

Autoreferat

Spis treści

1. Wstęp	1
2. Przebieg pracy naukowej i zawodowej	1
3. Jednotematyczny cykl publikacji	3
4. Omówienie publikacji stanowiących osiągnięcie	5
4.1. Wprowadzenie do obszaru i problematyki badań	5
4.2. Określanie preferencji społecznych dotyczących stanu zdrowia w Polsce	7
4.3. Metody wspomagania decyzji regulatora publicznego	11
4.4. Wykorzystanie zbiorów rozmytych do modelowania braku precyzji w preferencjach dotyczących stanu zdrowia	16
4.5. Podsumowanie wyników naukowych cyklu publikacji	18
5. Omówienie najważniejszych pozostałych publikacji	20
6. Podsumowanie dorobku naukowego	22
7. Skróty i akronimy stosowane w autoreferacie	24
8. Referencje	25

1. Wstęp

Celem autoreferatu jest zaprezentowanie mojego dorobku naukowego na tle pracy zawodowej. W części drugiej przedstawiłem przebieg mojej pracy naukowej i zawodowej. W części trzeciej wskazałem cykl publikacji powiązanych tematycznie i stanowiących osiągnięcie naukowe przedstawione do oceny i stanowiące podstawę do ubiegania się o stopień naukowy doktora habilitowanego w dziedzinie nauk ekonomicznych w dyscyplinie ekonomia, zaś w części czwartej scharakteryzowałem wyniki uzyskane w ramach tego cyklu. W kolejnej, piątej części omówiłem moje pozostałe najważniejsze publikacje naukowe. Dalej, w części szóstej, przedstawiłem podsumowanie mojego dorobku naukowego. W siódmej części zebrałem skróty i akronimy wykorzystane w autoreferacie, zaś ostatnia część zawiera referencje do przywołanych tu prac.

Mój pozostały dorobek naukowy (w tym realizowane projekty naukowe i wystąpienia konferencyjne) oraz dorobek dydaktyczny i organizacyjny scharakteryzowałem bardziej szczegółowo w Załączniku 4 do wniosku. Wykaz opublikowanych prac naukowych przedstawia dodatkowo analiza bibliometryczna wykonana przez Bibliotekę Szkoły Głównej Handlowej w Warszawie, również załączona do wniosku.

2. Przebieg pracy naukowej i zawodowej

Po ukończeniu IV L.O. im. H. Sawickiej w Kielcach studiowałem w latach 1996–2001 w Szkole Głównej Handlowej w Warszawie (SGH) na kierunku Metody Ilościowe i Systemy Informacyjne. Po przedstawieniu pracy magisterskiej „*Stopy zwrotu a dynamika chaotyczna – WGPW*”, której promotorem był prof. dr hab. T. Szapiro, ukończyłem studia z wynikiem celującym, otrzymując dyplom z wyróżnieniem i tytuł zawodowy magistra 13 czerwca 2001 r. W latach 1999–2003 równoległe studiowałem na Wydziale Matematyki, Informatyki i Mechaniki

Uniwersytetu Warszawskiego, gdzie 24 kwietnia 2003 r. uzyskałem tytuł zawodowy licencjata – promotorem mojej pracy pt. „*Paradoksy w rachunku prawdopodobieństwa*” był dr E. Stachowski.

Początki mojej pracy naukowej sięgają trzeciego roku studiów w SGH: wziąłem udział w badaniach w ramach grantu Komitetu Badań Naukowych „*Konstrukcja rozwiązań integracyjnych w hierarchicznych problemach negocjacyjnych o niekoherentnych strukturach preferencji*” (nr H02D 00510, kier.: prof. dr hab. T. Szapiro). Na czwartym roku studiów w SGH, dzięki stypendium Socrates-Erasmus, studiowałem jeden semestr na *Universidade Lusofona de Humanidades e Tecnologias* w Lizbonie w Portugalii. Na ostatnim roku studiów magisterskich odbyłem staż w Instytucie Ekonometrii, Kolegium Analiz Ekonomicznych (KAE) w SGH. W 2001 r. rozpocząłem studia doktoranckie w KAE, do moich zadań poza udziałem w badaniach naukowych należało prowadzenie ćwiczeń. W czasie studiów doktoranckich moje badania naukowe koncentrowały się na zagadnieniach wspomagania i modelowania decyzji sekwencyjnych z uwzględnieniem zjawiska niespójności czasowej, w kontekście zarówno dyskontowania hiperbolicznego, jak i wielokryterialnej oceny wariantów decyzyjnych. W latach 2004–2005 w ramach otrzymanego stypendium *Marie Curie* prowadziłem badania nad doktoratem w *Center for operations research and econometrics* w Louvain-la-Neuve w Belgii. W latach 2006–2013 pracowałem jako asystent w Zakładzie Farmakoekonomiki Warszawskiego Uniwersytetu Medycznego. W 2007 r. obroniłem w SGH pracę doktorską pt. „*O mechanizmach utraty efektywności w wielokryterialnej analizie sekwencyjnych problemów decyzyjnych*”, której promotorem był prof. dr hab. T. Szapiro.

W 2007 r. rozpocząłem pracę na stanowisku adiunkta w Zakładzie Wspomagania i Analizy Decyzji Instytutu Ekonometrii w KAE (od momentu rozpoczęcia było to moje pierwsze miejsce pracy, pracę tę kontynuuję do dziś). Prowadzę badania na styku teorii decyzji i jej zastosowań w obszarze ekonomii zdrowia oraz regularnie prowadzę wykłady w obszarze analizy i teorii decyzji w j. polskim i angielskim, na studiach stacjonarnych i niestacjonarnych, na poziomie licencjackim, magisterskim oraz MBA (łącznie prowadziłem wykłady na czterech programach MBA, współorganizowałem i współkieruję programem MBA oferowanym we współpracy z Warszawskim Uniwersytetem Medycznym), prowadziłem także w ramach zespołu wykłady na studiach doktoranckich. Promuję prace licencjackie (113) i magisterskie (57), jestem promotorem pomocniczym dwojga doktorantów.

W trakcie mojego zatrudnienia w SGH odbyłem kilkakrotnie staże i wizyty naukowe: w University of York, York, Zjednoczone Królestwo (2012 r., tydzień), University of Glasgow, Glasgow, Zjednoczone Królestwo (2013 r., tydzień), Università degli Studi di Milano-Bicocca, Mediolan, Włochy (2013 r., tydzień). W latach 2015–2016 w ramach stypendium *Fulbright Senior Award* przebywałem sześć miesięcy w Tippie College of Business w University of Iowa, Iowa City, Iowa, Stany Zjednoczone, wybranym przez mnie jako możliwość naukowej kontynuacji kontaktów nawiązanych jeszcze w trakcie stypendium *Marie Curie* (z prof. R. Amirem).

Od czasu uzyskania stopnia doktora moje zainteresowania naukowe koncentrowały się wokół metod wspomagania i analizy decyzji, a zwłaszcza decyzji podejmowanych w obszarze rynku ochrony zdrowia. W szczególności moje badania dotyczyły problematyki modelowania i szacowania preferencji społecznych dotyczących stanów zdrowia. Charakteryzowałem własności istniejących i proponowałem nowe, posiadające lepsze własności metody wspierania decyzji w kontekście ryzyka i wielokryterialności (m.in. tzw. krzywe akceptowalności), wykorzystywałem i rozwijałem przede wszystkim dwie metody szacowania preferencji – tzw. metodę handlowania czasem (ang. *time trade-off*) oraz metodę wyborów dyskretnych (ang. *discrete choice experiments*), zaproponowałem też i rozwijałem metody wykorzystania zbiorów rozmy-

tych do opisywania preferencji dotyczących stanów zdrowia — zarówno w ujęciu aksjomatycznym, jak i proponując empiryczne metody szacowania preferencji z uwzględnieniem inherentnego braku precyzji.

Obecnie kontynuuję badania naukowe, m.in. jako kierownik w ramach grantów finansowanych przez Narodowe Centrum Nauki (OPUS10, pt. „Metody analizy decyzji w problemach wielokryterialnych i szacowania skłonności do zapłaty/akceptacji przy wykorzystaniu modelowania rozmytego”) oraz przez stowarzyszenie naukowe EuroQol Group (pt. „A fuzzy approach to time trade-off experiment in EQ-5D-3L valuation” i “Building values sets based on TTO results by averaging model predictions and actually observed means”) oraz w ramach współpracy międzynarodowej jako partner naukowy w projekcie POLONEZ (“Evidence use in the allocation of health care budgets: A comparative study of drug reimbursement in Poland, Canada, and Germany”, kierownik: dr D. Wranik, uczelnia macierzysta: Dalhousie University, Kanada).

3. Jednotematyczny cykl publikacji

Efektorem badań prowadzonych przeze mnie po otrzymaniu stopnia doktora jest cykl publikacji powiązanych tematycznie, pod zbiorczym tytułem „**Modelowanie preferencji i wspomaganie decyzji na rynku ochrony zdrowia w ujęciu ekonomicznym**”:¹

1. **M. Jakubczyk** (2009): „**Impact of complementarity and heterogeneity on health related utility of life**”, *Central European Journal of Economic Modelling and Econometrics*, 1, 139-156;
2. **M. Jakubczyk**, B. Kamiński (2010): “**Cost-effectiveness acceptability curves – caveats quantified**”, *Health Economics*, 19 (8), 955–963, **IF = 1,946, MNiSW = 32**;
3. D. Golicki, **M. Jakubczyk**, M. Niewada, W. Wrona, J.J.V. Busschbach (2010): “**Valuation of EQ-5D Health States in Poland: First TTO-based Social Value Set in Central and Eastern Europe**”, *Value in Health*, 13 (2), 289–297, **IF = 2,342, MNiSW = 32**;
4. **M. Jakubczyk** (2015a): “**Using a Fuzzy Approach in Multi-Criteria Decision Making with Multiple Alternatives in Health Care**”, *Multiple Criteria Decision Making*, 10, 65–81, **MNiSW = 12**;
5. **M. Jakubczyk** (2015b): „**Rekurencyjne i konserwatywne reguły decyzyjne przy wyborze technologii medycznych w warunkach ryzyka**” w: D. Kopańska-Bródka (red.) „*Studia ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*”, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, ISSN 2083-8611, 235, 85–99, **MNiSW = 10**;
6. **M. Jakubczyk**, D. Golicki, M. Niewada (2016): “**The impact of a belief in life after death on health-state preferences: True difference or artifact?**”, *Quality of Life Research*, 25 (12), 2997–3008, **IF = 2,344, MNiSW = 35**;
7. **M. Jakubczyk**, B. Kamiński (2017): “**Fuzzy approach to decision analysis with multiple criteria and uncertainty in health technology assessment**”, *Annals of Operations Research*, 251 (1–2), 301–324, **IF = 1,709, MNiSW = 30**;

¹ Część projektów badawczych realizowana była ze współautorami ze względu na interdyscyplinarny charakter badań. Mój autorski wkład w pracach dotyczących zastosowań metod ekonomicznych do badania preferencji w obszarze zdrowia nie obejmował zagadnień czysto medycznych. Punktację *impact factor* (IF) i Ministerstwa Nauki i Szkolnictwa Wyższego (MNiSW) przypisano wg wyników analizy bibliometrycznej załączonej do wniosku.

8. **M. Jakubczyk, D. Golicki (2018): “Estimating the Fuzzy Trade-Offs Between Health Dimensions with Standard Time Trade-Off Data”** w: J. Kacprzyk, E. Szmidt, S. Zadrozny, K. Atanassov, M. Krawczak (red.) “Advances in Fuzzy Logic and Technology 2017. IWIFSGN 2017, EUSFLAT 2017.”, *Advances in Intelligent Systems and Computing*, vol 642. Springer, ISSN: 2194-5357; ISBN: 978-3-319-66823-9, DOI: 10.1007/978-3-319-66824-6, 266–277, **MNiSW = 5**;
9. **M. Jakubczyk, B. Kamiński, M. Lewandowski (2018): “Eliciting Fuzzy Preferences Towards Health States with Discrete Choice Experiments”**, w: C. Berger-Vachon, A. Gil Lafuente, J. Kacprzyk, Y. Kondratenko, J. Merigó, C. Morabito (red.) “Complex Systems: Solutions and Challenges in Economics, Management and Engineering”, *Studies in Systems, Decision and Control*, vol 125. Springer, ISBN 978-3-319-69988-2, DOI: 10.1007/978-3-319-69989-9, 131–148, **MNiSW = 5**;

Cykl wymienionych powyżej publikacji wskazuję jako osiągnięcie naukowe w rozumieniu ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (art. 16, ust. 1 i 2), stanowiące znaczny wkład w rozwój dyscypliny ekonomia. Prace te poruszają następujące, sekwencyjnie powiązane ze sobą i uzupełniające się problemy badawcze:

- określenie preferencji społecznych w Polsce dotyczących stanu zdrowia (z uwzględnieniem jakości i czasu trwania życia) w kontekście złożoności zdrowia jako dobra ekonomicznego i uwarunkowań demograficznych i światopoglądowych;
- określenie własności istniejących i konstrukcję nowych metod wspomagania decyzji regulatora publicznego dotyczących wyboru technologii medycznych na rynku ochrony zdrowia w ujęciu ekonomicznym, z wykorzystaniem miar skuteczności leczenia bazujących na miarach preferencji społecznych i z uwzględnieniem niepewności;
- uwzględnienie braku precyzji przy modelowaniu preferencji społecznych dotyczących stanu zdrowia z wykorzystaniem teorii zbiorów rozmytych.

Zaproponowany powyżej podział tematyczny wykorzystałem w dalszej części w podsumowaniu głównych wyników wymienionych badań. Wkład własny w powyższych pracach, które powstały we współautorstwie, podałem w Załączniku 4 (zgodnie ze stosownymi oświadczeniami współautorów, także załączonymi do wniosku). Wiele prac miało charakter interdyscyplinarny i wymagało odniesienia uzyskanych wyników do obszaru ochrony zdrowia i charakteryzowania jakości życia związanej ze zdrowiem, stąd współpraca z pracownikami naukowymi Warszawskiego Uniwersytetu Medycznego (D. Golickim i M. Niewadą). W szczególności, w badaniu Jakubczyk, Golicki i Niewada (2016) całkowicie moim wkładem było opracowanie koncepcji badania, zaprojektowanie i przeprowadzenie analiz ekonometrycznych i statystycznych oraz omówienie uzyskanych wyników w kontekście metod szacowania preferencji. W badaniu Golicki *et al.* (2010) głównie moim wkładem była przeprowadzona analiza ekonometryczna, zaś całkowicie moim wkładem była analiza statystyczna konsekwencji zwiększenia liczby eksperymentów u jednego respondenta. Z kolei w badaniu Jakubczyk i Golicki (2018) całkowicie moim wkładem było zaproponowanie i stworzenie koncepcji badania, skonstruowanie modelu preferencji oraz opracowanie i zaimplementowanie metody szacowania tych preferencji, jak również wykonanie obliczeń. Podobnie, całkowicie moim wkładem było opracowanie modelu preferencji wykorzystanego w badaniu Jakubczyk, Kamiński i Lewandowski (2018).

