



UNIwersytet Jagielloński
COLLEGIUM MEDICUM

WYDZIAŁ FARMACEUTYCZNY
KATEDRA TECHNOLOGII POSTACI LEKU I BIOFARMACJI

Aleksander Mendyk

Modelowanie empiryczne w ocenie dostępności farmaceutycznej i biologicznej
substancji leczniczych ze stałych postaci leku

ZAŁĄCZNIK 2A DO WNIOSKU HABILITACYJNEGO:
AUTOREFERAT

Kraków 2016

Spis treści

1. Dane osobowe.....	2
2. Osiągnięcie naukowe zgłoszone do postępowania habilitacyjnego.....	3
2.1. Publikacje wchodzące w skład osiągnięcia naukowego.....	3
2.2. Cel badawczy.....	5
2.3. Wprowadzenie.....	6
2.4. Wyniki badań związanych z osiągnięciem naukowym w postępowaniu habilitacyjnym.....	9
2.4.1. Modelowanie zależności parametrów oceny <i>in vitro</i> stałych postaci postaci leku od ich składu jakościowego i ilościowego oraz technologii sporządzania.....	9
2.4.2. Empiryczne modelowanie zależności <i>in vitro in vivo</i> (IVIVR).....	16
2.4.3. Konstrukcja i publikacja programów komputerowych do zautomatyzowanego modelowania empirycznego.....	17
3. Podsumowanie.....	26
4. Pozostałe osiągnięcia naukowe.....	28
4.1. Działalność naukowo-badawcza przed uzyskaniem stopnia doktora.....	28
4.2. Działalność naukowo-badawcza po uzyskaniu stopnia doktora.....	28
4.2.1. Technologie informatyczne.....	28
4.2.2. Ocena <i>in silico</i> kardiotoksyczności substancji leczniczych.....	29
4.2.3. Optymalizacja depozycji <i>in vitro</i> substancji leczniczych z systemów nośnikowych do podania wziewnego.....	31
4.2.4. Niestandardowe modele korelacji i zależności <i>in vitro in vivo</i> (IVIVC/IVIVR).....	34
4.2.5. Próby wyjaśniania mechanizmów uwalniania substancji leczniczych z matrycowych postaci leku na podstawie analizy metodą obrazowania magnetycznego rezonansu jądrowego (MRI).....	35
4.3. Podsumowanie dorobku naukowego.....	36
4.4. Prace programistyczne.....	36
5. Piśmiennictwo.....	37

1. Dane osobowe

Aleksander Mendyk

Katedra Technologii Postaci Leku i Biofarmacji UJCM w Krakowie, ul. Medyczna 9,
30-688 Kraków, e-mail: mfmendyk@cyf-kr.edu.pl; aleksander.mendyk@uj.edu.pl

Dyplomy i stopnie naukowe

- 1997**
- stopień magistra farmacji
 - praca magisterska pt „Próba zastosowania sztucznych sieci neuronowych w modelowaniu farmakokinetyczno-farmakodynamicznym” wykonana w Zakładzie Farmacji Fizycznej i Farmakokinetyki UJCM pod kierunkiem dr Wojciecha Jawienia i opieką naukową śp. prof. dr hab. Joanny Szymury-Oleksiak została nagrodzona przez Polskie Towarzystwo Farmaceutyczne w 1998 za najlepszą pracę magisterską w roku 1997
- 2004**
- stopień doktora nauk farmaceutycznych
 - praca doktorska pt. „Sztuczne sieci neuronowe jako uniwersalne narzędzia modelowania w technologii postaci leku i biofarmacji” wykonana w Katedrze Technologii Postaci Leku i Biofarmacji UJCM pod opieką naukową prof. dr hab. Renaty Jachowicz, wyróżniona przez Radę Wydziału Farmaceutycznego UJ CM oraz nagrodzona indywidualną nagrodą Ministra Edukacji Narodowej i Sportu

Przebieg zatrudnienia

- 1997 – 2005**
- asystent w Katedrze Technologii Postaci Leku i Biofarmacji Collegium Medicum Uniwersytetu Jagiellońskiego w Krakowie.
- 2005 – teraz**
- adiunkt w Katedrze Technologii Postaci Leku i Biofarmacji Collegium Medicum Uniwersytetu Jagiellońskiego w Krakowie
- 1998 – 2005**
- praca w aptece otwartej
- 2006 – 2009**
- współzałożyciel i współwłaściciel firmy pharm-info.net
- 2005 – 2015**
- wykonawca 19 prac eksperckich w 6 firmach farmaceutycznych

- 2015 – teraz • ekspert ds. biorównoważności w USP Zdrowie sp. z o. o.

Ukończone kursy

- 1998 • “Principles of Pharmacokinetic/Pharmacodynamic Modeling”. Kurs dotyczył modelowania PK/PD i prowadzony był przez prof. W.J. Jusko z University of Buffalo, USA
- 2001 • Leonardo da Vinci JOINT „Joint International Training Module on Communication Techniques for SMEs”. Kurs prowadzony w Katedrze Telekomunikacji Wydziału Informatyki, Elektroniki i Telekomunikacji Akademii Górniczo-Hutniczej w Krakowie dotyczył nowoczesnych technologii komunikacyjnych wdrażanych do obsługi informatycznej małych i średnich przedsiębiorstw.

2. Osiągnięcie naukowe zgłoszone do postępowania habilitacyjnego

Osiągnięcie naukowe zgłoszone do postępowania habilitacyjnego stanowi cykl 9 publikacji o łącznym współczynniku oddziaływania *IF* równym 16,588 i punktacji MNiSW równej 203 punkty. Temat osiągnięcia: „Modelowanie empiryczne w ocenie dostępności farmaceutycznej i biologicznej substancji leczniczych ze stałych postaci leku”

2.1. Publikacje wchodzące w skład osiągnięcia naukowego

P.1 Aleksander Mendyk, Renata Jachowicz and Przemysław Dorożyński. Artificial neural networks in the modeling of drugs release profiles from hydrodynamically balanced systems Acta Pol. Pharm. 2006, 62, 75-80. (*MNiSW = 6, udział własny 90% w tym: koncepcja pracy, zebranie i opracowanie danych, wykonanie obliczeń, stworzenie modeli i napisanie tekstu*)

P.2. Aleksander Mendyk, Peter Kleinebudde, Markus Thommes, Angelina Yoo, Jakub Szlęk, Renata Jachowicz. Analysis of pellets properties with use of artificial neural networks, European Journal of Pharmaceutical Sciences, 2010, 41, 421-429 (*IF = 3.291, MNiSW = 32, udział własny*

75% w tym: koncepcja pracy, zebranie i opracowanie danych, wykonanie obliczeń, stworzenie modeli i napisanie tekstu)

P.3. Sinan Güres, **Aleksander Mendyk**, Renata Jachowicz, Przemysław Dorożyński Peter Kleinebudde. Application of artificial neural networks (ANN) and genetic programming (GP) for prediction of drug release from solid lipid matrices. International Journal of Pharmaceutics, 2012, 436, 877-879. (*IF = 3.458, MNiSW = 35, udział własny 40% w tym: koncepcja pracy, opracowanie danych, wykonanie obliczeń, stworzenie modeli i napisanie tekstu*)

P.4. **Aleksander Mendyk**, Renata Jachowicz, Kamil Fijorek, Przemysław Dorożyński, Piotr Kulinowski and Sebastian Polak. KinetDS: an Open Source software for dissolution test data analysis, Dissolution Technologies, 2012, 19, 6-11 (*IF = 0.612, MNiSW = 15, udział własny 75% w tym: koncepcja pracy, zebranie i opracowanie danych, projekt i napisanie programu komputerowego KinetDS, wykonanie obliczeń i napisanie tekstu*)

P.5. **Aleksander Mendyk**, Paweł Tuszyński, Sebastian Polak, Renata Jachowicz. In vitro-in vivo relationship (IVIVR) generalized model based on artificial neural networks. Drug Design, Development and Therapy, 2013, 7, 223-232. (*IF = 3.026, MNiSW = 35, udział własny 85% w tym: koncepcja pracy, wykonanie części obliczeń, kodowanie struktur chemicznych polimerów, analiza wrażliwościowa modeli neuronowych i napisanie tekstu*)

P.6. **Aleksander Mendyk**, Adam Paclawski, Jakub Szlęk, Renata Jachowicz. PhEq_bootstrap: an Open Source software for simulation of f2 distribution in cases of a large variability in the dissolution profiles. Dissolut Technol, 2013, 20, 13 – 17. (*IF = 0.712, MNiSW = 15, udział własny 85% w tym: koncepcja pracy, projekt i napisanie programu komputerowego PhEq_bootstrap, wykonanie obliczeń i napisanie tekstu*)

P.7. Jakub Szlęk, Adam Paclawski, Raymond Lau, Renata Jachowicz, **Aleksander Mendyk**. Heuristic modeling of macromolecule release from PLGA microspheres Int J Nanomedicine. 2013;8:4601-11. doi: 10.2147/IJN.S53364. (*IF = 4,195, MNiSW = 35, udział własny 50% koncepcja pracy, wykonanie części obliczeń, napisanie części tekstu, kierownictwo projektu naukowego obejmującego badania zamieszczone w pracy i rola autora korespondencyjnego*)

P.8. Aleksander Mendyk, Sinan Güres, Renata Jachowicz, Jakub Szlęk, Sebastian Polak, Barbara Wiśniowska, and Peter Kleinebudde, “From Heuristic to Mathematical Modeling of Drugs Dissolution Profiles: Application of Artificial Neural Networks and Genetic Programming,” *Comput. Math. Methods Med*, 2015, Article ID 863874. doi:10.1155/2015/863874 (*IF = 0.766, MNiSW = 15, udział własny 50%, w tym: koncepcja pracy, wykonanie części obliczeń, stworzenie części modeli i napisanie tekstu*)

P.9. Aleksander Mendyk, Paweł Konrad Tuszyński, Mohammad Hassan Khalid, Renata Jachowicz and Sebastian Polak. How-To: Empirical IVIVR Without Intravenous Data. *Dissolution Technologies*, 2015, 22(2), 12 – 18, dx.doi.org/10.14227/DT220215P12 (*IF = 0,528, MNiSW = 15, udział własny 80%, w tym: koncepcja pracy, projekt i napisanie programu komputerowego RIVIVR, wykonanie części obliczeń i napisanie tekstu*)

Łącznie:

IF = 16,588; MNiSW = 203, średni udział własny 70%

2.2. Cel badawczy

Celem badawczym było stworzenie empirycznych modeli opartych na narzędziach inteligencji obliczeniowej mających zastosowanie w ocenie zależności łączących parametry postaci leku z ich składem jakościowym i ilościowym oraz technologią sporządzania. Parametryzacja postaci leku rozumiana jest jako uwzględnienie badań uwalniania substancji leczniczej *in vitro*, jak też i profili farmakokinetycznych leków uzyskanych *in vivo* w badaniach klinicznych.

Prace realizowano w trzech głównych wątkach:

1. modelowanie narzędziami inteligencji obliczeniowej zależności parametrów oceny *in vitro* stałych postaci leku od ich składu jakościowego i ilościowego oraz technologii sporządzania [P.1, P.2, P.3, P.7, P.8]
2. empiryczne modelowanie narzędziami inteligencji obliczeniowej zależności *in vitro in vivo* (IVIVR) dla stałych postaci leku [P.5, P.9]
3. konstrukcja i publikacja programów komputerowych do zautomatyzowanego modelowania

empirycznego:

- modelowanie matematyczne procesu uwalniania substancji leczniczej w warunkach *in vitro* [P.4]
- ocena równoważności profili uwalniania substancji leczniczych z postaci leku w warunkach dużej zmienności wyników uwalniania [P.6]
- korelacja *in vitro in vivo* (IVIVC/IVIVR) [P.9]

2.3. Wprowadzenie

W świetle aktualnie obowiązujących zaleceń dotyczących racjonalizacji i kontroli procesu wytwarzania poprzez jego zrozumienie (ang. *process analytical technologies*, PAT) [15] należy wziąć pod uwagę wszystkie dostępne informacje pozwalające na racjonalną i ilościową ocenę zależności wpływających na jakość produktu. Inicjatywa FDA zmierzająca do wprowadzenia jako obowiązującego systemu wytwarzania leków z jakością „wbudowaną w produkt” (ang. *quality by design*, QbD) [19] wymusza zastosowanie szeregu technik analitycznych w celu zgromadzenia niezbędnych danych oraz ich oceny. Stanowi to wyzwanie dla przemysłu z uwagi na znaczną złożoność procesu produkcyjnego, która jest m.in. konsekwencją dostępności szerokiego zakresu substancji pomocniczych oraz technik sporządzania postaci leku. Spośród badań potwierdzających odpowiednią jakość leku, badanie dostępności farmaceutycznej dostarcza istotnych informacji charakteryzujących opracowaną formułację. Wynika to z faktu, że szybkość uwalniania substancji leczniczej zależy od składu formułacji, doboru substancji pomocniczych i zastosowanego procesu technologicznego. Wynik tego badania prowadzonego w warunkach *in vitro* podawany zazwyczaj jako profil uwalniania substancji leczniczej w czasie. Stwarza to komplikacje natury matematycznej związane z wielowymiarowym charakterem profilu uwalniania substancji leczniczej. Analogicznie jest w przypadku matematycznego opisu dostępności biologicznej, gdzie AUC, C_{\max} i t_{\max} to minimalny zakres parametrów opisujących tę wielkość [41]. Wyrazem tendencji do kwantyfikowania biologicznych konsekwencji zmian w charakterystyce produktu i procesu jego wytwarzania jest pojęcie korelacji *in vitro in vivo* (IVIVC), w którym dostępność biologiczna powiązana jest zależnością matematyczną z dostępnością farmaceutyczną. Należy również przywołać wytyczne dotyczące zmian porejestacyjnych [16], gdzie wśród czynników krytycznie wpływających na bezpieczeństwo stosowania leku wymienia się obok składu ilościowego i

jakościowego postaci leku również technologie i miejsce wytwarzania leku.

