

**PRACE NAUKOWE**

Uniwersytetu Ekonomicznego we Wrocławiu

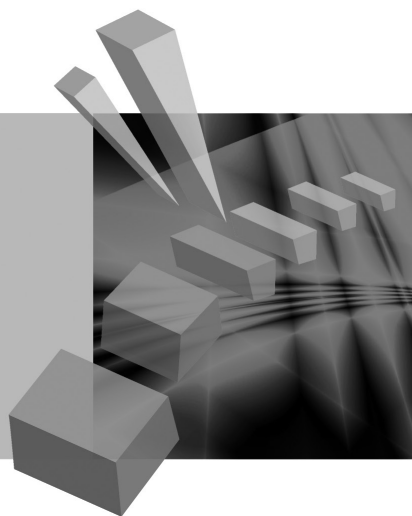
**RESEARCH PAPERS**

of Wrocław University of Economics

**279**

# Taksonomia 21

## Klasyfikacja i analiza danych – teoria i zastosowania



Redaktorzy naukowi

**Krzysztof Jajuga**

**Marek Walesiak**



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu  
Wrocław 2013

Redaktor Wydawnictwa: Aleksandra Śliwka

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Małgorzata Czupryńska

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:

[www.ibuk.pl](http://www.ibuk.pl), [www.ebscohost.com](http://www.ebscohost.com),

The Central and Eastern European Online Library [www.ceeol.com](http://www.ceeol.com),

a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon

[http://kangur.uek.krakow.pl/bazy\\_ae/bazekon/nowy/index.php](http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php)

Informacje o naborze artykułów i zasadach recenzowania znajdują się

na stronie internetowej Wydawnictwa

[www.wydawnictwo.ue.wroc.pl](http://www.wydawnictwo.ue.wroc.pl)

Tytuł dofinansowany ze środków Narodowego Banku Polskiego

oraz ze środków Sekcji Klasyfikacji i Analizy danych PTS

Kopiowanie i powielanie w jakiegokolwiek formie

wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu

Wrocław 2013

**ISSN 1899-3192** (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)

**ISSN 1505-9332** (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

## Spis treści

<b>Wstęp</b> .....	9
<b>Sabina Denkowska, Kamil Fijorek, Marcin Salamaga, Andrzej Sokolowski:</b> Sejm VI kadencji – maszynka do głosowania .....	11
<b>Barbara Pawelek, Adam Sagan:</b> Zmienne ukryte w modelach ekonomicznych – respecyfikacja modelu Kleina I .....	19
<b>Jan Paradysz:</b> Nowe możliwości badania koniunktury na rynku pracy .....	29
<b>Krzysztof Najman:</b> Samouczące się sieci GNG w grupowaniu dynamicznym zbiorów o wysokim wymiarze .....	41
<b>Kamila Migdał-Najman:</b> Zastosowanie jednowymiarowej sieci SOM do wyboru cech zmiennych w grupowaniu dynamicznym .....	48
<b>Aleksandra Matuszewska-Janica, Dorota Witkowska:</b> Zróżnicowanie płac ze względu na płeć: zastosowanie drzew klasyfikacyjnych .....	58
<b>Iwona Foryś, Ewa Putek-Szeląg:</b> Przestrzenna klasyfikacja gmin ze względu na sprzedaż użytków gruntowych zbywanych przez ANR w województwie zachodniopomorskim .....	67
<b>Joanna Banaś, Małgorzata Machowska-Szewczyk:</b> Klasyfikacja internetowych rachunków bankowych z uwzględnieniem zmiennych symbolicznych.....	77
<b>Marta Jaročka:</b> Wpływ metody doboru cech diagnostycznych na wynik porządkowania liniowego na przykładzie rankingu polskich uczelni .....	85
<b>Anna Zamojska:</b> Badanie zgodności rankingów wyznaczonych według różnych wskaźników efektywności zarządzania portfelem na przykładzie funduszy inwestycyjnych.....	95
<b>Dorota Rozmus:</b> Porównanie dokładności taksonomicznej metody propagacji podobieństwa oraz zagregowanych algorytmów taksonomicznych opartych na idei metody <i>bagging</i> .....	106
<b>Ewa Wędrowska:</b> Wrażliwość miar dywergencji jako mierników niepodobieństwa struktur.....	115
<b>Katarzyna Wójcik, Janusz Tuchowski:</b> Wpływ automatycznego tłumaczenia na wyniki automatycznej identyfikacji charakteru opinii konsumenckich ...	124
<b>Małgorzata Misztal:</b> Ocena wpływu wybranych metod imputacji na wyniki klasyfikacji obiektów w modelach drzew klasyfikacyjnych.....	135
<b>Anna Czapkiewicz, Beata Basiura:</b> Badanie wpływu wyboru współczynnika zależności na grupowanie szeregów czasowych .....	146
<b>Tomasz Szubert:</b> Czynniki różnicujące poziom zadowolenia z życia oraz wartości życiowe osób sprawnych i niepełnosprawnych w świetle badań „Diagnozy społecznej” .....	154