4. Omówienie publikacji stanowiących osiągnięcie

4.1. Wprowadzenie do obszaru i problematyki badań

Zdrowie może być traktowane jako przedmiot badań ekonomii zgodnie z jej klasycznymi definicjami, jako nauki badającej, „w jaki sposób społeczeństwo gospodarujące decyduje o tym, co, jak i dla kogo wytwarzać” (Begg, Fischer i Dornbusch, 1999), czy też jako „[nauki] o społecznych prawach rządzących produkcją i podziałem dóbr w społeczeństwie; [nauki] o racjonalnym gospodarowaniu” (Słownik języka polskiego). Zdrowie można przy tym traktować jako dobro konsumpcyjne (bycie zdrowym jest pożądane) lub kapitałowe (umożliwia pracę, zużywa się w czasie, można w nie inwestować, por. Grossman, 1972). Poprawa poziomu zdrowia w społeczeństwie ma korzystny wpływ na funkcjonowanie całej gospodarki (por. Suhrcke *et al.*, 2006, oraz Figueras i McKee, 2012).

Z rynkowego punktu widzenia przedmiotem transakcji nie jest zdrowie, lecz usługi zdrowotne (np. konsultacje lekarskie) lub technologie medyczne (np. leki). Rynek ten stanowi w krajach rozwiniętych istotną część całej gospodarki: w Polsce w roku 2015 wydatki bieżące na ochronę zdrowia wyniosły 6,34% Produktu Krajowego Brutto (PKB), przy czym wydatki publiczne stanowiły 4,44% PKB (dane Głównego Urzędu Statystycznego). W wielu krajach udział wydatków na opiekę zdrowotną w PKB jest wyższy i wynosi na przykład 6,6% na Litwie, 7,4% w Czechach i na Węgrzech, 8,1% na Słowacji, 9,1% w Zjednoczonym Królestwie, 11,3% w Niemczech i aż 17,1% w Stanach Zjednoczonych (dane Banku Światowego z 2014 r).

Rynek ochrony zdrowia posiada wiele wyróżniających go cech, będących przedmiotem badań ekonomicznych, por. Arrow (1963). Na przykład, nieodłącznym elementem dostarczania usług zdrowotnych jest asymetria informacji między konsumentami i dostawcami (lekarzami) – wiedza medyczna wymaga wieloletnich studiów i praktyki. Ponieważ efekt leczenia jest często odłożony w czasie i obciążony niepewnością o charakterze stochastycznym², konsument usług medycznych często nie ma możliwości zweryfikowania jakości tych usług. Dużego znaczenia nabiera zatem kwestia zaufania do dostawców i wymogów etycznych (np. przysięga Hipokratesa). Popyt na usługi zdrowotne jest stochastyczny (w odróżnieniu od np. popytu na żywność) i rozkłada się nierównomiernie między osobami. Potrzeba skorzystania z usług medycznych średnio narasta z wiekiem i zazwyczaj koncentruje się w obrębie jedynie części całego życia: np. wg badań około 31% wydatków ponoszonych jest w wieku średnim (40–64 lata), zaś 48,6% po 65. roku życia (por. Alemayehu i Warner, 2004); aż ok. 25% wydatków ponoszonych jest w ostatnim roku życia (por. Hogan *et al.*, 2001). Z uwagi na tę koncentrację wydatków, istotnym elementem rynku ochrony zdrowia są ubezpieczenia zdrowotne pozwalające na rozłożenie ryzyka zarówno w czasie życia, jak i między różne osoby.

Na rynek usług związanych ze zdrowiem wpływają nie tylko regulacje czysto rynkowe, dotyczące np. cen. Opieka zdrowotna traktowana bywa często jako szczególne prawo obywateli. Na przykład artykuł 68 Konstytucji Rzeczypospolitej Polskiej stwierdza: „każdy ma prawo do ochrony zdrowia”. Takie podejście do opieki zdrowotnej obecne jest także w systemach filozoficznych polityk społecznych. W teorii zdolności (ang. *capabilities*) Sena (1985) ocena dobrobytu społeczeństwa i kierunków polityki wymaga określenia możliwości dążenia przez obywateli do pożądanych sposobów funkcjonowania w życiu (ang. *functionings*). Nussbaum (2013) wskazuje w takim ujęciu, że troska o zdrowie powinna być zadaniem współczesnego państwa. W konsekwencji, zdrowie (oczekiwaną długość życia) włączono jako jeden z elementów miernika stopnia rozwoju społeczeństw, HDI (ang. *human development index*).

² Występuje tzw. niepewność I rzędu (ang. *first-order uncertainty*) – np. wyleczenia pacjenta można spodziewać się jedynie z określonym prawdopodobieństwem, por. Briggs *et al.* (2012).

Ze względu na powyższe uzasadnienia, w krajach rozwiniętych opieka zdrowotna jest w dużej mierze regulowana przez państwo i w znacznej części finansowana ze środków publicznych. Udział wydatków publicznych na opiekę zdrowotną w wydatkach całkowitych wynosi 71% w Polsce i choć jedynie 48,3% w Stanach Zjednoczonych, to już 66% na Węgrzech, 67,9% na Litwie, 72,5% na Słowacji i aż 77% w Niemczech, 83,1% w Zjednoczonym Królestwie i 84,5% w Czechach (dane Banku Światowego z 2014 r.). Jednym z obszarów regulacji jest ustalanie, które technologie medyczne są finansowane przez państwo. Ograniczone zasoby finansowe wymuszają dokonywanie wyborów, przy uwzględnieniu społecznych preferencji wyrażających się za pomocą kryteriów medycznych i ekonomicznych, a odzwierciedlanych w ramach tzw. analiz kosztów-efektywności³ (CEA, ang. *cost-effectiveness analysis*) lub analiz kosztów-użyteczności⁴ (CUA, ang. *cost-utility analysis*). Tworzenie metod tej analizy wymaga korzystania ze zbudowanych uprzednio modeli indywidualnych i grupowych preferencji podmiotów podejmujących decyzje. Co istotne, szczególnie z ekonomicznego punktu widzenia, analizowane technologie medyczne konkurują o środki finansowe z jednego budżetu publicznego, a jednocześnie generują efekty różnego rodzaju — np. wydłużają życie lub poprawiają jakość życia związaną ze zdrowiem (HRQoL, ang. *health-related quality of life*) — tym samym konieczne jest sprowadzenie tych różnych efektów do miary umożliwiającej porównywanie.

W praktyce porzuca się próbę maksymalizacji dobrobytu społecznego (tzw. podejście dobrobytowe, ang. *welfarist approach*) z uwagi np. na trudności praktyczne i wątpliwości natury etycznej (różnicowanie opieki zdrowotnej w zależności od oszacowanych różnic w preferencjach). Wciąż, w pragmatycznym podejściu decydenckim (ang. *extra-welfarist* lub *decision maker's approach*, por. Brouwer i Koopmanschap, 2000) miarę zdrowia definiuje się przez odwołanie do preferencji indywidualnych członków społeczeństwa w odniesieniu do stanów zdrowia (jakości i długości życia), aby decyzje podejmowane w oparciu o tę miarę nawiązywały do potrzeb społeczeństwa mierzonych preferencjami. Najczęściej wykorzystuje się tzw. model teoretyczny QALY (ang. *quality-adjusted life years*, tj. lat życia skorygowanych o jakość), którego aksjomatyczne podstawy przedstawili Pliskin, Shepard i Weinstein (1980), Bleichrodt, Wakker i Johannesson (1997) oraz Miyamoto *et al.* (1998). Każdy stan zdrowia, Q , ma przypisaną liczbę $u(Q)$ interpretowaną jako użyteczność spędzenia jednego roku w tym stanie. W przyjętej aksjomatyzacji użyteczność można traktować w sensie von Neumanna-Morgenssterna (1947), tj. postulować maksymalizację wartości oczekiwanej tej użyteczności przy wyborach w warunkach ryzyka. Wartości QALY normalizuje się (tj. zgonowi przypisuje się wartość 0, zaś pełnemu zdrowiu wartość 1).

Powyższe rozważania mają w wielu krajach istotny walor praktyczny, w tym także w Polsce. Ustawa refundacyjna (2011) i towarzyszące akty prawne nakazują przy decyzjach o refundacji technologii lekowych wyrażanie korzyści zdrowotnych z wykorzystaniem QALY i obliczenie kosztu uzyskania tych efektów. Przy szacowaniu faktycznych korzyści dodatkowo pojawia się problem niepewności oszacowań, wynikający z estymacji w oparciu o dane z prób losowych w badaniach klinicznych (kwestie te omawiam bardziej szczegółowo nieco dalej).

³ Określenie „kosztów-efektywności” jest często używane w polskiej literaturze w kontekście oceny technologii medycznych i dlatego tu przywołane (np. Wytyczne Agencji Oceny Technologii Medycznych i Taryfikacji idą dalej i korzystają z określenia „analiza efektywności kosztów”), choć z punktu widzenia ekonomii efektywność jest także traktowana jako sposób ustalenia procesu wytwórczego, aby przy danych nakładach uzyskać maksymalny możliwy efekt, czyli samo pojęcie „efektywność” uwzględnia już kwestię kosztów działania. Dodatkowo angielskie „*effectiveness*” należy raczej powiązać ze słowem „efekt”, zaś słowu „efektywność” odpowiada angielski termin „*efficiency*”. W dalszej części tekstu wykorzystywane są akronimy CEA lub CUA.

⁴ Określenie to stosuje się niekiedy, gdy jednostkę efektu można interpretować jako miarę użyteczności w rozumieniu teorii decyzji.

W tak zdefiniowanym obszarze prowadziłem badania mające na celu:

- określenie preferencji społecznych w Polsce dotyczących stanu zdrowia (z uwzględnieniem jakości i czasu trwania życia) w kontekście złożoności zdrowia jako dobra ekonomicznego i uwarunkowań demograficznych i światopoglądowych;
- określenie własności istniejących i konstrukcję nowych metod wspomagania decyzji regulatora publicznego dotyczących wyboru technologii medycznych w ujęciu ekonomicznym, z uwzględnieniem niepewności i z wykorzystaniem miar skuteczności leczenia bazujących na miarach preferencji społecznych;
- uwzględnienie braku precyzji w modelach preferencji społecznych dotyczących stanu zdrowia, wykorzystując teorię zbiorów rozmytych.

Otrzymane w powyższych trzech nurtach wyniki omówiłem odpowiednio w częściach 4.2, 4.3 i 4.4.

4.2. Określanie preferencji społecznych dotyczących stanu zdrowia w Polsce

Pierwszym badaniem w obszarze pomiaru preferencji dotyczących zdrowia, w którym uczestniczyłem, było badanie **Golicki et al. (2010)**. Celem było określenie preferencji społecznych dotyczących stanów zdrowia w Polsce i przypisanie poszczególnym stanom zdrowia średnich w populacji wartości użyteczności w ramach modelu QALY — tj. zbudowanie tzw. zestawu wartości użyteczności (ang. *value set*). Dodatkowo nowatorskim wynikiem była ocena konsekwencji powiększenia liczby eksperymentów myślowych stosowanych u jednego respondenta.

W badaniu do zdefiniowania i opisanie HRQoL wykorzystano system opisowy EQ-5D-3L, opracowany przez grupę EuroQol Group (por. Brooks i De Charro, 1996, Brazier et al., 2017). Stan zdrowia jest opisany przez pięć tzw. domen (stąd 5D): poruszanie się (MO, ang. *mobility*), samoopiekę (SC, ang. *self-care*), możliwość wykonywania codziennych czynności (UA, ang. *usual activities*), odczuwanie bólu/dyskomfortu (PD, ang. *pain/discomfort*) i odczuwanie niepokoju/przygnębienia (AD, ang. *anxiety/depression*). Każda domena jest opisana przez jeden z trzech poziomów (stąd 3L): 1 oznacza brak problemów (np. z poruszaniem się), zaś 3 oznacza duże problemy (np. bycie przykutym do łóżka). Kwestionariusz EQ-5D-3L można wykorzystać dla różnych chorób (jest tzw. kwestionariuszem generycznym) i jest preferowaną metodą pomiaru HRQoL we wnioskach ocenianych przez *National Institute for Clinical Excellence* w Anglii.

Badaniem objęto 321 respondentów o strukturze płci i wieku reprezentatywnej dla populacji dorosłych Polaków. W literaturze najczęściej wskazuje się, że respondenci powinni rekrutować się z populacji generalnej — ponoszącej koszty ubezpieczenia zdrowotnego, a dodatkowo z powodów pragmatycznych (trudno znaleźć osoby w niektórych, rzadkich stanach zdrowia), por. Whitehead i Ali (2010) oraz Drummond et al. (2009). W omawianym badaniu respondentami byli odwiedzający w ośmiu szpitalach w Polsce. Każdy respondent w zasadniczej części eksperymentu oceniał użyteczność 23 stanów zdrowia (wybranych na podstawie Lamers et al., 2006, por. Tabela 1 w oryginalnej publikacji). W badaniu wykorzystano metodę handlowania czasem⁵ (TTO, ang. *time trade-off*) do pomiaru użyteczności, por. Attema et al. (2013). W metodzie TTO dla danego stanu Q respondent w serii pytań proszony był o wybór między dwoma wariantami: 10 lat w stanie Q (po których następuje zgon) i T lat w stanie pełnego zdrowia (po których następuje zgon). W modelu QALY, indyferencja wskazuje $u(Q) = T/10$, dla

⁵ Takie tłumaczenie zostało powszechnie przyjęte w polskiej literaturze przedmiotu.

znormalizowanych użyteczności: $u(\text{pełne zdrowie})=1$ i $u(\text{zgon})=0$. W omawianej pracy zmodyfikowaną metodę TTO wykorzystano dla stanów Q ocenianych jako gorsze od zgonu, tj. dla których natychmiastowy zgon był preferowany w porównaniu do 10 lat w stanie Q .

Po oczyszczeniu danych otrzymano 6777 obserwacji pochodzących od 305 respondentów. Uzyskane oceny użyteczności ekstrapolowano na pełen zbiór 243 stanów z wykorzystaniem modelowania ekonometrycznego wykorzystującego podejście panelowe w modelu efektów losowych (por. Dolan *et al.*, 1996).⁶ Wykorzystano zmienne zero-jedynkowe przypisane kombinacjom domena-poziom, aby określić addytywny wpływ pogorszenia poszczególnych domen na użyteczność, por. Tabela 5 w oryginalnej publikacji.

Nowatorskim elementem badania było istotne zwiększenie liczby stanów ocenianych przez jednego respondenta do 23. Uprzednio zazwyczaj wykorzystywano 13 (por. Shaw, Johnson i Coons, 2005, Dolan *et al.*, 1996, Badia *et al.*, 2001, Greiner *et al.* 2005) lub 7 stanów (por. Jelsma *et al.*, 2003); okazjonalnie liczbę zwiększano do 16 lub 17 (por. Wittrup-Jensen *et al.*, 2002, Tsuchiya *et al.*, 2002, Lamers *et al.*, 2006). Zbadano wpływ rozszerzenia badania na wiarygodność otrzymywanych ocen, testując dwie hipotezy: (H1) na średnią wartość użyteczności przypisaną stanowi zdrowia nie wpływa, czy stan jest oceniany w środkowej (na pozycjach 6–17) czy w końcowej części eksperymentu (na pozycjach 18–23); (H2) wariancja ocen użyteczności nie różni się dla pozycji 6–17 i 18–23. Testowanie hipotezy H1 miało na celu określenie, czy wydłużenie eksperymentu TTO może wprowadzić obciążenie wyników (np. respondenci zmęczeni eksperymentem przypisują większe wartości użyteczności, nie chcąc korzystać z bardziej skomplikowanej wersji TTO dla stanów gorszych od zgonu). Testowanie hipotezy H2 – sprawdzenie, czy wydłużenie eksperymentu może zmniejszyć precyzję estymatorów np. przez zwiększenie rozrzutu udzielanych odpowiedzi z uwagi na zmniejszoną koncentrację. W analizie tej pominięto pozycje 1–6 ze względu na możliwość postępującego procesu uczenia się decydenta. Z uwagi na wielokrotne testowanie (44 stany zdrowia) i wspólne wnioskowanie dla całej rodziny hipotez (że wydłużenie eksperymentu jest problematyczne) wykorzystano korektę Hölma-Bonferroniego. Nie odrzucono żadnej z dwóch hipotez, wykazując możliwość rozszerzenia eksperymentu i potwierdzając wiarygodność uzyskanych wyników (por. Tabela 7 w oryginalnej publikacji).

