



prof. dr hab. n. farm. Łukasz Komsta  
Uniwersytet Medyczny w Lublinie  
Wydział Farmaceutyczny  
Katedra i Zakład Chemii Leków  
ul. Jaczewskiego 4, 20-090 Lublin, tel. 81 4487387, fax 81 4487381

## Recenzja pracy doktorskiej mgr Patryka Własiuka

### *„Evaluation and interpretation of analytical data*

### *- the likelihood ratio and chemometrics for forensic and medical purposes”*

PRZEDŁOŻONA do recenzji praca doktorska składa się z trzech publikacji opublikowanych w *Drug Testing and Analysis* (IF 2.9, 70 pkt.), *Analytica Chimica Acta* (IF 6.2, 140 pkt.) oraz *Chemometrics and Intelligent Laboratory Systems* (IF 3.9, 100 pkt.). Do publikacji dołączony jest blisko 100-stronicowy obszerny komentarz teoretyczny. Pod względem bibliometrycznym pracę należy ocenić bardzo wysoko, a komentarz pod względem treści przekracza o rząd wielkości współczesne wymagania dotyczące dysertacji stanowiących zbiór publikacji. Wkład Doktoranta nie jest poparty oświadczeniami, jednakże można ocenić go na wiodący i wystarczający.



W pierwszej publikacji Doktorant podjął próbę analizy zmian stężeń glukuronidu etylu we włosach, zarejestrowanych we Włoszech w okresie siedmiu lat (2011 — 2018). Ze względu na zauważoną cykliczność zmian stężenia w ciągu roku, dane zostały dopasowane do modelu  $x_i = \beta_0 + \beta_1 I_i + \sum_{j=1}^{d/2} \left[ \beta_{1+j} \sin\left(\frac{2\pi j}{d} i\right) + \beta_{2+j} \cos\left(\frac{2\pi j}{d} i\right) \right] + \varepsilon_i$ , co w zasadzie jest równoważne z wykonaniem regresji liniowej oraz dodatkową niepełną transformacją Fouriera otrzymanych reszt :  $x_i = \beta_0 + \beta_1 I_i + \sum_{j=1}^{d/2} \left[ \beta_j e^{-\frac{2\pi j}{d} i} \right] + \varepsilon_i$ , gdzie  $\beta_j$  jest liczbą zespoloną, zawierającą część rzeczywistą odpowiadającą  $\beta_{1+j}$  i część urojoną odpowiadającą  $\beta_{2+j}$ .

Analiza istotności poszczególnych składowych transformacji pokazała, że trend roczny jest jedynym istotnym trendem cyklicznym w badanej serii danych, a Autor podjął się kompleksowej dyskusji i postawił wiele hipotez dotyczących tej zmienności, niekoniecznie związanych z konsumpcją alkoholu.

Chociaż założenie kształtu sinusoidalnego ma swoje umocowanie teoretyczne w przypadku tych

danych, interesujące byłoby porównanie otrzymanego modelu z modelem liniowym o wyrazie wolnym zależnym od miesiąca (czyli model liniowy bez transformacji Fouriera, ale z wprowadzoną dodatkową zmienną typu *factor* zawierającą miesiąc). Pozwoliłoby to modelowi dopasować się do nieco innego, asymetrycznego kształtu, który może np. wynikać z kompleksowości procesów stanowiących o tej zależności. Porównanie istotności, wyjaśnionej wariancji oraz AIC czy BIC byłoby dodatkowym czynnikiem diagnostycznym, czy proces zmian jest czysto sinusoidalny, czy stoi za nim jakaś bardziej złożona tajemnica.



W drugiej publikacji Doktorant podjął próbę stworzenia jedno- oraz wielorakich modeli opartych o iloraz wiarygodności, służących do klasyfikacji pochodzenia geograficznego oliwy na podstawie zawartości kwasów tłuszczowych. Praca jest o tyle cenna, że dane w niej wykorzystane służyły za zestaw testowy w wielu wcześniejszych publikacjach, czytelnik może zatem łatwo odnaleźć próby stosowania innych algorytmów do oceny tych danych.

Na uwagę zasługuje kompleksowe podejście do tworzenia modeli z trzema strategiami — modele oparte na jednej zmiennej, modele oparte na wszystkich, jak również selekcja zmiennych w oparciu o test statystyczny.

Publikacja zakłada rozkład normalny analizowanych danych — zarówno poprzez zastosowanie testu parametrycznego do oceny różnic, jak również do tworzenia modeli oznaczonych  $LR_{NRM}$ . Tu interesujące byłoby sprawdzenie, które grupy wyników odbiegają od takiego rozkładu i jak to porównanie zmieniłoby się w przypadku stosowania testów nieparametrycznych, albo czy można byłoby zastosować inne rozkłady do tworzenia modeli opartych o konkretny rozkład, a nie o KDE.



Ostatnia publikacja porównuje różne metody wygładzania rozkładu danych do modelowania LR. Wątek ten wnosi istotne nowości do teorii modelowania oraz tworzy rekomendacje dotyczące wyboru parametrów przy analizie podobnych danych w przyszłości. Podobnie jak w przypadku publikacji poprzedniej, dane wykorzystane w eksperymentach są znanymi zestawami testowymi, na których próbuje się często nowe idee algorytmów.

Za ważne osiągnięcie tej publikacji należy na pewno uznać wykazanie, że metoda  $RT_1$  jest w stanie zapewnić optimum pomiędzy minimalizacją  $\Delta C_{ur}$  a maksymalizacją CR. W poprzedniej, cytowanej publikacji, nieco lepszą metodą okazywała się ta oparta o walidację krzyżową i redukcję wymiarowości, jednakże — jak piszą Autorzy — nie brana była wtedy pod uwagę entropia krzyżowa.

Tutaj pojawia się pytanie: czy redukcja wymiarowości danych z tej publikacji, a następnie tworzenie modeli np. na wartościach głównych składowych, zmieniłaby cokolwiek w trafności klasyfikacji lub w innych parametrach diagnostycznych otrzymanych modeli?



Wszystkie publikacje stanowiące rozprawę przeszły rygorystyczny proces edytorski w prestiżowych czasopismach. Ja również nie widzę w nich niczego, co wzbudzałoby wątpliwości, a moje rozważania i pytania nie wynikają z uchybień lub braków, lecz z zasygnalizowania możliwych tematów do dyskusji podczas obrony doktorskiej.



**Podsumowując, praca doktorska mgr Patryka Własiuka spełnia wszystkie wymagania stawiane pracom doktorskim, zawiera istotne elementy nowości naukowej i wnioskuję o dopuszczenie Doktoranta do dalszych etapów przewodu doktorskiego.**

**Dodatkowo, biorąc pod uwagę wyraźny i znaczny wkład badań w metodologię modelowania LR, zarówno od strony teorii, jak i zastosowań, z przyjemnością wnioskuję o wyróżnienie przedłożonej mi do recenzji pracy.**

Katedra i Zakład Chemii Leków  
Uniwersytetu Medycznego w Lublinie  
prof. dr hab. Łukasz Komsta

Lublin, 9 października 2023.