<b>Marcin Szymkowiak:</b> Konstrukcja estymatorów kalibracyjnych wartości globalnej dla różnych funkcji odległości .....	164
<b>Wojciech Roszka:</b> Szacowanie łącznych charakterystyk cech nieobserwowanych łącznie .....	174
<b>Justyna Brzezińska:</b> Metody wizualizacji danych jakościowych w programie <b>R</b> .....	182
<b>Agata Sielska:</b> Regionalne zróżnicowanie potencjału konkurencyjnego polskich gospodarstw rolnych w województwach po akcesji do Unii Europejskiej .....	191
<b>Mariusz Kubus:</b> Liniowy model prawdopodobieństwa z regularyzacją jako metoda doboru zmiennych .....	201
<b>Beata Basiura:</b> Metoda Warda w zastosowaniu klasyfikacji województw Polski z różnymi miarami odległości .....	209
<b>Katarzyna Wardzińska:</b> Wykorzystanie metody obwiedni danych w procesie klasyfikacji przedsiębiorstw .....	217
<b>Katarzyna Dębowska:</b> Modelowanie upadłości przedsiębiorstw oparte na próbach niezbilansowanych .....	226
<b>Danuta Tarka:</b> Wpływ metody doboru cech diagnostycznych na wyniki klasyfikacji obiektów na przykładzie danych dotyczących ochrony środowiska ..	235
<b>Artur Czech:</b> Zastosowanie wybranych metod doboru zmiennych diagnostycznych w badaniach konsumpcji w ujęciu pośrednim .....	246
<b>Beata Bal-Domańska:</b> Ocena relacji zachodzących między inteligentnym rozwojem a spójnością ekonomiczną w wymiarze regionalnym z wykorzystaniem modeli panelowych .....	255
<b>Mariola Chrzanowska:</b> <i>Ordinary kriging</i> i <i>inverse distance weighting</i> jako metody szacowania cen nieruchomości na przykładzie warszawskiego rynku .....	264
<b>Adam Depta:</b> Zastosowanie analizy wariancji w badaniu jakości życia na podstawie kwestionariusza SF-36v2 .....	272
<b>Maciej Beręsewicz, Tomasz Klimanek:</b> Wykorzystanie estymacji pośredniej uwzględniającej korelację przestrzenną w badaniach cen mieszkań .....	281
<b>Karolina Paradysz:</b> Benchmarkowa analiza estymacji dla małych obszarów na lokalnych rynkach pracy .....	291
<b>Anna Gryko-Nikitin:</b> Dobór parametrów w równoległych algorytmach genetycznych dla problemu plecakowego .....	301
<b>Tomasz Ząbkowski, Piotr Jałowiecki:</b> Zastosowanie reguł asocjacyjnych do analizy danych ankietowych w wybranych obszarach logistyki przedsiębiorstw przetwórstwa rolno-spożywczego .....	311
<b>Agnieszka Przedborska, Małgorzata Misztal:</b> Zastosowanie metod statystyki wielowymiarowej do oceny wydolności stawów kolanowych u pacjentów z chorobą zwyrodnieniową leczonych operacyjnie .....	321
<b>Dorota Perło:</b> Rozwój zrównoważony w wymiarze gospodarczym, społecznym i środowiskowym – analiza przestrzenna .....	331

<b>Ewa Putek-Szeląg, Urszula Gieraltowska, Analiza i diagnoza wielkości produkcji energii odnawialnej w Polsce na tle krajów Unii Europejskiej..</b>	342
--	-----

## Summaries

<b>Sabina Denkowska, Kamil Fijorek, Marcin Salamaga, Andrzej Sokolowski: VIth-term Sejm – a voting machine .....</b>	18
<b>Barbara Pawelek, Adam Sagan: Latent variables in econometric models – respecification of Klein I model .....</b>	28
<b>Jan Paradysz: New possibilities for studying the situation on the labour market .....</b>	40
<b>Krzysztof Najman: Self-learning neural network of GNG type in the dynamic clustering of high-dimensional data.....</b>	47
<b>Kamila Migdał-Najman: Applying the one-dimensional SOM network to select variables in dynamic clustering .....</b>	57
<b>Aleksandra Matuszewska-Janica, Dorota Witkowska: Gender wage gap: application of classification trees.....</b>	66
<b>Iwona Foryś, Ewa Putek-Szeląg: Spatial classification of communes by usable land traded by the APA in the Zachodniopomorskie voivodeship...</b>	76
<b>Joanna Banaś, Małgorzata Machowska-Szewczyk: Classification of Internet banking accounts including symbolic variables .....</b>	84
<b>Marta Jarocka: The impact of the method of the selection of diagnostic variables on the result of linear ordering on the example of ranking of universities in Poland.....</b>	94
<b>Anna Zamojska: Empirical analysis of the consistency of mutual fund ranking for different portfolio performance measures.....</b>	105
<b>Dorota Rozmus: Comparison of accuracy of affinity propagation clustering and cluster ensembles based on bagging idea.....</b>	114
<b>Ewa Wędrowska: Sensitivity of divergence measures as structure dissimilarity measurements .....</b>	123
<b>Katarzyna Wójcik, Janusz Tuchowski: Machine translation impact on the results of the sentiment analysis .....</b>	134
<b>Małgorzata Misztal: Assessment of the influence of selected imputation methods on the results of object classification using classification trees ...</b>	145
<b>Anna Czapkiewicz, Beata Basiura: Simulation study of the selection of coefficient depending on the clustering time series.....</b>	153
<b>Tomasz Szubert: Factors differentiating the level of satisfaction with life and the life's values of people with and without disabilities in the light of the "Social Diagnosis" survey .....</b>	162
<b>Marcin Szymkowiak: Construction of calibration estimators of totals for different distance measures .....</b>	173

