

RACJONALNE MODELE ROZUMOWANIA KLINICZNEGO¹

Wojciech Rutkiewicz*
Uniwersytet Warszawski

Streszczenie: *Debata dotycząca medycyny opartej na dowodach (EBM) w obecnym stanie coraz częściej skupia się na filozoficznych założeniach paradygmatu ogłoszonego w 1992 roku przez Grupę Roboczą EBM. Jednym z obszarów zainteresowania badaczy jest problem rozumowania klinicznego, prowadzącego do wydania diagnozy oraz podjęcia decyzji klinicznej. W polskiej literaturze temat ten wciąż nie doczekał się wyczerpującego zaprezentowania. W niniejszym artykule dokonuję przeglądu modeli rozumowania klinicznego. Zaczynam od debaty o pojęciu racjonalności, w której wyróżniam dwa stanowiska: melioryzm i panglozjanizm. Następnie przechodzę do omówienia trzech typów ujęcia rozumowania klinicznego w perspektywie meliorystycznej. Omówienie modeli rozumowania lekarskiego z perspektywy debaty o racjonalności daje możliwość przyjrzenia się regułom rozumowania przyświecającym decyzjom klinicznym, co może mieć ciekawe konsekwencje badawcze i edukacyjne.*

Słowa kluczowe: *medycyna oparta na dowodach, diagnoza, rozumowanie, racjonalność, melioryzm.*

RATIONAL MODELS OF CLINICAL REASONING

Abstract: *Debate regarding evidence based medicine in current state focuses on basic philosophical assumptions of paradigm announced in 1992 by EBM Working Group. One of the topic in this debate is clinical reasoning in light of medical decision-making. Polish literature still lacks of comprehensive presentation in this theoretical field. In current article I present overview of models of clinical reasoning. I begin from the debate of rationality distinguishing two positions called meliorism and panglosianism.*

¹ Chciałbym podziękować p. prof. UW dr hab. Annie Wójtowicz za cenne komentarze i wsparcie w tworzeniu niniejszego artykułu. Dziękuję również anonimowym recenzentom za wszelkie uwagi merytoryczne i techniczne.

* Wojciech Rutkiewicz, Instytut Filozofii Uniwersytetu Warszawskiego, ul. Krakowskie Przedmieście 3, 00-927 Warszawa, e-mail: rutkiewicz.w@gmail.com

Next I proceed to elaboration of three models for clinical reasoning in context of meliorists position. It gives an opportunity to look closer for inferential rules of clinical decision-making and as a result may be precious perspective for future re-search and medical education.

Keywords: *evidence-based medicine, diagnosis, reasoning, rationality, meliorism.*

1. MEDYCYNĄ OPARTA NA DOWODACH

Nurt *evidence-based medicine* (EBM) w ciągu 27 lat zdominował współczesny obraz medycyny. Jako początek nowego paradygmatu w leczeniu uatarło się wskazywać rok 1992, kiedy opublikowano kanoniczny dla EBM artykuł autorstwa zespołu kanadyjskich lekarzy, tzw. Grupy Roboczej EBM (1992). Ma to jednak wymiar czysto symboliczny, ponieważ nurt EBM jest wynikiem sporów i zmian dokonywanych w dziedzinie medycyny co najmniej od 1946 roku. Wtedy to przeprowadzono i opisano pierwsze randomizowane badanie kliniczne, którego formuła z czasem stała się złotym standardem nabywania wiedzy klinicznej (Feinstein, Horwitz, 1982; Bhat, 2010; Jones, Podolsky, 2015). W pierwszej dekadzie popularyzowania EBM, nurt ten ogłoszono jedną z najbardziej wpływowych idei (Hitt, 2001), trwale wpisując jego założenia w wyobrażenia na temat leczenia ludzi – także w Polsce. W roku 2003 powstał Polski Instytut Evidence Based Medicine zajmujący się promowaniem zasad EBM, a przede wszystkim wsparciem w podejmowaniu decyzji zgodnie z tymi zasadami (Statut PI-EBM, 2003).

EBM zrodziło się ze sporów dotyczących statusu badań nauk podstawowych (Paul, 1938; Thomas, 1980; Sackett, 2002), metodologii badawczej (Edwards, 1974; Lehman 1993) czy źródeł wiedzy klinicznej (Concato, 2000; Miles, Loughlin, Polychronis, 2008). W pewnym sensie nurt ten stał się odpowiedzią na nadzieje specjalistów, chcących nieść pomoc przy jak najmniejszym ryzyku porażki. Spory jednak nie zostały zażegnane. Wręcz przeciwnie – ujęcie ich w konkretne ramy teoretyczne doprowadziło do okopania się na swoich dotychczasowych pozycjach oraz powstania nowych sporów. I być może taki jest właśnie charakter siły napędzającej rozwój EBM.

Pierwotnym zamysłem Grupy Roboczej EBM było oparcie decyzji lekarskich na najlepszej jakości dostępnych dowodach empirycznych (Guyatt, 1990; Sackett, Rosenberg, 1995; Sackett, Rosenberg, Gray, Haynes, Richardson, 1996). Podkreślano, że w dotychczasowej praktyce lekarskiej – w paradygmacie tradycyjnym – dominowały dwa założenia (EBM Working Group, 1992):

- Znajomość etiologii i patofizjologii chorób jest wystarczające dla postępowania klinicznego.
- Opinia ekspercka oraz doświadczenie lekarza są wystarczające dla tworzenia wiarygodnych wytycznych postępowania klinicznego.

Założenia te sankcjonowały tzw. rozumowanie mechanistyczne (*mechanistic reasoning*) w działalności lekarskiej i stanowią główne osie sporu między paradygmatem tradycyjnym a EBM. W obliczu porażek rozumowania mechanistycznego (Andersen, 2012; Howick, Glasziou, Aronson, 2010), zwolennicy EBM wysuwają zarzut, że w indywidualnym postępowaniu z pacjentem powyższe założenia nie są warunkami wystarczającymi, choć koniecznymi. David Sackett i William Rosenberg (1995) przekonywali, że decyzje lekarza powinny opierać się na wiarygodnych danych pochodzących z badań przedmiotowych, laboratoryjnych oraz populacyjnych, a wiedza i doświadczenie lekarskie powinny być zintegrowane z rygorystycznymi zasadami myślenia. Kładzie się przez to nacisk na konieczność ustawicznego aktualizowania wiedzy poprzez pozyskiwanie dowodów naukowych. Poddanie ich przez lekarza krytycznej ocenie pod kątem wiarygodności i użyteczności umożliwia włączenie najnowszej wiedzy klinicznej w przebieg leczenia (Rosenberg, Donald, 1995). Kluczowa jest przy tym znajomość tzw. hierarchii dowodów EBM, na szczycie której znajdują się dane z randomizowanych badań klinicznych (*randomized controlled trials* – RCT). Niebanalną rolę w tym procesie odgrywają również wartości wyznawane przez pacjenta (Haynes, Deveraux, Guyatt, 2002b). Jednym ze sztandarowych przykładów na rzecz przewagi zasad zalecanych w ramach EBM jest postępowanie wobec pacjentów zgłaszających się do lekarza z bólami dolnego odcinka kręgosłupa. Zgodnie z rozumowaniem mechanistycznym temu rodzajowi bólu miałyby zarządzić odpoczynek. Jednak według wskazań EBM to nie odpoczynek, lecz lekkie ćwiczenia szybciej postawią pacjenta na nogi (Hitt, 2001).

Z wielu dostępnych opisów zasad EBM wybrzmiewa przede wszystkim prostota. Jest tak nieprzypadkowo, ponieważ z założenia nurt ten ma zaopatrzyć specjalistów w intuicyjne i skuteczne narzędzie działania w bardzo zindywidualizowanych warunkach leczenia². W kolejnych latach, po ogłoszeniu EBM, nowy paradygmat stanowił przedmiot rozległej krytyki (Duggan, 1992; Charlton, Miles, 1998; Williams, 2010; Horwitz, Singer, 2017; Fava, 2017). Czołowi proponenci EBM przypomnieli o niej z okazji 25-lecia nurtu, wskazując na kilka najczęściej wysuwanych zarzutów. Zaliczyli do nich m.in. redukcjonizm metodologiczny oraz promowanie „medycyny z gotowców” (*cookbook medicine*) i myślenia wykluczającego eksperckie doświadczenie lekarza (Djullbegovic, Guyatt, 2017). Mimo wielu zmian zachodzących w obrę-

² Jest to szczególnie widoczne w seriach artykułów publikowanych przez Grupę Roboczą EBM. Przybliżono w nich podstawowe założenia nurtu, których struktura opiera się na opisie i rozwiązaniu konkretnych przypadków, tzw. „case studies”. Patrz: *User's Guide to the Medical Literature*.

bie EBM (warto zwrócić uwagę na zmianę w formie hierarchii ważności dowodów, patrz: GRADE Working Group, 2004; Howick, Glasziou, Aronson, 2009), znaczna część tej krytyki zasadza się na stałych motywach (Djulbegovic, Guyatt, Ashcroft, 2009; Tonelli, 1998; Miles, Loughlin, Polychronis, 2008) z pytaniem o dowód na skuteczność metody EBM na czele (Saad, 2007; Straus, McAlister, 2000).