Omawiana praca była pierwszym badaniem użyteczności w systemie EQ-5D-3L w Europie Centralnej i Wschodniej wykorzystującym metody oparte na wyborach (ang. *choice-based*) do szacowania użyteczności stanu zdrowia.⁷ Porównanie zgodności otrzymanych wartości użyteczności względem innych krajów (Niemiec, Zjednoczonego Królestwa i Słowenii) wykazało różnice w strukturze preferencji (np. większe względne znaczenie domeny PD w Polsce, por. Tabela 8 i Rysunek 1 w oryginalnej publikacji), a przez to zasadność stosowania polskich wartości przy podejmowaniu decyzji przez regulatora publicznego w Polsce, aby odzwierciedlić właściwą dla kraju strukturę preferencji. Opracowane normy użyteczności (por. Tabela 6 w oryginalnej publikacji), tj. przyporządkowanie stanom zdrowia wartości użyteczności, jest

⁶ Modelowania tego nie trzeba przeprowadzać w tzw. badaniach nasyconych (ang. *saturation studies*), w których wykorzystuje się wszystkie stany zdrowia (por. np. Santos *et al.*, 2016). Wadą badań nasyconych jest duży koszt badania (są 243 stany zdrowia w systemie EQ-5D-3L) oraz trudność wyobrażenia sobie przez respondenta niektórych stanów zdrowia (np. bycie przykutym do łóżka i odczuwanie krąpcowego bólu bez żadnego przygnębienia).

⁷ Wcześniejsze badanie w Słowenii (Prevolnik Rupel i Rebojl, 2000) wykorzystywało metodę wizualnej skali analogowej (VAS, ang. *visual analogue scale*), niemającej podstaw teoretycznych, np. w ramach modelu QALY.

oficjalnie rekomendowane w analizach ekonomicznych przez wytyczne Agencji Oceny Technologii Medycznych (2009) oraz nowsze wytyczne Agencji Oceny Technologii Medycznych i Taryfikacji (AOTMiT, 2016).⁸

Omówione badania kontynuowałem w ramach pracy **Jakubczyk (2009)**⁹, której celem było, po pierwsze, zrozumienie heterogeniczności preferencji dotyczących stanów zdrowia, tj. poznanie czynników demograficznych wpływających na te preferencje, a po drugie poznanie mechanizmów kształtujących strukturę preferencji u poszczególnych respondentów — określenie interakcji między atrybutami charakteryzującymi stan zdrowia (domenami EQ-5D) i wpływającymi na przypisaną użyteczność, tj. zrozumienie ekonomicznego charakteru współzależności (czy atrybuty są względem siebie komplementarne czy substytucyjne). Inspiracją do podjęcia badań były współczesne przemiany demograficzne i rosnąca świadomość znaczenia tych przemian dla procesów gospodarowania.

W badaniu wykorzystałem dane z badania Golicki *et al.* (2010), testując inną procedurę czyszczenia danych (ostatecznie w zbiorze danych znalazło się 6701 pojedynczych ocen użyteczności dla 295 respondentów). Zbudowałem cztery modele regresji liniowej, oznaczane M1–M4. W M1 wykorzystano jedynie zmienne opisujące domeny (tj. zmienne typu $d_{i,j}$ wskazujące, czy domena i jest na poziomie j , $i = 1, \dots, 5$, $j = 1, 2, 3$), nie uwzględniono żadnych cech demograficznych ani możliwych interakcji między domenami; M1 stanowił punkt odniesienia dla pozostałych modeli. W modelu M2 dodatkowo uwzględniono zmienne opisujące interakcje między domenami we wpływie na użyteczności — zmienne N2, N3, I2, I3, D1 (i kwadraty trzech ostatnich) występujące w literaturze (por. Tsuchiya *et al.*, 2002, Greiner *et al.*, 2004, Lamers *et al.*, 2006), a także iloczyny par zmiennych $d_{i,3}$. Uwzględnienie tej drugiej grupy zmiennych było nowatorskim elementem badania i miało na celu bezpośrednie określenie interakcji między konkretnymi aspektami HRQoL (ze względu na umiarkowaną wielkość próby ograniczono się do interakcji na trzecim poziomie wartości domen). W modelu M3 uwzględniono zmienne demograficzne (np. płeć, wiek, własny stan zdrowia) — zarówno indywidualnie, jak i w iloczynach ze zmiennymi $d_{i,2}$ oraz $d_{i,3}$. Modele M1–M3 szacowano jako modele panelowe przy założeniu efektów losowych. Z modeli usuwano zmienne nieistotne (przy poziomie 0,05), testowano także łączną nieistotność usuwanych zmiennych.

Model M4 zbudowano jako model losowych parametrów (por. np. Maddala, 2006), tj. założono, że wpływ zmiennych $d_{i,j}$ na użyteczność różni się dla poszczególnych respondentów i jest zmienną losową generowaną z rozkładu prawdopodobieństwa (założono rozkład normalny) opisującego daną populację; estymacji podlega wówczas średnia tego rozkładu. Model M4 szacowano w podejściu bayesowskim metodą MCMC (podejście bayesowskie testowano już w badaniu Golicki *et al.*, 2010, lecz w modelu stałych parametrów).

Najważniejsze wnioski uzyskane z badania Jakubczyk (2009) przedstawiają się następująco. Model M1 potwierdził wyniki badania Golicki *et al.* (2010), że PD (ból/dyskomfort) jest

⁸ Kontynuuję omówiony nurt badań — obecnie realizujemy projekt mający na celu m.in. zbudowanie norm użyteczności dla Polski z wykorzystaniem nowo wprowadzonego systemu opisowego EQ-5D-5L. Wstępne wyniki prezentowano już na konferencjach międzynarodowych: D. Golicki, M. Jakubczyk, M. Niewada, K. Graczyk (2017): plakat pt. „Polish Tariff for the EQ-5D-5L”, 34th EuroQol Plenary Meeting, Barcelona, Hiszpania, oraz D. Golicki, M. Niewada, K. Graczyk (2017): plakat pt. “The First EQ-5D-5L Value Set in Central and Eastern Europe”, International Society for Pharmacoeconomics and Outcomes Research Annual European Congress, Glasgow, Wielka Brytania, i krajowych: np. M. Jakubczyk (2017): referat pt. „Modelowanie użyteczności stanu zdrowia — aspekty ekonometryczne (czyli piki, triki i wyniki)”, XV Konferencja Polskiego Towarzystwa Farmakoekonomicznego, Warszawa, Polska.

⁹ Wcześniejsza data wynikała z procesu wydawniczego dla publikacji Golicki *et al.* (2010), dostępnej w Internecie już w 2009 r.

domeną o największym wpływie (na poziomie 3) na użyteczność stanu zdrowia w Polsce, zaś domena AD (niepokój/przygnębienie) – o najmniejszym (pogorszenie tej domeny na poziomie 2 nie miało statystycznie istotnego wpływu na użyteczność).

Wyniki oszacowań modelu M2 (por. Tabela 5 w oryginalnej publikacji) wskazują na współzależności w kształtowaniu użyteczności stanu zdrowia między domeną PD i wszystkimi pozostałymi domenami. Ujemne wartości współczynników (zmienną objaśnianą była tzw. dysużyteczność, tj. różnica między użytecznością stanu a 1 – znormalizowaną użytecznością stanu pełnego zdrowia) wskazują na komplementarność domen – pogorszenie kolejnej domeny ma mniejszy negatywny wpływ na użyteczność. Jedną z możliwych interpretacji jest, że respondenci wyobrażają sobie (ew. wiedzą z doświadczenia), że odczuwanie znacznego bólu utrudnia korzystanie z życia, mimo nieograniczonej możliwości np. poruszania się.

Wyniki modelu M3 (por. Tabela 6 w oryginalnej publikacji) pokazały po pierwsze, że zmienne demograficzne wpływają na odczuwaną użyteczność głównie poprzez interakcje ze zmiennymi $d_{i,3}$, tj. poprzez kształtowanie wpływu pogorszenia domen na poziom 3. Najważniejsze wnioski to: mężczyźni przywiązują większą wagę do (pogorszenia) domeny SC i mniejszą do domen UA oraz PD; osoby wykształcone przywiązują większą wagę do domen SC i PD; osoby, które same odczuwają problemy w domenie MO, przypisują pogorszeniu tej domeny mniejszą dysużyteczność, zaś przeciwny efekt obserwujemy dla domeny AD (co może tłumaczyć mały wpływ tej domeny w modelu – respondenci z populacji ogólnej, zwłaszcza z uwagi na sposób doboru próby w badaniu Golicki *et al.*, 2010, nie doceniają faktycznego wpływu problemów psychologicznych).

Nowatorskim i nieoczekiwanym wynikiem badania Jakubczyk (2009) było wykazanie wpływu wiary w życie po śmierci (zmienna *faith* w Tabeli 6 w oryginalnej publikacji) – jako przybliżenie religijności – na odczuwaną użyteczność. Zmienna *faith* w ostatecznej specyfikacji wystąpiła w największej liczbie interakcji spośród zmiennych demograficznych, zaś jedna z tych interakcji charakteryzowała się największym (co do modułu) wpływem na użyteczność (interakcja ze zmienną $d_{1,3}$ opisującą pogorszenie domeny UA). Osoby religijne (dokładniej – wierzące w życie po śmierci) przypisywały mniejszą utratę użyteczności stanom zdrowia, zwłaszcza stanom, w których pogorszona jest możliwość wykonywania codziennych czynności. Wnioski z modelu M3 wskazują, że prowadzone badania preferencji powinny zadbać o reprezentatywność próby w zakresie ww. zmiennych demograficznych.

Wreszcie wyniki dla modelu M4 pokazały (por. Tabela 6 w oryginalnej publikacji), że respondenci różnią się istotnie między sobą co do odczuwanej ważności poszczególnych domen. Wykorzystanie podejścia losowych parametrów było nowatorskim elementem badania (nieznane mi są wcześniejsze publikacje wykorzystujące takie podejście do szacowania użyteczności stanów zdrowia w EQ-5D-3L w literaturze). W literaturze przedmiotu dopiero od kilku lat modele parametrów losowych stają się standardowym podejściem do specyfikowania modelu (np. Rowen *et al.*, 2014, Shirowa *et al.*, 2016, Feng *et al.*, 2017).

Wyniki uzyskane w badaniu Jakubczyk (2009) stanowiły impuls do przeprowadzenia dokładniejszej analizy związków między religijnością i preferencjami dotyczącymi stanów zdrowia. I tak, celem badania **Jakubczyk, Golicki i Niewada (2016)** było określenie wpływu religijności na użyteczność stanów zdrowia szacowaną w eksperymentach TTO i na końcowy zestaw użyteczności oraz przedyskutowanie wyników w kontekście konsekwencji dla wykorzystania tych zestawów w analizach typu CEA i ew. modyfikacji metod szacowania preferencji (wskazanie, na ile różnice użyteczności powinny być traktowane jako inherentny element preferencji, a na ile efekt konkretnej metody szacowania preferencji).

Wykorzystano dane z badania Golicki *et al.* (2010). Religijność określono na podstawie zmiennej mierzącej wiarę w życie po śmierci na trzech poziomach: silna religijność (*zdecydowanie wierzę*), umiarkowana religijność (*zdecydowanie wierzę* lub *raczej wierzę*) i brak religijności (*raczej nie, zdecydowanie nie* i *nie wiem*). W badaniu uwzględniono zmienne demograficzne, aby uwzględnić ewentualny wpływ innego czynnika (ang. *confounding*) jednocześnie oddziałującego na religijność i użyteczność (np. wieku lub wykształcenia).

Najważniejsze wyniki uzyskane w badaniu przedstawiają się następująco. Osoby religijne częściej nie były skłonne zrezygnować z czasu życia w eksperymencie TTO w zamian za poprawę HRQoL (osoby takie w literaturze anglojęzycznej są określane jako tzw. *non-traders*): iloraz szans wynosi ok. 2 dla osób silnie religijnych (w porównaniu do osób pozostałych, tj. umiarkowanie religijnych i niereligijnych) i ok. 1,5 dla osób umiarkowanie/silnie religijnych (w porównaniu do niereligijnych). Wyniki te uzyskano dla analizy jednoczynnikowej i dla analizy z wykorzystaniem testu Cochran-Mantel-Haenszela oraz wieloczynnikowej regresji logistycznej pozwalającej na wyeliminowanie wpływu innych zmiennych. Osoby religijne (zarówno silnie, jak i umiarkowanie) rzadziej uznają stan za gorszy od zgonu. Badanie potwierdziło zatem wpływ religijności na wyniki eksperymentu TTO.

Wpływ religijności objawia się głównie gotowością do oddania jakiegokolwiek czasu życia w eksperymencie TTO, a w dużo mniejszym stopniu wielkością oddanego czasu (wśród osób, które zdecydowały się na oddanie choć części czasu), por. Tabela 4 w oryginalnej publikacji. Sugeruje to, że osoby religijne częściej nie akceptują eksperymentu myślowego wykorzystanego w metodzie TTO (decydowania o wybraniu lepszego lecz krótszego życia), z uwagi na (być może podświadome) skojarzenia z decydowaniem w obszarach ostatecznych, zarezerwowanych (wg osób religijnych, m.in. w wierze chrześcijańskiej) dla Stwórcy. Taka interpretacja jest uwiarygodniona wynikami jakościowymi z przebiegu innych eksperymentów TTO, por. Au Eong *et al.* (2012) i Wee *et al.* (2008) oraz związkami religijności i wyników TTO i VAS raportowanymi w innych badaniach (por. Augestad *et al.*, 2013, i komunikacja osobista z K. Rand-Hendriksenem). Z tego względu w konkluzji w badaniu rekomendowano usunięcie wpływu religijności na oszacowania użyteczności¹⁰

Osoby silnie religijne przypisują średnio stanom użyteczność o 0,136 większą (w modelu addytywnym) lub przypisują dysużyteczność mniejszą o czynnik 2,1 (w modelu multiplikatywnym). W badaniu zbudowano skorygowany zestaw użyteczności (por. Tabela 3 w oryginalnej publikacji); średnia różnica wartości wyniosła 0,046.

