Warszawa, 25-26.09.2025

Nowoczesne metody statystyczne w badaniach medycznych

Polska Grupa Narodowa Międzynarodowego Towarzystwa Biostatystyki Klinicznej
Zakład Metod Statystycznych i Analiz Biznesowych Instytutu Statystyki i Demografii

PROGRAM I ABSTRAKTY

Redakcja książeczki abstraktów i projekt okładki:

dr Izabela Miechowicz (Uniwersytet Medyczny im. Karola Marcinkowskiego w Poznaniu)
dr inż. Anna Sowińska (Uniwersytet Medyczny im. Karola Marcinkowskiego w Poznaniu)

Fotografia na okładce:

Fragment Placu Zamkowego z kolumną Zygmunta w Warszawie – fot. Rudy Balasko

Wydawca:

Katedra i Zakład Informatyki i Statystyki Uniwersytetu Medycznego w Poznaniu

ISBN 978-83-962849-1-4

Warszawa 2025

Książka abstraktów zamieszczona na stronie: https://iscb.pl/o-konferencji-2/

Spis treści:

Organizatorzy

Zakład Metod Statystycznych i Analiz Biznesowych Instytutu Statystyki i Demografii Szkoły Głównej Handlowej w Warszawie (SGH)

oraz

Polska Grupa Narodowa Międzynarodowego Towarzystwa Biostatystyki Klinicznej, która zrzesza członków International Society for Clinical Biostatistics (ISCB), którzy pracują lub uczą się na uczelniach, w instytutach naukowych, firmach farmaceutycznych oraz innych instytucjach mających swoją siedzibę w Polsce. ISCB zostało założone w 1978 roku w celu stymulowania badań nad regułami i metodologią stosowaną w projektowaniu i analizie badań klinicznych, w celu zwiększenia znaczenia teorii statystycznej i wnioskowania statystycznego w dziedzinie medycyny klinicznej oraz promowania lepszego zrozumienia i interpretacji biostatystyki przez profesjonalistów innych dziedzin niż statystyka oraz ogół społeczeństwa. Towarzystwo jest otwarte dla wszystkich osób zainteresowanych rozwojem i zastosowaniem metod biostatystycznych w szeroko pojętych badaniach medycznych. Członkami ISCB są klinicyści, statystycy oraz naukowcy i praktycy innych dyscyplin: epidemiologii, farmakologii klinicznej, chemii klinicznej. (http://www.iscb.pl).

Członkowie ISCB spotykają się na organizowanych corocznie konferencjach naukowych.

Komitet Organizacyjny Konferencji:

Przewodniczący:

- dr hab. Wioletta Grzenda, prof. SGH
- mgr Kinga Sałapa (PGN ISCB)

Członkowie:

- dr Izabela Miechowicz
- dr Anna Sowińska
- dr Justyna Marcinkowska

Komitet Naukowy Konferencji:

Członkowie:

- dr hab. Katarzyna Buszko, prof. UMK
- dr hab. Agnieszka Chłoń-Domińczak, prof. SGH
- dr hab. Beata Czarnacka-Chrobot, prof. SGH
- dr hab. Wioletta Grzenda, prof. SGH
- dr hab. Michał Michalak (UMP)
- dr hab. Barbara Więckowska (UMP)
- dr hab. Elżbieta Chełmecka (SUM)

Program konferencji

Dzień 1: czwartek, 25 września

Otwarcie konferencji 11:00-11:10

1	dr hab. Andrzej Zawistowski, prof. SGH, prorektor ds. dydaktyki i studentów	
2	mgr Kinga Sałapa, Prezes zarządu ISCBPL	
3	dr hab. Wioletta Grzenda, prof. SGH	

Sesja 1, 11:10-12:50, Przewodniczący sesji: dr hab. Wioleta Grzenda, prof. SGH

1	dr Dawid Majcherek	Różnice w dostępie do opieki zdrowotnej w Europie w kontekście cyfryzacji
2	dr hab. Elżbieta Chełmecka	Trzeźwy czy nietrzeźwy, oto jest pytanie. Statystyka w toksykologii post mortem
3	dr Tomasz Burzykowski	Użycie informacji o wczesnych wynikach leczenia w analizach przejściowych w próbach klinicznych z długoterminowym kryterium oceny skuteczności leczenia
4	dr Artur Czech	Opieka zdrowotna w Polsce – taksonomiczne ujęcie regionalne i czasowe

Sesja 2, 13:10-14:25, przewodniczący sesji: mgr Kinga Sałapa

1	dr hab. Barbara Więckowska	Dekompozycja testu ilorazu wiarygodności na składowe dla klas zdarzeń i braku zdarzeń: Wizualna prezentacja metodą U-smile
2	lekarz Katarzyna B. Kubiak	Czy kryterium informacyjne Akaikego może się uśmiechnąć?
3	dr hab. Katarzyna Buszko	Zastosowanie entropii do oceny złożoności sygnałów biomedycznych rejestrowanych w teście pochyleniowym u pacjentów z omdleniami wazowagalnymi

Program konferencji

Dzień 2: piątek, 26 września

Sesja 3, 10:00-11:40, przewodniczący sesji: dr hab. Barbara Więckowska

1	dr Justyna Marcinkowska	Interdyscyplinarne kształcenie studentów medycyny z wykorzystaniem interaktywnych raportów – nowe wyzwania dydaktyczne
2	dr hab. Michał Michalak	Interpretowalność modeli uczenia maszynowego w diagnostyce migotania przedsionków: rola wykresów SHAP w selekcji cech i odkrywaniu nieliniowych relacji
3	mgr Przemysław Talar	Zastosowanie różnych metod statystycznych do oceny skuteczności programu KOS-Zawał
4	mgr Jan Jurkowski	Ocena użyteczności wstępnej selekcji genów do sieci neuronowych dla wielowymiarowych danych genomicznych

Sesja 4, 12:00-13:15, przewodniczący sesji: dr hab. Elżbieta Chełmecka

1	dr Mateusz Lejawa	Ocena wpływu telomerów na zwapnienie tętnic wieńcowych: Randomizacja Mendlowska
2	dr Adam Korczyński	Validating the safety profile thought laboratory data in clinical trial
3	mgr Dominik Nowakowski	Missing value knockoffs: a methodology for controlled variable selection in omics datasets with missing data

Plakaty

1	dr hab. Agnieszka Micek	Antocyjany a zdrowie mózgu
2	dr hab. Agnieszka Micek	Zastosowanie meta-analizy do syntezy danych weryfikujących związek niedożywienia z objawami depresji u pacjentów onkologicznych

Abstrakty referatów

Ocena użyteczności wstępnej selekcji genów do sieci neuronowych dla wielowymiarowych danych genomicznych

Jan Jurkowski [1] & Małgorzata Ćwiklińska-Jurkowska [2]

[1] Inkubator Uniwersytetu Warszawskiego,
[2] Katedra Biostatystyki, Uniwersytet Mikołaja Kopernika, Toruń