<b>Wojciech Roszka:</b> Joint characteristics' estimation of variables not jointly observed.....	181
<b>Justyna Brzezińska:</b> Visualizing categorical data in $\mathbf{R}$ .....	190
<b>Agata Sielska:</b> Regional diversity of competitiveness potential of Polish farms after the accession to the European Union .....	200
<b>Mariusz Kubus:</b> Regularized linear probability model as a filter .....	208
<b>Beata Basiura:</b> The Ward method in the application for classification of Polish voivodeships with different distances.....	216
<b>Katarzyna Wardzińska:</b> Application of Data Envelopment Analysis in company classification process.....	225
<b>Katarzyna Dębowska:</b> Modeling corporate bankruptcy based on unbalanced samples .....	234
<b>Danuta Tarka:</b> Influence of the features selection method on the results of objects classification using environmental data.....	245
<b>Artur Czech:</b> Application of chosen methods for the selection of diagnostic variables in indirect consumption research.....	254
<b>Beata Bal-Domańska:</b> Assessment of relations occurring between smart growth and economic cohesion in regional dimension using panel models	263
<b>Mariola Chrzanowska:</b> Ordinary kriging and inverse distance weighting as methods of estimating prices based on Warsaw real estate market .....	271
<b>Adam Depta:</b> Application of analysis of variance in the study of the quality of life based on questionnaire SF-36v2 .....	280
<b>Maciej Beręsewicz, Tomasz Klimanek:</b> Using indirect estimation with spatial autocorrelation in dwelling price surveys.....	290
<b>Karolina Paradysz:</b> Benchmark analysis of small area estimation on local labor markets .....	300
<b>Anna Gryko-Nikitin:</b> Selection of various parameters of parallel evolutionary algorithm for knapsack problems .....	310
<b>Tomasz Ząbkowski, Piotr Jałowiecki:</b> Application of association rules for the survey of data analysis in the selected areas of logistics in food processing companies .....	320
<b>Agnieszka Przedborska, Małgorzata Misztal:</b> Using multivariate statistical methods to assess the capacity of the knee joint among the patients treated surgically for osteoarthritis .....	330
<b>Dorota Perło:</b> Sustainable development in the economic, social and environmental dimensions – spatial analysis.....	341
<b>Ewa Putek-Szeląg, Urszula Gieraltowska:</b> Analysis and diagnosis of the volume of renewable energy production in Poland compared to EU countries .....	352

**Agnieszka Przedborska**

Uniwersytet Medyczny w Łodzi

**Małgorzata Misztal**

Uniwersytet Łódzki

---

## ZASTOSOWANIE METOD STATYSTYKI WIELOWYMIAROWEJ DO OCENY WYDOLNOŚCI STAWÓW KOLANOWYCH U PACJENTÓW Z CHOROBA ZWYRODNIENIOWĄ LECZONYCH OPERACYJNIE

---

**Streszczenie:** Choroba zwyrodnieniowa stawu kolanowego zajmuje trzecie miejsce na liście chorób zwyrodnieniowych za zmianami zwyrodnieniowymi kręgosłupa i biodra, obejmując około 5% populacji. Zasadniczym celem jej leczenia jest zahamowanie jej postępu, zniesienie dolegliwości bólowych, poprawa wydolności czynnościowej oraz jakości życia pacjentów. W pracy poddano analizie informacje dotyczące pacjentów z chorobą zwyrodnieniową stawów kolanowych leczonych operacyjnie osteotomią nadwładkową kości piszczelowej. Celem przeprowadzonych badań było wyodrębnienie grup pacjentów podobnych z punktu widzenia skuteczności i efektów klinicznych zastosowanej metody leczenia.

**Słowa kluczowe:** choroba zwyrodnieniowa stawu kolanowego, statystyczna analiza wielowymiarowa, analiza składowych głównych, analiza skupień.

### 1. Wstęp

Choroba zwyrodnieniowa stawu kolanowego zajmuje trzecie miejsce za zmianami zwyrodnieniowymi kręgosłupa i biodra, obejmując około 5% populacji [Kubacki i in. 2005]. Okresowy lub stały ból, sztywność poranna, upośledzenie funkcji i niekształcenia obrysu stawu rysują pełny obraz kliniczny gonartrozy. Proces chorobowy polega na przedwczesnym zużyciu i zwyrodnieniu tkanek tworzących staw. Destrukcja w większości przypadków rozpoczyna się od przedziału przysródkowego stawu kolanowego, co prowadzi do zniekształcenia osi kończyny.

Zasadniczym celem leczenia choroby zwyrodnieniowej stawu kolanowego jest zahamowanie jej postępu, zniesienie dolegliwości bólowych, poprawa wydolności czynnościowej oraz jakości życia pacjentów.

Osteotomia nadwładzłowa kości piszczelowej z korekcją osi kończyny jest jedną z metod leczenia operacyjnego w chorobie zwyrodnieniowej stawu kolanowego. Zabieg ten umożliwia równomierne rozłożenie sił nacisku na powierzchnie stawowe, odciążenie chorego przedziału stawu i zmniejszenie dolegliwości bólowych [Kubacki i in. 2005; Przedborska i in. 2010].

W pracy poddano analizie informacje dotyczące pacjentów z chorobą zwyrodnieniową stawów kolanowych leczonych operacyjnie osteotomią nadwładzłową kości piszczelowej. Celem przeprowadzonych badań było wyodrębnienie grup pacjentów podobnych z punktu widzenia skuteczności i efektów klinicznych zastosowanej metody leczenia.

## 2. Materiał i metody

Materiał kliniczny stanowiła grupa 24 pacjentów, u których w latach 1992-2006 w Klinice Chirurgii Urazowej i Ortopedii Szpitala Klinicznego im. WAM w Łodzi wykonano osteotomię nadwładzłową Maqueta kości piszczelowej z powodu choroby zwyrodnieniowej stawu kolanowego (zabieg jednostronny). Pacjenci ci w roku 2007 wzięli udział w badaniu kontrolnym<sup>1</sup>.

W celu kompleksowej oceny wpływu zabiegu na wydolność stawu kolanowego każdy pacjent został scharakteryzowany zestawem zmiennych obejmujących dane z dokumentacji medycznej i wywiadu chorobowego, wyniki badań ortopedycznych, ocenę radiologiczną zmian zwyrodnieniowych, wyniki badań siły napięcia izokinetycznego mięśni zginaczy i prostowników stawu kolanowego, ocenę wydolności pacjenta na podstawie dwóch skal czynnościowych oraz subiektywną ocenę wyniku leczenia.

Ze względu na małą liczbę pacjentów oraz występowanie dużej liczby silnie skorelowanych zmiennych w pierwszym kroku analizy dokonano redukcji wymiaru przestrzeni cech za pomocą analizy głównych składowych. Następnie zidentyfikowano grupy pacjentów podobnych, wykorzystując (1) hierarchiczne metody aglomeracyjne oraz (2) metodę Czekanowskiego. W kolejnym kroku określono charakterystyczne cechy pacjentów należących do poszczególnych wyróżnionych grup.

