Recenzja pracy doktorskiej mgr Łukasza Kaczyńskiego pt. *„Ocena efektywności praktycznej jako kluczowego elementu raportów oceny technologii medycznych. Pragmatyczne badanie kliniczne"*

Praca doktorska mgr Łukasza Kaczyńskiego obejmuje 161 stron wraz z piśmiennictwem i spisem tabel. Na tych stronach mgr Kaczyński ocenia efektywność praktyczną stanowiącą podstawowy element raportu technologii medycznej. Przedmiotem takiego raportu jest kompleksowa ocena przydatności danej technologii dla opieki zdrowotnej określonego, w tym przypadku polskiego, społeczeństwa. Stąd też wynika duże społeczne znaczenie raportu technologii medycznej i podjęcie właśnie takiego tematu przez mgr Kaczyńskiego świadczy o jego dojrzałości naukowej.
Ten zamysł stawia jednak - przed autorem - trudne wyzwania tłumaczące potknięcia na jakie naraził się Autor, mgr Łukasz Kaczyński. Zanim przyjrzyjmy się kilku definicjom, które tworzy Autor na str. 20 doktoratu przy okazji klasyfikowania abstraktów, trzeba odnotować, że chociaż badanie kliniczne jest przedmiotem pracy to nigdzie nie znalazłem (może źle szukałem, biję się w piersi) definicji badania klinicznego. Wracając jednak do definicji te, które znalazłem w ocenianej pracy (cytuję):

- „Rodzaj badania" - *badanie, które nie spełniało kryteriów metodyki badań poszukiwanych w ramach wyszukiwania (w przypadku wyszukiwania na temat pRCTs oznaczało ono poszczególne pragmatyczne badania z randomizacją, a nie artykuł prezentujący teoretyczne informacje na temat projektowania i oceny jakości pRCTs, a w przypadku wyszukiwania odnośnie do narzędzi oceniających, jakość badań obserwacyjnych oznaczało opracowanie wtórne z analizą badań z randomizacją).*
  [pRCT, w cytowanym fragmencie, oznacza wg. Autora (patrz wyjaśnienie skrótów) pragmatyczne badanie z randomizacją i jest skrótem nazwy angielskiej (pragmatic randomised controlled trials, str. 6)].

Jestem zdania, że być może łatwiej by było szukać definicji „pozytywnej", a nie przez negację. Byłoby też bardziej logiczne gdyby powyższa definicja była jednoznaczna i nie oznaczała raz *„poszczególne pragmatyczne badania z*

randomizacją" a następnie (w tej samej definicji) *„opracowanie wtórne z analizą badań z randomizacją".*
Nie ulega wątpliwości, że Autor postawił przed sobą zadanie bardzo trudne. Dość powiedzieć, że właściwie dopiero Karl Popper dał właściwą wykładnię badania eksperymentalnego.

Podobne uwagi do wyżej wymienionych można by skierować pod adresem punktów: „Populacja", „Komparator" czy „Punkt końcowy".
„Punkt końcowy" to - zdaniem Autora - *„badanie nie zawierające punktów końcowych poszukiwanych w ramach wyszukiwania (w przypadku wyszukiwania odnośnie do narzędzi oceniających jakość badań obserwacyjnych oznaczało badanie pierwotne na temat inny niż wspomniane skale wtórne bez sprecyzowania, że wykonano ocenę jakości takich badań)"*
Biję się w piersi i chcę podkreślić, że mam bardzo konserwatywne podejście do pisania prac naukowych. Wydaje mi się, że nie zawsze niezrozumienie tekstu jest wynikiem ułomności czytającego (w tym przypadku moim) ale czasem, tak jak w przytoczonym materiale może być wynikiem nadmiernej zawiłości autora.