Analiza danych mikromacierzowych stanowi szczególne wyzwanie z powodu skrajnej dysproporcji między liczbą zmiennych (genów) a liczbą obserwacji klinicznych, zjawisko niekiedy zwane „przekleństwem wymiarowości".
Celem pracy było porównanie trzech podejść do redukcji wymiarów: klasycznej analizy głównych składowych (PCA), metody wyboru genów za pomocą metody PAMcv (cross-validate Prediction Analysis for Microarrays) oraz redukcji niebezpośredniej (poprzez usuwanie parametrów modelu), opartej na regularyzacji końcowego klasyfikatora.
Materiał i metody

Badanie przeprowadzono na danych genomicznych z dychotomicznym wektorem klasyfikacyjnym, dotyczącym rozpoznawania nowotworu białaczki. Zastosowane dane mikromacierzowe zawierają wyniki badania wielkiej liczby ekspresji genów (3571 genów) dla 72 pacjentów, z proporcjami przynależności do 2 grup 65.3% i 34.7%). Tak wysoki wymiar wymagał zastosowania procedur redukcji wymiarów oraz odpowiedniego sposobu oceny jakości klasyfikacji. W każdej konfiguracji ostateczny model klasyfikacyjny opierał się sieciach neuronowych, których skuteczność oceniano za pomocą bootstrapu, aby ograniczyć ryzyko przeuczenia wynikające z niewielkiej liczby obserwacji. Sprawdzano różne ilości genów, od 20 do tysiąca.

Spośród badanych sieci neuronowych sieci głębokie wykazywały większe przeuczenie (ang. overfitting), lepszy wynik uzyskano dla jednokierunkowej sieci neuronowej z jedną warstwą ukrytą,
Największe różnice między wstępną redukcją wymiaru za pomocą PAMcv i PCA na korzyść PCAcv wykazano dla 20 genów. Dla większej ilości genów błąd generalizacji był zaniedbywalnie mniejszy, wówczas również różnica błędu między PAM cv i PCA stawała się zaniedbywalna.
SVM (Support Vector Machines) dało wyniki porównywalne do sieci neuronowych.
Wnioski
Uzyskane wyniki wskazują, że metoda PAM, mimo iż generowała wyższy błąd uczący w porównaniu z PCA czy regularyzacją samej sieci neuronowej, zapewniała niższy błąd

bootstrapowy. Sugeruje to większą zdolność do generalizacji i unikania przeuczenia w kontekście danych mikromacierzowych o wysokiej wymiarowości. PCA okazała się skuteczna w redukcji liczby cech, lecz nie zabezpieczała dostatecznie przed nadmiernym dopasowaniem, natomiast regularyzacja wag w sieci miała ograniczony efekt przy tak dużym stosunku liczby genów do liczby próbek.
PCA wykazywało większy overfitting. Ta metoda maksymalizuje wariancję w danych, co nie musi skutkować dobrymi predykcjami.
Model sieci neuronowej poprzedzony wstępną redukcją genów metodą PAMcv charakteryzuje się zarówno mniejszym przeuczeniem, jak i większą wyjaśnialnością (ang. explainability), niż model redukujący wymiar metodą PCA- ze względu na oparcie na kilkudziesięcioleciu konkretnych genach a nie ich linowej kombinacji.

Słowa kluczowe: sieci neuronowe, wstępna redukcja wymiaru, ekspresja genów, genomiczne dane o nowotworach, wielowymiarowa klasyfikacja

Interdyscyplinarne kształcenie studentów medycyny z wykorzystaniem interaktywnych raportów – nowe wyzwania dydaktyczne

Justyna Marcinkowska

Katedra i Zakład Informatyki i Statystyki, Uniwersytet Medyczny im Karola Marcinkowskiego w Poznaniu, Polska

Celem wystąpienia jest zaprezentowanie roli interdyscyplinarnego podejścia w kształceniu studentów medycyny oraz omówienie potencjału interaktywnych narzędzi raportowania (na przykładzie RShiny, Power BI i innych środowisk open source) w procesie dydaktycznym. Wystąpienie ma także charakter otwarcia dyskusji nad nowymi wyzwaniami edukacyjnymi, jakie niesie łączenie nauk ścisłych z naukami medycznymi w kontekście nauczania analizy danych.

Przedstawiono doświadczenia dydaktyczne związane z wykorzystaniem języka R i pakietu RShiny, wskazując ich zalety w kontekście otwartego dostępu i interaktywności. Jako przykład zaproponowano scenariusz zajęć oparty na analizie danych klinicznych (np. parametry laboratoryjne, wyniki badań obrazowych, dane epidemiologiczne), które mogą być wizualizowane w formie interaktywnych dashboardów. Rozważania obejmują również możliwości zastosowania alternatywnych narzędzi, takich jak Power BI czy Tableau, w tworzeniu dynamicznych raportów służących dydaktyce i analizie medycznej.

Zastosowanie narzędzi interaktywnych w dydaktyce medycznej sprzyja rozwijaniu praktycznych kompetencji analitycznych i cyfrowych u studentów, pozwala na lepsze zrozumienie złożonych procesów klinicznych oraz zwiększa zaangażowanie w proces uczenia się. Powiązanie nauk ścisłych (statystyka, informatyka, analiza danych) z naukami medycznymi otwiera nowe możliwości w edukacji, czyniąc ją bardziej nowoczesną i interdyscyplinarną.

Wystąpienie ma charakter zaproszenia do dyskusji nad przyszłością dydaktyki akademickiej na kierunkach medycznych. Podkreślono konieczność rozszerzenia oferty edukacyjnej o kursy z analizy danych w środowisku open source oraz potrzebę interdyscyplinarnej współpracy między statystykami, informatykami a klinicystami. Narzędzia takie jak RShiny czy Power BI mogą stać się ważnym elementem dydaktyki na uniwersytetach medycznych, odpowiadając na rosnące zapotrzebowanie studentów na praktyczne umiejętności w zakresie analizy i interpretacji danych klinicznych.

Słowa kluczowe: RShiny, Power BI, analiza danych klinicznych, wizualizacja, raport interaktywny, dydaktyka medyczna, interdyscyplinarność, nowe wyzwania

Różnice w dostępie do opieki zdrowotnej w Europie w kontekście cyfryzacji

Dawid Majcherek[1] & Scott William Hegerty[2] & Arkadiusz Michał Kowalski[1] & Małgorzata Stefania Lewandowska[1] & Desislava Dikova[3]

[1]Szkoła Głowna Handlowa, Warszawa, Polska
[2]Northeastern Illinois University in Chicago, USA,
[3]Vienna University of Economics and Business, Austria

Rozwój technologii cyfrowych w ochronie zdrowia otwiera nowe możliwości ograniczania różnic w dostępie do usług medycznych na terenie Unii Europejskiej. Pandemia COVID-19 unaoczniła, że narzędzia cyfrowe mogą skutecznie wspierać systemy zdrowotne w sytuacjach kryzysowych i poprawiać ich wydajność. Wykorzystując to doświadczenie, przygotowaliśmy zbiór danych obejmujący 27 państw członkowskich, oparty na European Health Interview Survey (EHIS), wskaźniku DESI oraz innych źródłach Eurostatu. Dzięki zastosowaniu metody k-średnich przeprowadziliśmy analizę skupień, która pozwoliła na pogrupowanie krajów według dwóch wymiarów: poziomu cyfryzacji i skali nierówności w dostępie do opieki zdrowotnej.