Wykorzystanie RCT jako źródeł wiedzy klinicznej jest od samego początku nurtu EBM uznawane przez jego proponentów za absolutną podstawę wiarygodnej i rzetelnej praktyki klinicznej. Nawet jeśli dochodziło do pewnych ustępstw wobec alternatywnych podejść, RCT stanowią punkt odniesienia dla praktyka EBM (Howick, Glasziou, Aronson, 2010). To zdecydowane stanowisko lekarzy wywodzących się z nurtu medycyny opartej na dowodach spotykało się z ostrą krytyką (Grossman, Mackenzie, 2005; Cartwright, 2007; Worall, 2007; Cartwright, Munro, 2010; Cartwright, 2011). Charakterystyczną cechą RCT jest ściśle określona struktura eksperymentu, z precyzyjnie sparametryzowanymi celami oraz odpowiednio scharakteryzowaną próbą badawczą³. Aby w jak największym stopniu zniwelować ryzyko wpływu czynników zakłócających na ostateczny wynik, do badania należy zrekrutować odpowiednio wielu uczestników. Pozyskane w wyniku przeprowadzonych RCT dowody naukowe ze względu na wielkość próby odnoszą się do pewnej populacji pacjentów. Generują sylwetkę „pacjenta statystycznego”. Zgodnie z założeniami EBM, taka praktyka badawcza, mówiąc metaforycznie, przerzuca most między łóżkiem chorego a stanowiskiem badacza (Sackett, 2002) oraz wyznacza między nimi „ścieżkę tam i z powrotem” (Tonelli, 1998). Pojawia się jednak pytanie, w jaki sposób na podstawie danych dotyczących grupy osób można podejmować decyzje skierowane do pojedynczego pacjenta, nawet jeśli wystarczająco przypomina on uczestników tej grupy (Feinstein, 1999). Proponenci EBM już od samego początku przedstawiali pewne techniczne rozwiązania tego problemu, takie jak próba na pojedynczym przypadku (*n-of-1 research*) lub modele określania stosunku korzyści do ryzyka podejmowanych terapii (Guyatt, Keller, Jaeschke, Rosenbloom *et al.*, 1990; Glasziou, Irwig, 1995). Jednak bez względu na to, jak bardzo pacjent będzie podobny do innych osób z próby badawczej badania randomizowanego, lekarz będzie posługiwał się rozumowaniem zaledwie uprawdopodobniającym wnioski o charakterze klinicznym. Biorąc pod uwagę, że EBM jest lansowane jako nurt medycyny nastawionej na indywidualną sytuację pacjenta (Guyatt, Busse, 2006), wydaje się, że to zbyt mało. Sytuację dodatkowo utrudnia fakt, że mimo zmian w obrębie EBM, opinia eksperta lekarza jest postrzegana jako dowód słabej jakości. Jednocześnie okazuje się, że to w dużej mierze od opinii eksperta zależy poprawna diagnoza stanu pacjenta

³ Szczegółowy opis wymogów metodologicznych randomizowanych badań klinicznych można znaleźć np. w: Rosenberg, Lachin, 2016.

i określenie prognozy klinicznej (Tonelli, 1999). W tej sytuacji kluczowe staje się to, w jaki sposób przebiega rozumowanie kliniczne lekarza.

Jest to wątek odrębny od rozważań nad technicznymi aspektami EBM, który w literaturze przedmiotu odwołuje nas do filozoficznych podstaw medycyny opartej na dowodach⁴ (Silva, Wyer, 2009; Loughlin, Upshur, Goldenberg, Bluhm, Borger-son, 2010). Jednym z obszarów badawczych, na którego zgłębienie kładzie się coraz większy nacisk, jest rozumowanie kliniczne (*clinical reasoning*), obejmujące proces diagnozowania pacjenta i formułowania prognoz (*diagnosis/prognosis*). W jego ramach podejmuje się próby sformułowania odpowiedzi m.in. na następujące pytania o charakterystykę rozumowania klinicznego, istnienie logiki diagnozy czy związku rozumowania klinicznego z myśleniem naukowym i deontycznym (Mebius, Kennedy, Howick, 2016).

Dotychczas w ramach założeń EBM pytania te postrzegano jako poboczne dla technicznego zagadnienia przechodzenia od generalnych danych populacyjnych pozyskiwanych w badaniach eksperymentalnych do rozwiązań problemów zdrowotnych dotyczących indywidualnego pacjenta (Tonelli, 1998; Norman, 1999; Tonelli, 2006; Silva, Wyer, 2009).

W niniejszym artykule za cel stawiam sobie przegląd modeli rozumowania dostępnych lekarzowi w praktyce klinicznej. Nurt medycyny opartej na dowodach stanowi dla mnie punkt wyjścia. Daje bowiem konkretny *modus operandi* działań lekarza. Zacznę od prezentacji debaty poświęconej pojęciu racjonalności i wyróżnienia dwóch stanowisk: melioryzmu oraz panglozjanizmu. Następnie przejdę do przedstawienia trzech metod modelowania rozumowania klinicznego zgodnych ze stanowiskiem melioryzmu. Będą to kolejno:

- model integracyjny EBM,
- model progowy,
- model bayesowski.

Omówienie modeli rozumowania lekarskiego z perspektywy debaty o racjonalności daje możliwość przyjrzenia się regułom rozumowania przyświecającym decyzjom klinicznym. W ten sposób przełamaniu ulega tendencja do rozważania problematyki rozumowania klinicznego wyłącznie z pozycji normatywnych i deskryptywnych, co może mieć ciekawe konsekwencje badawcze i edukacyjne.

⁴ Nie istnieje powszechna zgoda co do tłumaczenia wyrażenia *evidence-based*. Zbigniew Szawarski posługiwał się określeniem „medycyna wiarygodna” (Wulff, Göttsche, 2005), jednak w polskim piśmiennictwie naukowym częściej występuje tłumaczenie literalne „medycyna oparta na dowodach”. W niniejszym artykule będę posługiwał się tą ostatnią wersją.

2. MELIORYZM VS PANGLOZJANIZM

W medycynie występuje rozziw (*gap*) między normatywnym i deskryptywnym ujęciem decyzji klinicznych (Rutkiewicz, 2018). Najogólniej rzecz ujmując, w ramach podejścia normatywnego określa się, w jaki sposób *powinien być* przeprowadzany proces decyzyjny. Zazwyczaj odwołuje się tutaj do abstrakcyjnych reguł postępowania, których wykorzystanie ma prowadzić do wykrycia najlepszego scenariusza działania i osiągnięcia celu. Złotym standardem normatywnego ujęcia procesu decyzyjnego jest teoria użyteczności oczekiwanej zaproponowana przez Johna von Neumanna i Oskara Morgensterna (1944). Odniosę się do niej przy okazji omawiania modelu progowego rozumowania klinicznego. Natomiast w ramach podejścia deskryptywnego określa się, w jaki sposób *jest* przeprowadzany proces decyzyjny (Elqayam, Evans, 2011; Djelbegovic, Elqayam, Dale, 2018). W tym ujęciu chodzi o identyfikację metod i mechanizmów faktycznie stosowanych przez ludzi w procesie decyzyjnym. Dystynkcja na podejście normatywne i deskryptywne jest silnie związana z dwoma wątkami. Pierwszy z nich dotyczy sporu w psychologii między podejściem poznawczym i ewolucyjnym do analizy nad naturą procesów decyzyjnych (Evans, 2003; Kahneman, 2003; Osman 2004; Evans, Stanovich, 2013; Stanovich, 2016). Drugi odnosi się do sporu wokół kategorii racjonalności ograniczonej (*bounded rationality*) zaproponowanej przez Herberta Simona (1955; Meyer, 2007; Campitelli, Gobet, 2010) oraz związanego z tym sporu o heurystyki (Samuels, Stich, Bishop, 2002; Wójtowicz, Winkowski, 2013).

Debata o kryteria racjonalności jest bardzo ważna dla problemu rozumowania klinicznego. Odniosę się do niej w kontekście dwóch stanowisk zaprezentowanych przez Keitha Stanovicha i Richarda Westa (2000). Rozważają oni zagadnienie zdolności ludzkiego umysłu do rozwiązywania problemów decyzyjnych. Wyróżniają w tym kontekście dwa stanowiska: meliorizm (*meliorism*) i panglozjanizm. Zgodnie z pierwszym z nich, człowiek nie jest istotą z natury racjonalną, lecz może taką być, jeśli podda się odpowiedniemu treningowi (Stanovich, West, 2000; Stupple, Ball, 2014). Natomiast według przeciwstawnego mu panglozjanizmowi człowiek jest z natury racjonalny, o czym rozstrzygają jego zdolności adaptacyjne (Stanovich, West, 2000; Elqayam, Evans 2011). Nazwa ostatniego poglądu upamiętnia filozofa Panglossa znanego z Wolterowskiego *Kandyda*. Przynależność do jednego lub drugiego stanowiska decyduje o tym, w jaki sposób interpretowane są sformułowanie sądów (np. na temat miar prawdopodobieństwa wystąpienia danego zdarzenia) oraz natura różnicy między normatywnym a deskryptywnym podejściem do podejmowania decyzji. W ujęciu meliorystycznym wyjaśnienie takie daje program heurystyk i błędów poznawczych (Tversky, Kahneman, 1974; Kahneman, Tversky 1979; Tversky, Kahneman, 1983; Tversky, Kahneman, 1986; Kahneman, Tversky, 1996; Kahneman, 2012).

Jego inicjatorami byli Amos Tversky i Daniel Kahneman. W rezultacie zaprojektowanych eksperymentów wykryli oni mechanizmy psychologiczne ujawniające się w sytuacji wyboru między opcjami obciążonymi ryzykiem (efekt pewności), reakcji na możliwość poniesienia straty (efekt odbicia) lub przypisywania wyższej wartości przedmiotom, które posiadamy (efekt posiadania) (Kahneman, Tversky, 1979; Secchi, 2011; Gluzicka, 2013). Okazuje się, że obecność takich mechanizmów w procesach mentalnych przyczynia się do niespójności decyzji z wymogami teorii użyteczności oczekiwanej. Jednocześnie stanowi to najsilniejszy dowód na rzecz stanowiska meliorystycznego. Decydenci nie postępują racjonalnie w sensie teorii użyteczności oczekiwanej ze względu na konstrukcję ludzkiego umysłu, choć mogliby to zrobić pod warunkiem nabycia odpowiednich umiejętności (Kahneman, 2012: 552). Eksperymenty Tversky'ego i Kahnemana⁵ wskazują, że ludzie postawieni przed problemem decyzyjnym (w warunkach ryzyka lub niepewności) polegają na bezwysiłkowym myśleniu intuicyjnym⁶. W rezultacie badacze zaproponowali program heurystyk i błędów poznawczych (H&B). Jego główna teza brzmi, że sądy o prawdopodobieństwie mogą być niespójne z formalnymi warunkami formułowania tego typu ocen. Dzieje się tak, ponieważ ludzie częściej polegają na heurystykach, czyli uproszczonych rozumowaniach prowadzących do efektywnego, lecz niedokładnego rozwiązania problemu decyzyjnego (Kahneman, 2012; Wójtowicz, Winkowski, 2018). W efekcie ludzie podatni są na popełnianie błędów systematycznych, polegających na otrzymywaniu wyników niezgodnych z prawdziwymi (w sensie formalnym) wartościami. Pierwotnie Tversky i Kahneman (1974) wymieniali trzy rodzaje heurystyk:

- **Rozpoznawalność** (*representativeness*) – tendencja do oceny prawdopodobieństwa na podstawie stopnia, w jakim przypadek *X* przypomina (reprezentuje) pewną klasę *Y*.
- **Dostępność** (*availability*) – tendencja do oceny prawdopodobieństwa na podstawie łatwości przypomnienia sobie lub wyobrażenia danego przypadku.
- **Zakotwiczenie** (*anchoring*) – tendencja do oceny prawdopodobieństwa na podstawie dostosowania do początkowej liczby lub wartości (kotwicy).