W świetle powyższych faktów, w chwili obecnej obserwuje się tendencję wdrażania w przemyśle farmaceutycznym szeregu technik modelowania ze szczególnym naciskiem na identyfikację zmiennych kluczowych i redukcję złożoności powstających modeli.

Modelowanie i symulacje (ang. *modeling and simulation, MS*) to dynamicznie rozwijająca się dzisiaj dziedzina wiedzy, która służy wielu zastosowaniom zarówno w nauce jak i technice, a także nawet w codziennej aktywności społeczeństwa. Notowany aktualnie gwałtowny postęp w technikach komputerowych jest podstawą do stworzenia bezprecedensowych w historii naszej cywilizacji możliwości ilościowej oceny otaczającej nas rzeczywistości. Stanowi to, obok równie znaczącego rozwoju metod matematycznych i języków programowania, główną siłę napędową modelowania jako dziedziny wiedzy i stosowanych praktycznie rozwiązań.

Z punktu widzenia podejścia do modelowania wyróżnia się jego dwa rodzaje:

- mechanistyczne,
- empiryczne.

W modelowaniu mechanistycznym niezbędna jest teoria opisująca analizowane zagadnienie, dzięki której znane są matematyczne podstawy budowanego modelu, a także fizyczne znaczenie jego elementów tj. stałych i zmiennych [40]. Model oparty jest więc o zidentyfikowane już wcześniej mechanizmy rządzące analizowanym problemem, dzięki czemu jego budowa często sprowadza się tylko do ustalenia wartości jego parametrów. Takie podejście, za którym stoi zdobyta już wiedza, jest szeroko stosowane zwłaszcza tam, gdzie potrzebna jest możliwość dokładnego prześledzenia sposobu działania modelu i wyjaśnienia zasad leżących u podstaw otrzymywanych wyników. Tego typu modele, zwane białymi skrzynkami (ang. *white box*), są dzisiaj uznawane za podstawowe narzędzia modelowania dzięki możliwości ścisłego kontrolowania ich zachowania oraz algorytmicznemu, tj. powtarzalnemu charakterowi pracy. Z punktu widzenia zasad walidacji obowiązujących w przemyśle farmaceutycznym [11] modele mechanistyczne są więc preferowane, czego wyrazem jest chociażby dążenie do liniowego charakteru korelacji *in vitro in vivo* na poziomie A [14].

Mimo oczywistych zalet modele mechanistyczne nie zawsze są w stanie dostarczyć wystarczającego obrazu analizowanego zagadnienia. Do głównych przyczyn niepowodzeń modelowania mechanistycznego należy zaliczyć:

- brak dobrze ugruntowanych teorii opisujących badany problem,
- mnogość potencjalnych zmiennych sterujących,
- nieliniowość.

Wymienione wyżej przyczyny często są ze sobą powiązane i stanowią poważne wyzwanie dla dostępnych narzędzi i technik analizy danych. Znalazło to wyraz m.in. w powstaniu wyspecjalizowanych jednostek badawczych, jak np. Instytut Santa Fe [20], którego główną misją są badania nad złożonymi systemami fizycznymi, biologicznymi czy społecznymi. Poszukiwania narzędzi oceny systemów złożonych doprowadziły do konkluzji, że modele mechanistyczne są zdecydowanie niewystarczające z uwagi na swoje uzależnienie od aktualnego poziomu wiedzy w naukach podstawowych. Tak więc zwrócono się w stronę technik empirycznych, wśród których znaczącą rolę odgrywają narzędzia heurystyczne. Heurystyki są definiowane jako twórcze sposoby odkrywania rozwiązania problemu, często nie podlegające ujęciu algorytmicznemu czy też dowodowi matematycznemu ich działania. Mimo, iż na pierwszy rzut oka podejście „intuicyjne” wydaje się nieefektywne w rozwiązywaniu skomplikowanych problemów naukowych, dzięki gwałtownemu rozwojowi technik obliczeniowych w latach 80 i 90 ubiegłego wieku, wykazano że modele heurystyczne mogą być z powodzeniem stosowane w zagadnieniach, w których modelowanie mechanistyczne nie pozwala na uzyskanie zadowalających wyników [45, 47]. Rozwój technik obliczeniowych doprowadził do szerokiego zastosowania metod heurystycznych zaliczanych do sztucznej inteligencji (ang. *artificial intelligence*, AI) a dokładniej rzecz mówiąc do tzw. inteligencji obliczeniowej (ang. *computational intelligence*, CI):

- sztucznych sieci neuronowych,
- obliczeń ewolucyjnych ,
- logiki rozmytej.

Wszystkie trzy wyżej wymienione narzędzia posiadają wspólną cechę, tj. zdolność do samoadaptacji do danych. Wprawdzie logika rozmyta jako odrębny system logiczny takich możliwości nie posiada, jednak aktualnie najczęściej jest ona spotykana w układach hybrydowych,

tw. sieciach neuronowo-rozmytych, które są systemami samoadaptacyjnymi [47]. Zdolność do automatycznej reprezentacji analizowanego zbioru danych jest rezultatem odkrycia technik uczenia maszynowego, które w swoich oryginalnych założeniach miały pozwolić na stworzenie maszyn myślących. Jak wiadomo, sztuczna inteligencja pozostaje wciąż domeną fantastyki naukowej mimo wciąż pojawiających się doniesień o zdaniu przez kolejną maszynę testu Turinga [2], czy też prognoz o rychłym wynalezieniu myślących maszyn dzięki inicjatywie Elona Muska pod nazwą OpenAI [29]. Mimo to samoadaptacja do danych realizowana automatycznie przez komputery wyposażone w odpowiedni program stała się kamieniem milowym w modelowaniu empirycznym. Zdolności samoadaptacyjne systemów inteligencji obliczeniowej manifestują się zarówno podczas automatycznej budowy modelu analizowanego zjawiska, jak też i w procesie decyzyjnym, w którym ustalana jest ważność zmiennych wejściowych dla uzyskiwanego przez model wyjścia. Ta ostatnia cecha pozwala na przeprowadzenie procedury analizy wrażliwościowej modelu, dzięki czemu można wyłonić zmienne kluczowe dla analizowanego zjawiska, co stanowi m.in. domenę tzw. drążenia danych (ang. *data-mining*), a w przemyśle farmaceutycznym może być utożsamiane z identyfikacją krytycznych parametrów jakościowych (ang. *critical quality attributes*, CQA) oraz krytycznych parametrów procesu (ang. *critical process parameters*, CPP) [19]. Samoadaptacja jest istotą empirycznego charakteru modelowania za pomocą narzędzi inteligencji obliczeniowej. Model budowany jest automatycznie przez system komputerowy, tylko na podstawie dostępnego zbioru danych, bez żadnych założeń co do jego charakteru. Obecność nieliniowych elementów strukturalnych takich jak np. funkcje aktywacji w sieciach neuronowych, determinuje nieliniowy charakter budowanych modeli, tak trudny nieraz do osiągnięcia za pomocą klasycznych metod.

2.4. Wyniki badań związanych z osiągnięciem naukowym w postępowaniu habilitacyjnym

2.4.1. Modelowanie zależności parametrów oceny *in vitro* stałych postaci leku od ich składu jakościowego i ilościowego oraz technologii sporządzania

Znalezienie ilościowej zależności pomiędzy składem jakościowym i ilościowym postaci leku a profilem uwalniania z niej substancji leczniczej wymaga wzięcia pod uwagę szeregu czynników, wśród których oprócz składu ilościowego i jakościowego istotne są również np. warunki badania dostępności farmaceutycznej. Stanowi to więc problem, który nadaje się do zastosowania narzędzi inteligencji obliczeniowej (CI), wśród których za jedno z najbardziej efektywnych uważa się sztuczne sieci neuronowe (ANNs) [18, 47]. ANNs to przykład konektywistycznego podejścia do

inteligencji obliczeniowej. Zostały one stworzone jako próba naśladowania biologicznych układów neuronowych zarówno pod względem strukturalnym jak i funkcjonalnym. „Kodowanie” wiedzy w sieci neuronowej jest realizowane w trakcie uczenia maszynowego – procesu automatycznej akwizycji wiedzy przez sieć neuronową z wykorzystaniem dostępnych źródeł danych. Dzięki temu modele powstają autonomicznie, bez żadnych wstępnych założeń teoretycznych, na podstawie samego procesu uczenia maszynowego. ANNs zostały zastosowane do przewidywania profilu uwalniania substancji leczniczej z systemów zrównoważonych hydrodynamicznie (HBS) przygotowanych z wykorzystaniem hydroksypropylometylocelulozy [P.1]. Baza danych powstała z wykorzystaniem wyników dla 66 formuacji przygotowanych dla 6 modelowych substancji leczniczych: metamizolu sodowego, metforminy, L-dopy, ketoprofenu, papaweryny i prednizolonu. Do opisu składu jakościowego formuacji użyto narzędzi informatyki chemicznej w celu wyliczenia deskryptorów molekularnych. Sztuczne sieci neuronowe zastosowano zarówno w funkcji narzędzia do modelowania jak też i do analizy wrażliwościowej. Tę ostatnią funkcję wykorzystano do selekcji 7 zmiennych kluczowych spośród 77 dostępnych w oryginalnej bazie danych. Sprawność modeli predykcyjnych potwierdzono w procedurze 10-krotnego wzajemnego sprawdzania i wykazano że modele neuronowe są w stanie integrować różne źródła informacji i prawidłowo przewidywać profil uwalniania substancji leczniczej dla nieznanymi formuacji. Informacji na temat badanego problemu dostarczyła analiza wyselekcjonowanych zmiennych kluczowych, wśród których obok zmiennych opisujących skład ilościowy znalazły się informacje na temat ilości atomów azotu i tlenu, stałej Henry'ego, lipofilności cząsteczki oraz jej budowy elektronowej. Modele neuronowe okazały się być niezwykle elastycznymi narzędziami integrującymi zmienne o zróżnicowanym charakterze w celu stworzenia reprezentatywnego modelu procesu uwalniania substancji leczniczej z systemów HBS. Wyniki analizy klasteryzacyjnej zredukowanego wektora wejściowego pozwoliły na wyłonienie dwóch grup substancji leczniczych o zbliżonej do siebie charakterystyce opisanej przez znalezione przez modele neuronowe zmienne kluczowe. Dominującą cechą różnicującą wspomniane klastry jest rozpuszczalność substancji leczniczej – jedna grupa substancji charakteryzowała się dobrą a druga niską rozpuszczalnością w wodzie. Można więc wnioskować, że za pomocą modelowania neuronowego wyselekcjonowano deskryptory dla substancji leczniczej, które w sposób bezpośredni korespondowały z jej rozpuszczalnością w wodzie. Potwierdza to obecność w zbiorze zmiennych kluczowych stałej Henry'ego. Rozważania te są przykładem zastosowania modelowania empirycznego w procedurach drażenia danych (ang. *data-mining*), gdzie modelowanie empiryczne staje się narzędziem do odkrywania zależności w dostępnych

danych w procesie automatycznej akwizycji wiedzy.