Badanie ujawniło pięć wyodrębnionych klastrów – dwa z wysokim, dwa z umiarkowanym i jeden z niskim poziomem niezaspokojonych potrzeb zdrowotnych. Najwyższy stopień przygotowania cyfrowego odnotowano jedynie w krajach nordyckich, Hiszpanii i na Cyprze. Część państw Europy Zachodniej, pomimo dobrze rozwiniętych systemów opieki zdrowotnej, znalazła się w grupie o przeciętnym poziomie cyfryzacji i umiarkowanych barierach w dostępie do usług. Większość krajów UE nadal wymaga znaczących inwestycji w infrastrukturę cyfrową, aby zwiększyć wykorzystanie rozwiązań e-zdrowia przez pacjentów i pracowników ochrony zdrowia.

Nasze wnioski wskazują, że aby zapewnić bardziej sprawiedliwy dostęp do świadczeń medycznych, konieczne jest traktowanie innowacji cyfrowych jako kluczowego elementu współczesnych systemów zdrowotnych.

Użycie informacji o wczesnych wynikach leczenia w analizach przejściowych w próbach klinicznych z długoterminowym kryterium oceny skuteczności leczenia

Tomasz Burzykowski[1,2,3]

[1]International Drug Development Institute, Louvain la Neuve, Belgia
[2]Hasselt University, Hasselt, Belgia
[3]Uniwersytet Medyczny w Białymstoku, Białystok, Polska

W próbach klinicznych z długoterminowym kryterium oceny skuteczności leczenia T (np. czasem przeżycia), niezbędny w celu oceny T okres obserwacji chorych może być długi. Ogranicza to możliwość przeprowadzenia wczesnych analiz przejściowych, które mogłyby skrócić czas trwania próby w przypadku stwierdzenia pozytywnych efektów leczenia na wczesnym etapie próby. Atrakcyjnym rozwiązaniem w takiej sytuacji może być użycie analiz przejściowych wykorzystujących informację o wczesnym wyniku leczenia S (np. wartości biomarkera). Garcia Barrado i Burzykowski (2024) opracowali metodologię, która pozwala na uwzględnienie tego typu analiz przejściowych dla S i T dowolnego typu (ciągłych, binarnych, itp.). Okazuje się, że dzięki temu można znacząco skrócić oczekiwany czas trwania i/lub liczebność próby. Możliwe jest to przede wszystkim w przypadku, gdy efekt leczenia dla S jest mocno skorelowany z efektem leczenia dla T, tzn. jeśli S jest dobrym zastępczym kryterium oceny skuteczności leczenia („surogatem") dla T.

Metodologia zaproponowana przez Garcia Barrado i Burzykowskiego (2024) zakłada, że współczynniki modelu definiującego własności S jako zastępczego kryterium oceny skuteczności leczenia są znane. W praktyce, współczynniki te muszą być szacowane na podstawie danych z prób klinicznych. W niniejszej pracy przedstawiona jest modyfikacja wspomnianej wyżej metodologii uwzględniająca użycie oszacowań współczynników modelu.

Materiał i metody:

By uwzględnić użycie oszacowań współczynników modelu, wariancja statystki testowej używanej w analizie przejściowej wymaga zwiększenia poprzez dodanie wyrazu będącego funkcją macierzy wariancji kowariancji oszacowanych współczynników. W pracy przedstawiona jest postać tego wyrazu. Wynik zastosowania poprawki zilustrowany hest przy pomocy przykładu rzeczywistej próby klinicznej w onkologii.

Wyniki:
Zgodnie z oczekiwaniem, uwzględnienie błędu szacowania redukuje potencjalne zyski dotyczące oczekiwanego czas trwania i/lub liczebności próby. Jednak w przypadku, gdy S jest dobrym surogatem dla T, redukcja może być niewielka.

Wnioski:
Przedstawiona metoda pozwala na realną ocenę zysku z włączenie analiz przejściowych wykorzystujących informację o wczesnym wyniku leczenia w proces planowania prób klinicznych z długoterminowym kryterium oceny skuteczności leczenia.

Bibliografia:
Garcia Barrado L, and Burzykowski T (2024). Using an early outcome as a sole source of information of interim decisions regarding treatment effect on a long-term endpoint: the non-Gaussian case. Pharmaceutical Statistics, 23(6), 928–938.

Słowa kluczowe: wczesne wyniki leczenia; długoterminowe kryterium oceny skuteczności leczenia; analizy przejściowe; randomizowana próba kliniczna

Opieka zdrowotna w Polsce – taksonomiczne ujęcie regionalne i czasowe

Prof. Teresa Słaby & dr Artur Czech

Szkoła Głowna Handlowa w Warszawie, Polska

Zdrowie jest postrzegane jako jeden z podstawowych czynników wpływających na jakość życia ludzi. Prawie każde działanie człowieka można uznać za potencjalnie wpływające na ochronę własnego zdrowia. Przykładowo, zmiana klimatu powoduje wzrost poziomu odpowiedzialności za własne zdrowie, co z kolei powoduje większe zainteresowanie stopniem zaspokojenia potrzeb zdrowotnych. Ocena tego zaspokojenia staje się trudna do kwantyfikacji. W związku z czym przyjęto, iż pojęcie ochrony zdrowia wiązane jest z całą populacją ludności obejmującą zarówno ludzi chorych jak i zdrowych oraz uwzględnia system opieki zdrowotnej. Istniejący w Polsce system opieki zdrowotnej nieustanne zmaga się z niedoborem środków finansowych, które w około 72% pochodzą ze środków publicznych (wydatki budżetu państwa oraz jednostek samorządu terytorialnego). Ponadto, obszar Polski jest zróżnicowany pod względem rozwoju społeczno-gospodarczego i zanieczyszczenia środowiska, co również wpływa na poziom opieki zdrowotnej w poszczególnych regionach. W związku z czym diagnoza poziomu opieki zdrowotnej jest kluczowa w kontekście przyjęcia jasnej i jednoznacznej polityki zdrowotnej, która byłaby skuteczna pod względem racjonalnego wykorzystania ograniczonych środków finansowych.

Celem pracy jest próba oceny stanu opieki zdrowotnej w poszczególnych województwach jako podstawy w procesie poprawy poziomu opieki zdrowotnej w poszczególnych regionach Polski w warunkach ograniczenia środków finansowych.

