa się, co jednocześnie oznacza, że zmniejszają się częstości hospitalizacji tych kobiet z przyczyn „rozrodczych" (różnica jednak była nieistotna statystycznie).

Tabela 3. Reguły asocjacji dotyczące grupy wiekowej (40-50 lat) pacjentek pochodzących z woj. podlaskiego hospitalizowanych w klinice w latach 1996, 2000, 2004.

Rok 1996						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
2	wiek_kod == 40-50	==>	grupa == g	24,83180	92,2727	57,0261
Rok 2000						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
2	wiek_kod == 40-50	==>	grupa == g	20,99323	94,4162	53,5897
Rok 2004						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
2	wiek_kod == 40-50	==>	grupa == g	21,68897	95,5285	52,6288

Podczas analizy elementów dotyczących procesu hospitalizacji warto przyjrzeć się długim hospitalizacjom. Tab. 4 przedstawia asocjacje dotyczące pobytów w szpitalu trwających od 8 do 14 dni. Widzimy, że w każdym analizowanym roku, jeżeli leczenie trwało 8-14 dni, to w ponad 90% przypadków dotyczyło ono hospitalizacji z powodów ginekologicznych – współczynnik zaufania wyniósł odpowiednio 94,7%, 90,7%, 96,3%. Oczywiste jest, że długie pobyty pacjenta w szpitalu generują duże koszta. Na przestrzeni lat 1996-2004 zaszły istotne zmiany w systemie ochrony zdrowia. Czy wpłynęło to na długość czasu hospitalizacji w klinice ginekologii? Popatrzmy na zmiany współczynnika wsparcia dla kolejnych lat. Początkowo (w roku 1996) długie, ginekologiczne pobyty stanowiły 18,9% wszystkich hospitalizacji, w roku 2000 ich odsetek zmniejszył się do poziomu 14,3%, a w roku 2004 wyniósł 13,2%. Zmniejszenie się odsetka długich hospitalizacji było istotne statystycznie ($p<0,0001$) pomiędzy latami 1996 i 2004. Widzimy więc, że reformy systemu ochrony zdrowia wpłynęły na skrócenie czasów hospitalizacji. Uzyskane zostało to przez wprowadzenie konieczności finansowego rozliczania się oraz częściowo pod wpływem rozwoju metod leczenia.

Tabela 4. Reguły asocjacji dotyczące długich hospitalizacji (8-14 dni) pacjentek z woj. podlaskiego leczonych w klinice w latach 1996, 2000, 2004.

Rok 1996						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
6	dlugosc_kod == 1-2 tyg	==>	grupa == g	18,89908	94,7853	50,4224
Rok 2000						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
8	dlugosc_kod == 1-2 tyg	==>	grupa == g	14,27765	90,6810	43,3117
Rok 2004						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
9	dlugosc_kod == 1-2 tyg	==>	grupa == g	13,24412	96,3087	41,2936

Analiza przebiegu hospitalizacji pacjentek migrujących z innych województw do kliniki ginekologii pozwoliła odpowiedzieć na pytania o wiek przyjezdnych kobiet, ich miejsce zamieszkania oraz przyczynę i długość hospitalizacji. Tabela 5 przedstawia asocjacje opisujące najliczniejszą grupę przyjezdnych pacjentek – w wieku 30-40 lat. Widzimy, że jeżeli pacjentki należały do grupy wiekowej 30-40 lat to w 89,2% przypadków pozostawały w szpitalu 1 dzień (rok 1996). W roku 2000 jednodniowe hospitalizacje stanowiły 92,6%, natomiast w 2004 dotyczyły 85,6% kobiet. Współczynnik wsparcia jest wysoki i mówi, że wśród wszystkich pacjentek przyjeżdżających do kliniki z innych województw co druga kobieta pochodziła z grupy wiekowej 30-40 lat i pozostawała na oddziale 1 dzień (dla poszczególnych lat 1996, 2000, 2004 odpowiednio uzyskano 54,4%, 54,7%, 46,3%).