Nowatorskim elementem badania było systematyczne zbadanie wpływu religijności na użyteczność dotyczącą stanu zdrowia z wykorzystaniem palety technik analitycznych (w celu wyeliminowania np. wpływu innych zmiennych) w dużej próbie losowej. Dodatkowo w pracy przedstawiono metodę modyfikacji TTO, pozwalającej na zbadanie, czy efekt obserwowany w danych jest spowodowany różnicą w samych preferencjach, czy też konsekwencją stosowania klasycznej TTO jako metody szacowania preferencji (ilustracja na Rysunku 3 i dyskusja w tekście w oryginalnej publikacji).

4.3. Metody wspomaganie decyzji regulatora publicznego

Badania preferencji dotyczących stanów zdrowia służą nie tylko celom poznawczym. Jak przedstawiono we wstępie, motywacją do ich przeprowadzania jest chęć wykorzystania tych oszacowań do wspomaganie decyzji o wyborze technologii medycznych. Zamierzeniem jest określenie

¹⁰ Tego typu podejście zapowiedziano w kontekście toczącego się badania szacowania preferencji dla stanów zdrowia określonych wg EQ-5D-5L, w czasie prezentacji konferencyjnej, por. M. Jakubczyk (2017): referat pt: „Modelowanie użyteczności stanu zdrowia — aspekty ekonometryczne (czyli piki, triki i wyniki)”, XV Konferencja Polskiego Towarzystwa Farmakoekonomicznego, Warszawa, Polska.

korzyści klinicznych stosowania danej technologii medycznej, które zestawia się z kosztami stosowania technologii (por. np. Drummond *et al.*, 2015).

Jak wspomniano, w procesie leczenia pacjentów występuje niepewność I rzędu (tj. wynik leczenia w konkretnym przypadku jest niepewny). W analizach ekonomicznych niepewność tę zazwyczaj pomija się, gdyż QALY jako miara oparta na użyteczności von Neumanna-Morgensterna, inkorporuje w sobie stosunek decydentów do ryzyka wobec niepewnego stanu zdrowia, zaś koszt leczenia jednego pacjenta rozkłada się na całe społeczeństwo, co skutkuje neutralnością względem ryzyka z uwagi na twierdzenie Arrowa-Linda (1970). W analizach ekonomicznych istotne staje się zatem oszacowanie wartości oczekiwanej korzyści (w QALY) i kosztu. Decyzje publiczne często wykorzystują stosunek korzyści ze stosowania danej technologii (dodatkowa oczekiwana liczba QALY przy stosowaniu rozważanego wariantu w miejsce dotychczas stosowanego) do kosztu – tzw. inkrementalny współczynnik kosztów-efektywności (ICER, ang. *incremental cost-effectiveness ratio*). W literaturze wskazano warunki optymalności takiego podejścia, por. Weinstein i Zeckhauser (1973) oraz Laska *et al.* (1999). Wykorzystanie współczynnika ICER wskazano jako jedno z kryteriów w ustawodawstwie polskim (Ustawa Refundacyjna, art. 12., ust. 13., oraz Rozporządzenie Ministra Zdrowia w sprawie minimalnych wymagań, par. 5., ust. 2., pkt. 2.).

Oszacowania wartości oczekiwanych w praktyce wykorzystują dane o charakterze prób losowych (np. randomizowane badania z grupą kontrolną), więc uzyskane estymatory obarczone są błędem stochastycznym (występuje tzw. niepewność II rzędu, por. Briggs *et al.*, 2012). Dodatkowo istotnym elementem jest brak precyzyjnie określonej skłonności decydenta do płacenia (WTP, ang. *willingness-to-pay*) za efekty kliniczne, tj. monetarna wycena QALY (podejmowane są próby określenia prognozy na podstawie obserwowanych decyzji, por. Devlin i Parkin, 2004, lub Niewada *et al.*, 2013).¹¹ Jedną z metod oceny niepewności związanej z dostępnymi wariantami decyzyjnymi (technologiami medycznymi) najczęściej wykorzystywanych w decyzjach publicznych są tzw. krzywe akceptowalności (CEAC, ang. *cost-effectiveness acceptability curves*), por. van Hout *et al.* (1994), Löthgren i Zethraeus (2000) oraz Fenwick *et al.* (2004). Krzywe CEAC pokazują (przy interpretacji bayesowskiej) prawdopodobieństwo tego, że dany wariant jest optymalny (z uwagi na kryterium kosztu i efektu) w zależności od wartości WTP, por. Briggs (1999). Krzywe CEAC są rekomendowane do stosowania w Polsce (wytyczne AOT-MiT, 2016).

W literaturze wskazywano nieintuicyjne własności krzywych CEAC. Fenwick *et al.* (2001) pokazali, że przy wyborze między dwoma wariantami dla danego poziomu WTP wybór wariantu o większym prawdopodobieństwie maksymalizacji tzw. korzyści netto (NB, ang. *net benefit*) i wybór wariantu maksymalizującego oczekiwaną wartość NB są równoważne, tylko jeśli niepewność charakteryzująca oba warianty (efekty i koszty) ma symetryczny rozkład. Fenwick *et al.* (2004) podali przykład niemonotonicznych krzywych CEAC przy porównywaniu dwóch wariantów (tj. naruszających intuicyjne oczekiwanie, że wzrost wartości WTP powinien faworyzować jeden z dwóch wariantów). Barton *et al.* (2008) zauważyli, że w przypadku wyboru spośród więcej niż dwóch wariantów, także dla symetrycznych rozkładów prawdopodobieństwa opisujących niepewność oszacowania charakterystyk wariantów, wariant maksymalizujący oczekiwaną wartość NB dla pewnego poziomu WTP może nie maksymalizować prawdopodobieństwa oferowania maksymalnego NB dla żadnego poziomu WTP.

¹¹ Obserwacja ta była także motywacją do badań Jakubczyk i Kamiński (2017) oraz Jakubczyk (2015a) włączonych do cyklu publikacji i omówionych w dalszej części.

Celem badania **Jakubczyk i Kamiński (2010)**, umiejscowionego w omówionym powyżej nurcie w literaturze, było formalne scharakteryzowanie niepożądanych własności metody CEAC w różnych klasach sytuacji decyzyjnych (rozpatrywano wybór między dwiema technologiami oraz spośród $n > 2$ technologii; rozważano niepewność scharakteryzowaną rozkładem normalnym lub niesprecyzowanymi klasami rozkładów; uwzględniono możliwość wystąpienia korelacji między niepewnością oszacowań dla kosztów i efektów). Badanie miało charakter teoretyczny. Rozważano lokalne (tj. dla zadanego poziomu WTP) oraz globalne (tj. przebieg dla różnych WTP) własności krzywych CEAC.

Przy wyborze z $n = 2$ wariantów pokazano, jakim ograniczeniom podlega możliwa rozbieżność między oceną atrakcyjności wariantu obliczoną jako standaryzowana (tj. podzielona przez odchylenie standardowe) oczekiwana różnica NB oraz prawdopodobieństwo maksymalizacji NB przez rozważany wariant (s. 957 wg numeracji oryginalnej publikacji). Wskazano także oszacowanie górne dla wartości oczekiwanej doskonałej informacji, gdy decydent dokonuje wyboru kierując się maksymalizacją prawdopodobieństwa maksymalizacji NB.

W badaniu wskazano konsekwencje i ograniczenia stosowania CEAC przy porównywaniu wariantów w przypadku wystąpienia dominacji stochastycznej pierwszego rzędu (FOSD, ang. *first-order stochastic dominance*). W sytuacji wyboru spośród dwóch wariantów, FOSD gwarantuje zgodność podejść opartych na maksymalizacji wartości oczekiwanej oraz prawdopodobieństwa. W przypadku wielu wariantów (gdy błędy oszacowań ich atrybutów są niezależne) krzywa CEAC dla wariantu FOSD-zdominowanego musi znajdować się poniżej (wskazywać mniejszą atrakcyjność dla decydenta) krzywej dla wariantu dominującego (nie przedstawiono dowodu twierdzenia, tę lukę uzupełniłem w badaniu Jakubczyk, 2015b, opisanym w dalszej części).

Wskazano, że dla $n > 2$ wariantów, porównania parami z wykorzystaniem CEAC (tj. porównując warianty parami z uwagi na prawdopodobieństwo, że dany z nich oferuje większy NB) może generować cykliczną relację ostrych preferencji, a jednocześnie przedstawiono oszacowanie górne najmniejszego prawdopodobieństwa w takim cyklu (75% w ogólności, tj. dla dowolnej liczby wariantów).

Kolejnym nowatorskim wynikiem uzyskanym w badaniu było pokazanie, że stosowanie CEAC jako reguły wyboru (tj. dla danego poziomu WTP wybór wariantu maksymalizującego spośród wszystkich prawdopodobieństwa tego, że oferowany NB jest największy) narusza własność α funkcji wyboru, tj. dodanie nowego elementu do zbioru wariantów może spowodować, że zostanie wybrany uprzednio dostępny lecz niewybrany element (inaczej – niespełniona jest własność niezależności od nieistotnych wariantów). Podobny wynik pokazali wcześniej Sاداتsafavi *et al.* (2008), ale przedyskutowali tę własność jedynie w kontekście statycznego zbioru wariantów (tj. nie w kontekście własności α reguły wyboru) i w przypadku skorelowanych charakterystyk porównywanych wariantów, podczas gdy w badaniu Jakubczyk i Kamiński (2010) pokazano, że istotnym elementem są wariancje oszacowań NB dla porównywanych wariantów (s. 958 wg numeracji oryginalnej publikacji).

Wreszcie, w badaniu uogólniono wyniki Fenwick *et al.* (2004), pokazując wpływ parametrów opisujących warianty (w przypadku $n = 2$ wariantów) na monotoniczność CEAC (por. rozdz. 3 w publikacji, w szczególności Tabela II). W szczególności pokazano, że dla symetrycznych rozkładów opisujących niepewność oszacowań niemonotoniczność CEAC jest ograniczona: krzywa CEAC może przeciąć linię 50% (stanowiącą naturalny próg przy podejmowaniu decyzji dla $n = 2$ wariantów) jedynie raz. Jednocześnie pokazano, że w przypadku skośnych rozkładów własność ta nie zachodzi, tj. możliwe jest kilkukrotne przecinanie wartości 50% przez krzywą CEAC. Podsumowanie najważniejszych wyników badania Jakubczyk i Kamiński (2010) przedstawiono w tabeli III w oryginalnej publikacji.

Opisane badanie kontynuowałem w pracy **Jakubczyk (2015b)**. Celem było zaproponowanie modyfikacji CEAC jako reguły decyzyjnej, w świetle niepożądanych własności klasycznych CEAC scharakteryzowanych powyżej. W badaniu rozważano przypadek symetrycznych rozkładów prawdopodobieństwa opisujących niepewność oszacowań charakterystyk porównywanych ($n > 2$) wariantów. W pracy Jakubczyk (2015b) przedstawiłem dowód twierdzenia zasygnalizowanego lecz niedowiedzianego w badaniu Jakubczyk i Kamiński (2010) o relacji między krzywymi CEAC w przypadku FOSD (Twierdzenie 1. w badaniu Jakubczyk, 2015b). Nowym, istotnym wynikiem pracy było wskazanie konsekwencji występowania korelacji między niepewnościami oszacowań charakterystyk wariantów (Tabela 2. w oryginalnej publikacji).

W badaniu zaproponowałem dwie nowe funkcje wyboru spośród dowolnie (skończenie) wielu wariantów oparte na CEAC: rekurencyjną i konserwatywną regułę wyboru (odpowiednio Definicja 1 i 2 w oryginalnej publikacji). Idea polega na porównywaniu każdego wariantu z wykorzystaniem CEAC jedynie z jednym innym wariantem, odpowiednio: wybranym rekurencyjnie spośród pozostałych lub skutkującym najmniej korzystnym wynikiem porównania. Analizę własności zaproponowanych metod przeprowadziłem dla ustalonej wartości WTP (czyli badałem lokalne własności CEAC, zgodnie z terminologią wprowadzoną powyżej). Wykazano (Twierdzenie 2 w oryginalnej publikacji), że obie zaproponowane metody mają pożądane i intuicyjne własności w przypadku niepewności opisu wariantów zadanej wielowymiarowym (także skorelowanym) rozkładem normalnym: są zgodne z kryterium opartym na maksymalizacji wartości oczekiwanej NB, spełniają własność α (w szczególności odporne są na klonowanie wariantu), są odporne na zmianę struktury korelacji i zgodne z kryterium FOSD.

Wykazałem zgodność reguły konserwatywnej z FOSD w przypadku niezależnych zmienionych losowych charakteryzujących porównywane warianty (Twierdzenie 3 w oryginalnej publikacji). Dowiodłem, że w przypadku przechodniości relacji przy porównaniach wariantów parami z wykorzystaniem CEAC, obie proponowane reguły wyboru wskazują element maksymalny ww. relacji (Twierdzenie 4 w oryginalnej publikacji), tj. wielość wariantów sama w sobie nie prowadzi do wewnętrznej niespójności proponowanych reguł. W badaniu przeprowadziłem także empiryczne studium przypadku, ilustrując obie wprowadzone metody danymi wykorzystanymi w cytowanej powyżej pracy Barton *et al.* (2008), za Goeree *et al.* (2002).

Badania dotyczące sposobu wykorzystania CEAC kontynuowałem w kontekście trudności określenia wartości WTP. Punktem wyjścia była obserwacja, że oszacowania wartości statystycznego życia (ang. *value of statistical life*) przedstawione w literaturze są bardzo zróżnicowane, por. Bellavance *et al.*, (2009). Trudność określenia WTP w kontekście zdrowia wynika z braku doświadczeń rynkowych konsumentów przy tego typu transakcjach, co skutkuje np. wyjątkowo dużą rozbieżnością między WTP a skłonnością do akceptacji (WTA, ang. *willingness-to-accept*), por. Horowitz i McConnell (2002). Trudność określenia wartości WTP (WTA) nie ma charakteru (jedynie) statystycznego (tj. estymacji na podstawie próby losowej) — pojedynczy respondent zazwyczaj miałby trudność określenia precyzyjnie wartości życia (na ile jego zdaniem regulator publiczny w imieniu społeczeństwa powinien wycenić jedno QALY).

Do opisu świata z uwzględnieniem braku precyzji zaproponowano w literaturze m.in. zbiory rozmyte (Zadeh, 1965), tj. zbiory umożliwiające częściową przynależność elementów, wyrażoną tzw. funkcją przynależności o wartościach z przedziału $[0,1]$. Takie podejście zastosowano w badaniu **Jakubczyk i Kamiński (2017)**. Celem badania było: (1) zaproponowanie aksjomatycznego ujęcia rozmytej relacji preferencji w kontekście dwukryterialnych porównań technologii medycznych (z uwzględnieniem efektu i kosztu); (2) powiązanie rozmytej relacji preferencji i wartości WTP i WTA zredefiniowanych jako zbiory rozmyte oraz wskazanie możliwości empirycznej oceny tych parametrów; (3) określenie możliwości jednoczesnego

uwzględnienia braku precyzji i niepewności stochastycznej oszacowań (kosztu i efektu) we wspomaganie decydenta.