W przeprowadzonym badaniu wykorzystano pozycyjną metodę taksonomiczną z mediną Webera. Analizę oparto o dane zaczerpnięte z Banku Danych Lokalnych Głównego Urzędu Statystycznego w Polsce. W wyniku przeprowadzonej analizy otrzymano uporządkowanie województw w postaci ich rankingów pod względem stanu opieki zdrowotnej. Uwzględniono ponadto w diagnozie wpływ upływu czasu. Zastosowanie taksonomicznej metody pozycyjnej pozwoliło na uodpornianie prowadzonej analizy na wpływ asymetrii w rozkładach empirycznych wybranych obiektywnych cech diagnostycznych. Ponadto, zastosowanie wielowymiarowej mediany Webera pozwoliło na uwzględnienie w przeprowadzonym badaniu wzajemnych i bezpośrednio nieobserwowalnych związków w zbiorze cech diagnostycznych, co jest niezwykle istotne w tego typu badaniach.

Słowa kluczowe: ochrona zdrowia, opieka zdrowotna, analiza taksonomiczna, miernik syntetyczny, województwo.

missKnockoffs: a methodology for controlled variable selection in omics datasets with missing data

Dominik Nowakowski[1] & Julie Josse[2] & Szymon Majewski[3] & Asaf Weinstein[4], Małgorzata Bogdan[5]

[1]Medical University of Bialystok, Department of Biostatistics and Medical Informatics, Poland
[2] Inria-Inserm Centre in Montpellier, PreMeDICaL team, France
[3] University of Warsaw, Faculty of Mathematics, Informatics and Mechanics, Poland
[4]Hebrew University of Jerusalem, Department of Statistics and Data Science, Israel
[5]Lund University, Department of Statistics, Sweden and University of Wroclaw, Department of Mathematics, Poland

Over the past two decades, the problem of model selection in high-dimensional data has gained considerable attention in statistics and machine learning. Despite significant progress and the development of numerous methodologies, only a limited number of approaches adequately address the challenge of missing data—an issue that frequently arises in real-world datasets, particularly in omics research. To bridge this methodological gap, we introduce a novel statistical approach, missKnockoffs, which extends the Model-X knockoff framework to accommodate incomplete data.

The proposed method operates in two stages: the first involves imputing missing values using a variety of imputation strategies, while the second generates synthetic knockoff variables to control the false discovery rate (FDR). To mitigate the randomness associated with the knockoff generation process, we employ multiple knockoff copies in conjunction with aggregation techniques.

We further propose an innovative aggregation scheme for knockoff statistics that exhibits favorable theoretical properties. The performance of missKnockoffs was evaluated through extensive simulation studies, with statistical power and FDR control serving as the primary evaluation metrics. Additionally, the method's effectiveness was validated on real-world omics datasets.

The proposed method demonstrates higher statistical power compared to existing approaches and is among the few techniques specifically designed to identify relevant variables in the presence of missing data. It is also applicable to the analysis of high-dimensional data, which is particularly important in biomedical research.

Slowa kluczowe: Incomplete data, FDR control, multiple knockoffs, multiple imputation, omic data

Zastosowanie różnych metod statystycznych do oceny skuteczności programu KOS-Zawał

Martyna Kurzak, Przemysław Talar, Maciej Polak, Agnieszka Doryńska

Dział Przetwarzania Danych i Analiz, Agencja Oceny Technologii Medycznych i Taryfikacji w Warszawie, Polska

Głównym problemem badawczym jest ocena przeżycia pacjentów po zawale serca z wykorzystaniem różnych metod statystycznych w celu określenia skuteczności programu KOS-Zawał. Analiza ma na celu nie tylko stwierdzenie czy udział w programie wiąże się z istotną poprawą rokowań pacjentów, ale również porównanie uzyskiwanych wyników w zależności od zastosowanej metody.

Próbę badawczą stanowili pełnoletni pacjenci po zawale serca, hospitalizowani w latach 2018–2020 na oddziałach zajmujących się leczeniem chorób sercowo-naczyniowych w Polsce. Uczestnicy zostali zakwalifikowani do jednej z dwóch grup – objętych programem kompleksowej opieki specjalistycznej po zawale serca KOS-Zawał lub pozostających poza programem. Każdego pacjenta objęto roczną obserwacją w zakresie przeżycia, liczoną od momentu zakończenia hospitalizacji. W celu oceny przeżywalności zastosowano różne metody analizy przeżycia, w tym jednowymiarowy i wielowymiarowy model proporcjonalnego hazardu Coxa, a także techniki statystyczne ograniczające wpływ czynników zakłócających, takie jak dopasowanie na podstawie propensity score matching (PSM) oraz inverse probability weighting (IPW). W metodzie PSM zastosowano metodę najbliższego sąsiada.

Do analizy włączono łącznie 189 812 pacjentów (z czego 157 889 nie było objętych programem KOS-Zawał). Średni wiek badanych wynosił 67,5 lat, a odsetek mężczyzn wyniósł 64,4%. W trakcie rocznej obserwacji odnotowano 10,3% zgonów (z czego 9,5% dotyczyło pacjentów nieobjętych programem). Badani nieobjęci programem byli starsi (średnia wieku 68,02) i bardziej obciążeni chorobowo. Uzyskane wyniki dotyczące analizy przeżycia wskazują, że w przypadku jednowymiarowego modelu proporcjonalnego hazardu Coxa udział w programie wiązał się z 58,4% redukcją ryzyka zgonu w ciągu roku (HR = 0,416; 95% CI: 0,395-0,438). W modelu wielowymiarowym zaobserwowano redukcję siły efektu programu KOS-Zawał (HR = 0,594; 95% CI: 0,564-0,626). Po zastosowaniu metody PSM (dopasowanie 1:2) HR wyniósł 0,628 (95% CI: 0,594-0,665), natomiast przy wykorzystaniu metody IPW – 0,582 (95% CI: 0,547-0,619).

Wyniki analiz różniły się w zależności od zastosowanej metody oceny przeżycia, co podkreśla znaczenie świadomego doboru technik statystycznych oraz ich poprawnego i rzetelnego opisu.

Niezależnie jednak od przyjętej metody, uczestnictwo w programie KOS-Zawał wiązało się z istotnym zmniejszeniem ryzyka zgonu w porównaniu z pacjentami nieobjętymi programem.

Słowa kluczowe: analiza przeżycia, model proporcjonalego hazardu Coxa, propensity score matching, inverse probability weighting

Ocena wpływu telomerów na zwapnienie tętnic wieńcowych: Randomizacja Mendlowska

Mateusz Lejawa, Karolina Lejawa

Śląski Uniwersytet Medyczny w Katowicach, Polska

Zwapnienie tętnic wieńcowych (CAC) jest kluczowym biomarkerem miażdżycy i oceny ryzyka chorób sercowo-naczyniowych (CVD). Chociaż CAC jest ściśle powiązane ze starzeniem się, jego potencjał jako wskaźnik starzenia biologicznego pozostaje niejasny. Długość telomerów (TL) stanowi uznany wskaźnik starzenia biologicznego. Celem badania była ocena, czy skracanie telomerów (TL) jest przyczynowo związane z rozwojem CAC, a w konsekwencji - ustalenie, czy CAC może być wykorzystywany jako wskaźnik starzenia biologicznego, a nie jedynie marker procesów miażdżycowych.