Kontynuując ten wątek badawczy opracowano model uwalniania *in vitro* peptydowych substancji leczniczych z mikrosfer na bazie kopolimeru kwasu mlekowego i glikolowego (PLGA) [P.7]. W oparciu o dane z piśmiennictwa utworzono bazę wiedzy uwzględniającą 68 formułacji mikrosfer, która liczyła 745 rekordów danych i 320 zmiennych. Zmienną wyjściową była ilość uwolnionej substancji leczniczej w zadanym na wejściu czasie. Podobnie jak w poprzedniej publikacji [P.1] zastosowano informatykę chemiczną do kodowania zarówno substancji pomocniczych jak i leczniczych. Wykorzystano pakiet Marvin firmy ChemAxon do obliczenia deskryptorów molekularnych bezpośrednio dla struktur peptydowych bez żadnych ich uproszczeń. Procedura selekcji zmiennych kluczowych została rozbudowana w stosunku do poprzednio używanych ANNs o pakiet fscaret środowiska R [37]. Selekcja zmiennych kluczowych doprowadziła do redukcji wektora wejściowego do 17 zmiennych. Analiza charakteru wejść dostarczyła istotnej wiedzy na temat czynników wpływających na proces uwalniania substancji peptydowych z mikrosfer PLGA, wśród których obok składu ilościowego i jakościowego znaleziono również zmienne opisujące technologię sporządzania mikrosfer. Obok ANNs jako narzędzie modelowania zastosowano również programowanie genetyczne (GP), które zalicza się do grupy tzw. obliczeń ewolucyjnych. Unikatową cechą GP jest zdolność do automatycznej konstrukcji modelu zagadnienia w postaci klasycznego równania matematycznego, co zostało wykorzystane do budowy równania opisującego uwalnianie substancji leczniczych z mikrosfer PLGA. Uniknięto problemu czarnej skrzynki charakterystycznego dla modeli neuronowych. Równocześnie, aby zredukować zapotrzebowanie GP na moc obliczeniową, zastosowano ANNs do oszacowania minimalnego błędu generalizacji możliwego do uzyskania w analizowanym zagadnieniu. Zredukowano w ten sposób przestrzeń poszukiwań modelu GP do modeli które posiadają błąd generalizacji równy lub powyżej wartości osiągniętych przez ANNs. Znaleziony model w postaci równania matematycznego (równanie 1) posiadał podobną sprawność predykcyjną jak ANNs, ale ze zredukowaną do 9 liczbą wejść. Model ten w zadowalający sposób pozwala przewidywać uwalnianie różnych substancji białkowych i peptydowych z mikrosfer PLGA na podstawie właściwości substancji leczniczych (zmienna *VI*) i pomocniczych (zmienne *V3*, *V14*), składu ilościowego (zmienne *V5*, *V7*), metody sporządzania mikrosfer (zmienna *VII*) i rozmiaru cząstek mikrosfer (zmienna *V8*).

$$Q = \frac{\sqrt[4]{V17}}{C2 \cdot \sqrt{C1}} + e^{\sqrt{V11+C3+e^{V5}}} - \frac{\sqrt{V7} \cdot \ln(V7) - V17 \cdot (1 + \sqrt[4]{V8} \cdot C4)}{V11 \cdot e^{\frac{V1 \cdot V8}{V3 + \ln(V6)} \ln 11 - \frac{V6}{V8 \cdot V14}}} \quad (1)$$

gdzie:

- Q – ilość uwolnionej substancji leczniczej (%)
- $V1$ – index Hyper-Wiener substancji leczniczej
- $V3$ – lepkość PLGA
- $V5$ – ilość alkoholu poliwinylowego (PVA) w matrycy mikrosfery (%)
- $V6$ – ilość PVA w fazie wodnej (%)
- $V7$ – ilość substancji leczniczej w mikrosferach (%)
- $V8$ – średnia wielkość cząstek mikrosfer (μm)
- $V11$ – metoda sporządzania mikrosfer
- $V14$ – logD plastyfikatora przy pH=1
- $V17$ – czas (dni)
- $C1$ do $C4$ – stałe równania

We współpracy z zespołem prof. Petera Kleinebudde z Uniwersytetu w Düsseldorfie prowadziłem ocenę procesu uwalniania diprofiliny z ekstrudatów lipidowych, gdzie zastosowałem połączenie dwóch narzędzi modelowania empirycznego: sztucznych sieci neuronowych (ANNs) i programowania genetycznego (GP) [P.3]. Każdemu z ww. narzędzi przyporządkowano inną funkcję:

- ANNs służyły do oceny minimalnego błędu generalizacji oraz jako narzędzia do selekcji zmiennych kluczowych,
- GP zastosowano do stworzenia końcowego modelu matematycznego zależności profilu uwalniania substancji leczniczej od czasu oraz charakterystyki formulacji.

Podstawę do programowania genetycznego stanowiły wyselekcjonowane za pomocą modelowania neuronowego zmienne kluczowe: średnica ekstrudatu i zmienna czasowa. W celu uproszczenia zadania zastosowano znany empiryczny model Weibulla jako szkielet dla finalnego modelu matematycznego procesu uwalniania (równanie 2)

$$Q = 100 * \left(1 - \exp \left[\frac{-t^A}{K} \right] \right) \quad (2)$$

gdzie:

Q – ilość uwolnionej substancji leczniczej w czasie (%)

t – czas

A, K – stałe równania Weibulla

Zadanie stworzenia empirycznego modelu uwalniania diprofiliny z ekstrudatów rozbito na dwa mniejsze problemy w których zmienną czasową pozostawiono jako element równania Weibulla, zaś jego współczynniki korelowano bezpośrednio z średnicą ekstrudatów za pomocą programowania genetycznego. Otrzymane równania cząstkowe zintegrowano na końcu w ostateczny model (równanie 3).

$$Q = 100 * \left(1 - \exp \left[\frac{-t e_5^{\sqrt{d+c_6}}}{\exp(c_1 \cdot d + c_2 \cdot d^{\ln(c_3 \cdot d)} + c_4)} \right] \right) \quad (3)$$

gdzie:

d – średnica ekstrudatu

c_{1-4} – stałe

W ostatnim etapie prac przetestowano model matematyczny z użyciem takiej samej metodyki jak w przypadku sztucznych sieci neuronowych i wykazano, że posiada on porównywalne zdolności generalizacyjne jak ANNs. Świadczy to o dużym stopniu uogólnienia wiedzy zgromadzonej przez narzędzie programowania genetycznego, a w związku z tym o możliwości zastosowania modelu matematycznego do przewidywania profilu uwalniania diprofiliny z ekstrudatów w prospektywnych pracach formułacyjnych.

Kontynuując prace nad procesem uwalniania diprofiliny z ekstrudatów lipidowych doprowadziłem do uproszczenia równania 3 do postaci przedstawionej w równaniu 4 (oznaczenia jak w równaniu 3) [P.8]:

$$Q = 100 \cdot \left(1 - \exp \left(\frac{-t^{c_3}}{c_1 \cdot d^d + c_2 \cdot d} \right) \right) \quad (4)$$

Równocześnie dzięki znacznym nakładom obliczeniowym stworzono całkowicie empiryczny model matematyczny (równanie 5)

$$Q = \frac{(\ln(d) + d + c_1) \cdot (t + c_2)}{t - \sqrt{t} + c_3 \cdot d^2} + d^2 \quad (5)$$

Modele opisane równaniami 3, 4 i 5 przetestowano na zmodyfikowanych zbiorach danych uczących. Modyfikacja polegała na zagęszczeniu siatki wyników poprzez ich interpolację przez uprzednio wytrenowane ANNs. Interpolacja zbiorów danych przez ANNs jest równoznaczna ze wzbogaceniem zbiorów danych w uogólnioną wiedzę modeli neuronowych. Tak przygotowane dane uczące podstawiano w miejsce oryginalnych zbiorów w procedurze 5-krotnego wzajemnego sprawdzania i obserwowano różnice w błędach generalizacji. Wykazano, że modele bazujące na równaniu Weibulla zyskują na uczeniu na tak wzbogaconych danych i ich błąd generalizacji spada w porównaniu do otrzymanego na danych oryginalnych. W przypadku równania 5 nie zaobserwowano tego efektu, jako że model ten wykazywał niskie błędy generalizacji bez żadnych dodatkowych modyfikacji zbiorów uczących. Tak więc zastosowanie znacznych mocy obliczeniowych do modelowania za pomocą GP jest gwarancją wysokiej jakości otrzymanych modeli, jednak przy większych bazach danych i stopniu komplikacji problemu może okazać się że techniki preprocessingu danych z użyciem sieci neuronowych mogą pozwolić na zmniejszenie kosztu obliczeniowego budowy modelu bez utraty jego sprawności.

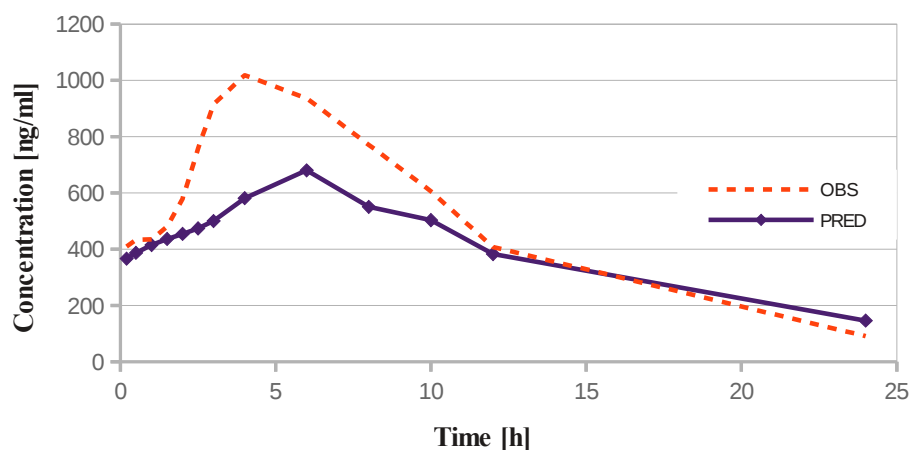
Narzędzia modelowania empirycznego zostały zastosowane również do oceny właściwości peletek sporządzonych metoda ekstruzji i sferonizacji z zastosowaniem różnych substancji leczniczych i pomocniczych [P.2]. Wyniki badań dla 227 formułacji peletek stanowiły podstawę do stworzenia bazy danych zawierającej ok. 3000 zmiennych wejściowych (cech) oraz dwie zmienne wyjściowe: współczynnik kształtu (ang. *aspect ratio* AR) i średni czas uwalniania (ang. *mean dissolution time*, MDT). Tak duża początkowa ilość cech była wynikiem zastosowania narzędzi informatyki chemicznej do opisu składu jakościowego formułacji poprzez wyliczenie deskryptorów molekularnych dla substancji leczniczych i pomocniczych. W celu zredukowania liczby zmiennych wejściowych zastosowano analizę wrażliwościową grupy modeli neuronowych metodą pochodnych

wg Engelbreck i wsp. [9]. Z uwagi na obecność w bazie danych ponad 10-krotnie większej liczby cech niż rekordów danych, napotkałem na trudności w znalezieniu punktów redukcji liczby cech w otrzymanym rankingu zmiennych wejściowych. Jako rozwiązanie zaproponowałem modyfikację metody Engelbreck i wsp. polegającą na wykorzystaniu znajomości klasy deskryptorów molekularnych i wyborze tylko jednego przedstawiciela danej klasy do ostatecznego wektora wejściowego. Modelowanie predykcyjne prowadzono za pomocą klasycznych sieci neuronowych typu perceptronu wielowarstwowego (MLP) oraz systemów neuronowo-rozmytych (NF) uruchamianych z wykorzystaniem własnoręcznie napisanego oprogramowania (Nets2010). Systemy NF zostały wybrane w celu stworzenia bazy reguł logicznych jako próby ominięcia problemu czarnej skrzynki występującego w przypadku ANNs. Logika rozmyta to zaproponowany przez Lotfi Zadeha [46] kompletny system logiczny opierający się na teorii prawdopodobieństwa oraz zbiorów, który po implementacji jako hybryda w postaci struktury sieci neuronowej (NF) jest samoadaptacyjnym narzędziem analizy danych zdolnym do identyfikacji badanego zagadnienia w postaci tabeli reguł logicznych. Jako, że systemy NF posiadają pewne ograniczenia związane z ich klasyfikacyjnym charakterem pracy, znacznie bardziej od nich elastyczne ANNs typu MLP zostały znów zastosowane w roli narzędzi odniesienia w celu stwierdzenia czy NF są w stanie osiągnąć wystarczająco dobre zdolności generalizacyjne. W przypadku modeli dla MDT wykazano porównywalne zdolności generalizacyjne modeli MLP i NF. W procedurze modelowania predykcyjnego wyselekcjonowano optymalne architektury modeli MLP i NF, a następnie zastosowano je w procedurach drążenia danych (ang. *data-mining*) do odkrywania zależności pomiędzy zmiennymi wejściowymi a wyjściowymi. W przypadku sieci MLP zastosowano specjalnie przygotowane zbiory testowe które pozwalały na identyfikację wpływu wybranych zmiennych wejściowych na zmienne wyjściowe. Modele NF, analizowano poprzez bezpośredni odczyt reguł logicznych w postaci JEŻELI ... TO W celu uproszczenia tabeli reguł dużych modeli NF posłużono się metodą hierarchicznej klasteryzacji do agregacji zbliżonych do siebie reguł a następnie trenowano modele NF na tak stworzonych prototypach reguł logicznych. Zidentyfikowano szereg zależności pomiędzy MDT i AR a składem i technologią sporządzania formulacji. Wykazano, że MDT jest ujemnie skorelowane z temperaturą i czasem suszenia oraz że zależność ta ma charakter asymptotyczny, tj. po przekroczeniu pewnych wartości progowych zmiany MDT nie są już znaczące. Znalaziono również dodatnią korelację MDT z zawartością fosforanu wapnia w formulacji.