Program heurystyk i błędów poznawczych jest krytykowany w ujęciu panglozjańskim za to, że opiera się na założeniu o istnieniu określonego zestawu norm, w odniesieniu do których stwierdza się, czy sąd jest racjonalny, czy też nie (Gigerenzer, 1996; Vranas 2000). Tymczasem nie wiadomo, czy istnieje tylko jeden taki zestaw. Panglozjanie skłonni są raczej rekonstruować normy wykorzystane do wydania sądu w kontekście konkretnego problemu decyzyjnego niż sprawdzać zgodność odpowie-

⁵ Później także innych, m.in. Richarda Thaler, Eldara Shafira czy Dana Ariely'ego.

⁶ Pod wpływem badań Stanovicha program H&B wzbogacono o kategorie Systemu 1 i Systemu 2. Są to zespoły funkcji poznawczych umysłu człowieka, umożliwiające generowanie wniosków. System 1 generuje je bezwysiłkowo, ale i mniej trafnie, natomiast System 2 wymaga większego zaangażowania na rzecz większej trafności wniosków.

dzi z uznanym przez ekspertów kanonem (Cohen, 1982). Stanovich i West (2000: 646) stwierdzają bowiem, że istnieją przypadkowe okoliczności niezgodności sądów z ich normatywnymi wzorcami:

- błędy w wykonaniu lub przedstawieniu rozwiązania (*performance errors*),
- obliczeniowe ograniczenia ludzkiego umysłu wobec problemów z wieloma zmiennymi,
- przyjęcie niewłaściwego modelu normatywnego przy ocenie wyprowadzonego sądu,
- odmienność interpretacji problemu od tej przyjętej przez ekspertów.

Spór między zwolennikami melioryzmu i panglozjanizmu występuje także na gruncie medycyny. Adaptuje się w jej ramach wiele teorii i modeli racjonalności, które posługują się pojęciami normy zarówno w sensie pierwszego, jak i drugiego stanowiska (Djulbegovic, Elqayam, 2017). Mnogość modeli racjonalności doprowadziła do tego, że w literaturze poświęconej racjonalności decyzji klinicznych pojawiło się pojęcie „normatywnego pluralizmu” (Djulbegovic, Elqayam, Dale, 2018). W medycynie kładzie się jednak nacisk na transparentność działań lekarza, a więc także podejmowanych przez niego decyzji. W dużej mierze dzięki wpływowi nurtu EBM, a także rosnącemu zainteresowaniu zagadnieniem związanym z autonomią pacjenta, odchodzi się od polegania w procesie klinicznym wyłącznie na autorytecie lekarskim. W dobie medycyny zinstytucjonalizowanej niezbędny stał się dostęp do metod weryfikacji działań lekarza. Niezbędne jest również jasne określenie korpusu wiedzy, w który jest wyposażony i który może zostać od niego wyegzekwowany (Redelmeier, Koehler, Liberman, Tversky, 1995; Goodman, 1999; Davidoff, 1999; Elstein, 1999; Switzer, Horton, 2007; Kestin, 2015). Z tych względów w niniejszej analizie modeli rozumowania klinicznego przyjmuję stanowisko meliorystyczne. Zakładam, że lekarz może nauczyć się formułować sądy i podejmować decyzje w sposób racjonalny, jeśli działa zgodnie z pewnym modelem. Pozwala to na zawężenie przedmiotu analizy i jej uporządkowanie. Nie będę jednak całkowicie ignorował stanowiska panglozjańskiego, ponieważ na koniec postaram się wskazać słabości omawianych modeli.

3. MODELE NORMATYWNE W MEDYCYNIE

Rozumowanie kliniczne może kończyć się wyprowadzeniem sądu o charakterze diagnostycznym lub prognostycznym. Ogłoszenie EBM „nowym paradygmatem” medycznym oznaczało zmianę w formułowaniu tego typu sądów (EBM Working Group, 1992). Jeśli przyjmuje się EBM jako wzorzec, należy dokładniej przybliżyć reguły formułowania sądów klinicznych.

3.1. Model integracyjny *evidence based medicine*

Racjonalność działania lekarskiego w ramach medycyny opartej na dowodach stanowi funkcję wiarygodności dostępnych dowodów oraz zasadności procesu ich wykorzystania w praktyce klinicznej (Djulbegovic, Elqayam, Dale, 2018). W istocie model EBM nastawiony jest na ustalenie prawdziwości sądów (Djulbegovic, Guyatt, Ashcroft, 2009). Z czasem sformułowania te ulegały modyfikacjom podporządkowanym aktualnemu etapowi debaty nad podstawami EBM. Skupię się na punktach wspólnych dla tych różnych opracowań, obecnych w tzw. integracyjnym modelu ekspertyzy klinicznej EBM (Haynes, Deveraux, Guyatt, 2002a). Można je przedstawić w postaci następujących postulatów:

- **Dążenie do realizacji wartości i preferencji pacjenta.** Medycyna oparta na dowodach jest promowana jako nurt skupiony na indywidualności pacjentów. Podstawą tego jest wzięcie pod uwagę wartości i preferencji pacjenta podczas określania opcji terapeutycznych przez lekarza. Przez wzgląd na zróżnicowane sytuacje pacjentów, te same metody leczenia tych samych schorzeń mogą być postrzegane przez pacjentów jako mniej lub bardziej atrakcyjne, co powinno wpływać na możliwość wykorzystania danej interwencji w leczeniu (Guyatt, Busse, 2006; Haynes, Deveraux, Guyatt, 2002b). Badania wskazują, że preferencje pacjentów bardzo często nie są zbieżne z preferencjami lekarzy. Wiele w tej kwestii zależy od sytuacji klinicznej pacjenta. Przykładowo, pacjenci są mniej chętni niż lekarze do prowadzenia leczenia obciążonego większym ryzykiem, natomiast chętniej sięgają po antybiotyki, nawet jeśli nie ma do tego bezpośrednich wskazań (Montgomery, Fahey, 2001).
- **Opieranie się na najlepszych dostępnych dowodach.** Praktyka EBM oznacza działanie integrujące doświadczenie lekarskie z aktualnymi dowodami naukowymi, pochodzącymi z wiarygodnych prac badawczych (Sackett, Rosenberg, Gray, Haynes, Richardson, 1996; Strauss, Sackett, 1998). Jakkolwiek nie istnieje pełna zgoda co do tego, w jaki sposób na gruncie EBM należy rozumieć pojęcie „dowodu” (Djulbegovic, Guyatt, 2009), wychodząc od ogólnego założenia, że jest to „empiryczna obserwacja związku zachodzącego między dwoma zdarzeniami” (Guyatt, Busse, 2006), dopuszczalne staje się wykorzystanie różnych metod obserwacji. Nie ma jednak wątpliwości, że medycyna oparta na dowodach w największym stopniu polega na rezultatach randomizowanych badań klinicznych.
- **Wykorzystanie doświadczenia lekarskiego.** Wbrew pojawiającej się krytyce (Tonelli, 1999), w ramach EBM nie dyskwalifikuje się w całości opinii eksperckiej lekarza. Kojarzona jest ona jednak z tradycyjnym nurtem medycyny opierającym się na autorytecie lekarza. Medycyna bazująca na dowodach

kładzie nacisk na przygotowanie lekarza pod kątem wiedzy patofizjologicznej, doświadczenia w pracy z pacjentem oraz określonych sposobów myślenia. Niezbędna jest w tym kontekście znajomość pojęć statystycznych i umiejętność ich wykorzystania w praktyce klinicznej (Tanenbaum, 1993; Sackett, Rosenberg, 1995; Wulff, Gøtzsche, 2005; Kestin, 2015). Jest to bowiem warunek poprawnej interpretacji wyników eksperymentów klinicznych. Jednocześnie jest to punkt sporny w dyskusji nad tym, jakie warunki powinno spełniać poprawne myślenie lekarskie (Croskerry, 2002).

Oprócz powyższych podstawowych zasad dla praktykowania medycyny opartej na dowodach, w działaniu lekarskim można wyróżnić kilka etapów. W celu efektywnej operacjonalizacji przypadków EBM odnosi się do indywidualnych przypadków w kategoriach problemów klinicznych. Lekarz powinien:

- precyzyjnie sformułować problem,
- wyczerpująco przeszukać literaturę (pod kątem postawionego problemu),
- dokonać krytycznej ewaluacji zebranych informacji pod względem ich wiarygodności i użyteczności,
- wykorzystać dowody w praktyce klinicznej pod warunkiem ich przydatności (Sackett, Rosenberg, 1995; Gajewski, Jaeschke, Brożek (red.), 2008).