Zbiorcze dane dotyczące wariantów genetycznych związanych z długością TL pozyskano z analizy bazującej na bazie UK Biobank. Dane dotyczące CAC uzyskano z metaanalizy badań asocjacyjnych całego genomu (GWAS). Do określenia związku pomiędzy długością TL a CAC zastosowano Randomizację Mendlowską (MR) - metodę statystyczną stosowaną w badaniach epidemiologicznych, wykorzystującą naturalne warianty genetyczne jako zmienne instrumentalne do badania zależności przyczynowo-skutkowych między czynnikami ryzyka a chorobami.

Genetycznie uwarunkowana dłuższa długość TL była istotnie związana z niższym poziomem CAC (metoda IVW - inverse variance weighted, p < 0,001), co wskazuje na ochronny wpływ TL na ryzyko rozwoju CAC. Analizy dodatkowe potwierdziły spójność tego wyniku (metoda MR-Egger, p = 0,03; metoda WME - weighted median, p = 0,004).

Uzyskane wyniki potwierdzają istnienie odwrotnej zależności przyczynowej pomiędzy długością TL a stopniem CAC, co sugeruje, że krótsze TL mogą zwiększać ryzyko rozwoju CAC. Wyniki te wzmacniają znaczenie CAC jako potencjalnego biomarkera starzenia biologicznego.

Słowa kluczowe: Telomery, Zwapnienie tętnic wieńcowych, Randomizacja Mendlowska

Zastosowanie entropii do oceny złożoności sygnałów biomedycznych rejestrowanych w teście pochyleniowym u pacjentów z omdleniami wazowagalnymi

Katarzyna Buszko

Katedra Biostatystyki i Teorii Układów Biomedycznych, Wydział Farmaceutyczny, Collegium Medicum im. Ludwika Rydygiera w Bydgoszczy, UMK w Toruniu, Polska

W tej pracy zostaną przedstawione możliwości zastosowania różnych rodzajów entropii do predykcji omdlenia w teście pochyleniowym. Analizowane sygnały zostały zarejestrowane w trakcie badania osób u których stwierdzono występowanie omdleń wazowagalnych. Omdlenie wazowagalne definiuje się jako krótkotrwałą odwracalną utratę przytomności na skutek, której często następuje upadek. Bezpośrednią przyczyną utraty przytomności jest krótkotrwała, odwracalna globalna hipoperfuzja mózgowa. Omdlenie wazowagalne może wystąpić na skutek długotrwałego przebywania w pozycji stojącej. Ten typ omdlenia może też wywołać silny stres, strach, ból i sytuacje związane z dużymi emocjami . W diagnostyce pacjenta z podejrzeniem omdlenia obok szczegółowego wywiadu i podstawowych badań (pomiar ciśnienia, EKG, badania krwi) wykonuje się test pochyleniowy.

W badaniach analizowano zapisy EKG, ciśnienia skurczowego, rozkurczowego i impedancji u 80 pacjentów poddanych testowi pochyleniowemu. U 57 pacjentów wystąpiło omdlenie w fazie biernej testu a u 23 w fazie prowokacji nitrogliceryną. Analizy koncentrują się na ocenie złożoności zapisów RRI, ciśnienia skurczowego (sBP), ciśnienia rozkurczowego (dBP) oraz objętości wyrzutowej (SV) w każdej fazie testu pochyleniowego przez wyznaczenie różnych miar entropii: Shannon Entropy (Sh), Sample Entropy (SE), Fuzzy Entropy (FE), Conditional Entropy (CE) i Permutation Entropy (PE)). A następnie ich porównanie w każdej fazie testu pomiędzy grupą z omdleniem w teście biernym HUTT(+) a grupą bez omdlenia (HUTT(-). Ze względu na nie spełnienie warunków stosowania testów parametrycznych, porównania pomiędzy fazami testu oraz grupami przeprowadzono wykorzystując testy nieparametryczne (Manna-Whitneya, Anova Friedmana z testem post-hoc Dunna). Analizy przeprowadzono na poziomie istotności α=0.05, wielokrotne testowanie hipotezy zerowej korygowano poprawką Bonfferoniego.

Szacowanie złożoności sygnałów RRI, SBP I DBP pokazało występowanie statystycznie istotnych różnic pomiędzy grupą HUTT(+) a grupą HUTT(-) w fazie leżącej testu dla przypadku Conditional Entropy wyznaczanej dla RRI (CE(RRI)). W trakcie pionizacji statystycznie istotne różnice otrzymano dla (SE(sBP), SE(dBP), FE(RRI), FE(sBP), FE(dBP), FE(SV), Sh(sBP), Sh(SV),

CE(sBP), CE(dBP)). W grupie pacjentów z grupy HUTT(+) częściej odnotowano statystycznie istotne zmiany w złożoności rejestrowanych sygnałów pomiędzy fazami testu niż w grupie pacjentów HUTT(-). W przypadku porównań entropii sygnałów w pozycji poziomej z entropią sygnałów podczas pionizacji w grupie HUTT(+) wszystkie analizowane entropie (SE(SV), FE(SV), Sh(SV), CE(SV), PE(SV)) różniły się istotnie.

Słowa kluczowe: entropia, omdlenie wazowagalne, analiza sygnałów biomedycznych

Dekompozycja testu ilorazu wiarygodności na składowe dla klas zdarzeń i braku zdarzeń: Wizualna prezentacja metodą U-smile

Barbara Więckowska[1], Przemysław Guzik[2]

[1]Katedra i Zakład Informatyki i Statystyki, Uniwersytet Medyczny im. Karola Marcinkowskiego w Poznaniu, Polska (https://ror.org/02zbb2597)
[2]Katedra i Klinika Intensywnej Terapii Kardiologicznej i Chorób Wewnętrznych (ITKC) Uniwersytetu Medycznego im. Karola Marcinkowskiego w Poznaniu, Polska (https://ror.org/02zbb2597)

Udostępnienie w ramach metody U-smile narzędzia statystycznej oceny modeli klasyfikacji binarnej oddzielnie dla każdej z klas (LR_1) i braku zdarzeń (LR_0) – dekompozycja statystyki testu ilorazu wiarygodności (LR)

Analizie poddano ogólnodostępny zbiór danych Heart Disease (n=331) zawierający 26 zmiennych, w tym parametry kliniczne oraz zmienne losowe (nieinformacyjne i informacyjne symetryczne/asymetryczne). Dla każdej zmiennej zbudowano modele regresji logistycznej. Statystyka LR została rozłożona na LR_0 i LR_1, testując ich istotność przy użyciu rozkładów Gamma: $LR_0 \sim Gamma(k/2, 1/2c)$; $LR_1 \sim Gamma(k/2, 1/2(1-c)$, gdzie c to proporcja klasy, k=1, to liczba zmiennych w modelu.