2.4.2. Empiryczne modelowanie zależności *in vitro in vivo* (IVIVR).

Zagadnienia zależności *in vitro in vivo* (IVIVR) oraz korelacji *in vitro in vivo* (IVIVC) są od wielu lat przedmiotem dyskusji ze względu na konsekwencje zwłaszcza w odniesieniu do opracowywanych produktów leczniczych. IVIVC zostało wprowadzone do praktyki farmaceutycznej przez FDA w latach 90 ubiegłego stulecia [14] i aktualnie zostało rozszerzone do IVIVR, w której dopuszczalne jest stosowanie modeli nieliniowych [35]. W publikacji P.5 przedstawiłem próbę stworzenia uogólnionego modelu IVIVR służącego wstępnej ocenie dostępności biologicznej substancji leczniczej ze stałych postaci leku na podstawie ich składu jakościowego i ilościowego oraz wyników badania dostępności farmaceutycznej. W celu uproszczenia zagadnienia założono, że model dotyczy będzie stałych postaci leku o natychmiastowym lub przedłużonym uwalnianiu substancji leczniczej. W tym ostatnim przypadku zawężono zakres prac do układów matrycowych. Kolejnym założeniem było wykorzystanie wyników badań *in vitro* prowadzonych w aparatach łopatkowym i koszyczkowym. Po przeglądzie piśmiennictwa zebrano bazę danych dla 93 formułacji spełniających ww. kryteria. Zakodowano skład ilościowy i jakościowy analizowanych form leku, a także uwzględniono wyniki badania uwalniania substancji leczniczych. Do kodowania składu jakościowego wykorzystano deskryptory molekularne obliczone za pomocą programu Marvin (Chemaxon, Węgry). Baza danych obejmowała 307 zmiennych wejściowych oraz jedną zmienną wyjściową reprezentującą stężenie leku we krwi pacjenta. Do modelowania wykorzystano ANNs zarówno jako finalne modele jak też i narzędzia analizy wrażliwościowej prowadzonej w celu selekcji zmiennych kluczowych. Wynikiem tej procedury była selekcja 28 zmiennych kluczowych, wśród których obok składu ilościowego i jakościowego formułacji znalazły się dwa punkty czasowe z profilu uwalniania *in vitro*. W celu poprawy zdolności generalizacyjnych modelu zastosowano komitet ekspertów złożony z 4 sieci neuronowych, wśród których jedna posiadała 7 warstw ukrytych. Pozwala to zaliczyć wyżej wymieniony model do narzędzi tzw. uczenia głębokiego (ang. *deep learning*) [3]. Selekcja tak skomplikowanej ANNs, jako jednego z najlepiej generalizujących modeli świadczy dobitnie o stopniu złożoności analizowanego zagadnienia. Wykazano, że model posiada zdolność do generalizacji posiadanej bazy wiedzy na nieznane mu struktury chemiczne, co potwierdzono na przykładzie postaci leku o natychmiastowym uwalnianiu montelukastu sodowego. Na rycinie 2 przedstawiono wyniki generalizacji profilu farmakokinetycznego montelukastu sodowego, którego nie było w oryginalnej bazie danych uczących ANNs. Widać dobre odwzorowanie ogólnego kształtu profilu, w tym zwłaszcza fazy eliminacji. System neuronowy niedoszacowuje C_{max} i całej

fazy wchłaniania. Większa precyzja przewidywania t_{max} i fazy eliminacji niż C_{max} i fazy absorpcji została również potwierdzona w trakcie 10-krotnego wzajemnego sprawdzania i stanowi charakterystyczną cechę otrzymanego modelu.

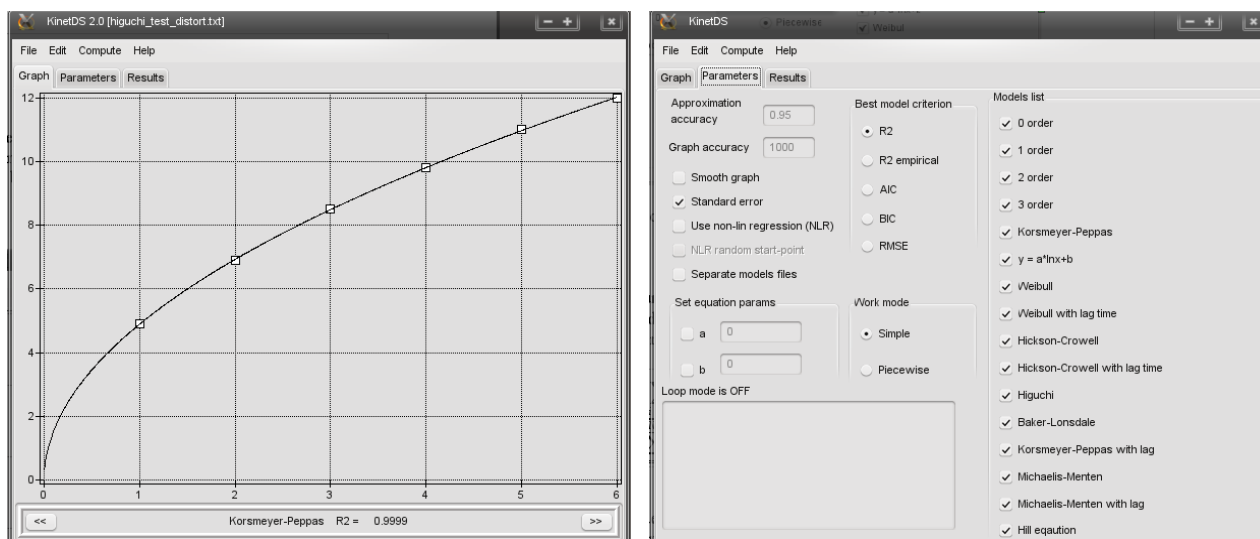


Ryc. 2. Predykcja profilu farmakokinetycznego montelukastu sodowego.

2.4.3. Konstrukcja i publikacja programów komputerowych do zautomatyzowanego modelowania empirycznego

Oprogramowanie do matematycznego opisu krzywych profili uwalniania

Program komputerowy *KinetDS* do matematycznego opisu krzywych profili uwalniania został przeze mnie napisany w języku ObjectPascal i opublikowany na licencji GNU GPL [P.4, 17].



Ryc. 3. Elementy graficznego interfejsu użytkownika programu KinetDS

Sposób pracy programu *KinetDS* opiera się na przeglądzie biblioteki dostępnych modeli i próbie znalezienia najlepiej dopasowanego modelu ze względu na określone przez użytkownika kryterium dopasowania:

- współczynnik determinacji (R^2),
- empiryczny współczynnik determinacji (R^2_{emp}),
- kryterium Akaike (AIC),
- kryterium Schwarza (BIC),
- spierwiastkowany błąd średniokwadratowy (RMSE).

W najnowszej wersji 3.0 programu *KinetDS* dostępnych jest 16 modeli matematycznych, a w tym modele:

- kinetyczne od 0 do 3 rzędu,
- Kormsmeiera-Peppasa,
- Higuchiego,
- Hicksona-Crowella,
- Weibulla,
- Hilla,
- Michaelisa-Menten.

Większość modeli wyposażono w dodatkowy parametr czasu opóźnienia (lag time) definiowany

jako czas od 0 do pierwszego zmierzonego punktu uwalniania.

O ile jest to możliwe dobór współczynników modeli odbywa się dwuetapowo:

1. za pomocą regresji liniowej w sytuacji gdy dostępna jest liniowa transformacja danego modelu,
2. za pomocą regresji nieliniowej prowadzonej metodą pełzającego sympleksu.

Punktami startowymi do regresji nieliniowej są wartości współczynników dobrane w regresji liniowej, o ile mogła ona być uprzednio przeprowadzona. Opisana sekwencja modelowania znacznie przyspiesza cały proces dzięki wstępnemu przybliżeniu optymalnych wartości parametrów przez regresję liniową. Parametr czasu opóźnienia dobierany jest iteracyjnie poprzez analizę dopasowania modelu przy danej wartości czasu opóźnienia, co stanowi kolejny element empirycznego modelowania zaimplementowany w *KinetDS*.

Oprócz metod z wykorzystaniem modeli matematycznych zaimplementowano również dwie bezmodelowe metody oceny profilu uwalniania:

- średni czas uwalniania (MDT),
- efektywność uwalniania (DE).

Dodatkową funkcjonalnością *KinetDS* jest tryb regresji przedziałowej realizowany również w sposób empiryczny poprzez próby identyfikacji punktów na profilu uwalniania, gdzie dochodzi do zmiany mechanizmu lub charakteru procesu uwalniania. Polega to na dobieraniu modeli do fragmentów profilu uwalniania, które to fragmenty powstają poprzez stopniowe dodawanie punktów na profilu i obserwację dopasowania dostępnych modeli. Rozbudowa każdego fragmentu profilu prowadzona jest zgodnie kierunkiem osi czasu, zaś decyzja o zakończeniu danego fragmentu i rozpoczęciu nowego podejmowana jest na podstawie kryterium dokładności aproksymacji będącego w swej istocie dopuszczalnym procentowym spadkiem wartości dopasowania najlepszego modelu po dodaniu kolejnego punktu czasowego na profilu. Najczęściej stosuje się wartości tego kryterium od 95% do 99,9%, co oznacza tolerancję wybranej miary dopasowania modelu na poziomie 95-99.9%. Dobór tego parametru odbywa się w sposób czysto empiryczny na podstawie eksperymentów numerycznych.

Program *KinetDS* wyposażono w procedury automatycznej analizy wielu profili w jednym przebiegu (tryb wsadowy) oraz rozbudowane raportowanie wraz ze statystykami dla każdego dopasowywanego modelu. Program dostępny jest dla środowisk Windows i Linuks i od początku swojej publikacji w sieci Internet został pobrany 5 976 razy (dane na 01.02.2016).

Oprogramowanie do oceny podobieństwa krzywych za pomocą współczynnika f_2 przy dużej zmienności profili uwalniania

Podstawową miarą podobieństwa profili uwalniania substancji leczniczej *in vitro* jest tzw. współczynnik podobieństwa f_2 (równanie 6), zaproponowany przez Moore'a i Flannera [28]. Jest to wygodna i prosta do interpretacji miara, której zakres mieści się pomiędzy 0 a 100. Wartości powyżej 50 uznawane są zarówno przez FDA jak i EMA za potwierdzenie podobieństwa profili uwalniania substancji leczniczej [7, 13].

$$f_2 = 50 \cdot \log \left[\frac{100}{1 + \frac{\sum_{i=1}^n (T_i - R_i)^2}{n}} \right] \quad (6)$$

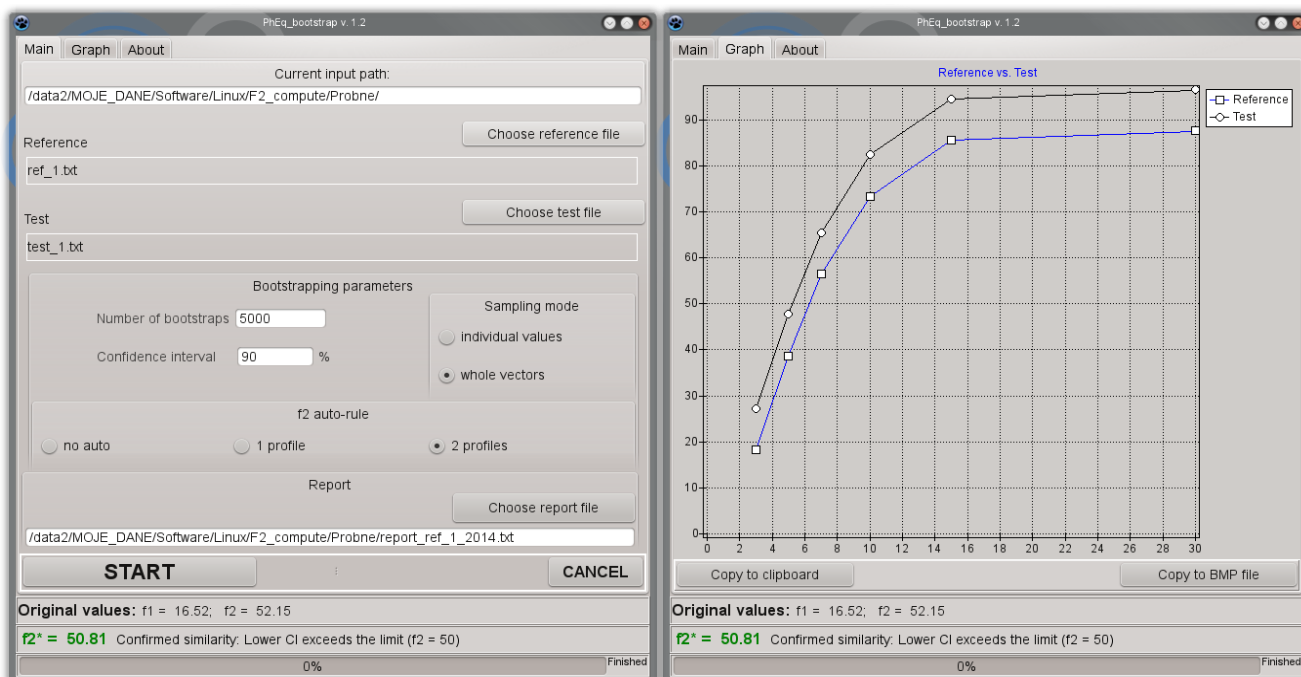
gdzie:

T – ilość uwolnionej substancji leczniczej z produktu testowego

R – ilość uwolnionej substancji leczniczej z produktu referencyjnego (oryginalnego)

i – punkt pomiarowy

n – liczba punktów pomiarowych



Ryc. 4. Elementy graficznego interfejsu użytkownika programu PhEq_bootstrap.