Obliczenia wykonano z wykorzystaniem pakietu STATISTICA 9.0, programu MacZek 3.0<sup>2</sup> oraz środowiska R (pakiet `clusterSim`).

---

<sup>1</sup> Zabieg osteotomii korekcyjnej wskazany jest u pacjentów z niezbyt zaawansowanymi zmianami zwyrodnieniowymi stawów kolanowych. W ciągu roku średnio wykonywane jest 8-10 zabiegów. Przedmiotem naszego zainteresowania były odległe wyniki (8-10-letni okres obserwacji). Wiek pacjentów oraz odległy czas po zabiegu, a co za tym idzie – wysoka śmiertelność pacjentów – były głównymi przyczynami trudności z zebraniem dużej grupy. Dodatkowo, część zoperowanych i zbadanych przez nas pacjentów została zdyskwalifikowana ze względu na fakt obustronnej operacji (lub obustronnych zmian zwyrodnieniowych). Ci pacjenci zostali pominięci w opracowaniu z powodu braku możliwości porównania kończyny chorej do zdrowej.

<sup>2</sup> Program autorstwa Piotra Jaskulskiego znajduje się na stronie: <http://eskimo73.republika.pl/maczek.html>.



### 3. Wyniki

Stan kończyny operowanej po zabiegu oceniono za pomocą zmiennych: różnica w długości kończyny operowanej i nieoperowanej (4 pomiary długości – względna, bezwzględna, udo, goleń – D1 – D4), różnica w wielkości obwodu kończyny – operowanej i nieoperowanej (6 pomiarów obwodu – pośladkowy, udowy I, udowy II, kolanowy, goleniowy I, goleniowy II – O1 – O6), zakres ruchu zgięcia stawu kolanowego (ZG), zakres ruchu wyprostowania stawu kolanowego (WY), siła mięśni prostowników i zginaczy według skali Lovetta (SP, SZ), siła napięcia izometrycznego mierzona aparatem F-30 dla prostowników, zginaczy części udowej, pasma biodrowo-piszczelowego i zginaczy części goleniowej (FP, FZ, FBP, FZG), wyniki skal czynnościowych Harrisa (H) i Merle D'Aubigne'a w modyfikacji Postela (M) oraz subiektywna ocena wyniku leczenia według skali Tappera (T).

Zmienne te są ze sobą silnie skorelowane (por. rys. 1).

Aby sprawdzić zasadność stosowania analizy głównych składowych, obliczono miary adekwatności doboru każdej indywidualnej zmiennej  $MSA_h$  [Walesiak, Gattar (red.) 2009, s. 326] oraz wartość współczynnika Kaisera – Meyera – Olkina (KMO).

Zmienna	D1	D2	D3	D4	O1	O2	O3	O4	O5	O6	ZG	WY	SP	SZ	FP	FZ	FBP	FZG	H	M	T
D1	1	0,78	0,61	0,65	0,09	0,10	0,23	0,00	0,23	0,14	0,07	0,17	0,36	0,35	0,21	0,32	0,43	0,18	0,06	0,01	-0,08
D2	0,78	1	0,73	0,82	-0,05	0,02	0,13	-0,28	-0,01	-0,01	0,38	0,37	0,53	0,32	0,36	0,42	0,07	0,32	0,32	0,31	0,24
D3	0,61	0,73	1	0,23	-0,12	-0,16	0,42	0,02	0,23	0,08	0,16	0,26	0,31	0,28	0,28	0,29	0,09	0,31	0,23	0,19	0,22
D4	0,65	0,82	0,23	1	0,02	0,17	-0,12	-0,38	-0,18	-0,04	0,42	0,29	0,55	0,29	0,34	0,40	0,06	0,28	0,24	0,26	0,12
O1	0,09	-0,05	-0,12	0,02	1	0,63	0,58	0,25	0,40	0,01	0,10	0,47	0,03	0,10	-0,06	0,03	-0,01	-0,25	-0,04	-0,01	-0,19
O2	0,10	0,02	-0,16	0,17	0,63	1	0,44	0,24	0,19	0,13	0,04	0,24	0,28	0,00	-0,15	0,09	0,04	-0,21	-0,13	-0,15	-0,32
O3	0,23	0,13	0,42	-0,12	0,58	0,44	1	0,60	0,74	0,36	0,05	0,33	0,04	0,06	0,05	-0,04	-0,02	-0,11	-0,13	-0,15	-0,19
O4	0,00	-0,28	0,02	-0,38	0,25	0,24	0,60	1	0,64	0,39	-0,34	-0,21	-0,40	-0,31	-0,30	-0,38	-0,09	-0,47	-0,52	-0,55	-0,48
O5	0,23	-0,01	0,23	-0,18	0,40	0,19	0,74	0,64	1	0,60	-0,16	-0,07	-0,22	-0,12	0,02	-0,06	0,09	-0,25	-0,36	-0,40	-0,38
O6	0,14	-0,01	0,08	-0,04	0,01	0,13	0,36	0,39	0,60	1	0,12	-0,24	-0,05	-0,09	0,13	-0,31	0,16	0,00	-0,20	-0,19	-0,31
ZG	0,07	0,38	0,16	0,42	0,10	0,04	0,05	-0,34	-0,16	0,12	1	0,48	0,50	0,52	0,58	0,14	0,03	0,43	0,56	0,65	0,39
WY	0,17	0,37	0,26	0,29	0,47	0,24	0,33	-0,21	-0,07	-0,24	0,48	1	0,48	0,41	0,33	0,31	-0,12	0,25	0,53	0,49	0,36
SP	0,36	0,53	0,31	0,55	0,03	0,28	0,04	-0,40	-0,22	-0,05	0,50	0,48	1	0,68	0,66	0,58	0,34	0,63	0,59	0,54	0,39
SZ	0,35	0,32	0,28	0,29	0,10	0,00	0,06	-0,31	-0,12	-0,09	0,52	0,41	0,68	1	0,65	0,45	0,49	0,49	0,68	0,65	0,51
FP	0,21	0,36	0,28	0,34	-0,06	-0,15	0,05	-0,30	0,02	0,13	0,58	0,33	0,66	0,65	1	0,61	0,48	0,72	0,57	0,51	0,46
FZ	0,32	0,42	0,29	0,40	0,03	0,09	-0,04	-0,38	-0,06	-0,31	0,14	0,31	0,58	0,45	0,61	1	0,27	0,44	0,39	0,26	0,34
FBP	0,43	0,07	0,09	0,06	-0,01	0,04	-0,02	-0,09	0,09	0,16	0,03	-0,12	0,34	0,49	0,48	0,27	1	0,37	0,11	0,03	-0,04
FZG	0,18	0,32	0,31	0,28	-0,25	-0,21	-0,11	-0,47	-0,25	0,00	0,43	0,25	0,63	0,49	0,72	0,44	0,37	1	0,58	0,51	0,53
H	0,06	0,32	0,23	0,24	-0,04	-0,13	-0,13	-0,52	-0,36	-0,20	0,56	0,53	0,59	0,68	0,57	0,39	0,11	0,58	1	0,96	0,92
M	0,01	0,31	0,19	0,26	-0,01	-0,15	-0,15	-0,55	-0,40	-0,19	0,65	0,49	0,54	0,65	0,51	0,26	0,03	0,51	0,96	1	0,88
T	-0,08	0,24	0,22	0,12	-0,19	-0,32	-0,19	-0,48	-0,38	-0,31	0,39	0,36	0,39	0,51	0,46	0,34	-0,04	0,53	0,92	0,88	1