Badanie miało charakter zarówno empiryczny, jak i teoretyczny. W części empirycznej (rozdz. 3.1. w oryginalnej publikacji) przedstawiono wyniki autorskiej ankiety wśród ekspertów z obszaru HTA w Polsce na temat odczuwanej wartości WTP i WTA, z uwzględnieniem braku precyzji co do tych wartości. Wyniki ankiety potwierdziły z jednej strony istotność obu kryteriów (tj. także finansowego) przy porównywaniu technologii medycznych, ale także nieprecyzyjne postrzeganie pożądanej wartości prognozy WTP i WTA. Uzyskane wyniki zasugerowały nowe możliwe wyjaśnienie obserwowanego w literaturze zjawiska rozbieżności między wartościami WTP i WTA (tzw. *WTP-WTA disparity*). O ile funkcje przynależności dla wartości WTP i WTA traktowanych jako zbiory rozmyte w podobnym miejscu osiągają wartość 0,5, o tyle respondenci w badaniach eksperymentalnych raportują (i być może w faktycznych decyzjach stosując jako prognozy) wartości WTP, do których czują silne przekonanie (mające funkcję przynależności wyraźnie przekraczającą 0,5), por. Fig. 4 w oryginalnej publikacji. Tak więc obserwowana rozbieżność może być wynikiem tendencji decydentów do asekuracyjnego interpretowania braku precyzji w introspektywnym poznaniu własnych preferencji.

W części teoretycznej zaproponowano aksjomatykę rozmytych preferencji między technologiami medycznymi porównywanymi parami (rozdz. 3.2). W ramach tej aksjomatyki zdefiniowano rozmyte WTP i WTA (Definicje 1 i 2), wskazując, jak wykorzystać badania ankietowe (na przykład takie, jak wykorzystane w artykule) do przybliżenia funkcji przynależności. Nowatorskim elementem badania było wykorzystanie zbiorów rozmytych do określenia braku precyzji w kontekście preferencji między technologiami medycznymi.

Łącząc niepewność stochastyczną oszacowań z brakiem precyzji (w zakresie wartości WTP i WTA), zdefiniowano rozmytą oczekiwaną korzyść netto (fENB, ang. *fuzzy expected net benefit*) oraz rozmytą oczekiwaną akceptowalność (fEA, ang. *fuzzy expected acceptability*) dla danego wariantu decyzyjnego jako alternatywne metody oceny atrakcyjności wariantu uwzględniające niepewność stochastyczną i brak precyzji w ocenie wartości WTP/WTA. W zależności od założonego poziomu pewności (w sensie wartości funkcji przynależności) pojęcia te pokazują odpowiednio: oczekiwaną wartość NB oferowaną przez dany wariant z daną pewnością oraz prawdopodobieństwo, że dany wariant oferuje większą wartość NB. Wykazano, że własności fENB oraz fEA są intuicyjne: funkcje przynależności są monotoniczne (*Proposition 2* w oryginalnej publikacji). Jednocześnie oba pojęcia są naturalnym rozszerzeniem klasycznych pojęć: oczekiwanego NB oraz wartości CEAC dla danego poziomu WTP, tj. są równe tym klasycznym odpowiednikom, jeśli usuniemy rozmycie (przyjmujemy ostre wartości WTP i WTA, por. *Proposition 3* w oryginalnej publikacji). Wykazano także intuicyjne zachowanie fENB i fEA w granicy przy zmniejszaniu ilości niepewności w problemie (*Proposition 4* w oryginalnej publikacji).

Scharakteryzowane powyżej badania kontynuowałem w pracy **Jakubczyk (2015a)**¹². Rozszerzyłem metodykę badania Jakubczyk i Kamiński (2017), uwzględniając możliwość więcej niż dwóch wariantów decyzyjnych (przyjmując, że wszystkie porównywane są do jednego wariantu stanowiącego *status quo*). Wykazałem, że rozmyta korzyść netto (fNB, ang. *fuzzy net benefit*) spełnia kryterium dominacji i dominacji rozszerzonej/uogólnionej (*Proposition 1* i *2* w oryginalnej publikacji), często stosowanych kryteriów w klasycznej analizie CEA. W ramach rozważanego modelu zaproponowałem funkcję wyboru spośród dowolnie (skończenie) wielu wariantów opartą na porównywaniu fNB (Definicja 3 w oryginalnej publikacji) i wykazałem

¹² Różnica w roku publikacji wynika z procesu wydawniczego pracy Jakubczyk i Kamiński (2017), która była opublikowana w Internecie już w 2015 r.

poprawność jej definicji (*Proposition 3*).¹³ Uzyskane wyniki pozwalają na praktyczne stosowanie podejścia rozmytego do rzeczywistych problemów decyzyjnych, w których często porównuje się więcej niż dwa warianty decyzyjne.

4.4. Wykorzystanie zbiorów rozmytych do modelowania braku precyzji w preferencjach dotyczących stanu zdrowia

Jak już przedstawiono w części 4.3, w badaniu Jakubczyk i Kamiński (2017) wskazano, jak można uwzględnić brak precyzji w zakresie preferencji między ilością zdrowia (wyrażoną ilością QALY) a kosztem. Kontynuacją tego wątku było badanie, jak uwzględnić brak precyzji w preferencjach dotyczących stanów zdrowia, tj. przy badaniu konieczności przyjmowania kompromisów między jakością i długością życia oraz między różnymi aspektami jakości życia.

W badaniu **Jakubczyk i Golicki (2018)** zaproponowano model rozmytych preferencji decydenta w kontekście systemu deskryptywnego EQ-5D-3L. W uproszczeniu — przyjęto, że dysużyteczność związana z pogorszeniem danej domeny na dany poziom jest opisana nie liczbą, a odcinkiem (tj. w omawianym badaniu rozważono normalne i prostokątne liczby rozmyte, tj. funkcja przynależności przyjmowała tylko wartości 0 i 1). Nowatorskim elementem było uwzględnienie braku precyzji, tj. niemożności stwierdzenia przez decydenta, która z wielu liczb o wartości funkcji przynależności równej 1 jest właściwszym opisem dysużyteczności. W tym sensie zbiory opisujące dysużyteczności należy w ramach omawianego badania interpretować w sensie tzw. epistemicznych zbiorów rozmytych (ang. *epistemic fuzzy sets*, por. Co-uso i Dubois, 2014).

W badaniu pokazano, jak parametry zaproponowanego modelu preferencji można oszacować na podstawie danych zebranych w standardowym eksperymencie TTO (tj. eksperymencie nieukierunkowanym wprost na poznanie braku precyzji respondenta): skonstruowano model procesu generującego dane TTO na podstawie danej struktury preferencji i metodę szacowania parametrów w podejściu bayesowskim. Zaletą zaproponowanego podejścia jest możliwość wykorzystania zebranych uprzednio danych TTO do szacowania w ramach nowego modelu preferencji (dane w standardowym eksperymencie TTO zebrano dotychczas w kilkudziesięciu krajach).

Możliwość wykorzystania zaproponowanej metody wykazano stosując ją do danych zebranych w badaniu Golicki *et al.* (2010), omówionym powyżej. Uzyskane wyniki pokazują (por. Tabela 1 w oryginalnej publikacji), że (intuicyjnie) większy brak precyzji dotyczy trzeciego poziomu dla wszystkich domen i że brak precyzji przekłada się w przybliżeniu na długość odcinków opisujących dysużyteczność poszczególnych domen (na poziomie trzecim) równą ok. 0,06 (na standardowej skali QALY). W przeprowadzonym badaniu mniejszą precyzją charakteryzowała się domena PD (ból/dyskomfort) — długość odcinka opisującego dysużyteczność wynosiła ok. 0,13.

¹³ Kontynuuję badania w omówionym nurcie. W 2016 przedstawiałem pracę podczas 6th World Conference on Soft Computing, Berkeley, USA, 2016, pt. “Estimating the membership function of the fuzzy willingness-to-pay/accept for health via Bayesian modelling” — oczekiwane jest wydanie publikacji przez wydawnictwo Springer z artykułami prezentowanymi podczas konferencji. Podczas konferencji 4th Meeting of the International Academy of Health Preference Research, Hamburg, Niemcy, 2016, prezentowałem plakat pt. “Estimating fuzzy willingness-to-pay/accept in healthcare: survey results from England” (współautor: Ł. Tanajewski), wykorzystujący metodę z badania Jakubczyk i Kamiński (2017). Na konferencji The IV AMMCS (Applied Mathematics, Modeling and Computational Science) International Conference, Waterloo, Kanada, 2017, prezentowałem referat: “Estimating the Crossover Point of a Fuzzy Willingness-to-Pay/Accept for Health to Support Decision Making”, zaś artykuł został zgłoszony do publikacji.

Badanie Jakubczyk i Golicki (2018) mieści się w aktualnie powstającym nurcie uwzględniania specyfiki danych TTO. Po pierwsze, dane te powstają jako wynik serii porównań parami, więc dostępne są obserwacje sprzed ustalenia indyferencji. Po drugie, z uwagi na ograniczoną protokołem badania liczbę możliwych finalnych odpowiedzi, stwierdzenie, że w finalnej odpowiedzi zidentyfikowano dokładnie punkt indyferencji, należy traktować jedynie jako przybliżenie. Obecnie prowadzone są badania, jak uwzględniać tę specyfikę przy szacowaniu preferencji, kiedy te rozumiane są w klasyczny sposób (tj. jako ostre liczby), por. Ramos-Goñi *et al.* (2017). Nowatorskim elementem w badaniu Jakubczyk i Golicki (2018) jest uwzględnienie, że same preferencje są rozmyte (a więc zmieniony jest estymand badania).¹⁴

Równoległe do powyższego, brałem udział w badaniu **Jakubczyk, Kamiński i Lewandowski (2018)**. Celem badania było zbudowanie modelu preferencji z uwzględnieniem braku precyzji, który można szacować na podstawie danych zebranych poprzez tzw. eksperyment wyboru dyskretnego (DCE, ang. *discrete choice experiment*). DCE jest alternatywną wobec TTO najczęściej stosowaną metodą szacowania użyteczności stanów zdrowia, por. Bansback *et al.* (2012).

W badaniu wykorzystano dane z próby reprezentatywnej dla populacji Stanów Zjednoczonych, pochodzące z konkursu na modelowanie danych DCE, tj. przewidywanie struktury wyborów między stanami zdrowia na podstawie obserwowanych wyborów w próbie uczącej.¹⁵ W danych stosowano system deskrypcyjny EQ-5D-5L, w którym każda domena (identyczne jak w EQ-5D-3L) opisana jest z wykorzystaniem pięciu poziomów, por. Herdman *et al.* (2011). W badaniu Jakubczyk, Kamiński i Lewandowski (2018) zaproponowano alternatywny model preferencji, w którym dysużyteczności poszczególnych domen/poziomów reprezentowane są jako zbiory rozmyte. W odróżnieniu od badania Jakubczyk i Golicki (2018), w omawianym badaniu Jakubczyk, Kamiński i Lewandowski (2018) funkcja przynależności przyjmowała w ogólności także wartości ze zbioru (0,1) (choć przyjęto dla uproszczenia, że funkcja przynależności jest przedziałami liniowa), zaś zbiory rozmyte bliższe są interpretacji w kategorii ontycznych zbiorów rozmytych (ang. *ontic fuzzy sets*, por. Couso i Dubois, 2014).

W badaniu przedstawiono model dokonywania wyboru przez decydenta spomiędzy par stanów (z trwaniem) oraz sposób szacowania parametrów modelu preferencji poprzez maksymalizację funkcji wiarygodności. Zaproponowana metoda umożliwiała także ocenę statystycznego błędu estymacji (wykorzystując podejście bootstrapowe). Wyniki badania (Tabela 1 w oryginalnej publikacji) wskazują na duży brak precyzji w preferencjach społecznych, szczególnie dla domen PD (ból/dyskomfort) i AD (niepokój/przygnębienie).

Istotną cechą zaproponowanego modelu jest uwzględnienie dominacji Pareto między porównywanymi wariantami względem standardowego podejścia do modelowania wyborów w eksperymentach DCE (np. wykładniczego modelu Bradleya-Terry'ego) – w omawianym tu badaniu, jeśli czasy trwania są równe, zaś jeden stan zdrowia dominuje drugi (tzn. wszystkie domeny ustalone są na nie gorszym poziomie) i zakresy utraty użyteczności dla różniących się

¹⁴ Przedstawione tu badania kontynuują obecnie w ramach projektu naukowego „*A fuzzy approach to time trade-off experiment in EQ-5D-3L valuation*” finansowanego przez EuroQol Reserach Foundation, w którym proponuję metodę zmodyfikowania badania typu TTO, aby bezpośrednio szacować rozmyte preferencje (zdefiniowane jako bardziej ogólne, trapezoidalne i normalne liczby rozmyte), a także w ramach grantu OPUS10 NCN pt. „Metody analizy decyzji w problemach wielokryterialnych i szacowania skłonności do zapłaty/akceptacji przy wykorzystaniu modelowania rozmytego”.

¹⁵ Więcej informacji o samym konkursie znajduje się na stronie <http://iahpr.org/eq-dce-competition/> [dostęp 1.2.2018] i w artykule Jakubczyk *et al.* (2017).

poziomów w poszczególnych domenach, dla których funkcja przynależności przyjmuje wartości ze zbioru $(0,1)$ są rozłączne, to decydent wybierze z prawdopodobieństwem 1 dominujący stan.

4.5. Podsumowanie wyników naukowych cyklu publikacji

Badania prowadzone przeze mnie i włączone do jednotematycznego cyklu publikacji miały charakter interdyscyplinarny: odnosiły rozważania z obszaru ekonomii (w tym teorii decyzji, badań operacyjnych, ekonometrii) do kwestii postrzegania znaczenia zdrowia (względnej ważności różnych atrybutów charakteryzujących zdrowie oraz ważności zdrowia w ujęciu finansowym), także z uwzględnieniem charakterystyk demograficznych i uwarunkowań psychologicznych (braku precyzji oraz wpływu czynników kulturowych, np. religijności, na preferencje).

Omówione badania miały charakter zarówno teoretyczny, jak empiryczny. Do badań teoretycznych zaliczyć należy Jakubczyk i Kamiński (2010), Jakubczyk (2015b, 2015a), Jakubczyk i Kamiński (2017). W dwóch pierwszych badano własności krzywych akceptowalności jako metody oceny wariantów decyzyjnych i proponowano modyfikacje wynikające z nieintuicyjnych (i niepożądanych) własności klasycznych krzywych akceptowalności; w ostatnich dwóch badaniach wskazano (w ujęciu aksjomatycznym) możliwość formalnego uwzględnienia braku precyzji w opisywaniu preferencji dotyczących zdrowia (poprzez wykorzystanie teorii zbiorów rozmytych).

Do badań empirycznych zaliczam publikacje: Jakubczyk (2009), Golicki *et al.* (2010), Jakubczyk, Golicki i Niewada (2016), Jakubczyk i Golicki (2018) oraz Jakubczyk, Kamiński i Lewandowski (2018). We wszystkich badano preferencje społeczne dotyczące stanów zdrowia, w ostatnich dwóch dodatkowo z uwzględnieniem braku precyzji charakteryzującego preferencje respondentów. Prace empiryczne zawierały także nowatorskie elementy metodologiczne wskazane powyżej (wykorzystanie teorii zbiorów rozmytych). I przeciwnie, elementy badań empirycznych znajdowały się także w pracach Jakubczyk i Kamiński (2017) – tu umieszczono badanie ankietowe dotyczące wartości WTP – i Jakubczyk (2015b) – tu przedstawiono studium przypadku dla możliwości wykorzystania zaproponowanych modyfikacji krzywych akceptowalności.