Zgodnie z wytycznymi FDA i EMA [7, 13] współczynnik f_2 stosowany jest dla profili uwalniania charakteryzujących się małą zmiennością wyników w obrębie badanej próby. Za wartości graniczne uznano poziom względnego odchylenia standardowego (RSD) równy 10% dla wszystkich punktów czasowych z wyjątkiem pierwszego, dla którego kryterium to wynosi 20%. Shah i wsp. [39] zaproponowali metodę pozwalającą na stosowanie współczynnika f_2 do porównywania profili uwalniania o dużej zmienności wykraczającej poza ww. kryteria RSD. Metoda ta oparta jest na scenariuszu „najgorszego przypadku”, w którym szacuje się dolną granicę przedziału ufności wartości oczekiwanej współczynnika f_2 na poziomie istotności 5%. W tym celu wykorzystuje się metodę bootstrappingu, tj. numerycznego zwielokrotniania wyników eksperymentów laboratoryjnych poprzez losowanie ze zwracaniem. Numerycznie rozbudowana pula profili uwalniania stanowi podstawę do analizy przedziału ufności dla wartości oczekiwanej f_2 . Zaprojektowałem i napisałem program komputerowy *PhEq_bootstrap* (ryc. 4) do implementacji przedstawionej metody [P.6]. Program ten podobnie jak *KinetDS* został napisany w języku ObjectPascal i opublikowany na licencji GNU GPLv3 [17, 31]. Oprogramowanie posiada zaimplementowaną metodę bootstrappingu do oceny rozkładu wartości oczekiwanej współczynnika f_2 wraz z analizą przedziału ufności z zadaniem przez użytkownika prawdopodobieństwem. Jako wartość do decyzji o podobieństwie profili podaje się dolną granicę przedziału ufności. Możliwa jest kontrola ilości powtórzeń procedury losowania, która wg różnych źródeł powinna wynosić od 500 do 5000 [39]. Wobec braku teoretycznych przesłanek co do ilości powtórzeń, parametr ten dobiera się empirycznie poprzez ocenę stabilności przedziału ufności w zależności od wybranej ilości powtórzeń. Jako dodatkową funkcjonalność *PhEq_bootstrap* zaproponowałem dwa tryby pracy procedury bootstrappingu. W pierwszym trybie numeryczne generowanie próby badanej oparte jest na losowaniu całych profili uwalniania, zaś w drugim trybie profile uwalniania konstruowane są od nowa poprzez wybór pojedynczych punktów losowanych z dostępnej puli. *PhEq_bootstrap* w swojej najnowszej wersji 1.2, wyposażylem w możliwość testowania różnych kryteriów przekraczania wartości 85% ilości uwolnionej substancji leczniczej w tym nawet wyłączenia tego kryterium. Program posiada rozbudowany moduł raportowania, gdzie każde losowanie jest przedstawione indywidualnie wraz z cząstkowymi profilami dla danego losowania. Zebrane w raporcie wyniki mogą być łatwo importowane do innych narzędzi obliczeniowych i stanowić podstawę do walidacji oprogramowania zgodnie ze standardami przemysłowymi. [12]. Program dostępny jest dla środowisk Windows i Linuks i od początku swojej publikacji w sieci

został pobrany 1 655 razy (dane z dnia 01.02.2016).

Oprogramowanie do modelowania zależności *in vitro in vivo* (IVIVR)

Osiągnięcie poziomu A korelacji *in vitro in vivo* (IVIVC) wiąże się z zastosowaniem metody konwolucji i/lub dekonwolucji:

- konwolucja: $r(t) = i(t) * w(t)$
- dekonwolucja: $i(t) = r(t) // w(t)$

gdzie:

$r(t)$ – profil *in vivo* po podaniu doustnym

$i(t)$ – uwalnianie substancji leczniczej w warunkach *in vivo*

$w(t)$ – profil *in vivo* po podaniu dożylnym

Ograniczeniem metod konwolucji i dekonwolucji jest konieczność posiadania profilu *in vivo* $w(t)$ po podaniu dożylnym. W celu ominięcia tego problemu zaproponowałem nową metodę budowania modeli IVIVR które odpowiadają IVIVC na poziomie A [P.9]. Istotą działania omawianego rozwiązania jest numeryczne generowanie krzywej $w(t)$ na podstawie posiadanych danych $r(t)$ i $i(t)$. Realizowane jest to poprzez proces optymalizacji hipotetycznej krzywej przebiegu zmian stężenia leku we krwi $w(t)$ po podaniu dożylnym. Kryterium optymalizacji $w(t)$ jest zbliżenie się do profilu farmakokinetycznego leku po podaniu doustnym $r(t)$ po konwolucji krzywej $w(t)$ z profilem uwalniania substancji leczniczej w warunkach *in vitro*. Ta ostatnia jest więc traktowana jako bezpośredni odpowiednik profilu uwalniania substancji leczniczej w warunkach *in vivo* $i(t)$. W tak zaprojektowanej procedurze optymalizowane są bezpośrednio punkty na krzywej $w(t)$, a sama procedura może przebiegać z wykorzystaniem dowolnego narzędzia zdolnego do nieliniowej optymalizacji wielowymiarowych funkcji celu. Narzędzie realizujące powyższe zadania nazwano *RIVIVR* i jest dostępne poprzez serwis sourceforge [38]. Zostało ono napisane w środowisku statystycznym R z wykorzystaniem mojej własnej, kaskadowej metody optymalizacji bazującej na sekwencyjnej pracy narzędzi optymalizacyjnych środowiska R: *GenSA*, *rgenoud* i *optim* (metody *NM* i *BFGS*) [37]. Ideą stworzonego algorytmu bezpośredniej optymalizacji krzywej $w(t)$ jest wykorzystanie narzędzi optymalizacyjnych o różnym charakterze w ściśle ustalonej kolejności, tak aby najpierw proces prowadzony był z wykorzystaniem optymalizatorów globalnych, bazujących na algorytmach bezgradientowych (*GenSA*, *NM*, *rgenoud*), a potem dopiero optymalizatorów

lokalnych gradientem funkcji celu (*rgenoud*, *BFGS*).

GenSA >> *rgenoud* >> *NM* >> *BFGS*

Narzędzie *rgenoud* zostało wymienione dwa razy z uwagi na to, że łączy w sobie cechy optymalizatora globalnego bazującego na algorytmie genetycznym i lokalnego na metodzie *BFGS*. W toku prac eksperymentalnych udało się uprościć powyższy system i zredukować liczbę optymalizatorów bez strat dla uzyskanego wyniku do dwóch:

GenSA >> *BFGS*

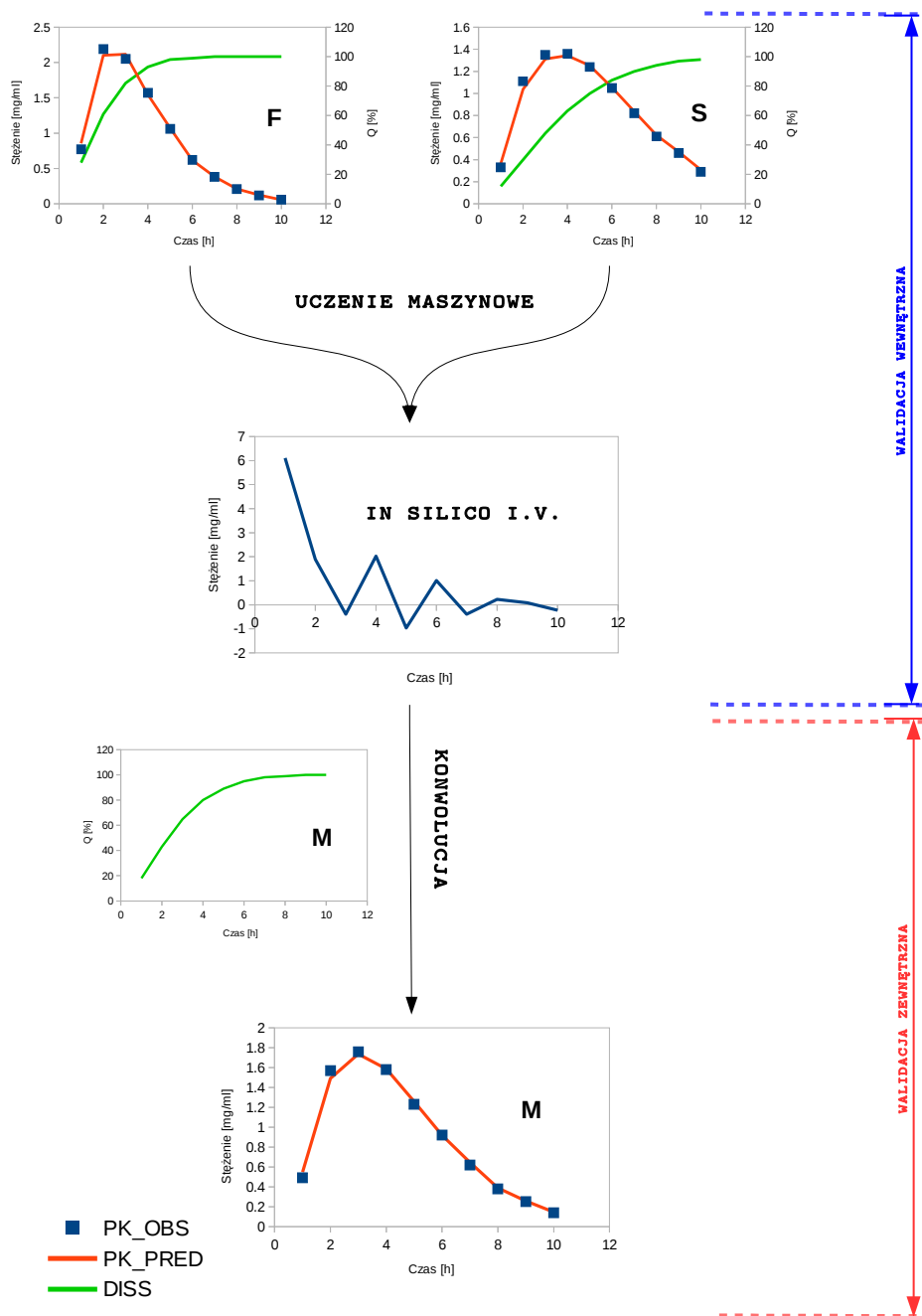
Dzięki temu przy standardowych ustawieniach czas pracy *RIVIVR* na nowoczesnym komputerze nie przekracza 5 min. co dla wspomnianych algorytmów iteracyjnych jest dużym osiągnięciem.

Ogólny schemat pracy *RIVIVR* (ryc. 5.) polega na wykorzystaniu technik uczenia maszynowego do wygenerowania sztucznego profilu i.v. i zastosowaniu go w procedurze konwolucji do przewidywania profilu farmakokinetycznego po podaniu p.o. dla nowej formułacji. *RIVIVR* został przetestowany na danych z piśmiennictwa, w których dostępne były wyniki badań *in vitro* i *in vivo* dla tej samej postaci leku [P.9]. Zastosowano taki sam schemat oraz kryteria walidacji jak zaleca FDA w swojej instrukcji [14]. Spośród ośmiu zestawów danych przedstawionych do analizy, *RIVIVR* uzyskał zadowalający wynik walidacji zewnętrznej w pięciu przypadkach tj. 62,5%, przy czasie pracy poniżej 4 min. na profesjonalnej stacji roboczej. Aby ocenić wartość tego wyniku porównano go z innymi narzędziami modelowania empirycznego zdolnymi do stworzenia bezpośredniej zależności profilu *in vivo* z *in vitro*. W szczególności wykorzystano programowanie genetyczne, którego wynikiem pracy były formuły matematyczne reprezentujące bezpośrednie mapowanie IVIV. Z użyciem GP udało się uzyskać wynik na poziomie 75% przypadków spełniających kryteria FDA, jednak wynik ten jest rezultatem ciągłej pracy 24-wątkowej stacji roboczej przez 2 tygodnie podczas kiedy dla *RIVIVR* wystarczyło 8 wątków przez 4 minuty. Należy również wspomnieć, że do modelowania GP zastosowano wzbogacony zbiór danych o parametry modeli kinetycznych (0-3 rząd), mechanistycznych (Korsmeyer-Peppas) i empirycznych (Weibull) opisujących krzywe uwalniania *in vitro*.

Tak znaczna dysproporcja zasobów niezbędnych do stworzenia modelu IVIVR w stosunku do GP skłania do stwierdzenia, że *RIVIVR* może stanowić przede wszystkim narzędzie do szybkiego skryningu możliwości budowy IVIVR w oparciu o posiadane dane. Niepowodzenie modelowania za pomocą *RIVIVR* jest wskazówką do szukania rozwiązań w trzech kierunkach:

- powrót do klasycznej IVIVC, jeśli to możliwe,
- stosowanie bezpośredniego mapowania np. za pomocą GP,
- zmiana metody badania uwalniania.

Ostatni przypadek nasuwa interpretację wyników z RIVIVR, w której mogą one stanowić wskazówkę do oceny biologicznego kontekstu badań *in vitro*, tj. dać szybką odpowiedź na pytanie czy prowadzone daną metodą badania *in vitro* mogą w jakikolwiek sposób być korelowane z profilem farmakokinetycznym. Uzasadnieniem powyższego twierdzenia jest fakt, że wygenerowany przez RIVIVR profil $w(t)$ nie ma już fizycznego charakteru profilu farmakokinetycznego (ryc. 5), a jest raczej zbiorem punktów mapujących w procedurze konwolucji profil uwalniania *in vitro* na profil $r(t)$. Czasem można wręcz zaobserwować, że niektóre punkty z $w(t)$ otrzymanego przez RIVIVR mają wartości ujemne. Elastyczność $w(t)$ otrzymanego przez RIVIVR zmniejsza wymagania do profilu *in vitro*, podczas gdy stosowanie zmierzonego profilu $w(t)$ zmusza do maksymalnego zbliżenia profilu *in vitro* do fizycznego $i(t)$. W związku z powyższym niepowodzenie modelowania RIVIVR wskazuje na to, że wyniki badań *in vitro* trudno będzie korelować innymi metodami i należałoby rozważyć ewentualną zmianę metody badań *in vitro* lub jej parametrów.