Rys. 1. Macierz korelacji między badanymi zmiennymi

Źródło: opracowanie własne.

Na podstawie obliczonych wartości  $MSA_h$  z badania wyeliminowano 9 zmiennych (D1, D2, D3, D4, O2, O5, O6, FZ, FBP), dla których wartości miary  $MSA_h$  były mniejsze niż 0,5. Dla pozostałych 12 zmiennych wartości miary  $MSA_h$  przedstawia tab. 1.

**Tabela 1.** Miary adekwatności doboru badanych zmiennych

Zmienna	O1	O3	O4	ZG	WY	SP	SZ	FP	FZG	H	M	T
MSA	0,564	0,516	0,715	0,695	0,705	0,861	0,896	0,823	0,851	0,758	0,745	0,788

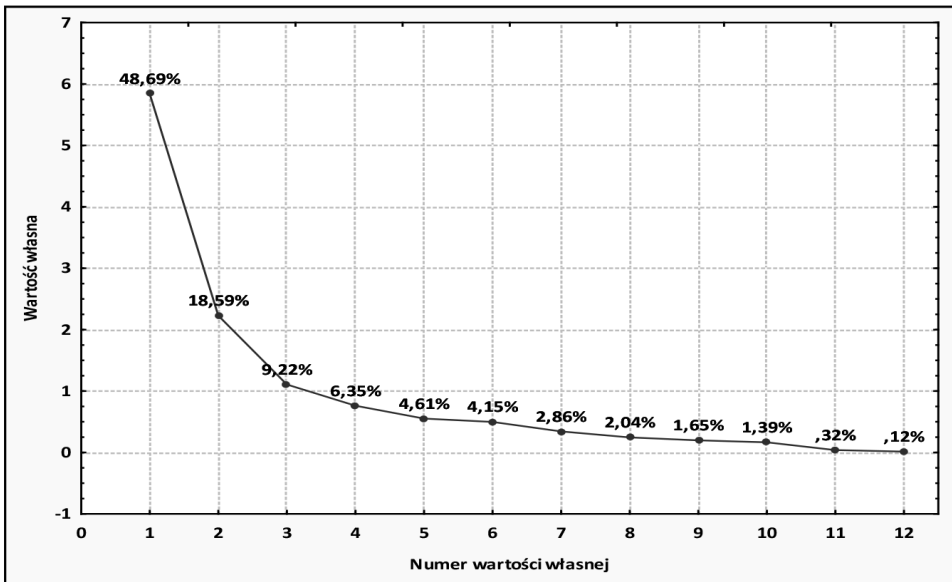
Źródło: opracowanie własne.

Zasadność zastosowania analizy składowych głównych potwierdza wartość współczynnika Kaisera – Meyera – Olkina:  $KMO = 0,7621$ . Obliczone wartości własne oraz procent wyjaśnianej przez nie wariancji przedstawiono w tab. 2.

**Tabela 2.** Wyniki analizy głównych składowych

Numer wartości	1	2	3	4	5	6	7	8	9	10	11	12
Wartość własna	5,84	2,23	1,11	0,76	0,55	0,5	0,34	0,25	0,2	0,17	0,04	0,01
Procent wyjaśnionej wariancji	48,69	18,59	9,22	6,35	4,61	4,15	2,86	2,04	1,65	1,39	0,32	0,12
Skumulowany procent wyjaśnionej wariancji	48,69	67,28	76,5	82,85	87,46	91,61	94,47	96,52	98,17	99,55	99,88	100

Źródło: opracowanie własne.

