



Michał Jakubczyk

Szkoła Główna Handlowa w Warszawie
Kolegium Analiz Ekonomicznych
Zakład Wspomagania i Analizy Decyzji
michal.jakubczyk@sgh.waw.pl

REKURENCYJNE I KONSERWATYWNE REGUŁY DECYZYJNE PRZY WYBORZE TECHNOLOGII MEDYCZNYCH W WARUNKACH RYZYKA¹

Streszczenie: W pracy rozważany jest problem wyboru technologii medycznych z uwzględnieniem wielokryterialności (ocena ze względu na efekt i koszt) oraz niepewności oszacowań ocen wariantów. Przedmiotem opracowania są metody wspierania decydenta, a w szczególności ilustrowania wpływu niepewności na ocenę wariantów. W opracowaniu zaproponowano modyfikację tzw. krzywych akceptowalności (*cost-effectiveness acceptability curves*, CEAC), tak aby zminimalizować wpływ nieintuicyjnych własności przedstawianych w literaturze (np. naruszeń własności α funkcji wyboru, naruszenia dominacji stochastycznej pierwszego rzędu). W szczególności zaproponowano stosowanie CEAC w oparciu o rekurencyjny lub konserwatywny dobór porównań parami do wykonania. Wykazano, że zaproponowane podejścia eliminują wiele niepożądanych własności CEAC.

Słowa kluczowe: ocena technologii medycznych, krzywe akceptowalności, funkcja wyboru, własność α , dominacja stochastyczna pierwszego rzędu.

Wprowadzenie

Przedmiotem opracowania są metody wspierania decydenta w ocenie technologii medycznych (*health technology assessment*, HTA). Celem – w koniecznym tu uproszczeniu – zaproponowanie modyfikacji stosowania tzw. krzywych akceptowalności (*cost-effectiveness acceptability curves*, CEAC), aby wyeliminować ich nieintuicyjne własności wskazywane w literaturze.

¹ Większość prezentowanych w opracowaniu wyników uzyskano w ramach projektu IP2010013870 finansowanego przez Ministerstwo Nauki i Szkolnictwa Wyższego (praca naukowa finansowana ze środków budżetowych na naukę w latach 2010-2011).

HTA to ocena konsekwencji stosowania technologii medycznej (np. leku, procedury medycznej) w miejsce innej i obejmuje metody z zakresu m.in. medycyny, ekonomii i statystyki [Gold i in., 1996]. W wielu krajach HTA stanowi zasadniczy element wspomagający politykę zdrowotną w zakresie refundacji leków. Ustawa o refundacji w Polsce wskazuje wyniki HTA jako jedno z kryteriów decyzji.

W problemach decyzyjnych w HTA występuje wielokryterialność – maksymalizacja efektów zdrowotnych musi uwzględniać koszty leczenia. Skutkuje to koniecznością uwzględnienia skłonności do płacenia (*willingness to pay*, WTP) za efekty zdrowotne. Wartość WTP w Polsce określono ustawą (w kwietniu 2015 r. wynosiła 119 577 zł za rok życia w pełnym zdrowiu). W wielu krajach wartości WTP nie ogłoszono oficjalnie i podejmuje się próby jej określenia na podstawie decyzji regulatora [Devlin i Parkin, 2004]. W HTA występuje też niepewność² – konsekwencje stosowania technologii są szacowane np. na podstawie badań klinicznych³.

W literaturze zaproponowano wiele metod wspierania decydenta, pozwalających na porównywanie wariantów, zrozumienie ryzyka związanego z wyborem, ocenę wpływu WTP na wynik [np. van Hout i in., 1994; Briggs i Fenn, 1998; Barton, Briggs i Fenwick, 2008; Eckermann i Willan, 2011]. Mnogość metod wynika ze złożoności problemów i trudności technicznych (niemożność obliczenia) lub dotyczących interpretacji (naruszenie aksjomatów decyzji) dla większości z nich. Jedną z metod są krzywe akceptowalności (CEAC), pokazujące prawdopodobieństwo⁴, że dany wariant jest optymalny (w sensie doprecyzowanym dalej) w zależności od WTP. CEAC jednocześnie ilustrują wpływ wielokryterialności i niepewności, toteż często stosowane są jako metoda analizy wrażliwości. Wady CEAC szeroko omówiono w literaturze, krótki przegląd zamieszczono w dalszej części niniejszego opracowania. W pracy zaproponowano modyfikację sposobu wykorzystania CEAC polegającą na ograniczeniu zbioru wariantów, z którymi następują porównania, poprzez wykonanie jedynie wybranych – określonych regułami formalnie zdefiniowanymi dalej – porównań parami. Zaproponowano w tym zakresie podejście określone jako rekurencyjne i konserwatywne, przedstawiono uzasadnienie przyjętych definicji i zbadano ich własności.