Ryc. 5. Schemat pracy RIVIVR. PK_OBS – obserwowane wyniki dla profilu farmakokinetycznego; PK_PRED – przewidywane wyniki dla profilu farmakokinetycznego; DISS – profil uwalniania substancji leczniczej; F/S/M – postać o szybkim/wolnym/pośrednim uwalnianiu substancji leczniczej

Testowanie tego typu hipotez ułatwia dodatkowa funkcjonalność *RIVIVR*, w której zamiast domyślnie stosowanych losowych punktów startowych dla $w(t)$ można wstawić zdefiniowaną przez użytkownika krzywą. Wykorzystując podany profil jako startowy, *RIVIVR* dokona próby jego modyfikacji w celu uzyskania lepszej korelacji IVIV. Porównanie oryginalnego profilu startowego z

wynikowym z *RIVIVR* może dostarczyć cennych informacji na temat posiadanych danych oraz skuteczności metody badania *in vitro*.

Podsumowanie podstawowych cech *RIVIVR*:

- zdolność do korelacji *in vitro in vivo* bez znajomości danych z podania dożylnego,
- brak wymagań dot. ilości formułacji niezbędnych do walidacji wewnętrznej – można używać nawet jednej formułacji,
- praca w środowisku R – uznanym i dostępnym za darmo środowisku statystycznym [37],
- dobra powtarzalność otrzymywanych wyników,
- licencja GPL [17], która zapewnia wolny dostęp do oprogramowania nawet do celów komercyjnych.

3. Podsumowanie

Modelowanie empiryczne pozwala na analizę zagadnień, w których zastosowanie innych metod jest trudne lub wręcz niemożliwe. W niniejszym opracowaniu przedstawiono modele umożliwiające ocenę wielowymiarowych zależności pomiędzy składem jakościowym i ilościowym formułacji stałych postaci leku a dostępnością farmaceutyczną i/lub biologiczną. Dysponując danymi zebranymi z piśmiennictwa i/lub prac eksperymentalnych wykonanych w Katedrze, bez żadnych założeń teoretycznych wyselekcjonowano zmienne kluczowe i zbudowano modele predykcyjne analizowanych zagadnień [P.1, P.2, P.3, P.5]. Podstawową rolę w budowie omawianych modeli odegrała informatyka chemiczna, dzięki której możliwe stało się przewidywanie charakterystyki formułacji zawierających zupełnie nowe dla systemu substancje lecznicze lub pomocnicze [P.1, P.2, P.3, P.5].

W celu uniknięcia problemu „czarnej skrzynki” zastosowano systemy neuronowo-rozmyte i programowanie genetyczne (GP). Łatwość interpretacji wiedzy systemu jest szczególnie wyrażona w automatycznej budowie równań matematycznych przez GP [P.3].

Mimo niewątpliwych zalet modelowania empirycznego, stopień komplikacji narzędzi do jego realizacji wciąż utrudnia szerokie stosowanie omawianych technik. Aktualnie trwają prace nad udostępnieniem maksymalnie zautomatyzowanych pakietów oprogramowania wymagających

minimalnej interwencji użytkownika jak np. moduły sztucznych sieci neuronowych z automatycznym doбором architektury obecne w dużych pakietach statystycznych (*Statistica*, *SPSS*). W powyższe tendencje dobrze wpisują się prace nad oprogramowaniem do automatycznego dopasowywania modelu matematycznego procesu uwalniania substancji leczniczej (*KinetDS*, **P.4**), oceny równoważności profili uwalniania dla dwóch leków w sytuacji dużych zmienności wyników badań *in vitro* (*PhEq_bootstrap*, **P.6**) oraz korelacji *in vitro in vivo* (*RIVIVR*, **P.9**). Wszystkie moje programy komputerowe cechuje wysoki stopień automatyzacji oraz łatwa dostępność dzięki publikacji na licencji GPL. Dostępność kodu źródłowego jest też podstawą do przeprowadzenia pełnej walidacji oprogramowania.

W prezentowanym cyklu prac stanowiących podstawę habilitacji można wskazać dwie grupy osiągnięć:

1. w zakresie nauk podstawowych:

- ocena czynników mających kluczowe znaczenie dla procesu uwalniania substancji leczniczych ze stałych postaci leku
- ocena czynników mających kluczowe znaczenie dla parametrów fizycznych stałych postaci leku
- stworzenie autorskiej metody numerycznej do dekonwolucji przez konwolucję
- stworzenie autorskiej metody do budowy modeli IVIVR odpowiadających IVIVC na poziomie A bez konieczności znajomości profilu farmakokinetycznego po podaniu dożylnym

2. w zakresie nauk stosowanych:

- opracowanie modeli predykcyjnych procesu uwalniania substancji leczniczych ze stałych postaci leku
- napisanie i publikacja narzędzi komputerowych do budowy ww. modeli
- stworzenie oprogramowania wykorzystującego metodę generowania profilu i.v. w celu realizacji IVIVR analogicznej do IVIVC poziomu A ale bez konieczności posiadania danych z podania dożylnego

4. Pozostałe osiągnięcia naukowe

4.1. Działalność naukowo-badawcza przed uzyskaniem stopnia doktora

W pracach naukowych prowadzonych przed uzyskaniem stopnia doktora skupiłem się przede wszystkim na opracowaniu metodyki zastosowania modelowania neuronowego w analizie wielowymiarowych zbiorów danych oraz na budowie bazy software'owej do modelowania neuronowego. Prace programistyczne były rozwinięciem projektu rozpoczętego jeszcze w trakcie koła naukowego przy Zakładzie Farmakokinetiki i Farmacji Fizycznej UJCM w roku 1995. i zaowocowały symulatorem sieci neuronowych, który stał się podstawowym narzędziem modelowania wykorzystanym zarówno w pracy magisterskiej, jak też i doktorskiej, po jego gruntownej rozbudowie. Symulator ten rozwijany był do roku 2013 i liczy ponad 15 000 linii kodu. Dodatkowo zaprojektowałem i napisałem szereg narzędziowych programów komputerowych służących automatyzacji procesu przygotowywania zbiorów danych do uczenia i testowania modeli, jak też i obróbki wyników modelowania, np. analizy zdolności klasyfikacyjnych modeli neuronowych. Stworzyłem modele oparte na sztucznych sieciach neuronowych do przewidywania szybkości rozpuszczania substancji leczniczych ze stałych rozproszeń [27]. Była to demonstracja zdolności sieci neuronowych do integracji w jednym modelu szeregu zróżnicowanych danych wejściowych, jak np. skład ilościowy i jakościowy formulacji – ten ostatni zakodowano używając indeksu konektywności (ang. *connectivity index*), tj. deskryptora molekularnego określającego stopień złożoności połączeń pomiędzy atomami w analizowanej cząsteczce [22].

ANNs zastosowałem również z powodzeniem w epidemiologii do prac nad przewidywaniem ryzyka cukrzycy ciężarnych na podstawie prostego kwestionariusza, które to narzędzie wbudowałem w opracowany przeze mnie formularz na stronie internetowej korzystając z technologii apletów Javy [33]. Prace te rozszerzone zostały na problem nadciśnienia z wykorzystaniem analogicznych narzędzi modelowania i internetowej publikacji stworzonych modeli [34].

4.2. Działalność naukowo-badawcza po uzyskaniu stopnia doktora

4.2.1. Technologie informatyczne

Po uzyskaniu stopnia doktora kontynuowałem prace nad modelowaniem komputerowym i empirycznym w naukach farmaceutycznych. Z uwagi na iteracyjny charakter pracy

wykorzystywanych narzędzi modelowania i ich zapotrzebowanie na duże moce obliczeniowe, rozpocząłem adaptację swoich narzędzi i metodyki pracy do wykorzystania w profesjonalnych środowiskach HPC (ang. *high performance computing*), gdzie jako podstawowy system operacyjny wykorzystywane są różne dystrybucje Linuksa [43]. W związku z tym przetłumaczyłem stworzoną w trakcie pracy doktorskiej bazę programową na oprogramowanie zgodne ze środowiskiem Linuks. Rozpocząłem również prace nad systemem zarządzania obliczeniami w środowisku gridowym wykorzystującym model obliczeń równoległych na zestawie dedykowanych stacji roboczych. Prace nad własnym systemem obliczeniowym zostały częściowo zainspirowane dzięki mojemu uczestnictwu w programie EGEE (Enabling Grids for E-science) zarządzanym przez ośrodek CERN [8].

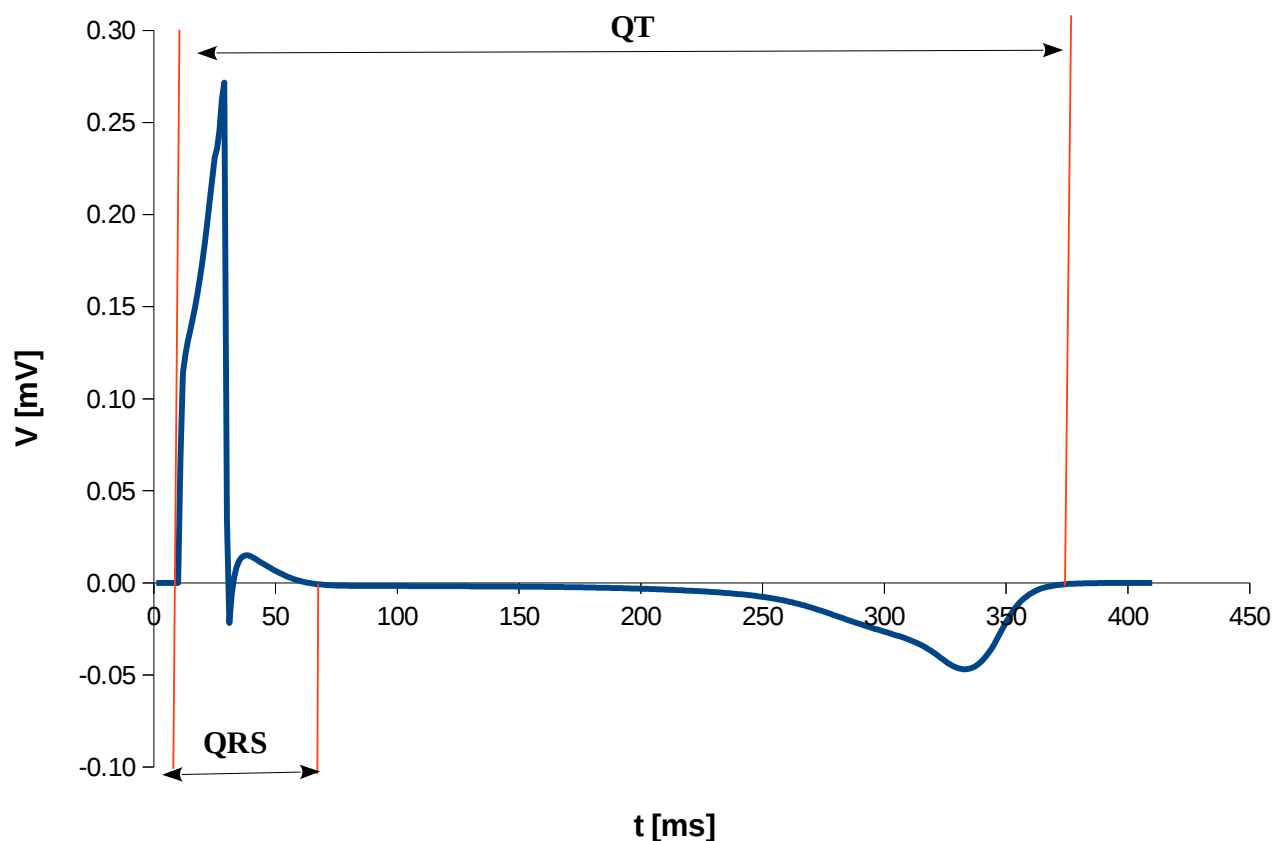
Kontakt ze środowiskiem Wolnego Oprogramowania (ang. *Open Source*) stał się również bodźcem do eksploracji szeregu narzędzi informatyki chemicznej poszerzających możliwości modelowania poza stosowany wcześniej indeks konektywności: *Marvin*, *Padel*, *CDK*, *EPI*, *Gamess*.

Podczas pracy z oprogramowaniem *Open Source* zwróciłem uwagę na środowisko statystyczne R [37], którego aktualnie jestem zarówno użytkownikiem jak i współtwórcą. Dzięki bardzo aktywnej społeczności R i dostępności ponad 8000 dodatkowych modułów rozszerzeń, znacząco powiększyłem zakres wykorzystywanych narzędzi modelowania począwszy od specyficznych typów sieci neuronowych (*monmlp*), przez drzewa decyzyjne (*randomForest*), systemy oparte na regułach logicznych (*Cubist*, *fugeR*), a na obliczeniach ewolucyjnych skończywszy (*rgp*). Okazało się również, że środowisko R znakomicie nadaje się do obliczeń w systemach gridowych i łatwo integruje się z rozwijanym przeze mnie systemem zarządzania obliczeniami rozproszonymi. Stanowiło to podstawę do włączenia R do aktualnej struktury systemu obliczeniowego poprzez konstrukcję własnych skryptów R implementujących ww. narzędzia [38].