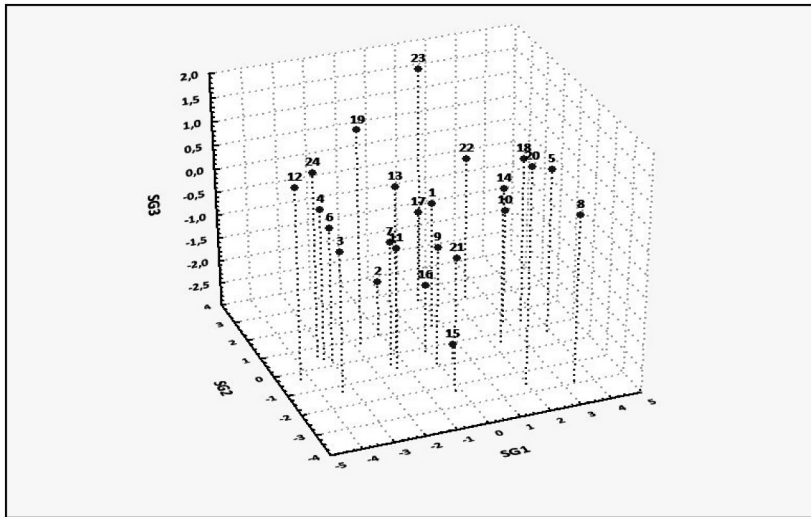
**Rys. 2.** Wykres osypiska

Źródło: opracowanie własne.

Trzy pierwsze wartości własne są większe od 1. Trzy pierwsze składowe główne wyjaśniają 76,5% całkowitej zmienności. Wybór do dalszej analizy trzech pierwszych głównych składowych potwierdza wykres osypiska (rys. 2).

Wykorzystując 3 pierwsze główne składowe, obliczono współrzędne czynnikowe przypadków<sup>3</sup>.

Wykres konfiguracji punktów reprezentujących pacjentów w trójwymiarowej przestrzeni głównych składowych przedstawia rys. 3. Można na nim zauważyć pewne zwarte skupienia punktów.



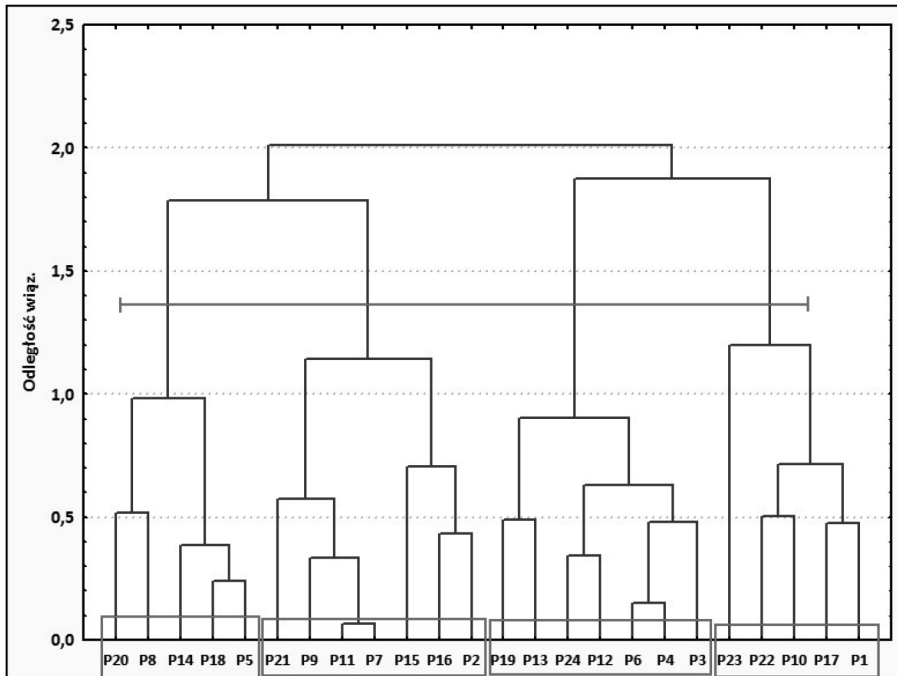
Rys. 3. Współrzędne czynnikowe przypadków w układzie trzech pierwszych głównych składowych

Źródło: opracowanie własne.

Obliczone współrzędne czynnikowe przypadków wykorzystano w analizie skupień do identyfikacji jednorodnych grup pacjentów.

Za pomocą pakietu `clusterSim` [Walesiak, Dudek 2012] wyznaczono optymalną procedurę analizy skupień dla badanego zbioru pacjentów. Zakładając minimalną liczbę skupień równą 2 i maksymalną liczbę skupień równą 4 oraz wykorzystując polecenie: `cluster.Sim(SG3, 2, 2, 4, "G1")`, jako optymalną wskazano procedurę wykorzystującą normalizację w przedziale  $[-1, 1]$ , odległość Euklidesa i metodę pełnego wiązania. Wyniki zastosowania tej procedury zaprezentowano na rys. 4. Wyodrębnione zostały 4 skupienia.

<sup>3</sup> Pominięto prezentację i interpretację uzyskanych ładunków czynnikowych.



**Rys. 4.** Wyniki grupowania pacjentów (odległość euklidesowa, metoda pełnego wiązania)

Źródło: opracowanie własne.

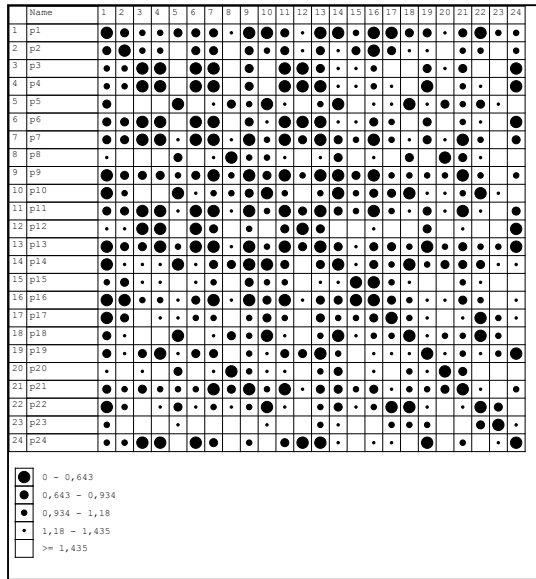
Dodatkowo do analizy wykorzystano także metodę Czekanowskiego (program MaCzek 3.0). Na rysunkach 5 i 6 przedstawiono nieuporządkowany i uporządkowany diagram Czekanowskiego.