Prace scharakteryzowane powyżej można umiejscowić w nurcie ekonomii zarówno pozytywnej, jak i normatywnej. Do pierwszego nurtu należą prace badające preferencje, tak jak są one zgłaszane przez respondentów w badaniach ankietowych: Jakubczyk (2009), Golicki *et al.* (2010), Jakubczyk, Golicki i Niewada (2016), Jakubczyk i Golicki (2018) oraz Jakubczyk, Kamiński i Lewandowski (2018). Do nurtu ekonomii normatywnej należą prace, w których proponowano nowe metody porównywania wariantów decyzyjnych o pożądanym własnościach: Jakubczyk (2015a, 2015b), Jakubczyk i Kamiński (2017).

Istotnym elementem moich badań są zbiory rozmyte jako narzędzie umożliwiające uwzględnienie braku precyzji w preferencjach; dotyczy to badań Jakubczyk (2015a), Jakubczyk i Kamiński (2017), Jakubczyk i Golicki (2018) oraz Jakubczyk, Kamiński i Lewandowski (2018). W badaniach proponowano różne wykorzystanie zbiorów rozmytych; w szczególności w badaniu Jakubczyk i Golicki (2018) zbiory rozmyte należy interpretować w sensie epistemicznym, zaś w pozostałych trzech – w sensie ontycznym. Za istotne uważam, że w badaniach wskazano możliwość praktycznego wykorzystania wprowadzonych pojęć teoretycznych, pokazując metody szacowania parametrów zbudowanych modeli preferencji.

W prowadzonych badaniach wykorzystywałem dane empiryczne pozyskane różnymi metodami – pochodzące z eksperymentów handlowania czasem (np. Jakubczyk, 2009), wyboru dyskretnego (Jakubczyk, Kamiński i Lewandowski, 2018) lub badań z wykorzystaniem skali Likerta (Jakubczyk i Kamiński, 2017). Badania prowadzono z wykorzystaniem danych

zebranych zarówno w Polsce (np. Jakubczyk, Golicki i Niewada, 2016), jak i za granicą, w Stanach Zjednoczonych (Jakubczyk, Kamiński i Lewandowski, 2018).

Za najważniejszy wkład do rozwoju dyscypliny ekonomia uzyskany w ramach przedstawionego osiągnięcia uważam:

1. określenie preferencji społecznych Polaków dotyczących stanów zdrowia, tj. ważności kryteriów odpowiadających różnym aspektom jakości życia i długości życia (w badaniach Golicki *et al.*, 2010; Jakubczyk, 2009), a dodatkowo wykazanie metodami analizy statystycznej odporności uzyskanych wyników na konstrukcję badania, tj. liczbę eksperymentów myślowych wykonywanych przez jednego badanego (Golicki *et al.*, 2010), i wskazanie charakteru powiązań między atrybutami opisującymi stan zdrowia we wpływie na użyteczność, tj. wykazanie ich komplementarności (Jakubczyk, 2009);
2. scharakteryzowanie heterogeniczności Polaków pod względem cech demograficznych w zakresie preferencji dotyczących stanów zdrowia (Jakubczyk, 2009), jak również wskazanie uwarunkowań światopoglądowych ograniczających możliwość wykorzystania metody TTO (handlowania czasem) do badania preferencji i — wynikające z tego — zaproponowanie metod korygowania wyników o wpływ światopoglądu (poprzez modelowanie ekonometryczne lub modyfikację metody TTO; badanie Jakubczyk, Golicki i Niewada, 2016);
3. wskazanie ograniczeń stosowania standardowych metod porównywania wariantów decyzyjnych w kontekście problemów decyzyjnych dotyczących wyboru technologii medycznych w warunkach ryzyka i dwukryterialności (tzw. krzywych akceptowalności; Jakubczyk i Kamiński, 2010), a następnie skonstruowanie nowych metod takiego porównywania (tzw. rekurencyjnych i konserwatywnych krzywych akceptowalności; Jakubczyk, 2015b) i wskazanie ich korzystniejszych własności w sytuacjach decyzyjnych różnych typów (dla dwóch lub więcej niż dwóch wariantów, z błędami estymacji danymi rozkładami symetrycznymi lub niesymetrycznymi, w przypadku skorelowania błędów oszacowań dla porównywanych wariantów lub nie);
4. zaksjomatyzowanie rozmytych preferencji w kontekście wyboru technologii medycznych (opisanych z uwzględnieniem dwóch kryteriów — kosztu i efektu), z uwzględnieniem ryzyka, w przypadku porównań parami (Jakubczyk i Kamiński, 2017), a następnie zaproponowanie metody porównywania wariantów z uwzględnieniem braku precyzji (rozmytych preferencji) w przypadku dokonywania wyboru spośród więcej niż dwóch wariantów i wykazanie posiadania pożądanych i intuicyjnych własności przez zaproponowane metody (Jakubczyk, 2015a); dodatkowo zaproponowanie definicji skłonności do płacenia i akceptacji (WTP i WTA) jako zbiorów rozmytych, co może stanowić jedno z wyjaśnień obserwowanej w literaturze rozbieżności między wartościami WTP i WTA (Jakubczyk i Kamiński, 2017);
5. wykorzystanie teorii zbiorów rozmytych do zbudowania modeli formalnych preferencji dotyczących stanów zdrowia z uwzględnieniem nieprecyzyjnego postrzegania ważności poszczególnych kryteriów (atrybutów opisujących jakość życia związaną ze zdrowiem w systemie EQ-5D) oraz opracowanie metod szacowania tak zdefiniowanych preferencji na podstawie danych empirycznych pozyskiwanych w standardowych metodach eksperymentalnych (zatem także zebranych w prowadzonych uprzednio badaniach): metody TTO (handlowania czasem) i DCE (eksperymentów wyborów dyskretnych) (Jakubczyk i Golicki, 2018; Jakubczyk, Kamiński i Lewandowski, 2018).

Co istotne, uzyskane wyniki mają także wymiar praktyczny. Wyniki badania Golicki *et al.* (2010) są oficjalnie wskazane jako zestaw norm użyteczności w Polsce do wykorzystywania w analizach CUA przy ocenianiu technologii medycznych przy decyzjach podejmowanych przez regulatora publicznego. Badania Jakubczyk (2015b), Jakubczyk i Kamiński (2017) oraz Jakub-

czyk (2015a) wprowadzają metody analizy wariantów decyzyjnych, które mogą być bezpośrednio wykorzystane w rzeczywistych problemach decyzyjnych porównywania technologii medycznych w warunkach niepewności.

Na koniec, chciałbym podkreślić osobistą satysfakcję z wybranego przed laty kierunku pracy naukowej¹⁶. Co ważne dla mnie jako ekonomisty, prowadzone badania pozwalają mi na przykład bardziej namacalnie zetknąć się z pojęciem dla wielu jedynie teoretycznym — użytecznością dóbr. Mam też przekonanie, że uzyskane przeze mnie wyniki pozwalają na podejmowanie lepszych decyzji, tj. bardziej racjonalne gospodarowanie ograniczonymi zasobami w obszarze rynku ochrony zdrowia. Jak stwierdził renomowany ekonomista Peter Wakker (2010): „obszar medyczny jest jednym z najważniejszych pól zastosowania teorii decyzji”.¹⁷

5. Omówienie najważniejszych pozostałych publikacji

Oprócz prac scharakteryzowanych powyżej, prowadzone przeze mnie badania zaowocowały także innymi publikacjami, z których najważniejsze omawiam w sposób syntetyczny w niniejszej części (pełną listę publikacji zawiera analiza bibliometryczna załączona do wniosku).

W czasie swojej pracy naukowej brałem udział w innych badaniach związanych z zagadnieniem opisywania preferencji dotyczących zdrowia lub opisywania decyzji regulatora publicznego dotyczących technologii medycznych. W artykule **Jakubczyk et al. (2017)** omówiono przebieg i wyniki konkursu na przewidywanie wyborów w DCE przez indywidualnych respondentów (w populacji amerykańskiej) organizowanego przez *International Academy of Health Preference Research*.¹⁸ W konkursie zwyciężyła drużyna *Discreetly Charming Econometricians*, pracami której kierowałem. W publikacji przedyskutowano zaproponowane modele prognostyczne, a w szczególności zwycięski model. Najistotniejsze wnioski z tego modelu są następujące: po pierwsze, istotność domen zmienia się w zależności od czasu trwania stanu zdrowia (np. istotność domeny UA zwiększa się przy rozpatrywaniu długotrwałych stanów zdrowia); po drugie, czas postrzegany jest nieliniowo (podobne wnioski wynikają z modeli wykorzystanych przez inne zespoły), tj. niezgodnie ze standardowym modelem QALY; po trzecie, różnica użyteczności między wariantami powinna być przekształcana na prawdopodobieństwo wybrania jednego z wariantów z wykorzystaniem dystrybuanty rozkładu Cauchy’ego (a nie np. rozkładu normalnego lub logistycznego, jak w podejściu probitowym lub logitowym); po czwarte, porównywanie stanu zdrowia do natychmiastowego zgonu wymaga innego modelu niż porównywanie dwóch niezdegenerowanych stanów zdrowia.

W artykule **Niewada et al. (2013)** badano czynniki związane z pozytywnymi lub negatywnymi decyzjami dot. finansowania technologii medycznych wydawanymi przez Agencję Oceny Technologii Medycznych (Radę Konsultacyjną/Przejrzystości) w Polsce. Badaniem objęto ponad 300 rekomendacji wydanych do roku 2011. W wyniku analizy nie stwierdzono statystycznie istotnego wpływu współczynnika ICER na decyzje (badaniem objęto czas sprzed obowiązywania obecnej Ustawy Refundacyjnej definiującej wprost próg WTP). Z kolei celem badania **Jakubiak-Lasocka i Jakubczyk (2014)** było porównanie wykorzystania różnych typów analiz ekonomicznych (analiz, w których jako miarę efektu wykorzystuje się QALY, oraz analiz, w których wykorzystuje się lata życia, nieskorygowane o jakość, tzw. LYG, ang. *life-ye-*

¹⁶ Jakkolwiek niełatwego gdyż interdyscyplinarnego i poruszającego także trudne kwestie choroby, bólu, śmierci, itp.

¹⁷ Woryginał (rozdz. 3.1): „the medical domain is one of the most important fields of application for decision theory”.

¹⁸ Por. <http://iahpr.org/eq-dce-competition/> [ostatni dostęp 1.2.2018].

ars gained). Analizą objęto wszystkie wnioski złożone do Agencji Oceny Technologii Medycznych w Polsce do 2011 roku. Wykazano, że analizy, w których uwzględnia się jakość życia (tj. wylicza się użyteczność), stanowiły większość. Co nieintuicyjne, chociaż uwzględnienie użyteczności stanu zdrowia wymaga przyjęcia w analizie większej liczby parametrów, oszacowania wartości współczynnika ICER w tego typu analizach obarczone było mniejszym błędem. Próba wykrycia stronniczego raportowania bardziej korzystnej miary nie wskazała na występowania tego typu (niepożądanego) zjawiska.

Prowadziłem także badania teoretyczne w obszarze teorii gier i wspomaganie decyzji, niezwiązane z konkretnym obszarem rynku. W badaniu **Amir, Jakubczyk i Knauff (2008)**¹⁹ rozważaliśmy własności symetrycznych gier supermodularnych (z komplementarnymi strategiami graczy). Wskazaliśmy warunki (różne zestawy warunków, np. przypadek pełnej zgodności interesów między graczami — *Proposition 2* w oryginalnej publikacji — lub własność tzw. *superjoin* z dodatkowymi technicznymi wymogami — *Proposition 4* w oryginalnej publikacji), kiedy niesymetryczne równowagi Nasha (rozważaliśmy równowagi w strategiach czystych) w tego typu grach mają wypłaty Pareto-zdominowane przez wypłaty jakiejś symetrycznej równowagi Nasha. Z kolei celem badania **Kamiński, Jakubczyk i Szufel (2018)** było zaproponowanie nowych metod (i zbadanie ich własności) prowadzenia analizy wrażliwości w problemach sekwencyjnych w warunkach ryzyka opisywanych z pomocą drzew decyzyjnych. Zaproponowano schemat konceptualny, w ramach którego można badać stopień stabilności rozwiązań (z uwagi na zaburzenia wartości prawdopodobieństw opisujących problem decyzyjny) oraz poszukiwać rozwiązań odpornych na pesymistyczne/optymistyczne/faworyzujące najbardziej prawdopodobne stany zaburzenia. Co istotne, publikacji towarzyszy darmowe oprogramowanie, umożliwiające wykorzystanie przedstawionych metod²⁰.

Istotnym nurtem w mojej pracy naukowej były badania teoretyczne i empiryczne w obszarze szacowania kosztów chorób, a w szczególności kosztów społecznych (tzw. koszty pośrednie, por. Jakubczyk *et al.*, 2010, oraz Koopmanschap i van Ineveld, 1992) i innych zjawisk związanych ze zdrowiem. Na przykład w pracy **Jakubczyk i Koń (2017)** przedstawiono model teoretyczny funkcjonowania firm, uwzględniający wpływ oczekiwań i dostosowań po stronie pracodawców na wielkość i poprawny sposób szacowania kosztów pośrednich. Uzyskane wyniki wskazały możliwość zredukowania kosztów pośrednich w wyniku dostosowań, w odróżnieniu od rezultatów uzyskanych w badaniu Pauly *et al.* (2002). W badaniu **Lasocka, Jakubczyk i Siekmeier (2013)** oszacowano koszty utraconej produktywności w Polsce wynikające z chorób odtytoniowych, wykorzystując metodę kapitału ludzkiego (por. Liljas, 1998). Oszacowany koszt roczny wyniósł ok. 15 miliardów złotych. Badałem również wpływ chorób na koszty ponoszone z perspektywy płatnika publicznego. W serii prac **Czech *et al.* (2015)**, **Jakubczyk *et al.* (2015)** i **Jakubczyk *et al.* (2016)** badano, odpowiednio, ryzyko epizodów hipoglikemii w zależności od stosowanej terapii przeciwcukrzycowej (na podstawie przeglądu systematycznego i meta-analizy z wykorzystaniem modelowania bayesowskiego przy szacowaniu ryzyka), strukturę zużycia zasobów w przypadku wystąpienia takiego epizodu (znów, na podstawie przeglądu literatury i meta-analizy) oraz koszty łączne hipoglikemii w dziewięciu krajach Europy środkowej i wschodniej. Uczestniczyłem w badaniach kosztowej-efektywności stosowania procedur medycznych. W pracy **Pelczarska *et al.* (2018)** badano różne scenariusze badań przesiewowych w kierunku wykrywania hipercholesterolemii rodzinnej, wykorzystując autorski model ekonomiczny opisujący przebieg diagnostyki i leczenia, i wykazano zasadność,

¹⁹ Badanie było rozpoczęte przed uzyskaniem przeze mnie stopnia naukowego doktora. Wyniki przedstawione w badaniu (ani temat badania) nie są związane z moją rozprawą doktorską.

²⁰ <https://github.com/pszufe/chondro/>, [ostatni dostęp 1.2.2018].

z punktu widzenia kryteriów ekonomicznych i medycznych (z uwzględnieniem wpływu leczenia na użyteczność stanów zdrowia), wdrożenia różnych, badanych strategii.