Wykresy U-smile wizualizowały względne współczynniki LR dla czterech podklas: poprawy/pogorszenia predykcji w każdej klasie. Linie ciągłe/przerywane wskazywały istotność statystyczną. Wygenerowano krzywe ROC i wykresy kalibracji jako punkt odniesienia dla wykresów U-smile.

Przy prawdziwości hipotezy zerowej statystyka LR ma rozkład Chi-kwadrat podczas gdy LR_0 i LR_1 mają rozkłady Gamma skalowane przez nierównowagę klas. Wykresy U-smile uwidoczniły asymetrię wydajności np. zmienna stde (exercise-induced ST depression relative to rest in mm) wykazała istotne „uśmiechy” tzn. dla obu klas linie ciągłe, podczas dla cholesterolu wynik był nieistotny tzn. linie przerywane na wykresie U-smile. Silnie asymetryczne zmienne losowe, były istotne tylko dla klasy zdarzeń (LR_1: p<0,05; linia ciągła), co nie było wykrywalne przez pole pod krzywą ROC.

Metoda U-smile uzupełnia tradycyjne metryki, rozkładając LR na składowe klasowe i identyfikując zmienne o istotnej statystycznie nierównowadze predykcyjnej.

Testowanie istotności LR_0 i LR_1 zwiększa interpretowalność, szczególnie w przypadkach asymetrycznych.

Metoda wpisuje się w zasady Explainable Machine Learning (wyjaśnialne uczenie maszynowe, EML), oferując intuicyjne narzędzie graficzne do oceny modeli.

Słowa kluczowe: test ilorazu wiarygodności, klasyfikacja binarna, wyjaśnialne uczenie maszynowe, metoda U-smile, istotność statystyczna

Trzeźwy czy nietrzeźwy, oto jest pytanie. Statystyka w toksykologii post mortem

Elżbieta Chełmecka[1] & Aleksandra Zorychta[1] & Małgorzata Głaz[2] & Joanna Nowicka[2] & Rafał Skowronek[2] & Marcin Tomsia[2]

[1]Zakład Statystyki Medycznej, Śląski Uniwersytet Medyczny w Katowicach, Sosnowiec, Polska
[2]Katedra i Zakład Medycyny Sądowej i Toksykologii Sądowo-Lekarskiej, Śląski Uniwersytet Medyczny w Katowicach, Polska

W postępowaniach/sekcjach prokuratorskich istotnym jest fakt ustalenia trzeźwości w chwili śmierci. Najczęściej wykorzystywana jest do tego krew lub mocz, natomiast w przypadkach degradacji zwłok używane są materiały alternatywne jak: płyn gałki ocznej, żółć, ślina, włosy, paznokcie, smółka i przychłonka oraz krążek miedzykręgowy czy chrząstka żebrowa. W sytuacji w której mamy do czynienia z daleko posuniętymi procesami rozkładu gnilnego a nawet prawie kompletnego zeszkieletowania chrząstka żebrowa może być jedynym materiałem, który pozwoli na określenie zarówno stężenia alkoholu etylowego jak i zażywania substancji psychoaktywnych antemortem.

Celem badań jest określenie po ilu dniach od śmierci będzie można określić trzeźwość w chwili śmierci oraz ustalenie, czy matryce alternatywne jak chrząstka żebrowa i krążek międzykręgowy mogą być uznane za materiał dowodowy.

W wyniku 57 sekcji pozyskano matryce do oznaczeń toksykologicznych: krew mocz, chrząstkę żebrową oraz dysk międzykręgowy. Próbki przechowywano w warunkach pokojowych, monitorowanych pod kątem temperatury. Dokonywano oznaczeń stężenia etanolu w kolejnych dniach. Badania na obecność etanolu przeprowadzono metodą chromatografii gazowej z detektorem płomieniowo-jonizacyjnym (GC-FID).

Wykazano silne korelacje pomiędzy wynikami uzyskanymi w moczu, chrząstce żebrowej oraz krążku międzykręgowym a krwią. W przeprowadzonych analizach, rozdzielono wyniki osób trzeźwych (stężenie etanolu <0,5‰; Ustawa z dnia 26 października 1982 r. o wychowaniu w trzeźwości i przeciwdziałaniu alkoholizmowi) i nietrzeźwych na podstawie oznaczenia z krwi i wykonano oddzielne analizy dla wszystkich rozpatrywanych matryc. Dla osób nietrzeźwych dokonano porównania metod analitycznych z chrząstki i krążka międzykręgowego z oznaczeniami z krwi będącymi „złotym standardem".

Na podstawie badań stężeń etanolu dla osób trzeźwych stwierdzono, że w żadnej z analizowanych matryc nie wytwarzał się alkohol endogenny. Dla osób nietrzeźwych w każdym przypadku: krew, mocz, chrząstka, dysk obserwowano spadek poziomu etanolu w okresie obserwacji. Dla krwi, począwszy od czwartej doby oraz dla chrząstki żebrowej i krążka

międzykręgowej od doby trzeciej nie można określić trzeźwości w chwili zgonu. Trzeźwość oceniana na podstawie badań stężenia etanolu w matrycach alternatywnych daje obiecujące wyniki, wprowadzenie tych metod do standardowych postępowań sekcyjnych wymaga dalszych badań.

Słowa klucze: Alkohol etylowy, trzeźwość, krew, chrząstka żebrowa, krążek międzykręgowy

Źródło finansowania badań: badania statutowe Śląskiego Uniwersytetu Medycznego w Katowicach BNW-1-010/K/4/F

Czy kryterium informacyjne Akaikego może się uśmiechnąć?

Katarzyna B. Kubiak[1], Barbara Więckowska[2]

[1] Zakład Zdrowia Publicznego i Medycyny Społecznej, Gdański Uniwersytet Medyczny, Polska
[2]Katedra i Zakład Informatyki i Statystyki, Uniwersytet Medyczny im. Karola Marcinkowskiego w Poznaniu, Polska
(https://ror.org/02zbb2597)

Celem pracy jest zastosowanie metody U-smile z kryterium informacyjnym Akaikego (AIC) do wyboru modelu w problemie klasyfikacji binarnej dla danych zbalansowanych.

Do referencyjnego modelu regresji logistycznej dodałyśmy jedną zmienną (modele zagnieżdżone). Dla każdej dodanej zmiennej, porównałyśmy zmiany predykcji w obu klasach odpowiedzi i otrzymałyśmy 4 grupy stratyfikowane według klasy odpowiedzi i zmiany predykcji: niezdarzenia z poprawą predykcji, niezdarzenia z pogorszeniem predykcji, zdarzenia z pogorszeniem predykcji, zdarzenia z poprawą predykcji. Dla każdej z czterech grup wyznaczyłyśmy stratyfikowaną różnicę AIC, które przedstawione obok siebie w wymienionej kolejności tworzą wykres U-smile stratyfikowanej różnicy AIC. Wykorzystałyśmy zbiór danych Heart Disease i syntetyczne dane z zadanych rozkładów prawdopodobieństwa (50% zdarzeń w zbiorze danych, n=300). Dodatkowo porównałyśmy model referencyjny z każdym nowym modelem testem ilorazu wiarygodności (poziom istotności 0.05).