Niewątpliwie bardzo korzystnie przedstawiają się inne rozdziały pracy. Do takich należą:

- Analiza pierwotnych punktów końcowych (str. 60), gdzie Autor w sposób jasny i logiczny przedstawia podstawy i znaczenie analizy ITT (intention to treat analysis). Autor podkreśla, co jest zasadne, że analiza ITT pozwala ocenić czy zastosowane leczenie w zwykłych warunkach, tj. z uwzględnieniem wszystkich występujących czynników zakłócających jest skuteczne.
- Kryteria włączenia (str. 67) Autor podkreśla, że w badaniu pragmatycznym powinno się uwzględnić wszystkich pacjentów z daną chorobą.
- Rekrutacja pacjentów (str. 68). W tej części Autor podkreśla, że pacjenci powinni być „zrekrutowani" (mam wątpliwości czy to słowo jest poprawne), czyli włączeni do badania w momencie przybycia do ośrodka, przy czym dodaje: *„bez żadnego dodatkowego wysiłku ze strony badacza".* Trochę to mylące bo przecież jakiś udział badacza, czyli „wysiłek" wydaje się potrzebny. W tym też ustępie Autor słusznie - jak się wydaje - ma wątpliwości związane z udziałem kampanii medialnych.
- Nawiązując nadal do rekrutacji pacjentów: Autor w ogóle nie odnosi się do liczby badanych w próbie. To bardzo ważny element. Można bowiem tak

dobrać tę liczebność aby zaspokoić bardzo specyficzne wymagania firmy farmaceutycznej finansującej próbę, co oczywiście ma się nijak do rodzaju chorych spotykanych w rzeczywistej praktyce. Charakterystycznym przykładem jest Próba Blacka'a i wsp. z Lancetu 1996;348:1535 w której do badania wpływu alendronianu na złamania kości wybrano staruszki ze złamaniem kręgosłupa, ale takie u których **nie** stwierdzano: wrzodu żołądka, wrzodu dwunastnicy, krwawienia z przewodu pokarmowego, zaburzeń trawienia, upośledzonej funkcji nerek, ciężkiej choroby, nadciśnienia, zawału serca, zaburzonej funkcji tarczycy, krwawień z przewodu pokarmowego, upośledzonej funkcji nerek, nadciśnienia, zawału serca, zaburzonej funkcji tarczycy, zaburzonej funkcji przytarczyc, i niedożywienia. Przepraszam za sarkazm: staruszek o takiej charakterystyce nie spotyka się w praktyce ! Stąd też Black i wsp. aby uzyskać efekt znamienny (chodziło o wpływ alendronianu na złamania kości) do badania włączyli 2027 kobiet na okres 3 lat!

- W następnym dziale *„Systemy edukacji zdrowotnej i rejestry medyczne"* Autor zwraca uwagę na elektroniczne rejestry medyczne (EMR) i ich pomocną rolę dla poprawienia rekrutacji pacjentów. W tym też dziale Autor dyskutuje poprawność i zasadność EMR-ów w świetle różnorodnych pytań klinicznych wymagających zarówno szczegółowego jak i specyficznego zbierania danych.
- W dziale *„Warunki przeprowadzenia badania"* Autor zabiera zdanie na temat „pragmatyczności" próby. Dość szczegółowo, ale bez jasnych, ostatecznych wniosków, Autor dyskutuje wybór ośrodka do przeprowadzenia badania. Słusznie też odnotowuje, że *„ Zmniejszenie stopnia pragmatyczności badania powoduje również włączenie tylko ośrodków wyspecjalizowanych lub akademickich, podczas gdy w rzeczywistości oceniania interwencja będzie stosowana w licznych i różnych typach placówek, w których sprawuje się opiekę nad pacjentami z daną chorobą"*. I słusznie stawia wniosek, że zmniejsza się również ryzyko iż wyniki próby nie będą mogły być *„zaimplementowane"* (daję cudzysłów, bo być może trzeba by napisać „zastosowane") w standardowych warunkach opieki nad pacjentem. W tymże dziale *„Warunki przeprowadzenia badania"* Autor nie wspomina, a szkoda, o różnych typach ośrodków w zależności od rodzaju badania. Zupełnie inaczej przedstawia się i ocenia ośrodek do badań pierwszej fazy rozwoju leku, a inaczej III i IV fazy. Przy tej okazji należałoby przedyskutować jakie

warunki muszą spełniać poszczególne typy ośrodków. Dość powiedzieć, że niektóre badania można przeprowadzić tylko w wysoce wyspecjalizowanych ośrodkach dysponujących odpowiednią - nieraz unikalną - aparaturą i sprzętem.