3.2. Słabości modelu integracyjnego EBM

Podstawową trudnością z integracyjnym modelem EBM jest jego pozorna prostota. W tym sensie należy zgodzić się z praktykami EBM, że ich nurt nie jest „medycyną z gotowców”. Mimo wyraźnej instrukcji postępowania klinicznego, lekarz musi być wyposażony w narzędzia wykraczające poza wiedzę biologiczną. Co więcej, w związku z osłabieniem jego profesjonalnego autorytetu, musi także w każdej chwili wykazywać się gotowością do uzasadnienia swoich decyzji. Reguły modelu EBM w rzeczywistości nie do końca rozwiązują problem rozumowania klinicznego w odniesieniu do indywidualnego pacjenta. Wciąż nie jest bowiem jasne, jakim typem/typami rozumowania powinien posługiwać się lekarz w obliczu indywidualnego problemu klinicznego. Wynikają z tego dwie często podnoszone trudności. Po pierwsze, jest to tzw. problem wiarygodności zewnętrznej RCT (*external validity problem*). Dotyczy to gwarancji przełożenia związku kauzalnego obserwowanego między grupami badawczymi na populację docelową (Cartwright, Munro, 2011). Okazuje się bowiem, że decyzje lekarzy w stosunku do grup różnią się od tych indywidualnych (Redelmeier, Tversky, 1990). Stawia to pod znakiem zapytania przekładalność zaleceń wynikających z RCT na jednostkowe sytuacje pacjentów. Po drugie, pojawia się coraz więcej doniesień o tzw. nieuprawnionym zróżnicowaniu (*unwar-*

ranted variation) w praktyce lekarskiej. Mimo powszechności i dostępności medycyny ukierunkowanej naukowo, pacjenci w podobnych sytuacjach zdrowotnych mogą nie otrzymać takiej samej pomocy. Różnica ta nie jest uzasadniona na gruncie występującej choroby, dostępnych dowodów lub preferencji pacjenta (Wennberg, 2010: 4; Corallo, Croxford, Goodman, Bryan, Srivastava, Stukel, 2014). Trudności te mogą sugerować, że lekarze posługują się różnymi typami rozumowań w zależności od napotkanego przypadku lub okoliczności. W obrębie samej EBM stwierdza się, że jej praktyk może wykorzystywać różne typy rozumowania (Goodman, 1995; Djulbegovic, Guyatt, 2009). Nie można jednak tej sprawy zbagatelizować w kontekście tak konkretnie ukierunkowanego nurtu. Brak ujednolicenia tej sprawy prowadzi do ryzyka wyciągania pochopnych wniosków (przeskoków do konkluzji) i ryzyka popełniania błędów (Elstein, 1999; Croskerry, 2002). W ujęciu meliorystycznym, EBM ma gwarantować racjonalność podejmowanych decyzji pod warunkiem postępowania zgodnego z regułami modelu. Już na etapie formułowania problemu klinicznego, praktyk ma do czynienia ze skomplikowanymi operacjami, takimi jak agregacja danych (podczas wywiadu lekarskiego) i wyprowadzanie na ich podstawie początkowych hipotez. Są to operacje logiczne, które bez odpowiedniego zrozumienia i włączenia do modelu mogą być przeprowadzane na wiele różnych sposobów, skutkując różnymi wnioskami w kontekście tego samego problemu. Takie nadmierne zróżnicowanie jest zaś sprzeczne z założeniami EBM.

3.3. Model progowy

Przypuśćmy, że lekarz rozważa, co należy zrobić z pacjentem, u którego podejrzewa gruźlicę płuc. W takiej sytuacji wykonuje się rozmaz mogący wykazać obecność prątków gruźliczych. Jeśli rozmaz wykazuje obecność prątków (pozytywny wynik testu), może oznaczać to, że ich populacja jest już bardzo duża. Wynik ten jednak nie jest pewny. Również przy uzyskaniu negatywnego wyniku rozmazu lekarz nie może być pewny, że pacjent nie choruje na gruźlicę: taki wynik może jedynie wskazywać na małą liczebność populacji bakterii lub na niedoskonałości techniczne badania (Dutt, Stead, 1994; Siddiqi, Lambert, Walley, 2003; Basinga, Moreira, Bisoffi, Bisig, van den Ende, 2007). Powstaje pytanie, jak lekarz może postąpić w sytuacji niepewnej diagnozy. Jak pokazuje tabela 1, opisany przypadek może zostać potraktowany jako problem decyzji z ryzykiem. Podejmowanie decyzji w sytuacji z ryzykiem charakteryzuje się tym, że działania dostępne lekarzowi prowadzą do różnych konsekwencji, których wystąpienia lekarz nie może być pewny, ale potrafi ocenić prawdopodobieństwa, z jakimi się zrealizują (Domurat, Zieliński, 2013: 23).

Tabela 1

Tabela dwudzielna dla wartości użyteczności możliwych działań lekarza w sytuacji, gdy pacjent jest lub nie jest chory

	Pacjent jest chory	Pacjent nie jest chory
Lekarz podejmuje leczenie	k_1	k_2
Lekarz wstrzymuje leczenie	k_3	k_4

Stephen G. Pauker i Jerome P. Kassirer na łamach *The New England Journal of Medicine* (1975) zaproponowali, że tego typu problemy należy rozwiązywać zgodnie z teorią oczekiwanej użyteczności von Neumanna i Morgensterna. Według tej teorii lekarz przed podjęciem decyzji powinien (Tyszka, 2010; Grzybowski, 2012):

- (1) uporządkować preferencje na zbiorze konsekwencji $k_1 - k_4$, a mówiąc dokładniej – wyrazić je za pomocą liczb tzw. użyteczności (U) przypisywanych konsekwencjom działań,
- (2) ocenić prawdopodobieństwo tego, że pacjent jest chory przy określonym wyniku testu rozmazu.

Na podstawie (1) i (2) lekarz powinien obliczyć wartość użyteczności oczekiwanych (EU) dostępnych opcji (podjęcia lub wstrzymania leczenia). Zgodnie z teorią, racjonalny będzie wybór opcji z wyższą wartością oczekiwaną (Cohen, 1996, Bekker, 2009).

Pierwotnym zamysłem Paukera i Kassirera było zapewnienie lekarzowi narzędzi określających punkt zwany progiem terapeutycznym (*treatment threshold*), w którym użyteczność oczekiwana obu decyzji, tj. o podjęciu lub wstrzymaniu leczenia, jest równoważna:

$$EU_{\text{leczenie}} = EU_{\text{wstrzymanie się.}}$$

Zaproponowali oni tzw. model progowy rozumowania klinicznego (*threshold model*). W odniesieniu do (1) lekarz dysponuje miarami użyteczności podjętych działań w zależności od tego, czy pacjent jest chory, czy nie. Tym sposobem lekarz przypisuje konsekwencji k_1 wartość użyteczności U_1 , k_2 – wartość U_2 , k_3 – wartość U_3 , a k_4 – wartość U_4 . Prawdopodobieństwo choroby natomiast opiera się na wynikach testów diagnostycznych, które stanowią dla lekarza źródło wiedzy na temat zgłaszającej się do niego osoby. Jeśli pacjent rzeczywiście znajduje się w stanie chorobowym i badania (testy diagnostyczne) to potwierdzają, lekarz dysponuje tzw. wynikiem prawdziwie dodatnim (*true positive* – $P(+/D)$). Może się jednak okazać, że test wskazuje występowanie choroby, a pacjent wcale na nią nie choruje – jest to wynik fałszywie dodatni (*false positive* – $P(+/ND)$). Natomiast gdy pacjent nie choruje i wynik badania jest negatywny, lekarz otrzymuje wynik prawdziwie ujemny (*true negative* – $P(-/ND)$). Wynik

falszywie ujemny (*false negative* – $P(-/D)$) występuje wówczas, gdy test nie wykrywa choroby, lecz w rzeczywistości pacjent na nią zapadł. W zależności od tego, w jakiej sytuacji znajduje się lekarz, tj. czy posiada wynik prawdziwie/falszywie dodatni, czy prawdziwie/falszywie ujemny, jego decyzje wiążą się z określonymi zagrożeniami (R) i korzyściami (B) dla pacjenta. Pauker i Kassirer definiują te zagrożenia i korzyści jako różnice między wartościami użyteczności konsekwencji podejmowanych działań klinicznych z uwzględnieniem możliwego stanu pacjenta (Djulbegovic, Hozo, Mayrhofer, van den Ende, Guyatt, 2019: 194):

$$B = U_1 - U_3$$

$$R = U_4 - U_2$$

Na tej podstawie określili równanie wyznaczające próg terapeutyczny. Wskazuje on wielkość prawdopodobieństwa, poniżej którego lekarz powinien wstrzymać się z podjęciem leczenia. Natomiast powyżej tego progu lekarz powinien natychmiast podjąć rekomendowaną terapię (Pauker, Kassirer 1975; Djulbegovic, Hozo, Mayrhofer, van den Ende, Guyatt, 2019; Djulbegovic, Elqayam, Reljic, Hozo, Miladinovic *et al.*, 2014):

$$Pt = \frac{1}{1 + \frac{B}{R}}$$

Pauker i Kassirer zakładali, że decyzje lekarskie staną się efektywniejsze i bardziej zindywidualizowane, jeśli ustalony zostanie terapeutyczny punkt odniesienia. Dzięki wyznaczeniu progu terapeutycznego lekarz dokładnie wie, kiedy podjęcie leczenia jest racjonalne, a kiedy nie. Według wyliczeń, w przypadku pacjentów z podejrzeniem gruźlicy i negatywnym wynikiem rozmazu leczenie powinno zostać podjęte, jeśli prawdopodobieństwo choroby przekracza 2,7% (Djulbegovic, van den Ende, Hamm, Mayrhofer, Hozo, Pauker, 2015: 487).

W kolejnych latach model progowy został udoskonalony (Pauker, Kassirer, 1980). Wprowadzono dwa rodzaje progów: próg testowy (*test threshold* – Tt) oraz próg terapeutyczny (*test-treatment threshold* – $Ttrx$ ⁷). Progi te wyznaczają takie prawdopodobieństwa występowania danej choroby, przy których użyteczności następujących działań lekarza są równe:

- wstrzymaniu się z leczeniem (poniżej Tt) i przeprowadzeniu dodatkowych badań, aby zweryfikować diagnozę (powyżej Tt i poniżej $Ttrx$),
- przeprowadzeniu dodatkowych badań, aby zweryfikować diagnozę (powyżej Tt i poniżej $Ttrx$), i wszczęciu natychmiastowego leczenia (powyżej $Ttrx$).

⁷ Posługuję się oryginalnymi skrótami zawartymi w artykule Paukera i Kassirera z 1980 roku.