 Copyright © StatSoft Polska 2014

Tabela 5. Reguły asocjacji dotyczące grupy wiekowej (30-40 lat) pacjentek pochodzących spoza woj. podlaskiego hospitalizowanych w klinice w latach 1996, 2000, 2004.

Rok 1996						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
2	wiek_kod == 30-40	==>	dlugosc_kod == 1 dzien	54,41667	89,2077	75,9070
Rok 2000						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
2	wiek_kod == 30-40	==>	dlugosc_kod == 1 dzien	54,72901	92,6259	75,0461
Rok 2004						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
1	wiek_kod == 30-40	==>	dlugosc_kod == 1 dzien	46,33028	85,5932	72,6775

Kolejna analiza wskazuje na główną przyczynę hospitalizacji pacjentek mieszkających poza województwem podlaskim, leczących się w klinice – tabela 6. Asocjacje dla każdego roku pokazują, że jeżeli przyczyną hospitalizacji było rozpoznanie Z31 (zapłodnienie IVF), to wszystkie pacjentki (zaufanie 100%) pozostawały w klinice jeden dzień. Możemy podsumować to, stwierdzając, że wszystkie pobyty w celu wykonania zapłodnienia in vitro (Z31) były jednodniowe. Współczynnik wsparcia pokazuje, jak liczna była to grupa wśród wszystkich pacjentek pochodzących z innych województw. W roku 1996 w tym celu przyjechało do kliniki 68,8% kobiet, w roku 2000 było to 82,3%, natomiast w roku 2004 współczynnik ten spadł do 58,2%. Zaobserwowane odsetki liczby hospitalizacji z powodu rozpoznania Z31 były istotnie statystycznie zależne od roku hospitalizacji ($p<0,001$).

Tabela 6. Reguły asocjacji opisujące przyczynę hospitalizacji Z31 dla pacjentek pochodzących spoza woj. podlaskiego hospitalizowanych w klinice w latach 1996, 2000, 2004.

Rok 1996						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
31	kod1 == Z31	==>	dlugosc_kod == 1 dzien	68,83333	100,0000	90,3888
Rok 2000						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
46	kod1 == Z31	==>	dlugosc_kod == 1 dzien	82,35919	100,0000	95,6553
Rok 2004						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
41	kod1 == Z31	==>	dlugosc_kod == 1 dzien	58,25688	100,0000	88,0890

Pacjentki kliniki migrujące spoza woj. podlaskiego pochodziły z obszaru całej Polski, a najliczniejszą grupę stanowiły kobiety podlegające pod NFZ 7 (woj. mazowieckie). Asocjacje zebrane w tabeli 7 pokazują, że jeżeli pacjentka przyjechała z woj. Mazowieckiego, to zazwyczaj pozostała na oddziale 1 dzień. W roku 1996 jednodniowe hospitalizacje stanowiły 83,3%, w roku 2000 – 89,2%, natomiast w 2004 roku 68,3%. Jednodniowa

hospitalizacja zazwyczaj oznacza wykonanie zapłodnienia in vitro, widzimy więc, że w 2004 roku odsetek innych przyczyn leczenia (dłuższe pobyty) wyraźnie się zwiększył. Hospitalizacje dłuższe niż jednodniowe wynosiły 16,7%, 10,8%, 31,7% odpowiednio dla poszczególnych analizowanych lat. Test chi-kwadrat potwierdził istotną zależność pomiędzy odsetkiem hospitalizacji jednodniowych a analizowanymi latami (p<0,0001).

Tabela 7. Reguły asocjacji opisujące pacjentki z NFZ 7 hospitalizowane w klinice w latach 1996, 2000, 2004.