² Zazwyczaj znane są rozkłady prawdopodobieństwa, tak więc terminologicznie można mówić o ryzyku. W dalszej części pracy oba określenia, niepewność i ryzyko, stosowane są zamiennie.

³ Rozróżnia się tutaj dwa typy niepewności. Typ I oznacza niepewność na poziomie jednego pacjenta i wynika ze specyfiki procesu leczenia. Typ II oznacza niepewność oszacowania parametrów (np. prawdopodobieństwa wyleczenia). W opracowaniu zajmujemy się niepewnością II rodzaju, czyli zakładamy, że ponieważ decyzje dotyczące wyboru terapii będą stosowane u wielu pacjentów, decydent interesuje się średnimi charakteryzującymi dostępne warianty.

⁴ Stosując podejście bayesowskie zamiast częstościowego.

1. Wspieranie decyzji w ocenie technologii medycznych

Poniżej wprowadzono notację, omówiono metody analizy decyzji stosowane w HTA, np. formalnie zdefiniowano CEAC i przedstawiono krótki przegląd trudności w ich stosowaniu.

1.1. Metody wspierania decydenta – formalizacja

Zakładamy dalej, że porównujemy n wariantów oznaczonych $i = 1, \dots, n$. Rozważany zbiór wariantów oznaczmy $X = \{1, \dots, n\}$. Przez i rozumiemy $X \setminus \{i\}$ (gdy X jest znany z kontekstu). Czasem dla zwiększenia czytelności warianty będą oznaczane T_1, \dots, T_n .

Pomijając niepewność, warianty decyzyjne można scharakteryzować jako (e_i, c_i) , tj. efekt i koszt. W praktyce HTA koszty najczęściej oblicza się z perspektywy płatnika publicznego, tj. zlicza się koszty refundowanych leków, hospitalizacji, dodatkowo płatnych procedur itd. Efekty najczęściej wyraża się jako tzw. lata życia w pełnym zdrowiu lub, inaczej, skorygowane o jakość lata życia (*quality-adjusted life years*, QALY). QALY mogą być, np. dla aksjomatyzacji, którą przedstawili Bleichrodt, Wakker i Johannesson [1997], interpretowane analogicznie do wartości użyteczności w ujęciu von Neumanna-Morgensterna.

Przy powyższym opisie dwa warianty można wówczas porównać, obliczając koszt uzyskania dodatkowej jednostki efektu przy stosowaniu droższej i bardziej skutecznej technologii (wykluczając trywialny przypadek dominacji, tj. np. $e_1 > e_2, c_1 < c_2$), czyli tzw. współczynnik koszty-efektywność (*incremental cost-effectiveness ratio*, ICER)⁵:

$$ICER_{2vs1} = \frac{c_2 - c_1}{e_2 - e_1}. \quad (1)$$

Podjęcie decyzji wymaga odniesienia ICER do WTP. Dla $n > 2$ wariantów liczba porównań szybko rośnie i wygodniej dla każdej technologii obliczyć tzw. korzyść netto (*net benefit*, NB), po monetarnym wyrażeniu efektu i odjęciu kosztu, i ją maksymalizować:

$$NB_i = WTP \times e_i - c_i. \quad (2)$$

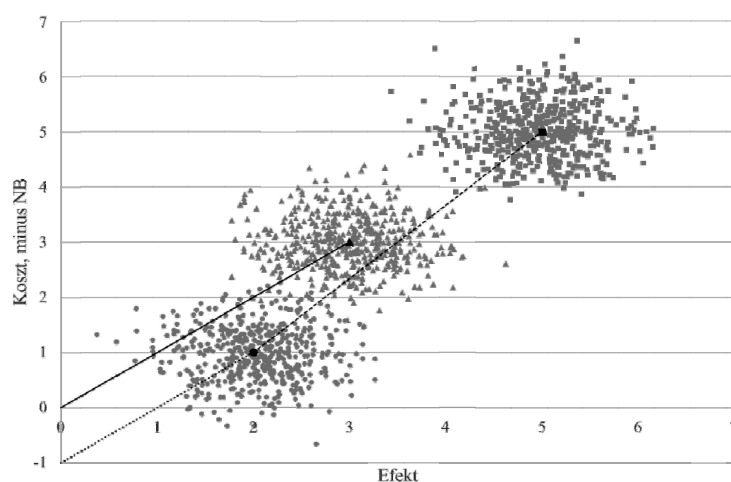
Niepewność powoduje, że zamiast (e_i, c_i) powinniśmy rozważać zmienne losowe (E_i, C_i) ⁶. Często dostępne badania kliniczne są duże i niepewność jest zadana rozkładem normalnym. Mniejsza ilość danych prowadzi np. do prawo-

⁵ W opracowaniu wykorzystano przyjęte powszechnie w literaturze przedmiotu polskie terminy, jakkolwiek niefortunne dla ekonomisty jest tłumaczenie angielskiego *effectiveness* (w sensie: skuteczność) jako efektywność (*efficiency*).

⁶ Dane w postaci parametrycznej w wyniku estymacji, a często także w postaci rozkładu empirycznego wynikające z procesu symulacji Monte Carlo.

stronnie skośnych rozkładów kosztów. Błędy ocen dla poszczególnych wariantów mogą być skorelowane, np. z uwagi na wykorzystanie wspólnego parametru (użyteczności stanu zdrowia)⁷. Tak więc wartości ICER i NB_i są zmiennymi losowymi. W HTA decyzje często podejmuje się, kierując się punktową oceną z wykorzystaniem wartości oczekiwanej korzyści netto (*expected net benefit*, ENB), ale dla decydenta istotne jest zrozumienie stopnia ryzyka związanego z decyzją. Próba opisanie niepewności za pomocą przedziałów ufności dla ICER i NB_i napotyka trudności: analiza rozkładów NB_i prowadzi do utraty informacji w przypadku korelacji, zaś ICER jest miarą ilorazową, co powoduje techniczne trudności [Briggs i Fenn, 1998]. Dlatego niepewność często ilustruje się za pomocą CEAC, czyli funkcji mierzących dla różnych WTP prawdopodobieństwo maksymalizowania korzyści netto przez daną technologię [van Hout i in., 1994; Löthgren i Zethraeus, 2000]:

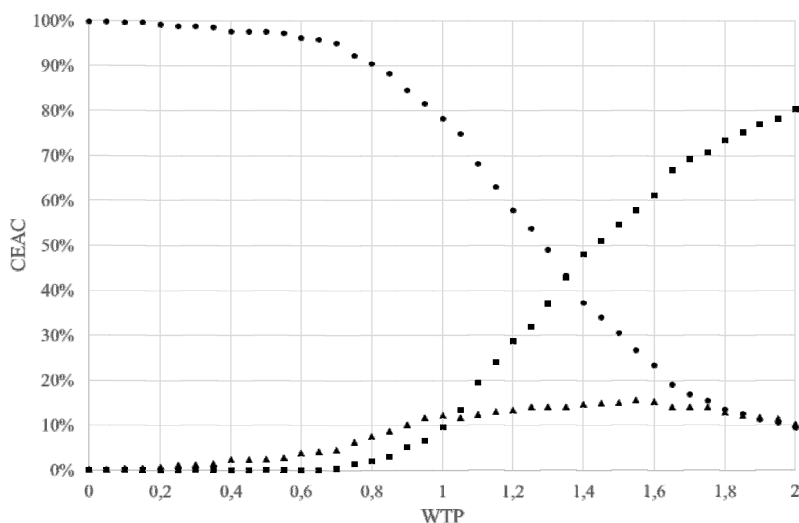
$$CEAC_i = P(NB_i = \max(NB_1, \dots, NB_n)). \quad (3)$$



Objaśnienia: Na osi rzędnych mierzony jest koszt technologii. W tych samych jednostkach wyrażana jest korzyść netto (NB), ponieważ większy koszt to mniejsza korzyść, więc oś rzędnych oznacza także minus NB. Oceny punktowe trzech wariantów (czarny) i niepewności tej oceny (szary). WTP = 1 jako nachylenie linii ciągłej, ICER w jednym z porównań – kreskowanej (tu ICER = 4/3). Linia kropkowana ilustruje obliczenie NB najtańszej technologii (NB = 1).