Uzyskane wyniki nie są jednoznaczne. Można wyodrębnić większą liczbę skupień niż w przypadku zastosowania metod aglomeracyjnych. Elementy skupień uzyskanych w obu metodach częściowo się pokrywają.

Jako ostateczne przyjęto 4 skupienia otrzymane z wykorzystaniem hierarchicznych metod aglomeracyjnych. Charakterystykę pacjentów wchodzących w skład każdego skupienia przedstawiono w tab. 3.

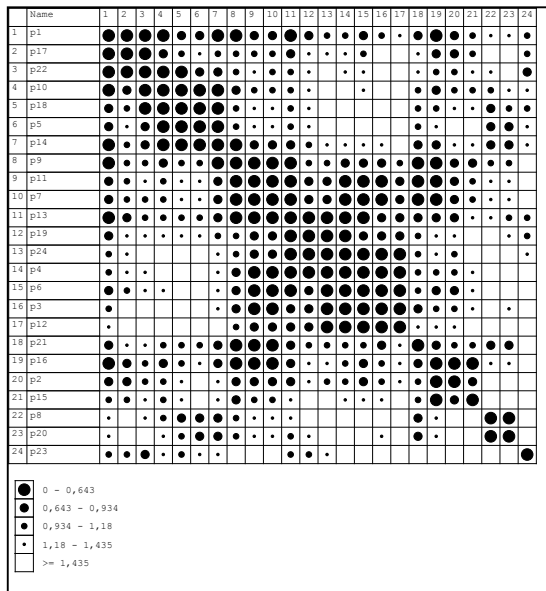
Analizując wyniki przedstawione w tab. 3, można zauważyć, że pacjenci skupienia 1 i skupienia 4 charakteryzują się gorszymi wynikami leczenia niż pacjenci skupienia 2 i skupienia 3. U pacjentów tych skupień odnotowano gorsze wyniki skal czynnościowych, występowanie bólu, ograniczenie sprawności i nieprawidłowy chód.

Gorszy wynik w tych grupach związany był również z techniką zaopatrzenia kończyny po zabiegu operacyjnym. Pacjenci ci oprócz metalowego implantu mieli kończynę unieruchomioną w opatrunku gipsowym, co niekorzystnie wpłynęło na jej funkcję.



Rys. 5. Nieuporządkowany diagram Czekanowskiego

Źródło: opracowanie własne.



Rys. 6. Uporządkowany diagram Czekanowskiego

Źródło: opracowanie własne.

Tabela 3. Charakterystyka wyodrębnionych grup pacjentów.

Cecha	Skupienie 1	Skupienie 2	Skupienie 3	Skupienie 4
Nr pacjentów	1, 10, 17, 22, 23	2, 7, 9, 11, 15, 16, 21	3, 4, 6, 12, 13, 19, 24	5, 8, 14, 18, 20
Płeć	wyłącznie kobiety	wyłącznie kobiety	3 kobiety, 4 mężczyźn	wyłącznie kobiety
Rodzaj zespolenia	metal + gips – 3 os., metal – 2 os.	metal + gips – 4 os., metal – 3 os.	metal + gips – 3 os., metal – 4 os.	metal + gips – 4 os., metal – 1 os.
Zmiany zwyrodnieniowe	wyłącznie III stopnia	II stopnia – 4 os., III stopnia – 3 os.	II stopnia – 4 os., III stopnia – 3 os.	I-II stopnia – 3 os., III stopnia – 2 os.
Występowanie bólu	ból wysiłkowy – 4 os., ból stały – 1 os.	ból wysiłkowy – 3 os., brak bólu – 4 os.	ból wysiłkowy – 2 os., brak bólu – 5 os.	ból wysiłkowy – 3 os., ból stały – 2 os.
Sprawność ogólna	ograniczona – 4 os., dobra – 1 os.	ograniczona – 3 os., dobra – 4 os.	ograniczona – 2 os., dobra – 5 os.	ograniczona – 5 os.
Chód	Nieprawidłowy – 4 os., prawidłowy – 1 os.	nieprawidłowy – 2 os., prawidłowy – 5 os.	nieprawidłowy – 2 os., prawidłowy – 5 os.	nieprawidłowy – 5 os.
Bolesność palpacyjna	nie występuje	nie występuje – 5 os., tak – 2 os.	nie występuje	nie występuje – 3 os., tak – 2 os.
Zakres ruchu zgięcia	70°-90° – 3 os., >=91° – 2 os.	<70° – 1 os., 70°-90° – 4 os., >=91° – 2 os.	70°-90° – 2 os., >=91° – 5 os.	<70° – 1 os., 70°-90° – 4 os.
Zakres ruchu wyprostu	bez ograniczeń	bez ograniczeń – 6 os., 6°-10° – 1 os.	bez ograniczeń	1°-5° – 1 os., 6°-10° – 4 os.
Wiek w chwili badania	61-70 lat – 1 os., 71-80 lat – 4 os.	61-70 lat – 1 os., 71-80 lat – 3 os., 81+ lat – 3 os.	61-70 lat – 5 os., 71-80 lat – 1 os., 81+ lat – 1 os.	do 60 lat – 1 os., 71-80 lat – 2 os., 81+ lat – 2 os.
Wiek w chwili zabiegu	do 60 lat – 1 os., 61-70 lat – 3 os., 71-80 lat – 1 os.	61-70 lat – 4 os., 71-80 lat – 3 os.	do 60 lat – 4 os., 61-70 lat – 2 os., 71-80 lat – 1 os.	do 60 lat – 1 os., 61-70 lat – 2 os., 71-80 lat – 2 os.
Średnie BMI	29,45±2,0	28,59±2,7	26,7±2,1	32,4±4,6
Średni czas od zabiegu	8,4±3,3 roku	10,1±3,2 roku	8,1±2,6 roku	9,0±1,0 roku
Ocena według skali Harrisa	zła – 2 os., dostateczna – 2 os., zadowolająca – 1 os.	zadowolająca – 2 os., dobra – 3 os., b. dobra – 2 os.	zadowolająca – 1 os., dobra – 1 os., b. dobra – 5 os.	zła – 3 os., dostateczna – 2 os.
Ocena według skali Merle d' Aubigne'a	dostateczna – 3 os., dobra – 2 os.	dobra – 2 os., b. dobra – 5 os.	dobra – 2 os., b. dobra – 5 os.	zła – 2 os., dostateczna – 2 os., dobra – 1 os.
Ocena według skali Tappera	zła – 2 os., dostateczna – 2 os., dobra – 1 os.	dobra – 2 os., b. dobra – 5 os.	dobra – 3 os., b. dobra – 4 os.	zła – 2 os.; dostateczna – 1 os.; zadowolająca – 1 os., dobra – 1 os.