Rok 1996						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
12	nfz == 7,	==>	dlugosc_kod == 1 dzien	21,58333	83,2797	46,1896
Rok 2000						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
23	nfz == 7,	==>	dlugosc_kod == 1 dzien	21,89160	89,1775	46,5714
Rok 2004						
	Poprzednik	**==>**	**Następnik**	**Wsparcie%**	**Zaufanie(%)**	**Korelacja(%)**
73	nfz == 7,	==>	dlugosc_kod == 1 dzien	21,10092	68,3168	43,8190

Analiza koszykowa pozwoliła na wykonanie dokładnej analizy częstości występowania różnych podgrup rozpatrywanych cech. Dzięki zaobserwowanym asocjacjom i wykonaniu testów statystycznych możemy powiedzieć, że wśród pacjentek z woj. podlaskiego istotnie zwiększył się odsetek osób najstarszych oraz zmniejszył się odsetek 1-2 tyg. hospitalizacji. Ponadto widzimy, że najliczniejszą grupę stanowiły kobiety w wieku 40-50 lat i zazwyczaj leczone były z podwodów ginekologicznych. Wśród pacjentek z innych województw najliczniejszą grupę stanowiły kobiety w wieku 30-40 lat, które zazwyczaj były hospitalizowane przez 1 dzień. Ponadto wystąpiły istotne zmiany w liczbie wykonywanych procedur IVF ET. Widzimy, że pacjentki z woj. mazowieckiego (NFZ 7) były hospitalizowane głównie 1 dzień oraz że istotnie zwiększył się dla nich w kolejnych latach odsetek dłuższych niż jednodniowe hospitalizacji.

Wnioski

Analiza korespondencji, analiza koszykowa i klasyczna analiza statystyczna w ujęciu praktycznym są metodami wzajemnie uzupełniającymi się. Pierwsza z nich pozwoliła na odkrycie powiązań pomiędzy zmiennymi, dotyczących procesu hospitalizacji. Umożliwiła ich graficzną prezentację w postaci mapy korespondencji. Ukazała obszary skupień, na podstawie których zostały przeprowadzone dalsze analizy w podgrupach.

Zaletą analizy koszykowej jest wykrywanie współistnienia wielu charakterystyk obiektu w oparciu o złożone asocjacje. Analiza koszykowa pozwoliła na głębszą i dokładniejszą eksplorację zgromadzonych informacji w określonych wcześniej podgrupach.

Obie metody ułatwiają formułowanie hipotez dotyczących powiązań pomiędzy zmiennymi, które następnie należy zweryfikować, wykorzystując klasyczną analizę statystyczną.

 Copyright © StatSoft Polska 2014

Bibliografia

1. Agrawal R., Imielinski T., Swami A., Mining association rules between sets of items in large databases. In Proceedings of the ACM SIGMOD Conference on Management of Data, Washington D.C., s. 207-216, 1993.
2. Greenacre M., Correspondence analysis in medical research, Statistical Methods in Medical Research, (1), s. 97-117, 1992.
3. Milewska A.J., Jankowska D., Gorska U., Milewski R., & Wolczynski S., Graphical representation of the relationships between qualitative variables concerning the process of hospitalization in the gynecological ward using correspondence analysis. Studies in Logic, Grammar and Rhetoric. Logical, Statistical and Computer Methods in Medicine, 29(42), s. 7-25, 2012.
4. Milewska A.J., Górska U., Jankowska D. et al., The use of the basket analysis in a research of the process of hospitalization in the gynecological ward, Studies in Logic, Grammar and Rhetoric, 25(38), s. 83-98, 2011.
5. Milewska A.J., Milewski R., Mnich S.Z., Karpińska M., Wołczyński S., Wpływ starzenia się społeczeństwa na strukturę chorobowości w ginekologii, Przegląd Menopauzalny, 5, s. 330-334, 2010.
6. Milewska A.J., Milewski R., Wołczyński S., Analiza zjawiska migracji pacjentów na Podlasie na przykładzie Kliniki Ginekologii, Polityka Zdrowotna, 7, s. 71-76, 2008.
7. Stanisz A., Przystępny kurs statystyki z zastosowaniem *STATISTICA* PL na przykładach z medycyny. Tom 3. Analizy wielowymiarowe, StatSoft Polska, Kraków, 2006.