4.2.2. Ocena *in silico* kardiotoksyczności substancji leczniczych

W trakcie prac nad systemem *Cardiac Safety SimulatorTM* (CSS, Certara) w zespole kierowanym przez dr Sebastiana Polaka opracowałem szereg modeli empirycznych wpływu substancji chemicznych na prądy jonowe w błonie kardiomiocytów [32, 44]. Modele te oparte zostały na informacji o strukturze chemicznej substancji chemicznych kodowanej w postaci deskryptorów

molekularnych i umożliwiają wstępną ocenę potencjalnego ryzyka arytmogennego dla substancji, które nawet nie zostały jeszcze zsyntetyzowane. W szczególności opracowałem i wdrożyłem modele predykcyjne dla kanałów potasowych (IKr, IKs), wapniowych (ICa) i sodowych (INa). W czasie powstawania oprogramowania było to jedyne tak kompletne rozwiązanie wdrożone do praktyki. Mój udział w powstawaniu systemu CSS polegał również na analizie sygnałów elektrofizjologicznych i zależności mechanicznych generowanych przez symulator. Prace dotyczyły analizy sygnału potencjału czynnościowego kardiomiocytu w celu zautomatyzowanej detekcji parametrów depolaryzacji APD50 oraz APD90. W przypadku sygnału pseudo-EKG skonstruowałem moduł automatycznej analizy długości odcinków QT, QRS (ryc. 6). Napisałem również moduł do połączenia sygnału pseudo-EKG z charakterystyką mechaniczną kardiomiocytu i obliczania tzw. okna elektromechanicznego (EMW). Do oceny ww. wielkości zastosowałem własnej konstrukcji empiryczne algorytmy topograficznej analizy sygnału, które działają autonomicznie, bez konieczności zakładania częstotliwości próbkowania sygnału, ani też jego prototypowania przez użytkownika, co jest zwykle standardowym podejściem stosowanym w komercyjnych rozwiązaniach. Omawiane wielkości liczbowe stanowią markery arytmogennego działania leków i są niezbędne zarówno do stwierdzenia faktu wpływu substancji chemicznej na rytm serca jak też i siły tego wpływu. Prace prowadzone były w języku Java pod nadzorem profesjonalnego programisty i z zastosowaniem rozwiązań pracy grupowej.



Ryc. 6. Parametry sygnału pseudo-EKG wygenerowanego przez model ten Tusscher 2004 [42].

4.2.3. Optymalizacja depozycji *in vitro* substancji leczniczych z systemów nośnikowych do podania wziewnego.

W zakończonym w 2015 r projekcie w ramach współpracy bilateralnej Polska-Singapur (2/3/POL-SIN/2012), którego byłem kierownikiem po stronie polskiej, prowadzone były prace nad modelowaniem depozycji *in vitro* cząstek z inhalacyjnych postaci leku. Analizowano zależność depozycji od charakterystyki cząstek uwzględniającej w szczególności ilościową ocenę topologii ich powierzchni [30]. Ocena ta była prowadzona na podstawie analizy zdjęć ze skaningowego mikroskopu elektronowego (SEM) i ich obróbce za pomocą technik analizy obrazu. Wykorzystano deskryptory topologii powierzchni cząstek nośników syntetyzowanych przez partnera singapurskiego i połączono wraz z informatyką chemiczną i wszelką dostępną fizyczną charakterystyką badanych układów w model depozycji cząstek *in vitro* w impaktorze kaskadowym. Modelowanie prowadzono z wykorzystaniem szeregu narzędzi inteligencji obliczeniowej jak sztucznych sieci neuronowych, drzew decyzyjnych, logiki rozmytej, ich różnorodnych hybryd oraz programowania genetycznego. W toku wyczerpujących prac obliczeniowych udało się za pomocą

regresji symbolicznej opracować model w postaci prostej formuły matematycznej zdolny do przewidywania depozycji cząstek substancji leczniczej (FPF) na podstawie charakterystyki formulacji, w tym głównie cząstek nośnika.

$$FPF(\%) = C_1 + \frac{Inp_{16}}{C_2 \cdot Inp_{17}} + \sqrt{Inp_{16}^{e^{Inp_3}} + Inp_6^{C_3 \cdot Inp_{14}} + C_4 \cdot Inp_{14} \cdot Inp_{12}} \quad (7)$$

gdzie:

$FPF(\%)$ – frakcja cząstek drobnych

Inp_3 – skośność profilu (Rsk)

Inp_6 – najwyższy szczyt (Rp)

Inp_{12} – powierzchnia właściwa

Inp_{14} – gęstość nośnika

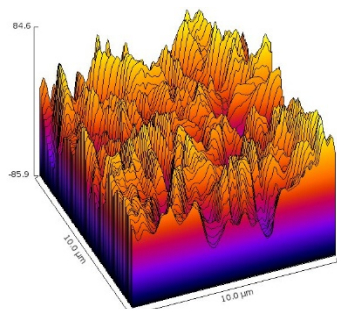
Inp_{16} – zawartość substancji leczniczej

Inp_{17} – szybkość przepływu

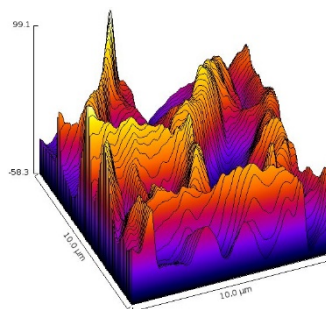
C_1 do C_4 - stałe

Na podstawie analizy równania 7 sformułowano szereg hipotez dotyczących zwiększenia FPF. Zwrócono uwagę na gęstość nośnika i skośność profilu (Rsk) powierzchni cząstki, gdzie maksymalizacja obydwu tych parametrów powinna przynieść zwiększenie FPF. Pożądane zmiany w charakterystyce powierzchni zilustrowano na rycinie 7 gdzie widoczne jest jak wzrost Rsk powiązany jest z wydłużeniem i rozrzedzeniem upakowania igieł stanowiących cechę charakterystyczną cząstek o topologii ziaren pyłku.

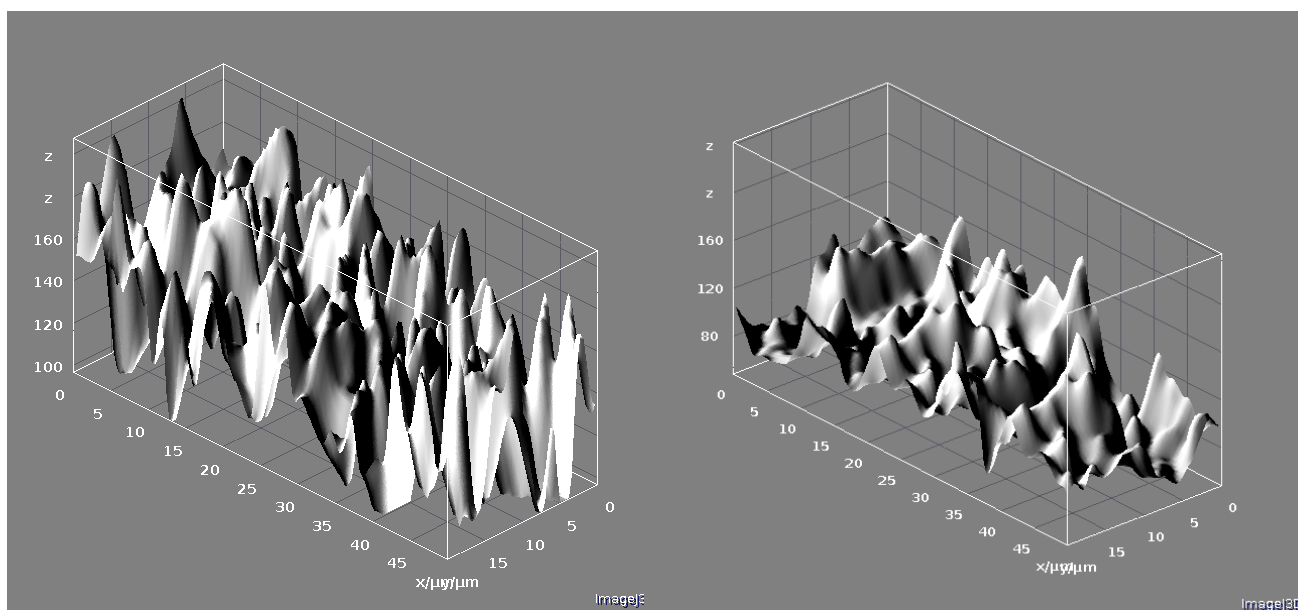
Rsk = 0,2



Rsk = 0,4



Ryc. 7 Postulowane zmiany w powierzchni cząstek w rekonstrukcji pseudo-3D w programie Fiji (ImageJ)



A

B

Ryc. 8 Rekonstrukcja pseudo-3D powierzchni nowo zsyntetyzowanych cząstek: A) HA; B) LA

Stworzona przez nas w programie *Fiji* wizualizacja (ryc. 7) posłużyła jako wzorzec do prac eksperymentalnych prowadzonych przez partnera singapurskiego projektu. W toku tych prac skupiono się na zwiększaniu gęstości i parametru Rsk (skośności rozkładu) i potwierdzono możliwości sterowania depozycją cząstek za pomocą modyfikacji ww. parametrów w sposób przewidywany w symulacjach komputerowych. W ośrodku singapurskim zsyntetyzowano cząstki hydroksyapatytu (HA) i laktozy (LA), na których następnie osadzono surowiczą albuminę wołową

(bovine serum albumin, BSA) jako substancję modelową. Cząstki HA posiadały gęstość 0.215 g/cm³ i parametr $R_{sk} = -0.3161$, a LA odpowiednio 0,586 g/cm³ i $R_{sk} = 0,7777$. Przy tak dobranych właściwościach wyznaczone eksperymentalnie wartości FPF wynosiły odpowiednio: dla HA 19,5% a dla LA 30,4%. Widać więc jasny trend w zwiększeniu FPF przy zwiększaniu gęstości cząstek i anizotropii ich powierzchni. Ilustracją tego ostatniego parametru są obok wzrastających wartości R_{sk} wizualizacje pseudo-3D powierzchni cząstek wykonane na podstawie zdjęć ze skaningowego mikroskopu elektronowego. Na ryc. 8A. przedstawiono fragment powierzchni cząstki HA zaś na ryc. 8B cząstki LA – widać wyraźnie nieregularność powierzchni cząstki LA co potwierdza powyższe rozważania.

4.2.4. Niestandardowe modele korelacji i zależności *in vitro in vivo* (IVIVC/IVIVR)

Pomysł na przedstawione w osiągnięciu naukowym narzędzie *RIVIVR* jest wynikiem moich poprzednich prac nad procedurą dekonwolucji numerycznej, którą zrealizowałem w trybie tzw. dekonwolucji przez konwolucję (ang. *deconvolution through convolution*, DTC) [25]. Jest to również procedura wieloparametrycznej optymalizacji krzywej, chociaż tutaj jest to krzywa uwalniania *in vivo* $i(t)$. Tak zrealizowana procedura dekonwolucji numerycznej została opublikowana w oficjalnym module środowiska R pod nazwą *Rivivc* służącym do stworzenia klasycznej, liniowej korelacji IVIVC na poziomie A [36]. Równocześnie moduł *Rivivc* stał się podstawą koncepcyjną jak też i elementem konstrukcyjnym systemu *RIVIVR* [P.9].

Oprócz prac programistycznych w dziedzinie korelacji *in vitro in vivo*, prowadziłem również prace nad budowaniem niestandardowych modeli korelacji i zależności *in vitro in vivo*. We współpracy z dr Martinem Culenem z University of Veterinary and Pharmaceutical Sciences, Brno, Czechy, z wykorzystaliśmy wyniki badań uwalniania substancji leczniczej *in vitro* otrzymanych z aparatu „Golem” będącego autorską konstrukcją czeskiego zespołu [1, 5]. Na przykładzie modelowania IVIVC/IVIVR dla atorwastatyny wykazaliśmy możliwość bezpośredniej korelacji profilu farmakokinetycznego z danymi *in vitro* z jednego z kompartmentów aparatu „Golem” symulującym jelito czcze w przewodzie pokarmowym człowieka [6]. W sposób pośredni potwierdzono tym samym zasadność konstrukcji aparatu „Golem” jako symulatora przewodu pokarmowego człowieka. Wynika to z faktu, że za pomocą kilku prostych operacji matematycznych doprowadzono do stworzenia modelu mapującego *in vitro* do *in vivo* bez konieczności stosowania metod konwolucji czy dekonwolucji. Łatwość tworzenia modelu stała się więc pośrednim dowodem na biologiczny kontekst metody badania dostępności farmaceutycznej. Równocześnie

zademonstrowano możliwość numerycznego wygenerowania profilu farmakokinetycznego atorwastatyny po podaniu dożylnym [21] i wykorzystania wyników takich symulacji komputerowych do budowy IVIVC w sposób klasyczny, metodą dekonwolucji numerycznej.