- Nadzwyczaj ważnym elementem badania jest wybór zastosowanej interwencji.
- Z kolei w dziale *„Ocena ośrodka badawczego"* Autor w zagadkowy sposób (na szczęście łatwy do odszyfrowania) stwierdza: *„W przypadku, kiedy podawanie chorym np. ocenianego leku wymaga dodatkowych szkoleń czy też zatrudnienia dodatkowego personelu, próba bądź wykonania kosztownych badań diagnostycznych wykraczających poza zakres działalności placówki, zaczyna być coraz bardziej wyjaśniającą, gdyż konieczność takich działań zmniejsza możliwość uogólnienia otrzymanych wyników"*. W tym świetle następny wniosek: *„Podobnie, stopień pragmatyczności badania zmniejsza konieczność przeprowadzenia go przez specjalistów z odpowiednim doświadczeniem czy certyfikacjami"* jest niezrozumiały. (Na marginesie: czy to należy rozumieć po prostu w ten sposób, że im mniejsza pragmatyczność tym większe zapotrzebowanie na specjalistów ?)
- Bardzo denerwujące jest pozornie błahe stwierdzenie: *„...przez specjalistów z odpowiednim doświadczeniem czy certyfikatami"*. Wiadomo ile fałszerstw wiąże się z „certyfikatami" i zupełnie zadawalające jest *„...z odpowiednim doświadczeniem"*.
- Kolejnym mało zrozumiałym stwierdzeniem jest: *„..w próbie pragmatycznej nie powinno definiować się, które interwencje są zabronione a które dozwolone.."* O tym jakie interwencje są stosowane w badaniu klinicznym decyduje w określonej próbie klinicznej komitet terapeutyczny.
- Na stronie 72 i dalszych Autor porusza zagadnienie adherence chorych (nie ma dobrego tłumaczenia: przyleganie chorych?) do stosowanej terapii. Z jednej strony należy bardzo wysoko ocenić zamiar Aurora dotyczący dyskusji znaczenia adherence, z drugiej wydaje się, że Autor nie potrzebnie ograniczał się w omawianiu tego bardzo ważnego zjawiska. Doniosłość tego zjawiska podkreślają takie oto dane. W USA odsetek chorych systematycznie stosujących lek wynosi zaledwie 52% (Anonim: Critical overview of antihypertensive therapies. Am J Managed Care 2000;6 (4 Suppl): S211-S221. Co więcej przy chronicznym leczeniu ten odsetek znacznie maleje.

- Bardzo korzystnie przedstawia się dział *„Ocena głównych punktów końcowych* a także dział *„Analiza danych"* z drobną uwagą, że uprzejmość w stosunku do czytającego nakazuje wprowadzane terminy takie jak adherence, compliance i inne tłumaczyć.
- W dziale *„Analiza danych"* Autor stwierdza: *„Podejściem pragmatycznym do analizy danych w badaniach z randomizacją jest zastosowanie analizy ITT, natomiast ocena as treated (czyli zgodnie z otrzymanym leczeniem) wskazuje na bardziej klasyczny charakter próby (w analizie uwzględnieni są jedynie pacjenci, którzy otrzymali przypisane w wyniku randomizacji leczenie).*