Jak widać, rozszerzona wersja modelu progowego uwzględnia wachlarz działań lekarskich, wśród których znajduje się również dalsza weryfikacja stanu pacjenta. Uwzględnienie wyników testów diagnostycznych oraz zagrożeń i korzyści, wynikających z możliwych do podjęcia działań lekarskich, generuje dwa równania wyznaczające próg testowy i terapeutyczny:

$$Tt = \frac{P(+/ND) \cdot R + Rt}{P(+/ND) \cdot R + P(+/D) \cdot B}$$

$$Ttrx = \frac{P(-/ND) \cdot R - Rt}{P(-/ND) \cdot R + P(-/D) \cdot B}$$

Wyznaczenie progów Tt oraz $Ttrx$ wskazuje graniczne wartości użyteczności poszczególnych decyzji przy określonym stanie pacjenta. Zauważmy, że powyższe równania wymagają od lekarza kompleksowego zrekonstruowania obrazu stanu pacjenta, z uwzględnieniem stosunku korzyści podejmowanych działań do zagrożeń. W ten sposób kształtuje się kryterium racjonalności możliwe do zastosowania nawet w nieoczywistej sytuacji klinicznej. Przypadek pacjenta z podejrzeniem gruźlicy, u którego rozmaz nie wykazuje obecności chorobotwórczych prątków, zostaje ujęty z perspektywy możliwości wystąpienia wyników fałszywych oraz zagrożeń płynących z (nie) podjęcia działań w takim wypadku. Jeśli ryzyko zachorowania na gruźlicę przekracza próg Tt , ale znajduje się poniżej $Ttrx$, lekarz powinien zlecić kolejne badania. Jeśli natomiast ryzyko zachorowania pacjenta przekracza $Ttrx$, pacjent powinien od razu zostać poddany farmakoterapii. Nie dziwi zatem, że model progowy jest postrzegany jako łącznik między EBM a teorią decyzji (Eisenberg, Hershey, 1983; Djulbegovic, van den Ende, Hamm, Mayrhofer, Hozo, Pauker, 2015: 486).

3.4. Słabości modelu progowego

Model progowy miał być odpowiedzią na problemy EBM w zakresie zagadnienia rozumowania klinicznego. Z czasem okazało się jednak, że w praktyce model ten wykazuje podobne ograniczenia co cała teoria użyteczności oczekiwanej (Djulbegovic, Hozo, Mayrhofer, van den Ende, Guyatt, 2019). Zdaniem Paukera i Kassirera, zastosowanie założeń EUT i modelu progowego stanowiło rozwiązanie dla kwestii niezgodności opinii specjalistów wobec danego problemu klinicznego w sytuacji dostępu do różnych opcji terapeutycznych (Kassirer, Pauker, 1981). Jak jednak sami zauważali, „ocena ryzyka i korzyści nie jest łatwa” (Pauker, Kassirer, 1980: 1112 [tłum. własne W.R.]). Podstawowa trudność zaproponowanego przez nich modelu polegała na kwantyfikacji ryzyka i korzyści. W praktyce stawia to lekarza przed pytaniem o to, w jakiej formie należy przedstawiać wartości użyteczności (U) zaprezentowane w tabeli 1, aby uzyskać jednolity i wiarygodny obraz kliniczny pacjenta. Dotychczas

instrumentarium teorii użyteczności oczekiwanej wykorzystywano do analizy problemów monetarnych. W praktyce badawczej najczęściej sprowadzono je do postaci loterii. W kontekście medycyny decydent ma do czynienia z problemami niemonetarnymi, tj. takimi, których rezultaty są trudne (o ile w ogóle możliwe) do wyrażenia w kategoriach wpłat i wypłat pieniężnych. Jest to duże ograniczenie badawcze (Redelmeier, Tversky, 1990; DeKay, Hershey, Spranca, Ubel, Asch, 2006). Można próbować poradzić sobie z tą kwestią, wykorzystując różnego rodzaju wskaźniki, np.: współczynnik śmiertelności, zapadalności czy QALY⁸. Pauker i Kassirer (1980: 1112) rekomendują nawet skonstruowanie „arbitralnych jednostek” na potrzeby konkretnego procesu diagnostycznego. W dużej mierze zatem wartości użyteczności oczekiwanej przypisywane poszczególnym decyzjom klinicznym mogą zależeć od uprzedniej decyzji lekarza (albo lekarza wraz z pacjentem) dotyczącej doboru wskaźników ryzyka i korzyści (Grimes, Schulz, 2005). Siła porównywanych miar użyteczności nie wynika wówczas z pozyskanych w procesie klinicznym danych. Chociaż nie stanowi to tak dużej trudności dla wewnętrznej spójności (zgodności z warunkami relacji preferencji) podejmowanych decyzji, konsekwencja ta wydaje się jednak kłopotliwa w kwestii uzyskania zgodności między zaleceniami różnych specjalistów dotyczących tego samego problemu klinicznego. Różni lekarze mogą w obliczu tego samego przypadku prezentować sprzeczne ze sobą funkcje użyteczności.

Należy odnieść się także do ograniczeń modelu progowego o charakterze praktycznym. Model progowy stanowi kompleksową metodę podejmowania decyzji klinicznych. W jej zakres wchodzi formułowanie sądów o prawdopodobieństwie zdarzeń, kalkulacja wartości użyteczności poszczególnych działań oraz określanie zależności między uzyskanymi wynikami. Takie operacje muszą zostać wykonane dla każdej hipotezy diagnostycznej stawianej przez lekarza. Przyjmując, że zespół objawów wykazywanych przez danego pacjenta sugeruje pięć różnych diagnoz, specjalista dla każdej z osobna określa wyniki prawdziwie dodatnie/ujemne i fałszywie dodatnie/ujemne, prawdopodobieństwa progowe Tt oraz $Ttrx$, wielkości ryzyka i korzyści dla możliwych działań i w końcu wartości użyteczności każdej z decyzji. Jest to bardzo żmudna praca. Nawet jeśli z pomocą przychodzą mu komputery i specjalistyczne oprogramowania (Kassirer, Moskowitz, Lau, Pauker, 1987), otwarta pozostaje jednak kwestia doboru wielkości poszczególnych miar prawdopodobieństwa. Ostatecznie podjęcie tej czy innej decyzji zależeć będzie od tego, czy prawdopodobieństwo wystąpienia danej choroby u pacjenta znajduje się poniżej progu Tt lub $Ttrx$ (wówczas należy wstrzymać się z leczeniem lub poddać pacjenta kolejnym testom), czy też powyżej Tt bądź $Ttrx$ (wówczas należy poddać pacjenta testom lub natychmiast rozpocząć leczenie). Czy przekroczenie progu o 0,01 lub nawet o 0,1 stanowi wystarczającą przesłankę do tego, aby porzucić leczenie lub od razu je zacząć? Jakkolwiek mode-

⁸ QALY, czyli *Quality Adjusted Life Years*, stanowi miarę stosunku długości życia pacjenta do jakości tego życia.

lom opartym na EUT od początku towarzyszyła nadzieja na zmniejszenie liczby nadużyć w praktyce klinicznej (Forst, 1974), dziś wciąż nierozwiązana pozostaje kwestia przesady – zarówno w kontekście podejmowania, jak i niepodejmowania leczenia (Djulbegovic, Elqayam, Dale, 2018), szczególnie w przypadkach chorób rzadkich lub przewlekłych. W perspektywie analizy rozumowania klinicznego wątpliwości te rodzą pytanie: Czy podejmowanie decyzji klinicznych w modelu progowym wymaga adaptacji jakichś dodatkowych zasad racjonalności poza zasadą maksymalizacji użyteczności oczekiwanej? Odpowiedź na to pytanie ma znaczenie w sytuacjach granicznych, w których ryzyko zachorowania znajduje się w okolicach progu Tt lub $Ttrx$, jest bliskie 0 albo wynosi niemal 1.

3.5. Model bayesowski

Nieco odmiennym i wcześniejszym od EBM naukowym obliczem medycyny jest procedura diagnozy różnicującej (*differential diagnosis*). Konstytuują ją dwie składowe: występujące u pacjenta objawy oraz tendencja (*propensity*) do zapadnięcia przez pacjenta na choroby zgodne z objawami. Na tych elementach kształtuje się profil diagnostyczny pacjenta (Miettinen, 2001), który w procesie klinicznym jest poddawany przez lekarza analizie i ewaluacji wraz z napływem kolejnych danych klinicznych. Specjalista formułuje hipotezy diagnostyczne dotyczące tego, co mogło spowodować aktualny stan pacjenta. Kolejnym krokiem jest pozyskanie informacji potwierdzających lub obalających te hipotezy, tak aby w końcowym etapie diagnozowania wyłonić hipotezę finałową (Wolf, Gruppen, Billi, 1985). W perspektywie diagnostyki różnicującej, proces myślowy lekarza angażuje typ rozumowania niededukcyjnego⁹ w kierunku formułowania hipotez (H) na podstawie występującego zestawu symptomów (S): $(S_1 \wedge \dots \wedge S_n) \rightarrow H_n$ (Goodman, 1999a). Jest to podejście przeciwstawiane w ramach EBM tradycyjnemu nurtowi medycyny opartemu na myśleniu, wychodzącym od założonej hipotezy i dążącym do przewidzenia odpowiednich symptomów chorobowych. Różnica między jednym a drugim stylem myślenia polega na tym, że podczas gdy w tym drugim można wskazać proste, mechaniczne reguły typu „jeśli H , to S ”, to w stylu rozumowania nazwanym tutaj niededukcyjnym określenie takich reguł okazuje się trudniejsze. Jedną z propozycji jest posłużenie się twierdzeniem Bayesa i związanym z tym współczynnikiem Bayesa.

⁹ Przy określeniu tego typu rozumowania nie posługuję się przymiotnikiem „indukcyjne” ze względu na odmienne rozumienie tego terminu w literaturze przedmiotu. Najczęściej stosuje się go jako określenie wyprowadzania wniosków na podstawie wyliczenia *symptomów*, co nie jest analitycznie równoważne z wyciąganiem wniosków na podstawie wyliczenia *przypadków*. To drugie rozumienie częściej występuje w kontekście dyskusji na temat metodologii badań klinicznych, lecz wówczas zastępuje się je typem rozumowania statystycznego.