6. Podsumowanie dorobku naukowego

W momencie składania wniosku mój punktowany dorobek naukowy (uwzględniając jedynie publikacje w całości przygotowane po uzyskaniu stopnia doktora²¹) obejmuje:

- 14 artykułów opublikowanych w czasopiśmie naukowych posiadających wskaźnik *impact factor* (IF),
- 14 artykułów opublikowanych w czasopiśmie z listy B Ministerstwa Nauki i Szkolnictwa Wyższego (MNiSW),
- 8 rozdziałów opublikowanych w monografiach naukowych,
- redakcję dwóch monografii naukowych.

Sumaryczny IF dla wymienionych powyżej publikacji wynosi 44,984, zaś sumaryczna wartość punktów MNiSW to 577. Łączna liczba cytowań moich prac (wg analizy bibliometrycznej, bez uwzględniania daty powstania publikacji) wynosi

- 121 (wg Web of Science Core Collection, bez autocytowań),
- 162 (wg Scopus),
- 366 (wg Google Scholar).

Cytowania te przekładają się na indeks Hirscha równy

- 5 (wg Web of Science Core Collection, bez autocytowań),
- 7 (wg Scopus),
- 9 (wg Google Scholar).

Wyniki moich badań prowadzonych po uzyskaniu stopnia doktora prezentowałem na wielu konferencjach w kraju i za granicą:

- 7 prezentacji ustnych na międzynarodowych konferencjach naukowych za granicą,
- 4 prezentacje ustne na międzynarodowych konferencjach naukowych w Polsce (po angielsku),
- prezentacja na seminarium naukowym za granicą (*University of Illinois at Chicago*),
- 2 prezentacje moich wyników przez zaproszonych dyskutantów z komentarzem autorów (dodatkowo przeprowadzałem na zaproszenie dyskusję cudzego artykułu),
- 20 prezentacji w formie plakatów na konferencjach międzynarodowych za granicą,
- 11 prezentacji ustnych na konferencjach krajowych.

Swoje wyniki prezentowałem m.in. na następujących konferencjach: *EuroQol Group Plenary Meeting* (dwukrotnie prezentacja ustna przez dyskutanta, jednokrotnie współautor prezentujący plakat), *International Academy of Health Preference Research* (trzykrotnie prezentacja ustna, jednokrotnie plakat), *International Workshop on Multiple criteria decision making* (dwukrotnie), *INFORMS Annual Meeting*, *28th European Conference on Operational Re-*

²¹ Analiza bibliometryczna przygotowana przez Bibliotekę SGH i załączona do wniosku obejmuje wszystkie opublikowane prace, niezależnie od daty. Cały dorobek punktowany scharakteryzowałem dodatkowo w Załączniku 4 do wniosku, wskazując mój udział procentowy w powstaniu poszczególnych prac. W załączniku tym scharakteryzowałem także bardziej szczegółowo mój dorobek naukowy oraz dorobek dydaktyczny i organizacyjny.

search, 6th World Conference on Soft Computing, The 10th Conference of the European Society for Fuzzy Logic and Technology, The IV AMMCS (Applied Mathematics, Modeling and Computational Science) International Conference oraz EuroScience Open Forum.

Uczestniczyłem w pracach komitetów organizacyjnych i naukowych konferencji (np. 32nd EuroQol Plenary Meeting oraz International Conference on Group Decision & Negotiation), moderowałem sesje na konferencjach krajowych i zagranicznych (podczas INFORMS Annual Meeting oraz International Society for Pharmacoeconomics and Outcomes Research Annual European Congress).

Recenzowałem artykuły naukowe nadesłane do redakcji dla wielu czasopism uwzględnionych w bazie Journal Citation Reports:

- *Annals of Operations Research* (2015, IF = 1,217, MNiSW = 25)²²,
- *The European Journal of Health Economics* (2013 i 2017, IF = 1,774, MNiSW = 35),
- *European Journal of Operational Research* (2012, IF = 2,358, MNiSW = 40),
- *Expert Review of Pharmacoeconomics & Outcomes Research* (2015, IF = 1,669, MNiSW = 20),
- *Health Economics* (2010 i 2012, IF = 2,227, MNiSW = 35),
- *Health and Quality of Life Outcomes* (2016, IF = 2,12, MNiSW = 35),
- *Kardiologia Polska* (2014, IF = 0,539, MNiSW = 15),
- *Pharmacoeconomics* (2010 i 2011, IF = 2,45, MNiSW = 30),
- *Quality of Life Research* (czterokrotnie w 2016 i 2017, IF = 2,486, MNiSW = 35).

Recenzowałem dwa wnioski zgłoszone do programu Fulbright Junior Research Award na zaproszenie komisji programu. Regularnie recenzuję wnioski i oceniam końcowe raporty zgłaszane do grupy naukowej EuroQol Group.

Jestem członkiem stowarzyszeń naukowych:

- *EuroQol Group* (<https://euroqol.org/>, zrzeszającej poniżej 90 osób na całym świecie; w ramach tej grupy członkiem podgrupy tzw. *Valuation Working Group* zajmującej się m.in. oceną wniosków grantowych),
- *International Academy of Health Preference Research* (<http://iahpr.org/>, zrzeszającej <40 tzw. *tenured members* i kolejnych <40 członków),
- *International Society for Pharmacoeconomics and Outcomes Research*,
- Polskiego Towarzystwa Farmakoekonomicznego,
- Polskiej Grupy Narodowej Międzynarodowego Towarzystwa Biostatystyki Klinicznej.

Jestem kierownikiem trzech realizowanych obecnie grantów naukowych:

- od 2016, w projekcie badawczym pt. „**Metody analizy decyzji w problemach wielokryterialnych i szacowania skłonności do zapłaty/akceptacji przy wykorzystaniu modelowania rozmytego**”, grant OPUS 10 przyznany przez Narodowe Centrum Nauki, umowa nr UMO-2015/19/B/HS4/01729;
- od 2016, w projekcie badawczym pt. „**A fuzzy approach to time trade-off experiment in EQ-5D-3L valuation**”, grant finansowany przez stowarzyszenie naukowe EuroQol Research Foundation (numer: EQ Project 2015320);

²² W tej sekcji punkty MNiSW i IF podano na najnowszy dostępny rok.

- od 2016, w projekcie badawczym pt. **“Building values sets based on TTO results by averaging model predictions and actually observed means”**, grant finansowany przez stowarzyszenie naukowe EuroQol Research Foundation (numer: EQ Project 2016640).

Dodatkowo jestem partnerem naukowym w grantcie POLONEZ (pt. **„Evidence use in the allocation of health care budgets: A comparative study of drug reimbursement in Poland, Canada, and Germany”**, kierownik: W. Wranik) oraz wykonawcą w grantcie Horizon 2020 (**„Raising Open and User-friendly Transparency-Enabling Technologies for Public Administrations (ROUTE-TO-PA)”**, kierownik: B. Kamiński) i dwóch grantach finansowanych przez *The EuroQol Foundation* (**“EQ-5D-5L valuation in Poland”** i **“EQ-5D-5L valuation in Poland: a methodological extension”**, kierownik: D. Golicki). Kierowałem grantem finansowanym w ramach programu MNiSW „Iuventus Plus” (**„Rekursywne krzywe akceptowalności w porównaniach technologii medycznych”**).

Otrzymałem stypendium MNiSW dla wybitnych młodych naukowców (2011); przydzielono mi stypendium *Fulbright Senior Award* (2015-2016). Otrzymywałem nagrody i listy gratulacyjne od Rektora SGH za działalność naukową i organizacyjną.

Mój dorobek naukowy wraz z dorobkiem dydaktycznym i organizacyjnym scharakteryzowałem bardziej szczegółowo w Załączniku 4 do wniosku.

7. Skróty i akronimy stosowane w autoreferacie

AD	Odczuwanie niepokoju/przygnębienia (ang. <i>anxiety/depression</i>), domena w systemie EQ-5D-3L/5L
AOTMiT	Agencja Oceny Technologii Medycznych i Taryfikacji
CEA	Analiza kosztów-efektywności (ang. <i>cost-effectiveness analysis</i>)
CEAC	Krzywa akceptowalności (ang. <i>cost-effectiveness acceptability curve</i>)
CUA	Analiza kosztów-użyteczności (ang. <i>cost-utility analysis</i>)
DCE	Eksperyment wyboru dyskretnego (ang. <i>discrete choice experiment</i>)
fEA	Rozmyta oczekiwana akceptowalność (ang. <i>fuzzy expected acceptability</i>)
fENB	Rozmyta oczekiwana korzyść netto (ang. <i>fuzzy expected net benefit</i>)
fNB	Rozmyta korzyść netto (ang. <i>fuzzy net benefit</i>)
FOSD	Dominacja stochastyczna pierwszego rzędu (ang. <i>first-order stochastic dominance</i>)
HDI	ang. <i>Human Development Index</i>
HTA	Ocena technologii medycznych (ang. <i>health technology assessment</i>)
HRQoL	Jakość życia związana ze zdrowiem (ang. <i>health-related quality of life</i>)
ICER	Inkrementalny współczynnik kosztów-efektywności (ang. <i>incremental cost-effectiveness ratio</i>)
IF	ang. <i>impact factor</i>
LYG	uzyskane lata życia (ang. <i>life years gained</i>)
MNiSW	Ministerstwo Nauki i Szkolnictwa Wyższego
MO	Zdolność poruszania się (ang. <i>mobility</i>), domena w systemie EQ-5D-3L/5L
NB	Korzyść netto (ang. <i>net benefit</i>)

PD	Odczuwanie bólu/dyskomfortu (ang. <i>pain/discomfort</i>), domena w systemie EQ-5D-3L/5L
PKB	Produkt Krajowy Brutto
QALY	Lata życia skorygowane o jakość (ang. <i>quality-adjusted life years</i>)
SC	Zdolność samoopieki (ang. <i>self-care</i>), domena w systemie EQ-5D-3L/5L
TTO	Metoda handlowania czasem (ang. <i>time trade-off</i>)
UA	Zdolność wykonywania zwykłych czynności (ang. <i>usual activities</i>), domena w systemie EQ-5D-3L/5L
VAS	Wizualna skala analogowa (ang. <i>visual analogue scale</i>)
WTA	Skłonność do akceptacji (ang. <i>willingness to accept</i>)
WTP	Skłonność do płacenia (ang. <i>willingness to pay</i>)

8. Referencje

1. Agencja Oceny Technologii Medycznych (2009): “**Wytyczne oceny technologii medycznych (HTA). Wersja 2.1**”, Warszawa.
2. Agencja Oceny Technologii Medycznych i Taryfikacji (2016): “**Wytyczne oceny technologii medycznych (HTA, ang. health technology assessment). Wersja 3.0**”, Warszawa.
3. Alemayehu B., Warner K.E. (2004): “**The Lifetime Distribution of Health Care Costs**”, *HSR: Health Services Research*, 39 (3), 627–642.
4. Amir R., Jakubczyk M., Knauff M. (2008): „**Symmetric versus asymmetric equilibria in symmetric supermodular games**”, *International Journal of Game Theory*, 37, 307–320.
5. Arrow K.J. (1963): „**Uncertainty and the Welfare Economics of Medical Care**”, *American Economic Review*, 53 (5), 941–973.
6. Arrow K.J., Lind R.C. (1970): “**Uncertainty and the Evaluation of Public Investment Decisions**”, *American Economic Review*, 60, 364–378.
7. Attema A.E., Edelaar-Peeters Y., Versteegh M.M., Stolk E.A. (2013): “**Time trade-off: one methodology, different methods**”, *European Journal of Health Economics*, 14 (Supp. 1), 53–64.
8. Au Eong K.G., Chan E.W., Luo N., Wong S.H., Tan N.W., Lim T.H., Wagle A.M. (2012): “**Validity of EuroQOL-5D, time trade-off, and standard gamble for age-related macular degeneration in the Singapore population**”, *Eye*, 26 (3), 379–388.
9. Augestad L., Rand-Hendriksen K., Stavem K., Kristiansen I. (2013): “**Time trade-off and attitudes toward euthanasia: Implications of using ‘death’ as an anchor in health state valuation**”, *Quality of Life Research*, 22, 705–714.
10. Badia X., Roset R., Herman M., Kind P. (2001): “**A comparison of GB and Spanish general population time trade-off values for EQ-5D health states**”, *Medical Decision Making*, 21, 7–16.
11. Bansback N., Brazier J., Tsuchiya A., Anis A. (2012): “**Using a discrete choice experiment to estimate health states utility values**”, *Journal of Health Economics*, 31, 306–318.
12. Barton G., Briggs A., Fenwick E. (2008): “**Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfect information (EVPI)**”, *Value in Health*, 11 (5), 886–897.

13. Begg D., Fischer S., Dornbusch R. (1999): “**Ekonomia. Makroekonomia**”, Polskie Wydawnictwo Ekonomiczne.
14. Bellavance F., Dionne G., Lebeau M. (2009): “**The value of statistical life: A meta-analysis with a mixed effects regression model**”, *Journal of Health Economics*, 28, 444–464.
15. Bleichrodt H., Wakker P., Johannesson M. (1997): “**Characterizing QALYs by Risk Neutrality**”, *Journal of Risk and Uncertainty*, 15, 107–114.
16. Brazier J., Ratcliffe J., Alomon J., Tsuchiya A. (2017): “**Measuring and Valuing Health Benefits for Economic Evaluation**”, Oxford University Press, ISBN: 978-0198725923.
17. Briggs AH. (1999): “**A Bayesian approach to stochastic cost effectiveness analysis**”, *Health Economics*, 8 (3), 257–261.
18. Briggs AH., Weinstein MC., Fenwick EAL., Karnon J., Sculpher MJ., Paltiel AD., ISPOR-SMDM Modeling Good Research Practices Task Force (2012): “**Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6**”, *Value in Health*, 15 (6), 835–842.
19. Brooks R., De Charro F. (1996): “**EuroQol: the current state of play**”, *Health Policy*, 37, 53–72.
20. Couso I., Dubois D. (2014): “**Statistical reasoning with set-valued information: Ontic vs. epistemic views**”, *International Journal of Approximate Reasoning*, 55, 1502–1518.
21. Czech M., Rdzanek E., Pawęska J., Adamowicz-Sidor O., Niewada M., Jakubczyk M. (2015): “**Drug related risk of severe hypoglycaemia in observational studies—a systematic review and meta-analysis**”, *BMC Endocrine Disorders*, 15, 57, DOI: 10.1186/s12902-015-0052-z.
22. Devlin N., Parkin D. (2004): “**Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis**”, *Health Economics*, 13, 437–452.
23. Dolan P., Gudex C., Kind P., Williams A. (1996): “**The time trade-off method: results from a general population study**”, *Health Economics*, 5, 141–154.
24. Drummond M., Brixner D., Gold M., Kind P., McGuire A., Nord E. (2009): “**Toward a Consensus on the QALY**”, *Value in Health*, 12 (S1), S31–S35.
25. Drummond M., Sculpher MJ., Claxton K., Stoddart GL., Torrance GW. (2015): “**Methods for the Economic Evaluation of Health Care Programmes**”, Oxford University Press, ISBN: 978-0199665884.
26. Goeree R., O’Brien BJ., Blackhouse G., Marshall J., Briggs A., Lad R. (2002): “**Cost-effectiveness and cost-utility of long-term management strategies for heartburn**”, *Value in Health*, 5 (4), 312–328.
27. Golicki D., Jakubczyk M., Niewada M., Wrona W., Busschbach J.J.V. (2010): “**Valuation of EQ-5D Health States in Poland: First TTO-based Social Value Set in Central and Eastern Europe**”, *Value in Health*, 13 (2), 289–297.
28. Greiner W., Claes C., Busschbach JJ., Graf von Schulenburg JM. (2005): “**Validating the EQ-5D with time trade off for the German population**”, *European Journal of Health Economics*, 6, 124–130.
29. Feng Y., Devlin N., Shah K., Mulhern B., van Hout B. (2017): “**New Methods for Modelling EQ-5D-5L Value Sets: An Application to English Data**”, *Health Economics*, doi: 10.1002/hec.3560.
30. Fenwick E., Claxton K., Sculpher M. (2001): “**Representing uncertainty: the role of cost-effectiveness acceptability curves**”, *Health Economics*, 7 (8), 723–740.