Rys. 1. Przestrzeń ocen

⁷ Należy częściej oczekiwać dodatnich korelacji, ale ujemne korelacje także są możliwe. Na przykład powiedzmy, że lek A jest skuteczny i drogi i może prowadzić do wystąpienia uleczalnego działania niepożądanego. Zdefiniujmy wariant 1 jako stosowanie leku A i leczenie działania niepożądanego, o ile wystąpi, co zwiększa koszt i pozostawia niezmienny efekt. Zdefiniujmy wariant 2 jako stosowanie leku A i zmianę leczenia na tańszy i gorszy lek B, jeśli wystąpi działanie niepożądane, co redukuje efekt i koszt. Niepewność związana z ryzykiem wystąpienia działania niepożądanego prowadzi do negatywnej korelacji kosztu obu wariantów.



Objaśnienie: Dla $WTP = 1$ technologia oznaczona kółkiem z największym prawdopodobieństwem maksymalizuje NB.

Rys. 2. CEAC dla wariantów z rys. 1

Powyższe rozważania zilustrowano na rys. 1 i 2. Na podstawie CEAC można zdefiniować regułę decyzyjną jako wybór technologii, dla której prawdopodobieństwo maksymalizowania NB jest największe:

$$C(X) = \arg \max_{i \in X} (P(NB_i = \max_{j \in X} (NB_j))). \quad (4)$$

Na rys. 1 i 2 występuje zgodność powyższej reguły i wyboru według ENB (dla $WTP = 1$), natomiast CEAC przeważnie wykorzystywane są nie do wybierania, a do wskazywania niepewności związanej z wyborem i mają co najmniej trzy zalety: 1) nie wymagają *explicite* zdefiniowania WTP (co często jest politycznie trudne), pokazując wynik dla wielu różnych wartości; 2) ich wynik jest zawsze w zakresie 0-100%, co ułatwia ich stosowanie w różnych problemach; 3) CEAC zawsze są zdefiniowane (a np. przedziały ufności ICER już nie). W literaturze wskazano także wiele wad CEAC omówionych w kolejnym podpunkcie.

1.2. Trudności w stosowaniu krzywych akceptowalności

Ograniczając się do porównywania dwóch wariantów, jeśli rozkłady niepewności mogą być asymetryczne, może zdarzyć się, że $E(NB_1) < E(NB_2)$ i jednocześnie $P(NB_1 > NB_2) > 50\%$ [Fenwick i in., 2001]. Oznacza to, że dla asyme-

trycznych rozkładów już dla $n = 2$ może wystąpić niezgodność wyboru dokonanego na podstawie oceny punktowej i wyniku analizy wrażliwości. Utrudnia to percepcję wyników analizy wrażliwości w kontekście jej niezgodności co do samego wyboru z analizą podstawową.

Dla asymetrycznych rozkładów niepewności możemy uzyskać nieprzechodnie wyniki porównań parami, np. $P(NB_1 > NB_2) > 50\%$, $P(NB_2 > NB_3) > 50\%$ i $P(NB_3 > NB_1) > 50\%$ [Jakubczyk i Kamiński, 2010].