  Dlaczego przeciwstawiając badanie z randomizacją z zastosowaniem ITT badaniu per protocol Autor twierdzi, że to drugie badanie jest *„bardziej klasyczne"* ?
- Część 4.1.8.4 *„Inne narzędzia służące do oceny stopnia pragmatyczności badań z randomizacją"* jest w całości przedstawiona prawidłowo, odpowiednio omówiona, a załączone piśmiennictwo właściwe.
- W części 4.1.8.5 *„Ocena stopnia pragmatyczności opracowań wtórnych"* Autor podkreśla wagę analizy czy uwzględnione poszczególne próby kliniczne mają charakter pragmatyczny.
- W 7-punktowym rozróżnieniu skuteczności i efektywności (część 4.1.8.5.2) Autor powołując się na Gartlehner'a i wsp. dyskutuje zagadnienie elementów badania, które odpowiadają na pytanie o skuteczności bądź efektywności badania.
- Po wymienieniu wspomnianych punktów i ich omówieniu Autor przechodzi do części zatytułowanej *„PRECIS-2 a aktualizacja wytycznych Agencji Oceny Technologii Medycznych i Taryfikacji.* Autor bardzo wnikliwie poddaje krytyce wytyczne AOTM i T i proponuje liczne uzupełnienia.
- Na stronach 85 i dalszych Autor omawia badania obserwacyjne podkreślając, że aktualnie rekomendowanymi skalami do oceny badań obserwacyjnych są NOS i NICE, Autor podaje odpowiednie piśmiennictwo i szczegóły.
- Na stronach 87 i dalszych Autor omawia skale: NOS i NICE
- Od strony 91 Autor omawia syntezę informacji o wyszukiwanych skalach, uwzględniając narzędzie AXIS, listę kontrolną Downs'a i Black'a
- Na stronach 94 i dalszych Autor omawia narzędzia: EPHPP, Hoy'a i Brooks'a, wytyczne JBI, narzędzia MINORS, RoBANS, ROIBINS-1 i

ROBINS-E. Te ostatnie okazały się jednym z bardziej istotnych narzędzi stworzonych przez ekspertów z Cochrane Collaboration. Wybrano 7 domen, które uznano za krytyczne w ramach oceny badania bez randomizacji zarówno przed zastosowaniem randomizacji jak i w momencie zastosowania interwencji i wreszcie po zastosowaniu interwencji. W tych przypadkach dokonano oceny błędu wynikającego z odchyleń od planowanej interwencji.

- W punkcie 4.2.5.9. *RTI Item Bank* to narzędzie powstało w oparciu o analizę 1492 pytań dotyczących metod z których 29 uznano za krytyczne.
- STARD. Autor omawia listę kontrolną STARD, która powstała na podstawie przeszukiwania baz Medliune, EmBase, Biosis i Cochrane Collaboration dotyczącą oceny jakości badań analizujących testy diagnostyczne. To narzędzie składa się z 25 punktów metodologicznych. Lista STARD została również wykorzystana do tworzenia innych narzędzi i tę możliwość omawia Autor.
- STROBE

  Lista kontrolna STROBE została stworzona dla autorów przedstawiających wyniki z trzech głównych rodzajów prób epidemiologicznych: kohortowej, kliniczno-kontrolnej i przekrojowej. Lista składa się z 22 punktów niezbędnych do prawidłowego raportowania wyników. Następnie Autor analizuje je szczegółowo w załączonej tabeli.
- QUIPS

  Skala QUIPS (Quality in Prognosis Studies Tool) została stworzona do oceny jakości badań prognostycznych ułatwiających zrozumienie klinicznego przebiegu choroby. Prawidłowa ocena jakości powinna obejmować 6 domen, które opisuje Autor. Wyróżniono trzy możliwe oceny każdej z domen. Autor analizuje każdą z nich i podkreśla, że próbę dobrej jakości charakteryzują domeny posiadające niskie ryzyko wystąpienia błędu systematycznego.
- W punkcie 4.2.6. *„Krytyczna analiza wyszukanych narzędzi"* Autor analizuje ocenę występowania potencjalnego ryzyka błędu systematycznego w opisanych próbach.
- Autor opowiada się za punktowaniem odpowiedzi w narzędziach stosując zakres punktacji 0-1, ponieważ wtedy zakres punktacji jest węższy i umożliwia lepszą, bardziej jednoznaczną kategoryzację jakości próby.