31. Fenwick E., O'Brien B., Briggs A. (2004): "**Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions**", *Health Economics*, 13, 405–415.
32. Figueras J., McKee M. (red.) (2012): "**Health Systems, Health, Wealth and Societal Well-being. Assessing the case for investing in health systems**", McGraw Hill, New York, ISBN: 78-0-335-24430-0.
33. Główny Urząd Statystyczny: **Narodowy Rachunek Zdrowia 2015**, <https://stat.gov.pl/obszary-tematyczne/zdrowie/zdrowie/narodowy-rachunek-zdrowia-2015,4,8.html>, dostęp 23.12.2017.
34. Grossman M. (1972): "**On the Concept of Health Capital and the Demand for Health**", *Journal of Political Economy*, 80 (2), 223–255.
35. Herdman M., Gudex C., Lloyd A., Janssen M., Kind P., Parkin D., Bonnel G., Badia X. (2011): "**Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L)**", *Quality of Life Research*, 20 (10), 1727–1736.
36. van Hout B., Gordon AM., Rutten F. (1994): "**Costs, effects and C:E-ratios alongside a clinical trial**", *Health Economics*, 3, 309–319.
37. Hogan C., Lunney J., Gabel J., Lynn J. (2001): "**Medicare Beneficiaries' Costs Of Care In The Last Year Of Life**", *Health Affairs*, 20 (4), 188–195.
38. Horowitz JK., McConnell KE. (2002): "**A Review of WTA/WTP Studies**", *Journal of Environmental Economics and Management*, 44, 426–447.
39. International Academy of Health Preference Research, <http://iahpr.org/eq-dce-competition/>, dostęp 29.12.2017.
40. Jakubczyk M. (2009): "**Impact of complementarity and heterogeneity on health related utility of life**", *Central European Journal of Economic Modelling and Econometrics*, 1, 139-156.
41. Jakubczyk M. (2015a): "**Using a Fuzzy Approach in Multi-Criteria Decision Making with Multiple Alternatives in Health Care**", *Multiple Criteria Decision Making*, 10, 65–81.
42. Jakubczyk M. (2015b): "**Rekurencyjne i konserwatywne reguły decyzyjne przy wyborze technologii medycznych w warunkach ryzyka**" w: D. Kopańska-Bródka (red.) „*Studia ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*”, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, ISSN 2083-8611, 235, 84–99.
43. Jakubczyk M., Craig BM., Barra M., Groothuis-Oudshoorn CGM., Hartman JD., Huynh E., Ramos-Goñi JM., Stolk EA., Rand K. (2017): "**Choice Defines Value: A Predictive Modeling Competition in Health Preference Research**", *Value in Health*, DOI: 10.1016/j.jval.2017.09.016.
44. Jakubczyk M., Golicki D. (2018): "**Estimating the Fuzzy Trade-Offs Between Health Dimensions with Standard Time Trade-Off Data**" w: J. Kacprzyk, E. Szmidt, S. Zadrożny, K. Atanassov, M. Krawczak (red.) "Advances in Fuzzy Logic and Technology 2017. IWIFSGN 2017, EUSFLAT 2017.", *Advances in Intelligent Systems and Computing*, vol 642. Springer, ISSN: 2194-5357; ISBN: 978-3-319-66823-9, DOI: 10.1007/978-3-319-66824-6, 266–277.
45. Jakubczyk M., Golicki D., Niewada M. (2016): "**The impact of a belief in life after death on health-state preferences: True difference or artifact?**", *Quality of Life Research*, 25 (12), 2997–3008.
46. Jakubczyk M., Kamiński B. (2010): "**Cost-effectiveness acceptability curves – caveats quantified**", *Health Economics*, 19(8), 955–963.

47. Jakubczyk M., Kamiński B. (2017): **“Fuzzy approach to decision analysis with multiple criteria and uncertainty in health technology assessment”**, *Annals of Operations Research*, 251 (1–2), 301–324.
48. Jakubczyk M., Kamiński B., Lewandowski M. (2018): **“Eliciting Fuzzy Preferences Towards Health States with Discrete Choice Experiments”**, w: C. Berger-Vachon, A. Gil Lafuente, J. Kacprzyk, Y. Kondratenko, J. Merigó, C. Morabito (red.) *“Complex Systems: Solutions and Challenges in Economics, Management and Engineering”*, Studies in Systems, Decision and Control, vol 125. Springer, ISBN: 978-3-319-69988-2, DOI: 10.1007/978-3-319-69989-9, 131–148.
49. Jakubczyk M., Kamiński B., Łatek M. (2013): **„Absencja chorobowa pracowników a funkcjonowanie przedsiębiorstw i rynku – wyniki analizy z użyciem symulacyjnego modelu wieloagentowego”**, w: T. Hermanowski (red.) *“Metody szacowania kosztów społecznych choroby i wpływu stanu zdrowia na aktywność zawodową i wydajność pracy”*, Wolters Kluwer ISBN: 978-83-264-4367-1, 48–70.
50. Jakubczyk M., Koń B. (2017): **“The impact of firms’ adjustments on the indirect cost of illness”**, *International Journal of Health Economics and Management*, 17, 377–394.
51. Jakubczyk M., Lipka I., Pawęska J., Niewada M., Rdzanek E., Zaletel J., Ramírez de Arrellano A., Doležal T., Chekorova Mitreva B., Nagy B., Petrova G., Šarić T., Yfantopoulos J., Czech M. (2016): **„Cost of severe hypoglycaemia in nine European countries”**, *Journal of Medical Economics*, 19 (10), 973–982.
52. Jakubczyk M., Rdzanek E., Niewada M., Czech M. (2015): **“Economic resources consumption structure in severe hypoglycemia episodes: a systematic review and meta-analysis”**, *Expert Review of Pharmacoeconomics & Outcomes Research*, DOI: 10.1586/14737167.2015.1076338.
53. Jakubczyk M., Wrona W., Macioch T., Golicki D., Niewada M., Hermanowski T. (2010): **„Koszty pośrednie w ocenie technologii medycznych”**, *Polski Merkurusz Lekarski*, XXVIII, 163, 42–45.
54. Jakubiak-Lasocka J., Jakubczyk M. (2014): **„Cost-effectiveness versus Cost-Utility Analyses: What Are the Motives Behind Using Each and How Do Their Results Differ?—A Polish Example”**, *Value in Health Regional Issues*, 4C, 66–74.
55. Jelsma J., Hansen K., De Weerd W., De Cock P., Kind P. (2003): **“How do Zimbabweans value health states?”**, *Population Health Metrics*, 1 (1), 11.
56. Kamiński B., Jakubczyk M., Szufel P. (2018): **„A framework for sensitivity analysis of decision trees”**, *Central European Journal of Operations Research*, 26 (1), 135–159, DOI: 10.1007/s10100-017-0479-6.
57. Konstytucja Rzeczypospolitej **Polskiej** (1997), Dz.U. 1997 nr 78 poz. 483, <http://prawo.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU19970780483>, dostęp 31.12.2017.
58. Koopmanschap M., van Ineveld B. (1992): **“Towards a new approach for estimating indirect cost of disease”**, *Social Science and Medicine*, 34, 1005–1010.
59. Lamers LM., McDonnell J., Stalmeier PF., Krabbe PF., Busschbach JJ. (2006): **“The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies”**, *Health Economics*, 15, 1121–1132.
60. Laska EM., Meisner M., Siegel C., Stinnett AA. (1999): **“Ratio-based and Net Benefit-based Approaches to Health Care Resource Allocation: Proofs of Optimality and Equivalence”**, *Health Economics*, 8, 171–174.

61. Lasocka J., Jakubczyk M., Siekmeier R. (2013): **“The annual costs of smoking-attributable productivity losses in Poland”**, *Advances in experimental medicine & biology*, 755, 179–187.
62. Liljas B. (1998): **“How to calculate indirect costs in economic evaluations”**, *Pharmacoeconomics*, 13, 1–7.
63. Löthgren M., Zethraeus N. (2000): **„Definition, interpretation and calculation of cost-effectiveness acceptability curves”**, *Health Economics*, 15, 363–372.
64. Maddala GS. (2006): **“Ekonometria”**, Wydawnictwo Naukowe PWN, Warszawa, ISBN: 8301146389.
65. Miyamoto J.M., Wakker P.P., Bleichrodt H., Peters H.J.M. (1998): **“The Zero-Condition: A Simplifying Assumption in QALY Measurement and Multiattribute Utility”**, *Management Science*, 44 (6), 839–849.
66. von Neumann J., Morgenstern O. (1947): **„Theory of Games and Economic Behavior”**, Princeton University Press.
67. Niewada M., Polkowska M., Jakubczyk M., Golicki D. (2013): **„What Influences Recommendations Issued by the Agency for Health Technology Assessment in Poland? A Glimpse Into Decision Makers’ Preferences”**, *Value in Health Regional Issues*, 2, 267–272.
68. Nussbaum M.C. (2013): **“Creating Capabilities. The Human Development Approach”**, Belknap Press, ISBN 9780674072350
69. Pauly M., Nicholson S., Xu P.D.J., Danzon P., Murray J., Berger M. (2002): **“A general model of the impact of absenteeism on employers and employees”**, *Health Economics*, 11, 221–231.
70. Pelczarska A., Jakubczyk M., Jakubiak-Lasocka J., Banach M., Myśliwiec M., Gruchała M., Niewada M. (2018): **“The cost-effectiveness of screening strategies for familial hypercholesterolaemia in Poland”**, *Atherosclerosis*, 270, 132–138, doi 10.1016/j.atherosclerosis.2018.01.036.
71. Pliskin J.S., Shepard D.S., Weinstein M.C. (1980): **“Utility Functions for Life Years and Health Status”**, *Operations Research*, 28 (1), 206–224.
72. Prevolnik-Rupel V., Rebolj M. (2000): **“The Slovenian VAS Tariff Based on Valuations of EQ-5D Health States from the General Population”**, The 17th Plenary Meeting of the EuroQoL Group, Pampeluna, Hiszpania, 23–46.
73. Ramos-Goñi JM., Craig BM., Oppe M., Ramallo-Fariña Y., Pinto-Prades JL., Luo N., Rivero-Arias O. (2017): **“Handling Data Quality Issues to Estimate the Spanish EQ-5D-5L Value Set Using a Hybrid Interval Regression Approach”**, *Value in Health*, DOI: <http://dx.doi.org/10.1016/j.jval.2017.10.023>.
74. Rowen DL., Brazier JE., van Hout B. (2014): **„A comparison of methods for converting DCE values onto the full health–dead QALY scale”**, *Medical Decision Making*, 35 (3), 328–340.
75. Rozporządzenie Ministra Zdrowia w sprawie minimalnych wymagań (2012), Dz.U. 2012 poz. 388, <http://prawo.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20120000388>, dostęp 26.12.2017.
76. Sadatsafavi M., Najafzadeh M., Marra C. (2008): **„Acceptability curves could be misleading when correlated strategies are compared”**, *Medical Decision Making*, 28 (3), 306–307.
77. Santos M., Cintra MA., Monteiro AL., Santos B., Gusmão-Filho F., Andrade MV., Noronha K., Cruz LN., Camey S., Tura B., Kind P. (2016): **“Brazilian Valuation of EQ-5D-3L”**

- Health States: Results from a Saturation Study**”, *Medical Decision Making*, 36 (2), 253–263.
78. Sen A. (1985): “**Commodities and capabilities**”, North-Holland, New York, ISBN: 9780444877307.
79. Severens J., Brunenberg D., Fenwick E., O’Brien B., Joore M. (2005): “**Cost-effectiveness acceptability curves and a reluctance to lose**”, *Pharmacoeconomics*, 23 (12), 1207–1214.
80. Shaw JW., Johnson JA., Coons SJ. (2005): “**US valuation of the EQ-5D health states: development and testing of the D1 valuation model**”, *Medical Care*, 43, 203–220.
81. Shiroiwa T., Ikeda S., Noto S., Igarashi A., Fukuta T., Saito S., Shimozuma K. (2016): “**Comparison of Value Set Based on DCE and/or TTO Data: Scoring for EQ-5D-5L Health States in Japan**”, *Value in Health*, 19 (5), 648–654.
82. „**Słownik języka polskiego**”, Szymczak M. (red.), PWN, Warszawa, 1988, ISBN: 83-01-00281-6.
83. Suhrcke M., McKee M., Stuckler D., Sauto Arce R., Tsoлова S., Mortensen J. (2006): “**The contribution of health to the economy in the European Union**”, *Public Health*, 120 (11), 994–1001.
84. Tsuchiya A., Ikeda S., Ikegami N., Nishimura S., Sakai I., Fukuda T., Hamashima C., Hisashige A., Tamura M. (2002): “**Estimating an EQ-5D population value set: the case of Japan**”, *Health Economics*, 11 (4), 341–353.
85. Ustawa Refundacyjna (2011), Dz.U. 2011 nr 122 poz. 696, <http://prawo.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20111220696>, dostęp 26.12.2017.
86. Wakker P. (2010): „**Prospect theory: for risk and uncertainty**”, Cambridge University Press, ISBN 978-0521748681.
87. Wee HL., Li SC., Xie F., Zhang XH., Luo N., Feeny D., Cheung YB., Machin D., Fong KY., Thumboo J. (2008): “**Validity, feasibility and acceptability of time trade-off and standard gamble assessments in health valuation studies: A study in a multi-ethnic Asian population in Singapore**”, *Value in Health*, 11, S3–S10.
88. Weinstein M., Zeckhauser R. (1973): “**Critical ratios and efficient allocation**”, *Journal of Public Economics*, 2, 147–157.
89. Whitehead S.J., Ali S. (2010): “**Health outcomes in economic evaluation: the QALY and utilities**”, *British Medical Bulletin*, 96 (1), 5–21.
90. Wittrup-Jensen KU., Lauridsen JT., Gudex C., Brooks R., Pedersen KM. (2002): “**Estimating Danish EQ-5D tariffs using TTO and VAS**” w: “Proceedings of the 18th Plenary Meeting of the EuroQol Group” Norinder A., Pedersen K., Roos P. (red.), The Swedish Institute for Health Economics 2002, 257–292.
91. The World Bank, **Health expenditure, total (% of GDP)** <https://data.worldbank.org/indicator/SH.XPD.TOTL.ZS>, dostęp 23.12.2017.
92. Zadeh L. (1965): “**Fuzzy sets**”, *Information and Control*, 8 (3), 338–353.

Michał Jakubczyk