Adaptacja modelu bayesowskiego w rozumowaniu klinicznym przebiegała stopniowo i w zasadzie towarzyszyła popularyzacji EBM, przede wszystkim przez wzgląd na skuteczną metodę analizowania dowodów i możliwość wykorzystania w testach diagnostycznych (Ashby, Smith, 2000; Ashby, 2006). Twierdzenie Bayesa pozwala na zrewidowanie początkowego prawdopodobieństwa wystąpienia choroby w świetle nowych informacji (Balla, Ianssek, Elstein, 1985). Rewizję tę przedstawia się jako prawdopodobieństwo warunkowe: $P(x|y)$, gdzie y stanowi warunek występowania x . Prawdopodobieństwo warunkowe przybiera formę następującego równania:

$$P(x|y) = \frac{P(x \cap y)}{P(y)}.$$

Omawiane wcześniej cztery możliwe wyniki testu diagnostycznego (prawdziwie dodatni, fałszywie dodatni, prawdziwie ujemny, fałszywie ujemny) można teraz przedstawić w formie prawdopodobieństwa warunkowego i tym samym ująć związek między ryzykiem wystąpienia choroby (A) a występowaniem symptomu (B):

- wynik prawdziwie dodatni: $P(B|A)$,
- wynik fałszywie dodatni: $P(B|\sim A)$,
- wynik fałszywie ujemny: $P(\sim B|A)$,
- wynik prawdziwie ujemny: $P(\sim B|\sim A)$.

Zauważmy, że wyniki te można pojmować teraz w nieco innym sensie niż w kontekście modelu progowego. Wyniki testu diagnostycznego, przedstawione jako prawdopodobieństwa warunkowe, mówią nam coś o samej chorobie, której ryzyko wystąpienia nas interesuje. Są to wówczas wskaźniki nozograficzne, odpowiadające stosunkom między pacjentami, którzy otrzymali dany wynik, a wszystkimi zdrowymi lub chorymi pacjentami (Wulff, Göttsche, 2005: 97). Twierdzenie Bayesa umożliwia odniesienie stanu pacjenta do szerszego obrazu prezentowanego przez dane i, co więcej, uwzględnia modyfikacje w obrębie diagnozy/prognozy pod wpływem napływających w procesie klinicznym informacji. Określenie prawdopodobieństwa zachorowania na A pod warunkiem występowania B wymaga uwzględnienia poszczególnych wskaźników w następujący sposób:

Twierdzenie Bayesa:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\sim A) \cdot P(\sim A)}$$

Zauważalna staje się silna zależność wartości $P(A|B)$ od fluktuacji w obrębie wskaźników oznaczających pacjentów z rozpoznaniem choroby wśród wszystkich chorych ($P(B|A)$), częstości występowania choroby, czyli prawdopodobieństwa apriorycznego zachorowania ($P(A)$), oraz tych ze stwierdzoną chorobą mimo jej rzeczywistego braku ($P(B|\sim A)$) (Westover, Westover, Bianchi, 2011). Te trzy wskaźniki stano-

wią także podstawę do wyznaczenia czynnika bayesowskiego (*Bayes Factor* – *BF*)¹⁰, czyli stosunku prawdopodobieństwa uzyskania wyników prawdziwie dodatnich w grupie wszystkich chorych do prawdopodobieństwa wyników fałszywie dodatnich wśród wszystkich, którzy nie zapadli na chorobę:

$$BF = \frac{P(B|A)}{P(B|\sim A)}$$

Wykorzystanie *BF* w praktyce klinicznej, podobnie jak wyżej, polega na modyfikacji prawdopodobieństwa apriorycznego dotyczącego zachorowania na daną chorobę pod wpływem informacji istotnych klinicznie, np. wyniku testu (Goodman, 1999b, Wójtowicz, 2018):

$$\frac{P(A)}{P(\sim A)} \cdot BF = \frac{P(A|B)}{P(\sim A|B)}$$

W ten sposób lekarz pozyskuje miarę prawdopodobieństwa aposteriorycznego. I chociaż *BF* opisuje zdolność testu diagnostycznego do odróżniania chorych od zdrowych i w tym sensie stanowi pewną jego własność, dostarcza bezpośrednią informację na temat różnicy ryzyka zachorowania w stosunku do częstości występowania danej choroby w populacji (Deeks, Altman, 2004), co może mieć decydujące znaczenie dla chorób występujących szczególnie rzadko (Pauker, Kopelman, 1992).

3.6. Słabości modelu bayesowskiego

Podobnie jak EBM, także modelowi bayesowskiemu krytyka towarzyszyła od samego początku. Kwestionowano same podstawy podejścia, zaczynając od tego, że w medycynie istotne są fakty, a nie wewnętrzne przekonania lekarza na temat stanu pacjenta. Nawiazywano tym samym do wykorzystania personalistycznej interpretacji prawdopodobieństwa w twierdzeniu Bayesa. Co więcej, określenie stopni przekonania dotyczących szeregu zdarzeń mogłoby przekraczać nawet możliwości eksperta w danej dziedzinie (Shortliffe, Buchanan, 1975: 242). To jeden z powodów nieadekwatności modelu bayesowskiego do realiów praktyki klinicznej. Rozumowania oparte na twierdzeniu Bayesa oraz czynnika bayesowskim w perspektywie diagnostyki różnicowej wymagają wyczerpującej listy wzajemnie wykluczających się hipotez klinicznych (Balla, Ianssek, Elstein, 1985: 327). Nie jest jednak jasne, czy w przypadku danego pacjenta listę taką da się stworzyć. Określony zestaw symptomów może prowadzić do zdiagnozowania różnych chorób, w tym rzadkich. Jak daleko można jednak posunąć się w konstruowaniu takiej listy? Aby nie popaść w absurd, pożądana wydaje się jakaś zasada stopująca. Pozostaje pytanie, jaka. Nie

¹⁰ W literaturze poświęconej rozumowaniu klinicznemu czynnik bayesowski (*BF*) określa się również jako *likelihood ratio* (*LR*). Mimo to, w celu uniknięcia nieporozumień, konsekwentnie będą posługiwał się nazwą „czynnik bayesowski” lub skrótem *BF*.

jest również jasne, czy lista ta będzie zawierać hipotezy wzajemnie się wykluczające. Słabością modelu bayesowskiego miałyby być jego nieskuteczność w sytuacji współwystępowania chorób. Postać mianownika twierdzenia Bayesa w kontekście diagnostyki różnicowej to suma prawdopodobieństw warunkowych występowania danych symptomów wobec wystąpienia danej choroby:

$$\sum_{i=1}^m P(B_1 \dots B_i | A_m) \cdot P(A_m).$$

W tej formie model bayesowski wymaga ogromnej ilości danych, które lekarz musiałby przetworzyć. To odsyła nas znowu do zarzutu zbyt dużego subiektywizmu w ocenie sytuacji klinicznej przy personalistycznej interpretacji prawdopodobieństwa, ponieważ lekarz mógłby być zmuszony do arbitralnej oceny miar prawdopodobieństw występowania zdarzeń klinicznych. Ratunkiem staje się założenie o niezależności symptomów, sprowadzające mianownik twierdzenia do formy iloczynu:

$$\sum_{i=1}^m P(B_1 | A_m) P(B_i | A_m) \cdot P(A_m).$$

To jednak założenie nierealistyczne, ponieważ przez wzgląd na kausalny charakter mechanizmów patofizjologicznych może być tak, że:

$$P(B_1 \cap B_i) > P(B_1) \cdot P(B_i) \text{ (Charniak, 1983: 71).}$$

Wszystkie trzy zaprezentowane powyżej modele wymagają od lekarza wykroczenia poza „bezpieczny teren” wiedzy medycznej. Dodatkowo operowanie na pojęciu prawdopodobieństwa w zderzeniu z trudnymi warunkami pracy naraża go na poleganie na myśleniu heurystycznym. Występowanie heurystyk wiąże się z ignorowaniem istotnych dla przypisywania wielkości prawdopodobieństwa pojęć statystycznych, takich jak: wartość bazowa, wielkość próby czy regres do średniej. Niewzięcie pod uwagę powyższych elementów zwiększa margines błędu w ocenie prawdopodobieństwa i szansę na decydowanie w oparciu o nieprawdziwe informacje. Ze względu na wielką, a czasami wręcz przytłaczającą ilość decyzji, którą musi podjąć lekarz w trakcie dyżuru, istnieje ryzyko, że osąd o prawdopodobieństwie występowania danego schorzenia będzie odbiegał od „formalnego ideału” (Croskerry, 2002; West, West, 2002). Przykładem tego jest diagnozowanie na podstawie rozpoznawania wzorca (*pattern recognition*). Zgodnie z tą metodą problem diagnozy może zostać przez lekarza ujęty jako problem dopasowania zbioru symptomów występujących u pacjenta do zbioru objawów charakterystycznych dla danej choroby lub chorób (Wolff, 2006). Rozpoznawanie wzorca przez wzgląd na skupienie na określonym zestawie symptomów wiąże się z ryzykiem wystąpienia heurystyki zakotwiczenia lub heurystyki rozpoznawania przy określaniu podobieństw zbiorów. Wskazuje się, że występowanie heurystyk prowadzących do błędów systematycznych jest nagminne w rozumowaniu lekarskim. Stanowi to duże ograniczenie dla normatywnych modeli wykorzystujących sady o prawdopodobieństwie zarówno przy ocenie częstości

występowania poszczególnych chorób, ocenie wpływu indywidualnych informacji o stanie pacjenta na ostateczne wnioski diagnostyczne (Redelmeier, Koehler, Liberman, Tversky, 1995; Elstein, 1999; Elstein 2002), jak również decydowaniu w sytuacjach granicznych, gdy lekarz ma do czynienia z niskim lub wysokim prawdopodobieństwem zdarzenia klinicznego (Kox, 2003).