- 4.2.7. Nowy kwestionariusz służący do oceny jakości badań obserwacyjnych - SQAROS. Autor opisuje narzędzie, które ma być nastawione na specyfikę pracy i potrzeby analityka farmakoekonomicznego. Załącza również kwestionariusz do oceny jakości badań obserwacyjnych - SQAROS. Na ten ***nowy*** kwestionariusz składa się 19 punktów z czego ostatnie trzy dotyczą analizy publikacji.
- W dziale 4.2.7.1. Autor dyskutuje kwestionariusz SQAROS w świetle wyszukiwanych narzędzi. Stworzony kwestionariusz odbiega strukturą od większości narzędzi stosowanych do oceny jakości prób bez randomizacji. Jak podkreśla Autor w tymże kwestionariuszu są punkty na podstawie których ocenia się jednolitość prezentowanych w publikacji wyników czy też ocenia obecność wnioskowania niepotwierdzonego rezultatami. W narzędziu SQAROS pytania zostały dobrane tak, aby zmniejszyć konieczność odpowiedzi niejednoznacznych.
- W części 4.3. *Zastosowanie odnalezionych/opracowanych narzędzi* Autor zaczyna od opisu przeprowadzonego wyszukiwania badań oceniających jednorazowe zastosowanie olanzapiny u chorych z schizofrenią doświadczających pobudzenia. Odnaleziono 353 „rekordy". Autor przedstawia strategię wyszukiwania i w tabelach liczbę trafień w bazach PubMed i EmBase. Na podstawie przeprowadzonego wyszukiwania odnaleziono 3 badania, które mogły być wykorzystane do oceny olanzapiny stosowanej domięśniowo lub doustnie u pacjentów z schizofrenią doświadczających pobudzenia w porównaniu z arypiprazolem. W tabelach 11 i 12 Autor przedstawia szczegółowy opis strategii wyszukiwania oraz charakterystykę wybranych badań Kinon'a, Kittipeerachon i Suzuki. Aby ocenić jakość badań Autor także stosuje skalę PRECIS-2 a wyniki przedstawia w tabelach 12-14. W tabeli 15 Autor ocenia pobudzenia w skali PANSS-EC (obejmującej 5 aspektów) pozwalających na ocenę olanzapiny w porównaniu do arypiprazolu.

W podsumowaniu trzeba podkreślić aktualność tematu ocenianej pracy doktorskiej, która w wielu elementach jest całkowicie nowatorska. Te stwierdzenia są spowodowane prostą przyczyną: tzw. „decision makers" czyli decydujący naukowcy (często urzędnicy) są coraz bardziej zainteresowani takimi danymi naukowymi, które charakteryzują się najwyższą jakością tak aby ich decyzje dotyczące leczenia mogły znamiennie poprawić leczenie chorych przy zawsze ograniczonych źródłach finansowania. Na takie zapotrzebowanie

odpowiadają badania kliniczne często nazywane pragmatycznymi. Charakteryzują się one: (1) wyborem do porównania klinicznie znaczących interwencji, (2) włączeniem do badania możliwie różnorodnej populacji pacjentów, (3) rekrutacją uczestników z możliwie heterogennych praktyk klinicznych, (4) zebraniem danych z uwzględnieniem najszerzej rozumianych wyników zdrowotnych. Niefortunnie takie badania są ograniczane przez brak odpowiednich funduszy do ich przeprowadzenia: zarówno narodowe instytucje (np. Ministerstwo Zdrowia), jak i przemysł farmaceutyczny rzadko wspierają takie badania. Jednakże z punktu widzenia pacjenta, a w szerszym sensie i społeczeństwa, właśnie prowadzenie takich badań i korzystanie z ich wyników są najbardziej istotne. Aby to osiągnąć wspomnianym badaniom należy przyznać priorytet, należy wprowadzić te badania do „krwioobiegu" praktyki klinicznej, uzmysłowić praktykom klinicznym znaczenie takich badań, wreszcie znaleźć publiczne lub prywatne fundusze na ich przeprowadzenie. Praca doktorska magistra Ł. Kaczyńskiego odpowiednio upowszechniona może stanowić świetny wstęp do powyższych działań.

Uznając więc jakość przedstawionej pracy doktorskiej i jej znaczenie dla systemu opieki zdrowotnej w kraju proponuję przejście do następnego etapu przewodu doktorskiego.

Prof. dr hab. med. Jacek Spławiński
ul. Modrzewiowa 5, 05-820 Piastów
tel.: 0-601 244 968,
prawo nr: 1707435; umowa nr: 9807/2001

prof. dr hab. med. Jacek Spławiński