Źródło: opracowanie własne.

Najlepsze wyniki leczenia zaobserwowano u pacjentów skupienia nr 3. Były to osoby operowane w młodszym wieku i o niższych wartościach wskaźnika BMI, a także o niższym stopniu nasilenia zmian zwyrodnieniowych.

Ponadto w skupieniu tym większość pacjentów po zabiegu operacyjnym miała ustabilizowane odłamy tylko metalem (bez opatrunku gipsowego), co umożliwiło wczesne podjęcie rehabilitacji i szybszy powrót do zdrowia.

Obserwacja ta prowadzi do wniosku, że zastosowanie opatrunku gipsowego ogranicza możliwość wczesnego podjęcia usprawniania, przez co wydłuża czas leczenia i związane jest z większymi jego kosztami. Dłuższe unieruchomienie wiąże się również z gorszą funkcją kończyny i gorszym wynikiem odległym.

#### 4. Podsumowanie

Głównym celem leczenia choroby zwyrodnieniowej stawu kolanowego jest zahamowanie jej postępu i zniesienie dolegliwości bólowych, a co za tym idzie – poprawa wydolności czynnościowej oraz jakości życia pacjentów.

W pracy zaproponowano wykorzystanie wybranych metod statystyki wielowymiarowej do oceny odległych wyników leczenia operacyjnego z wykorzystaniem metody osteotomii nadwięzadłowej Maqueta kości piszczelowej. Metody statystycznej analizy wielowymiarowej nie były dotychczas wykorzystywane przy analizie wyników klinicznych tego typu zabiegów w ortopedii.

Zasadniczym celem analiz było wyodrębnienie grup pacjentów podobnych z punktu widzenia skuteczności i efektów klinicznych zastosowanej metody leczenia. Chociaż badana grupa pacjentów nie była losowa i nie można uzyskanych wyników uogólnić na całą populację pacjentów, pewne sugestie z przeprowadzonych analiz można traktować jako wyniki badania wstępnego, pozwalające na sformułowanie konkretnych hipotez badawczych, których weryfikacja powinna się odbyć z wykorzystaniem próby losowej o odpowiedniej liczebności.

Profilowanie skupień pozwoliło wskazać grupy pacjentów z najgorszymi wynikami zabiegu. Pacjentów tych charakteryzuje przede wszystkim zaawansowany wiek w chwili zabiegu, otyłość (BMI powyżej 30) oraz technika zaopatrzenia kończyny po zabiegu (zespolenie łącznikiem metalowym i unieruchomienie gipsowe).

Uwzględnienie takiej wiedzy w praktyce lekarskiej pozwolić może na bardziej wnikliwe dobieranie techniki zabiegu do pacjenta. Może to także być sugestią do weryfikacji listy wskazań do tego typu zabiegu, bowiem brak odpowiedniego podejścia do pacjenta wpływa na opóźnienie rozpoczęcia rehabilitacji, wydłuża czas i koszt leczenia.

#### Literatura

- Kubacki J., Saulicz E., Gieremek K., *Możliwości terapii chirurgicznej we wczesnych i późnych okresach zmian zwyrodnieniowych stawu kolanowego (gonartroza)*, „Fizjoterapia Polska” 2005, vol. 3, no. 5, s. 359-363.
- Przedborska A., Waszczykowski M., Pruszyńska M., *Analiza wydolności czynnościowej stawów kolanowych u pacjentów z chorobą zwyrodnieniową leczonych operacyjnie osteotomią nadwięzadłową kości piszczelowej*, „Kwartalnik Ortopedyczny” 2010, vol. 4, s. 550-559.

- Walesiak M., Gatnar E. (red.), *Statystyczna analiza danych z wykorzystaniem programu R*, PWN, Warszawa 2009.
- Walesiak M., *Uogólniona miara odległości GDM w statystycznej analizie wielowymiarowej z wykorzystaniem programu R*, Wydawnictwo UE we Wrocławiu, Wrocław 2011.
- Walesiak M., Dudek A., *clusterSim* package, 2012, URL <http://www.R-project.org>.

## **USING MULTIVARIATE STATISTICAL METHODS TO ASSESS THE CAPACITY OF THE KNEE JOINT AMONG THE PATIENTS TREATED SURGICALLY FOR OSTEOARTHRITIS**

**Summary:** Osteoarthritis of the knee ranks third, after the degenerative changes of the spine and hips and affects 5% of the population. The main objective of the osteoarthritis in the knee treatment is to prevent the disease progression, to alleviate pain and to improve the quality of life for patients and their functional capacity. Some information of patients treated surgically for osteoarthritis in the knee with high tibial osteotomy were analyzed in the paper. The goal of the study was to identify groups of patients similar in terms of efficacy and clinical effects of this method of treatment.

**Keywords:** osteoarthritis of the knee, multivariate statistical analysis, principal components analysis, cluster analysis.