Dla każdego nowego modelu wykreśliłyśmy wykresy U-smile różnicy AIC. Stratyfikowana według klasy odpowiedzi różnica AIC pozwoliła odróżnić zmienne informacyjne (uśmiechnięte wykresy U-smile) od nieinformacyjnych (płaskie wykresy U-smile) dodane do modelu referencyjnego. Uzyskałyśmy całkowitą zgodność między wartościami stratyfikowanej różnicy AIC a testem ilorazu wiarygodności.

Zastosowanie metody U-smile z AIC umożliwiło ocenę wpływu nowej zmiennej na jakość modelu osobno w obu klasach odpowiedzi. Interpretacja wykresów U-smile stratyfikowanej różnicy AIC jest intuicyjna i szybka, a także umożliwia wybór zmiennych spośród grupy kandydatów, które w największym stopniu poprawiają model osobno w obu klasach odpowiedzi.

Słowa kluczowe: metoda U-smile, wykres U-smile, kryterium informacyjne Akaikego, klasyfikacja binarna, wybór modelu, wybór zmiennych, uczenie maszynowe

Antocyjany a zdrowie mózgu – analiza związku między zwyczajowym spożyciem a funkcjami poznawczymi

Agnieszka Micek

Pracownia Statystyczna, Wydział Nauk o Zdrowiu, Uniwersytet Jagielloński Collegium Medicum, Kraków, Polska

Wobec rosnącej liczby przypadków zaburzeń poznawczych, w tym demencji i choroby Alzheimera, wzrasta znaczenie skutecznej i wczesnej profilaktyki. Coraz większą uwagę poświęca się roli diety oraz jej składników bioaktywnych w utrzymaniu zdrowia mózgu. Szczególne zainteresowanie budzą antocyjany – naturalne związki polifenolowe – ze względu na ich właściwości przeciwutleniające, przeciwzapalne oraz potencjalnie neuroprotekcyjne. Ocena wpływu interwencji z udziałem antocyjanów na funkcje poznawcze w kontekście ich możliwego działania ochronnego przed pogorszeniem funkcji poznawczych – metaanaliza randomizowanych badań klinicznych (RCT).

Przeprowadzono systematyczne przeszukiwanie bazy danych PubMed w celu identyfikacji badań oceniających wpływ spożycia antocyjanów na funkcje poznawcze. Uwzględniono publikacje dostępne do kwietnia 2025 roku. Kryteriami włączenia objęto randomizowane badania interwencyjne zawierające ilościowe dane dotyczące funkcji poznawczych. Z wykorzystaniem modelu efektów losowych oraz estymatora DerSimonian'a i Laird'a dla wariancji, obliczono standaryzowane średnie różnice (SMD) poziomu funkcji poznawczych, analizując je w podziale na poszczególne domeny. Przeprowadzono również analizy wrażliwości w celu oceny stabilności wyników oraz analizy wpływu, pozwalające na identyfikację badań wywierających istotny wpływ na ogólny efekt. Ryzyko błędu publikacyjnego oceniono za pomocą wykresów lejkowych, testu Eggera oraz metody „trim-and-fill", umożliwiającej oszacowanie i korektę potencjalnie brakujących badań w metaanalizie.

We wszystkich analizowanych domenach funkcji poznawczych odnotowano statystycznie istotnie wyższe wyniki u osób z grupy interwencyjnej (spożywających antocyjany) w porównaniu z grupą kontrolną. Odpowiednie wartości SMD wyniosły: 0,46 dla ogólnego wyniku funkcji poznawczych, 0,37 dla koncentracji i myślenia przestrzennego, 0,24 dla funkcji wykonawczych, 0,30 dla pamięci epizodycznej, 0,19 dla zdolności psychomotorycznych oraz 0,21 dla pamięci werbalnej.

Uzyskane wyniki wskazują na korzystny wpływ suplementacji antocyjanami na różne aspekty funkcji poznawczych. Związki te mogą stanowić obiecujący element strategii prewencyjnej w kontekście starzenia się populacji i wzrastającej liczby chorób neurodegeneracyjnych.

Słowa kluczowe: polifenole, funkcje kognitywne, neuroprotekcja

Zastosowanie meta-analizy do syntezy danych weryfikujących związek niedożywienia z objawami depresji u pacjentów onkologicznych

Agnieszka Micek[1] & Ewa Błaszczyk-Bębenek[2] & Aneta Cebula[3] & Anna Wąż[4] & Kamil Konopka[5]

[1]Pracownia statystyczna, Wydział Nauk o Zdrowiu, Uniwersytet Jagielloński Collegium Medicum, Kraków, Polska
[2] Zakład Badań nad Żywieniem i Lekami, Wydział Nauk o Zdrowiu, Uniwersytet Jagielloński Collegium Medicum, Kraków, Polska
[3] Szkoła Doktorska Nauk Medycznych i Nauk o Zdrowiu, Uniwersytet Jagielloński Collegium Medicum, Polska
[4] Zespół ds. Żywienia Klinicznego, Szpital Uniwersytecki w Krakowie, Kraków, Polska;
[5]Katedra i Klinika Onkologii, Wydział Lekarski, Uniwersytet Jagielloński Collegium Medicum, Kraków, Polska

Zweryfikowanie związku między niedożywieniem a poziomem depresji u pacjentów onkologicznych. Materiał i metody: Przeprowadzono systematyczny przegląd baz danych PubMed i EMBASE pod kątem badań obserwacyjnych opublikowanych do grudnia 2024 r. Do przeglądu włączono 41 badań, a 29 kwalifikowało się do metaanalizy. W większości badań status odżywienia oceniany był za pomocą PG-SGA, MNA jak również NRS-2002 i GLIM. Do pomiaru poziomu depresji najczęściej wykorzystano następujące kwestionariusze: HADS, BDI, GDS, PHQ, HAMD. Dziewięć artykułów koncentrowało się na pacjentach w zaawansowanych stadiach choroby, podczas gdy pozostałe 32 obejmowały różne jej stadia. Częstość występowania zaburzeń odżywiania i objawów depresji oszacowano odpowiednio na 52% i 30%. W porównaniu z pacjentami z prawidłowym stanem odżywienia, niedożywieni pacjenci mieli 3,23 razy większe ryzyko wystąpienia depresji. Wykryto jedynie niewielką heterogeniczność (I2 = 22,5%, P = 0,229). Zarówno analiza w podgrupach, jak i analiza wrażliwości wykazały, że ani pominięcie poszczególnych badań, ani różnice w charakterystyce pacjentów nie wpływały istotnie na ogólne wyniki. Metaanaliza standaryzowanych średnich różnic wyników poziomu depresji między skrajnymi kategoriami stanu odżywienia wykazała znacznie wyższe wyniki u osób niedożywionych w porównaniu z osobami z prawidłowym stanem odżywienia (SMD = 0,69, 95% CI 0,50 do 0,88). W obu analizach – będących odpowiednio ilościową syntezą ryzyk względnych oraz standaryzowanych różnic w średnich, wykresy lejkowe oraz wyniki testu Egger'a ujawniły pewne oznaki asymetrii. Analiza trim-and-fill, korygująca potencjalne błędy publikacyjne poprzez uzupełnienie o brakujące badania (4 i 5 odpowiednio), potwierdziła wcześniejsze ustalenia, ukazując nieco słabszy wpływ niedoborów żywieniowych na depresję (RR = 2.71, 95% CI 1.94 to 3.77; SMD = 0.44, 95% CI 0.23 to 0.64). Wnioski: Istnieje pilna potrzeba włączenia oceny stanu odżywienia do rutynowej opieki nad pacjentami chorymi na raka. Wczesna identyfikacja i interwencja w przypadku niedożywienia może pomóc złagodzić rozwój lub zaostrzenie objawów depresyjnych.