4.2.5. Próby wyjaśniania mechanizmów uwalniania substancji leczniczych z matrycowych postaci leku na podstawie analizy metodą obrazowania magnetycznego rezonansu jądrowego (MRI)

We współpracy z dr Piotrem Kulinowskim z Instytutu Fizyki Jądrowej PAN prowadziłem prace nad próbami wyjaśniania mechanizmów uwalniania substancji leczniczych ze stałych postaci leku. Prace dotyczyły analizy procesu uwalniania substancji leczniczych z układów matrycowych z polimerami będącymi pochodnymi celulozy. Mój wkład polegał na ustaleniu matematycznych zależności pomiędzy profilem uwalniania substancji leczniczej a mierzonymi w obrazowaniu rezonansem magnetycznym (MRI) zmianami w wyróżnionych obszarach postaci leku jak suchy rdzeń tabletki, obszar interfejsu i hydrożelu [24]. Za pomocą metod modelowania empirycznego stworzyłem proste zależności liniowe oraz nieliniowe, które posłużyły do wskazania kluczowych czynników stanowiących elementy kontrolujące uwalnianie substancji leczniczej z układów matrycowych. Na podstawie analizy modeli empirycznych wykazaliśmy, że w niektórych przypadkach migracja wody do rdzenia tabletki i kinetyka tego procesu mogą być ważniejsze dla uwalniania substancji leczniczej niż warstwa hydrożelowa i jej ewolucja w czasie. Potwierdziliśmy w ten sposób istotne znaczenie wewnętrznej struktury tabletki dla procesu uwalniania substancji leczniczej. Prace te są tematycznie połączone z moimi wcześniejszymi badaniami prowadzonymi nad uwalnianiem substancji leczniczych z układów zrównoważonych hydrodynamicznie (HBS), przedstawionymi w osiągnięciu habilitacyjnym [P.1] oraz w innej publikacji [26]. W publikacjach tych zawarłem wyniki prac, w których za pomocą narzędzi informatyki chemicznej zakodowałem strukturę polimerów będących składnikami matrycy, po czym za pomocą sztucznych sieci neuronowych wyselekcjonowałem zmienne kluczowe dla procesu uwalniania substancji leczniczych z systemów HBS. W analizie wektora zmiennych kluczowych wykazano, że proces uwalniania jest zależny od geometrii cząsteczki polimeru budującego matrycę HBS. Wyniki te są zgodne z powyższymi rozważaniami strukturalnymi, chociaż w tym przypadku na poziomie molekularnym

4.3. Podsumowanie dorobku naukowego

Na mój całkowity dorobek naukowy składa się (dane na 01.04.2016):

- 45 publikacji oryginalnych o łącznym $IF = 70,435$, MNiSW = 796
- 5 publikacji pokonferencyjnych w czasopismach o łącznym $IF = 10,481$ (nie wliczone do sumarycznej wartości IF)
- 11 rozdziałów w pracach zbiorowych w tym 6 w wydawnictwach anglojęzycznych
- 80 doniesień zjazdowych na konferencjach polskich i zagranicznych, w tym 25 prezentacji ustnych na konferencjach międzynarodowych
- 12 publikacji popularnonaukowych
- 7 programów komputerowych opublikowanych jako *Open Source* oraz współautorstwo komercyjnego systemu komputerowego *Cardiac Safety SimulatorTM* firmy Certara
- 19 prac eksperckich dla przemysłu farmaceutycznego

Sumarycznie:

- $IF = 70,435$
- punkty KBN/MNiSW = 796
- indeks Hirscha = 7
- cytowania: 176

4.4. Prace programistyczne

Od roku 2010 opublikowałem, bądź byłem współtwórcą następujących dostępnych publicznie pakietów oprogramowania:

- *KinetDS 3.0* [23],
- *PhEq_bootstrap* [31],
- *Rivivc* [36] oficjalny pakiet środowiska R do dekonwolucji numerycznej dla korelacji *in vitro in vivo*,
- *R Tools for multivariate analysis* [38]: skrypty R do wielowymiarowej analizy danych,
- *Cardiac Safety SimulatorTM* (CSS) Certara [4].

Ponadto nadzorowałem projekt i wykonanie modułu selekcji zmiennych kluczowych *fscaret* [10].

5. Piśmiennictwo

1. Bárta J., Mudra P., Pešek V. et al., WO/2011/069472, Digestive tract simulator, 2011, (Dostęp 01.02.2016) <http://patentscope.wipo.int/search/en/WO2011069472>
2. BBC, Computer AI passes Turing test in 'world first' (Dostęp 01.02.2016) <http://www.bbc.com/news/technology-27762088>
3. Buduma N. Fundamentals of Deep Learning, O'Reilly Media, 2015
4. Cardiac Safety Simulator™ (CSS) Certara (Dostęp 01.02.2016) <http://www.certara.com/software/pbpbk-modeling/cardiac-safety-simulator/>
5. Čulen M., Rezacova A., Jampilek J., Dohnal J., Designing a dynamic dissolution method: a review of instrumental options and corresponding physiology of stomach and small intestine, Journal of Pharmaceutical Sciences, vol. 102, no. 9, pp. 2995–3017, 2013.
6. Čulen M., Tuszyński P.K., Polak S., Jachowicz R., Mendyk A., Dohnal J., Development of In Vitro-In Vivo Correlation/Relationship Modeling Approaches for Immediate Release Formulations Using Compartmental Dynamic Dissolution Data from “Golem”: A Novel Apparatus, BioMed Research International, vol. 2015, Article ID 328628, 2015. doi:10.1155/2015/328628
7. EMA. Guideline on the investigation of bioequivalence, CPMP/EWP/QWP/1401/98 Rev. 1 (Dostęp 01.02.2016) http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/01/WC500070039.pdf
8. Enabling Grids for E-science (Dostęp 01.02.2016) <http://eu-eggee-org.web.cern.ch/eu-eggee-org>
9. Engelbreck A.P., Cloete I. and Žurada J. “Determining the significance of input parameters using sensitivity analysis”, Form Natural to Artificial Neural Computations, J. Eds. Mira, F. Sandoval, IWANN, Malaga 1995, pp. 382-388.
10. fscaret (Dostęp 01.02.2016) <http://cran.r-project.org/web/packages/fscaret>
11. FDA. Analytical Procedures and Methods Validation for Drugs and Biologics. Guidance for Industry, U.S. Department of Health and Human Services. Food and Drug Administration. Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER). July 2015, CMC
12. FDA. General Principles of Software Validation; Final Guidance for Industry and FDA Staff. U.S. Department of Health and Human Services Food and Drug Administration Center for Devices and Radiological Health Center for Biologics Evaluation and Research, January 11, 2002. (Dostęp 01.02.2016) <http://www.fda.gov/RegulatoryInformation/Guidances/ucm085281.htm>
13. FDA. Guidance for Industry Bioavailability and Bioequivalence Studies Submitted in NDAs or INDs -General Considerations, U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), March 2014, Biopharmaceutics, (Dostęp 01.02.2016) <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm389370.pdf>
14. FDA. Guidance for Industry. Extended Release Oral Dosage Forms: Development, Evaluation, and Application of In Vitro/In Vivo Correlation. U.S. Department of Health and Human Services. Food and Drug Administration. Center for Drug Evaluation and Research (CDER). September 1997, BP 2
15. FDA, Guidance for industry: PAT – A framework for innovative pharmaceutical development, manufacturing

- and quality assurance, U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Veterinary Medicine (CVM) Office of Regulatory Affairs (ORA) September 2004, (Dostęp 01.02.2016) <http://www.fda.gov/downloads/Drugs/Guidances/ucm070305.pdf>
16. FDA. SUPAC-IR Immediate-Release Solid Oral Dosage Forms: Scale-Up and Post-Approval Changes: Chemistry, Manufacturing and Controls, In Vitro Dissolution Testing, and In Vivo Bioequivalence Documentation. Center for Drug Evaluation and Research (CDER), November 1995, CMC 5, (Dostęp 01.02.2016) <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070636.pdf>
 17. Free Software Foundation, GNU GPLv3, (Dostęp 01.02.2016), <http://www.gnu.org/licenses/gpl-3.0.en.html>
 18. Hertz J., Krogh A., Palmer R.G. Wstęp do teorii obliczeń neuronowych. wyd. II. WNT Warszawa 1995.
 19. ICH Q8R2 Pharmaceutical Development. Geneva, Switzerland, International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. 2009, (Dostęp 01.02.2016) http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q8_R1/Step4/Q8_R2_Guideline.pdf
 20. Instytut Santa Fe, (Dostęp 01.02.2016), <http://www.santafe.edu/>
 21. Jamei M., Marciniak S., Edwards D. et al., “The simcyp population based simulator: architecture, implementation, and quality assurance,” *Silico Pharmacology*, vol. 1, no. 1, p. 9, 2013.
 22. Kier L.B., Hall L. H., Murray W. J., Randic M. Molecular Connectivity I: Relationship to Nonspecific Local Anesthesia; *J. Pharm. Sci.* 64, 1975, 1971-1974.
 23. KinetDS3.0 (Dostęp 01.02.2016) <http://sourceforge.net/projects/kinetds>
 24. Kulinowski P., Hudy W., Mendyk A., Juszczyk E., Węglarz W.P., Jachowicz R., Dorożyński P. The Relationship Between the Evolution of an Internal Structure and Drug Dissolution from Controlled-Release Matrix Tablets. *AAPS PharmSciTech*, 2015 : Sep 4. [Epub ahead of print]
 25. Langenbucher F. Handling of computational in vitro/in vivo correlation problems by Microsoft Excel: III. Convolution and deconvolution. *Eur J Pharm Biopharm.* 56, (3), 2003, 429-37
 26. Mendyk A., Dorożyński P., Jachowicz R. Artificial Neural Networks as Modeling Tools in the Identification of Drugs Release Mechanisms from Hydrodynamically-Balanced Systems Formulated with Various Polymers. *Proceedings of International Joint Conference on Neural Networks*, Orlando, Florida, USA, August 12-17, 2007. *IJCNN Conference Proceedings 2007*, IEEE Catalog Number: 07CH37922C, ISBN 1-04244-1380-X, ISSN 1098-7576
 27. Mendyk A., Jachowicz R., “Neural network as a decision support system in the development of pharmaceutical formulation – focus on solid dispersions” *Expert Systems With Applications* 28, 2, 2005 pp. 285-294.
 28. Moore J.W., Flanner H.H., Mathematical Comparison of curves with an emphasis on in vitro dissolution profiles. *Pharm. Tech.* 20(6), : 64-74, 1996.
 29. OpenAI, (Dostęp 01.02.2016) <https://openai.com/about/>
 30. Paclawski A., Kazemi P., Szlęk J., Lau R., Jachowicz R., Mendyk A.: Modeling of fine particle fraction and optimization of formulation for pulmonary delivery. *Particulate Processes in the Pharmaceutical Industry IV*,

- Potsdam, Germany, 14th –18th September, 2014.
31. PhEq_bootstrap (Dostęp 01.02.2016) <http://sourceforge.net/projects/pheqbootstrap>
 32. Polak, S, Wiśniowska, B , Ahamadi, M , Mendyk, A. Prediction of the hERG potassium channel inhibition potential with use of artificial neural networks. Applied Soft Computing Journal, 2011, 11 (2), pp. 2611-261
 33. Polak S., Mendyk A., “Artificial Intelligence Technology as a Tool for Initial GDM Screening.”, Expert Systems with Applications, 26/4, 2004, pp. 455-460
 34. Polak S., Mendyk A., “Artificial Neural Networks as an Engine of Internet Based Hypertension Prediction Tool.” in Medical and Care Compunetics part 1 – Studies in Health Technology and Informatics, IOS Press, Amsterdam, The Netherlands 2004.
 35. Polli J. Dissolution Technologies, 7 (3), 2000, (Dostęp 01.02.2016), http://www.dissolutiontech.com/DTresour/800Articles/800_art1.html
 36. Rivive (Dostęp 01.02.2016) <http://cran.r-project.org/web/packages/Rivive/index.html>
 37. R Core Team. R: a language and environment for statistical computing. Vienna: The R Foundation for Statistical Computing. (Dostęp 01.02.2016) <http://www.R-project.org/>
 38. R Tools for multivariate analysis (Dostęp 01.02.2016) <http://sourceforge.net/p/rscripts/multivariate>
 39. Shah VP, Tsong Y, Sathe P, Liu JP. In vitro dissolution profile comparison-statistics and analysis of the similarity factor, f2. Pharm Res. 1998, 15, 889-96.
 40. Sugano K. Biopharmaceutics Modeling and Simulations: Theory, Practice, Methods, and Applications. 2012 John Wiley & Sons, Inc
 41. Sznitowska M., Kaliszan R., Biofarmacja, Elsevier Urban & Partner, Wrocław 2013, wyd.1
 42. ten Tusscher KH, Noble D, Noble PJ, Panfilov AV. A model for human ventricular tissue. Am J Physiol Heart Circ Physiol. 2004, 286, H1573-89
 43. top500.org (Dostęp 01.02.2016), <http://top500.org/>
 44. Wiśniowska B., Mendyk A., Szłęk J., Kołaczkowski M., Polak S. Enhanced QSAR models for drug-triggered inhibition of the main cardiac ion currents. J. Appl. Toxicol. 35, 2015, 1030-1039
 45. Yager R. R., Filev D. P., Podstawy modelowania i sterowania rozmytego. WNT Warszawa 1995.
 46. Zadeh L.A. Fuzzy Sets. Information and Control, 8, 1965, 338-353.
 47. Żurada J., Barski M., Jędruch W. Sztuczne sieci neuronowe. Podstawy teorii i zastosowania. PWN Warszawa 1996.