4. PODSUMOWANIE

Przypomnijmy, że stanowisko meliorystyczne w kwestii racjonalności wymaga znajomości i wykorzystania normatywnych teorii podejmowania decyzji. W tej perspektywie lekarz powinien być ekspertem w zakresie medycyny, lecz poza tym także całkiem niezłym matematykiem i psychologiem. Nic dziwnego, że zawód ten należy do najtrudniejszych. Według stanowiska panglozjańskiego to jednak absurd, aby twierdzić, że lekarz nie jest racjonalny, ponieważ nie zna podstaw rachunku prawdopodobieństwa lub nie posługuje się użytecznością oczekiwaną. Meliorystom nie chodzi jednak o to, aby egzaminować lekarzy ze znajomości pojęć. Jeśli podejmowane przez nich decyzje mają spełniać wymogi racjonalności, przede wszystkim muszą być ze sobą spójne (koherentne). Pacjent, który usłyszałby od swojego lekarza trzy konkurencyjne diagnozy i ocenę szansy każdej z nich na $\frac{1}{2}$, mógłby mieć słuszne podejrzenia wobec wiarygodności jego profesjonalnej opinii. Ludzie są omylni, lecz w medycynie dążymy do tego, aby błędy zdarzały się jak najrzadziej. A te, które i tak się zdarzają, niech przynajmniej nie będą wynikiem ignorancji.

Debata między stanowiskami melioryzmu i panglozjanizmu niewątpliwie ma ogromny walor edukacyjny. Nie da się jej streścić w jednym artykule, dlatego moim celem było przedstawienie modeli rozumowania klinicznego, odpowiadających założeniom melioryzmu. Chciałem również zasygnalizować słabe strony tych modeli, na które mogliby zwrócić uwagę panglozjanie. Nie roszczę sobie pretensji do pełności powyższego omówienia. Wiele wątków pomijam lub pozostawiam bez wyczerpującego rozwinięcia. Zaprezentowałem trzy modele rozumowania klinicznego: integracyjny model EBM, model progowy oparty na zasadach teorii użyteczności oczekiwanej oraz model bayesowski. Staralem się także wskazać role skróconych sposobów rozumowania jako istotny argument na rzecz stanowiska meliorystów, zgodnie z którym ludzie są skłonni do podejmowania decyzji nieracjonalnych. Medycyna oparta na dowodach charakteryzuje się wysoką zgodnością z tym podejściem. Niesystematyczne obserwacje dokonywane przez lekarza, kształtujące jego

specjalistyczne doświadczenie, są na gruncie EBM zastępowane weryfikowalnymi metodami wyprowadzania wniosków klinicznych. Lekarz musi się tych metod nauczyć, aby sprostać wymogom racjonalności EBM.

Racjonalność nie stanowi inherentnej własności ludzkiego umysłu. Racjonalność jest zewnętrzną wobec człowieka instancją, która poprzez stanowienie pewnych reguł daje uzasadnienie dla jego działań. W tym sensie racjonalność meliorystyczna – bez czynienia żadnych metafizycznych zobowiązań – wpisuje się w kartezjańską tradycję matematycznego *logosu* porządkującego sferę ludzkiej sprawczości. W obrębie medycyny działalność ta odbywa się w warunkach niepewności. Podstawowym zadaniem lekarza jest oswojenie tej niepewności. Podejmowane zaś przez niego decyzje mają prowadzić do wyzdrowienia pacjenta. Używanie pojęcia prawdopodobieństwa staje się wobec tego nieodzownym elementem pracy lekarza. Alan Hájek (2007: 563) cytował za biskupem Butlerem: „Prawdopodobieństwo jest życiowym przewodnikiem [tłum. własne W.R.]”, zaś Richard T. Cox stwierdzał (1946: 2): „Prawdopodobieństwo uznaje się także jako miarę rozsądnych oczekiwań dla pojedynczych zdarzeń [tłum. własne W.R.]”. W procesie klinicznym, w ujęciu melioryzmu, rozsądne oczekiwania rzeczywiście poprzedzają rozsądne działanie.

Modele rozumowania klinicznego w ujęciu racjonalności meliorystycznej konstruowane są na podstawie ściśle określonych reguł. Bez względu na dobór interpretacji prawdopodobieństwa, formułowanie sądów o wystąpieniu zdarzenia klinicznego jest zgodne z aksjomatyką Kołmogorowa. Miara użyteczności oczekiwanej nie może być liczona inaczej niż jako suma iloczynu prawdopodobieństw i wartości użyteczności dostępnych alternatyw wyboru. Rozumowanie bayesowskie nie będzie natomiast możliwe bez podstawienia wartości do poszczególnych komponentów twierdzenia Bayesa. To przywiązanie do reguł spełnia bardzo określoną funkcję. Ustanawia się w ten sposób kryterium prawomocności sądów klinicznych, indukujące poczucie kontroli i uporządkowania procesu klinicznego, bardzo ważnego dla pacjentów, którzy mogą wymagać konkretnej wiedzy i informacji, aby móc w spokoju poddać się terapii (Bailey, 2010). Dlaczego jednak reguły te muszą mieć charakter normatywny? Być może, wbrew EBM, profesjonalna intuicja i doświadczenie kliniczne lekarza mogą być wystarczającym probierzem racjonalności, o ile jego działania są skuteczne. Tym bardziej że w praktyce sprostanie wymogom EBM jest bardzo trudne, o ile w ogóle możliwe, jeśli zostaną uzupełnione o elementy modelu progowego lub bayesowskiego. Aby się o tym przekonać, do głosu należy dopuścić stronnictwo panglozjan w debacie poświęconej racjonalności rozumowania klinicznego i decyzji klinicznych.

BIBLIOGRAFIA

- Andersen, H. (2012). Mechanisms: what are they evidence for in evidence-based medicine? *Journal of Evaluation in Clinical Practice*, 18, 992–999.
- Ashby, D. (2006). Bayesian Statistics in medicine: A 25 year review. *Statistics in Medicine*, 25, 3589–3631.
- Ashby, D., Smith, M.F.A. (2000). Evidence-based medicine as Bayesian decision-making. *Statistic in Medicine*, 19, 3291–3305.
- Bailey, L. (2010). Strategies for Decreasing Patient Anxiety in the Perioperative Setting. *AORN Journal*, 92(4), 445–457.
- Basinga, P., Moreira, J., Bisoffi, Z., Bisig, B., van den Ende, J. (2007). Why Are Clinicians Reluctant To Treat Smear-Negative Tuberculosis? An Inquiry about Treatment Thresholds in Rwanda. *Medical Decision Making*, 27, 53–60.
- Bekker, L.H. (2009). Using decision-making theory to inform clinical practice. W: Edwards, A., Glyn, E. (red.), *Shared Decision-Making in Health Care* (45–52). Oxford: Oxford University Press.
- Cahan, A., Dilon, G., Manor, O., Paltiel, O. (2003). Probabilistic reasoning and clinical decision-making: do doctors overestimate diagnostic probabilities? *QJM: Monthly Journal of the Association of Physicians*, 96, 763–769.
- Campitelli G., Gobet, F. (2010). Herbert Simon's Decision-Making Approach: Investigation of Cognitive Processes in Experts. *Review of General Psychology*, 14(4), 354–364.
- Cartwright, N. (2007). Are RCTs the Gold Standard? *BioSocieties*, 2, 11–20.
- Cartwright, N. (2011). A philosopher's view of the long road from RCTs to effectiveness. *The Lancet*, 377, 1400–1401.
- Cartwright, N., Munro, E. (2010). The limitations of randomized controlled trials in predicting effectiveness. *Journal of Evaluation in Clinical Practice*, 16, 260–266.
- Charlton, G.B., Miles, A. (1998). The rise and fall of EBM. *QJM: Monthly Journal of the Association of Physicians*, 91, 371–374.
- Cheeseman, P. (1985). In Defense of Probability. *Proceedings of the 9th international joint conference on Artificial intelligence*, 2, 1002–1009.
- Concato J., Shah, N., Horwitz, I.R. (2000). Randomized, controlled trials, observational studies and the hierarchy of research designs. *The New England Journal of Medicine*, 342(25), 1887–1892.
- Corallo, N.A., Croxford, R., Goodman, C.D., Bryan, L.E., Srivastava, D., Stukel, A.T. (2014). A systematic review of medical practice variation in OECD countries. *Health Policy*, 114, 5–14.
- Corner, A., Hahn, U. (2013). Normative theories of argumentation: are some norms better than others? *Synthese*, 190, 3579–3610.
- Croskerry, P. (2002). Achieving Quality in Clinical Decision Making: Cognitive Strategies and Detection of Bias. *Academic Emergency Medicine*, 9(11), 1184–1204.
- Davidoff, F. (1999). Standing Statistics Right Side Up. *Annals of Internal Medicine*, 130(12), 1019–1021.
- Deeks, J.J., Altman, G. D. (2004). Diagnostic tests 4: likelihood ratios. *British Medical Journal*, 329, 168–169.