Słowa kluczowe: stan odżywienia, stan psychiczny, nowotwory

Interpretowalność modeli uczenia maszynowego w diagnostyce migotania przedsionków: rola wykresów SHAP w selekcji cech i odkrywaniu nieliniowych relacji.

Michał Michalak

Katedra i Zakład Informatyki i Statystyki, Uniwersytet Medyczny im. Karola Marcinkowskiego w Poznaniu, Polska (https://ror.org/02zbb2597)

Celem pracy była identyfikacja najistotniejszych parametrów analizy zmienności i asymetrii rytmu serca pozwalających na predykcję migotania przedsionków (AF) w krótkich, jednominutowych zapisach EKG. Szczególny nacisk położono na interpretację działania złożonych modeli uczenia maszynowego typu black-box z wykorzystaniem wykresów SHAP (SHapley Additive exPlanations).Analizowano 1-minutowe odprowadzenia EKG z oznaczonym występowaniem AF. Ekstrahowano szereg cech opisujących zmienność rytmu serca (HRV) oraz jego asymetrię, a także zmienne uzyskane z analizy spektralnej. Zbudowano dwa modele klasyfikacyjne: XGBoost oraz sieć neuronową. Ocena skuteczności odbywała się na podstawie wskaźników dokładność, AUC, czułości i swoistości. W celu identyfikacji kluczowych zmiennych oraz ich wpływu na wynik predykcji wykorzystano wartości SHAP. Analizowano zarówno wpływy indywidualne, jak i interakcje między cechami.

Zarówno model XGBoost, jak i sieci neuronowe wykazały bardzo wysoką skuteczność w klasyfikacji rytmu zatokowego i migotania przedsionków, osiągając dokładność odpowiednio 0.96 (dla modelu pełnego) i 0.95 (dla wersji uproszczonej), przy wskaźnikach AUC na poziomie 0.99 i 0.98. Redukcja liczby zmiennych wejściowych do pięciu najbardziej informatywnych cech, wyselekcjonowanych na podstawie wartości Shapleya (SHAP), pozwoliła znacząco uprościć modele bez istotnej utraty ich jakości predykcyjnej. W obu podejściach klasyfikacyjnych kluczowe znaczenie miały zmienne: SD2d (odchylenie standardowe odstępów RR wzdłuż linii l 2 wykresu Poincare dla zwolnień pracy serca), CSa (udział zmienności krótkoterminowej wynikającej z przyspieszeń HR w całkowitej HRV), SD1a (odchylenie standardowe różnic między kolejnymi odstępami RR dla przyspieszeń pracy serca) , PI (Nd) (wskaźnik Porty - względna liczba zwolnień do wszystkich zmian odstępów), lnLF (logarytm mocy niskich częstotliwości HRV), lnHF (logarytm mocy wysokich częstotliwości HRV) oraz lnTP (logarytm całkowitej mocy częstotliwości HRV). Modele SHAP ujawniły obecność złożonych, często nieliniowych interakcji między analizowanymi zmiennymi, w tym szczególnie istotne relacje pomiędzy lnTP a innymi predyktorami, co nie byłoby możliwe do uchwycenia tradycyjnymi metodami statystycznymi.

Zastosowane algorytmy uczenia maszynowego typu black-box nie tylko umożliwiły osiągnięcie wysokiej skuteczności predykcyjnej, ale także dostarczyły istotnych informacji o mechanizmach leżących u podstaw AF. Uproszczone modele, ograniczone do pięciu zmiennych, zachowały wysoką wartość diagnostyczną przy jednoczesnym obniżeniu kosztów obliczeniowych, co czyni je szczególnie atrakcyjnymi do zastosowań klinicznych oraz w rozwiązaniach mobilnych i telemetrycznych. Wykresy SHAP zapewniają przejrzystą interpretację działania modeli, identyfikację najważniejszych cech oraz zrozumienie mechanizmów decyzyjnych. Takie podejście może wspierać rozwój klinicznie użytecznych systemów wspomagania decyzji.

Słowa kluczowe: migotanie przedsionków, uczenie maszynowe, SHAP, zmienność rytmu, XGBoost, sieci neuronowe

Validating the safety profile thought laboratory data in clinical trial

A. Korczyński
Szkoła Główna Handlowa w Warszawie, Polska

Diagnostic and Screening Tests / Using statistics to improve the conduct of clinical trials

While searching for efficacious therapy safety remains the key objective. It is vital in phase I trials aimed at determining the dose level allowing to reach the expected treatment effect and, at the same time, ensuring safety of the patient. Safety data collection is a standard process in clinical trials and includes reporting of the adverse events (AEs) and laboratory assessments which are subject to medical review. There is however a significant difference between these two data sources as AEs are reported data outcomes whereas laboratory data can be considered as objective assessment of the patient status. The aim of this paper is to present a screening technique allowing early assessment of potential safety events through statistical method relying on laboratory data. It is considered as supportive tool for identification of outlier observations such us toxicities. In that sense the method is a form of an independent review of the adverse events reporting and can be utilized to monitor the rates of particular AEs as compared to the laboratory data outcomes. Clinical protocols with safety as primary objective would require summary of the AEs and an independent insight into the reporting rate would serve as sensitivity check on the occurrence and frequency of particular terms that are later described in the clinical study reports.

The idea is to evaluate the trajectories of individual laboratory assessments through quantile regression in order to allow for evaluation at intermediate steps, as we gain more data. Quantile regression is within the class of robust techniques which is essential as we would expect to see inconsistencies in the early data entries including those for units. The outcome of the assessment is then analysed through k-means clustering for identifying outliers in multidimensional settings. The method has to incorporate the assessment of the evolution over time and multidimensional aspect of the data. We are interested in analysing the absolute values as well as changes from baseline. The subset of parameters is assessed continuously over time for the rate of change to determine unusual data patterns. The technique can serve as additional tool supporting the medical review conducted throughout the clinical study.

Keywords: adverse events reporting, laboratory data evaluation, quantile regression, k-means clustering