- Djulbegovic, B., Elqayam, S. (2017). Many faces of rationality: Implications of the great rationality debate for clinical decision-making. *Journal of Evaluation in Clinical Practice*, 23, 915–922.
- Djulbegovic, B., Elqayam, S., Dale, W. (2018). Rational decision making in medicine: Implications for overuse and underuse. *Journal of Evaluation in Clinical Practice*, 24, 655–665.
- Djulbegovic, B., Elqayam, S., Reljic, T., Hozo, I., Miladinovic, B., Tsalatsanis, A., Kumar, A., Beckstead, J., Taylor, S., Canon-Bowers, J. (2014). How do physicians decide to treat: an empirical evaluation of the threshold model. *BMC Medical Informatics and Decision Making*, 14(47).
- Djulbegovic, B., Ende van den, J., Hamm, R., Mayrhofer, T., Hozo, I., Pauker, S. (2015). When is rational to order a diagnostic test, or prescribe treatment: the threshold model as an explanation of practice variation. *European Journal of Clinical Investigation*, 45(15), 485–493.
- Djulbegovic, B., Hozo, I., Mayrhofer, T., Ende van den, J., Guyatt, G. (2019). The threshold model revisited. *Journal of Evaluation in Clinical Practice*, 25, 186–195.
- Domurat, A., Zieliński, T. (2013). Niepewność i niejasność jako uwarunkowania decyzji ekonomicznych. *Decyzje*, 20, 21–47.
- Duggan, F.P. (1992). Time to abolish “gold standard”. *British Medical Journal*, 304, 1568–1569.
- Dutt, A.K., Stead, W.W. (1994). Smear-Negative Pulmonary Tuberculosis. *Seminars in Respiratory Infections*, 9(2), 113–119.
- Edwards, F.W.A. (1974). The History of Likelihood. *International Statistical Review*, 42(1), 9–15.
- Eisenberg, M.J., Hershey, C.J. (1983). Derived Thresholds: Determining the Diagnostic Probabilities at Which Clinicians Initiate Testing and Treatment. *Medical Decision Making*, 3 (2), 155–168.
- Elqayam, S., Evans, S.J. (2011). Subtracting “ought” from “is”: Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences*, 34, 233–290.
- Elqayam, S., Evans, S.J. (2013). Rationality in the new paradigm: Strict versus soft Bayesian approaches. *Thinking & Reasoning*, 19, 453–470.
- Elstein, S.A. (1999). Heuristics and Biases: Selected Errors in Clinical Reasoning. *Academic Medicine*, 74(7), 791–794.
- Evans, S.J. (2003). In two minds: dual-process accounts of reasoning. *TRENDS in Cognitive Sciences*, 7(10), 454–459.
- Evans, S.J., Stanovich, E.K. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, 8, 223–241.
- Fava, A.G. (2017). Evidence-based medicine was bound to fail: a report to Alvan Feinstein. *Journal of Clinical Epidemiology*, 84, 3–7.
- Gajewski, P., Jaeschke, R., Brożek, J. (red.). (2008). *Podstawy EBM, czyli medycyny opartej na danych naukowych: dla lekarzy i studentów medycyny*. Kraków: Medycyna Praktyczna.
- Gigerenzer, G. (1996). On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky. *Psychological Review*, 103(3), 592–596.
- Goldenberg, J.M. (2006). On evidence and evidence-based medicine: Lessons from the philosophy of science. *Social Science & Medicine*, 62, 2621–2632.
- Golik, J. (2016). Krótka historia paradoksu petersburskiego i jego wczesnych rozwiązań. W: Garnek, J. (red.), *Oblicze 2016* (65–76), Poznań: Koło Naukowe Matematyków UAM.
- Goodman, N.S. (1999). Toward Evidence-Based Medical Statistics. 1: The P Value Fallacy. *Annals of Internal Medicine*, 130(12), 995–1004.

- Goodman, N.S. (1999). Toward Evidence-Based Medical Statistics. 2: The Bayes Factor. *Annals of Internal Medicine*, 130(12), 1005–1013.
- Grimes, A.D., Schulz, F.K. (2005). Refining clinical diagnosis with likelihood ratios. *The Lancet*, 365, 1500–1505.
- Grossman, J., Mackenzie, J.F. (2005). The randomized controlled trial. Gold standard, or merely standard? *Perspectives in Biology and Medicine* 48(4), 516–534.
- Grzybowski, A. (2012). *Matematyczne modele konfliktu: wykłady z teorii gier i decyzji*, Częstochowa: Wydawnictwo Politechniki Częstochowskiej.
- Hájek, A. (2007). The reference class problem is your problem too. *Synthese*, 156, 563–585.
- Haynes, B.R. (2002). What kind of evidence is it that Evidence-Based Medicine advocates want health care providers and consumers to pay attention to? *BMC Health Services Research*, 2(3).
- Hitt, J. (2001). *The Year in Ideas A to Z: Evidence-Based Medicine*. Pobrane z: <https://www.nytimes.com/2001/12/09/magazine/the-year-in-ideas-a-to-z-evidence-based-medicine.html> [data dostępu: 18.09.2019].
- Horwitz, I.R., Singer, H.B. (2017). Why evidence-based medicine failed in patient care and medicine-based evidence will succeed. *Journal of Clinical Epidemiology*, 84, 14–17.
- Kahneman, D., Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–292.
- Kahneman, D., Tversky, A. (1996). On the Reality of Cognitive Illusions. *Psychological Review*, 103(3), 582–591.
- Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review*, 93(5), 1449–1475.
- Kahneman, D. (2012). *Pułapki myślenia. O myśleniu szybkim i wolnym*. Poznań: Media Rodzina.
- Kestin, I. (2015). Statistics in medicine. *Anaesthesia and Intensive Care Medicine*, 16(4), 200–207.
- Lehman, L.E. (1993). The Fisher, Neyman-Pearson Theories of Testing Hypotheses: One Theory or Two? *Journal of the American Statistical Association*, 88(424), 1242–1249.
- Meyer, M. (2007). Herbert Simon i jego idea ograniczonej racjonalności. *Decyzje*, 7, 111–115.
- Miles, A., Loughlin, M., Polychronis, A. (2008). Evidence-based healthcare, clinical knowledge and the rise of personalized medicine. *Journal of Evaluation in Clinical Practice*, 14, 621–649.
- Montgomery, A.A., Fahey, T. (2001). How do patients' treatment preferences compare with those of clinicians? *Quality in Health Care*, 10, i39–i43.
- Pauker, G.S., Kassirer, P.J. (1975). Therapeutic Decision Making: A Cost-Benefit Analysis. *The New England Journal of Medicine*, 293, 229–234.
- Pauker, G.S., Kassirer, P.J. (1980). The Threshold Approach To Clinical Decision Making. *The New England Journal of Medicine*, 302, 1109–1117.
- Paul, R.J. (1938). President's Address Clinical Epidemiology. *The Journal of Clinical Investigations*, 17(5), 539–541.
- Perezgonzalez, D.J. (2015). Fisher, Neyman-Pearson or NHST? A tutorial for teaching data testing. *Frontiers in psychology*, 6, 223.
- Polski Instytut EBM, (2003). *Statut Fundacji „Polski Instytut Evidence Based Medicine”*. Pobrane z: <http://ebm.org.pl/show.php?aid=160889> [data dostępu: 18.09.2019].
- Rawlins, D.M. (2018). Evidence-based medicine. *Medicine*, 46(7), 388–392.

- Rosenberg, F.W., Lachin, M.J. (2016). *Randomization in Clinical Trials*. New Jersey: John Wiley & Sons.
- Rutkiewicz, W. (2018). Normatywność i deskryptywność wnioskowania diagnostycznego w nurcie *evidence-based medicine*. *Filozofia Nauki*, 3(103), 26, 79–98.
- Rzepiński, T. (2009). Interpretacje pojęcia prawdopodobieństwa w sporze o randomizację. *Nowiny Lekarskie*, 78(5–6), 360–365.
- Sackett, L.D. (2002). Clinical epidemiology: what, who, and whither. *Journal of Clinical Epidemiology*, 55, 1161–1166.
- Sackett, L.D., Rosenberg, C.M.W. (1995). The need for evidence-based medicine. *Journal of the Royal Society of Medicine*, 88, 620–624.
- Secchi, D. (2010). *Extendable rationality. Understanding decision making in organizations*. Nowy Jork: Springer.
- Shortliffe, H.E., Buchanan, G.B. (1975). A model of inexact reasoning in medicine. *Mathematical Biosciences*, 23(3–4), 351–379.
- Siddiqi, K., Lambert, L.M., Walley, J. (2003). Clinical diagnosis of smear-negative pulmonary tuberculosis in low-income countries: the current evidence. *The Lancet. Infectious Diseases*, 3, 288–296.
- Stanovich, E.K. (2016). The Comprehensive Assessment of Rational Thinking. *Educational Psychologist*, 51(1), 23–34.
- Stanovich, E.K., West, F.R. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645–726.
- Stupple, N.J.E., Ball, J.L. (2014). The intersection between Descriptivism and Meliorism in reasoning research: further proposals in support of soft normativism. *Frontiers in psychology*, 5, 1269.
- Switzer, S.S., Horton, J.N. (2007). What Your Doctor Should Know About Statistics (but Perhaps Doesn't...). *Chance*, 20(1), 17–21.
- Thomas, L. (1980). Future Prospects for Basic Science in Medicine. *Bulletin of the American Academy of Arts and Sciences*, 34(1), 21–41.
- Tversky, A., Kahneman, D. (1974). Judgement under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131.
- Tversky, A., Kahneman, D. (1983). Extensional vs Intuitive Reasoning. The Conjunction Fallacy in Probability Judgement. *Psychological Review*, 90(4), 293–315.
- Tversky, A., Kahneman, D. (1986). Rational Choice and the Framing of Decisions. *The Journal of Business*, 59(4), S251–S278.
- Tyszka, T. (2010). *Decyzje. Perspektywa psychologiczna i ekonomiczna*. Warszawa: Wydawnictwo Naukowe Scholar.
- Wennberg, E.J. (2010). *Tracking Medicine. A researcher's quest to understand health care*. New York: Oxford University Press.
- Westover, B.M., Westover, D.K., Bianchi, T.M. (2011). Significance testing as perverse probabilistic reasoning. *BMC Medicine*, 9(20).
- Williams, A.B. (2010). Perils of Evidence-Based Medicine. *Perspectives in Biology and Medicine*, 53(1), 106–120.
- Wolf, M.F., Gruppen, D.L., Billi, E.J. (1985). Differential Diagnosis and the Competing-Hypotheses Heuristic. *Journal of American Medical Association*, 253(19), 2858–2862.

- Wolff, G.J. (2006). Medical diagnosis as pattern recognition in a framework of information compression by multiple alignment, unification and search. *Decision Support Systems*, 42(2), 608–625.
- Wolter (2014). *Powiatki filozoficzne*. Poznań: Wydawnictwo Zysk i S-ka.
- Worall, J. (2007). Evidence in Medicine and Evidence-Based Medicine. *Philosophy Compass*, 2/6, 981–1022.
- Wójtowicz, A. (2018). Jakie wnioski uznajemy za racjonalne? Wpływ prawdopodobieństwa apriorycznego na prawdopodobieństwo aposterioryczne. *Decyzje*, 29, 41–65.
- Wójtowicz, A., Winkowski, J. (2018). Heuristics: Daniel Kahneman vs Gerd Gigerenzer. W: Hetmański, M. *Rationality and Decision Making. From Normative Rules to Heuristics* (253-277). Poznań: Studies in the Philosophy of the Sciences and the Humanities 111.
- Wulff, R.H., Gøtzsche, C.P. (2005), *Racjonalna diagnoza i leczenie. Wprowadzenie do medycyny wiarygodnej, czyli Evidence-Based Medicine*. Łódź: AKTIS.
- Vranas, M.B.P. (2000). Gigerenzer's normative critique of Kahneman and Tversky. *Cognition*, 76, 179–193.