Dla $n > 2$ niezgodność z ENB może wystąpić także dla wielowymiarowych normalnych rozkładów NB_i . CEAC mogą (dla pewnego WTP) wskazywać na technologię, które nie maksymalizuje ENB dla żadnego WTP (jest zdominowana w sposób uogólniony⁸) [Barton, Briggs i Fenwick, 2008]. Wybór z użyciem CEAC narusza własność α (własność Chernoffa [Sen, 1971]) także dla gausowskich rozkładów [Jakubczyk i Kamiński, 2010], np. dla $n = 3$ wariantów T_1 , T_2 , T_3 , CEAC wskazuje na T_1 , natomiast po odrzuceniu T_3 – na T_2 (tabela 1). Wynik zależy zatem od restrukturyzacji problemu, pominięcie wariantu pozornie nieistotnego (T_1) zmienia wybór w sposób nieintuicyjny dla decydenta (z T_3 na T_2). W szczególności sklonowanie wariantu przegrywającego (T_2) może nieintuicyjnie sprawić, że stanie się (wraz z klonem T_3) wariantem wygrywającym.

Reguła decyzyjna oparta na CEAC jest wrażliwa na korelacje ocen wariantów [Sadatsafavi i in., 2008]. Jeśli oceny są niezależnymi zmiennymi losowymi, to opcja zdominowana w sensie dominacji stochastycznej pierwszego rzędu (*first-order stochastic dominance*, FOSD) ma nie wyższą wartość CEAC niż opcja ją dominująca, niezależnie od innych wariantów [Jakubczyk i Kamiński, 2010]. W przywołanej pracy nie podano dowodu, więc, uzupełniając, zapiszmy formalnie poniższe twierdzenie.

Twierdzenie 1⁹. Dla n niezależnych zmiennych losowych NB_1, \dots, NB_n , jeśli NB_1 FOSD NB_2 , to (możliwa jest równość, np. zero po obu stronach):

$$P(NB_1 \geq \max(NB_2, \dots)) \geq P(NB_2 \geq \max(NB_1, NB_3, \dots)). \quad (5)$$

Korelacja między ocenami wariantów zmienia tę implikację. W tabeli 2 przedstawiono przykłady. Ujemna korelacja sprawia, że warianty zdominowane wygrywają w porównaniach CEAC. Intuicyjnie: ujemna korelacja prowadzi do wspierania się T_1 i T_2 w porównaniach. Dodatnia korelacja między wariantami dominującymi sprawia, że warianty dominujące konkurują i przegrywają porównanie z wariantem zdominowanym.

⁸ Taką technologią jest np. technologia oznaczona trójkątami na rys. 1 (choć w naszym przykładzie nie jest ona wskazywana przez CEAC dla żadnego WTP, por. rys. 2).

⁹ Dowody wszystkich twierdzeń podanych w pracy przedstawiono w aneksie.

2. Rekurencyjne i konserwatywne krzywe akceptowalności

Przedstawione wady CEAC motywują do zaproponowania modyfikacji, które pozwoliłyby na usunięcie choć części nieintuicyjnych własności.

2.1. Definicje

Jak pokazano, np. sklonowanie nieoptymalnego wariantu lub wprowadzenie korelacji może prowadzić do utraty przez wariant rekomendacji z wykorzystaniem CEAC. Tymczasem to klonowanie nie daje nowej opcji decydentowi, identyczny wariant był już dostępny. Intuicyjnie – niepożądane własności CEAC wynikają z faktu, że porównujemy poszczególne warianty ze wszystkimi pozostałymi jednocześnie. Jest to także sprzeczne z operacyjnym rozumieniem procesu decydowania. W praktyce niewybranie danego wariantu oznacza, że wybrany zostanie inny (ale *jeden*, a nie *wszystkie*). W tym sensie istotne są porównania parami i warto rozważyć ograniczenie stosowania CEAC do takich porównań. Porównywanie wszystkich możliwych par prowadziłoby do bardzo wielu porównań i utrudniało odbiór przez decydenta, warto więc zdefiniować, które porównania należy wykonać. Zdefiniowano poniżej dwa podejścia. Rozważono *rekurencyjny* wybór komparatora jako najbardziej atrakcyjnego wariantu w porównaniach parami (ze swoimi, wybranymi analogicznie, komparatorami). Po drugie, rozważono *konserwatywny*, pesymistyczny, wybór komparatora – jako wariantu, względem którego analizowany wariant jest najmniej atrakcyjny.

Definicja 1. Zdefiniujemy *rekurencyjną regułę wyboru* $C^R(\cdot)$ dla $n = 2$:

$$C^R(\{k, l\}) = \arg \max_{i \in \{k, l\}} (P(NB_i = \max(NB_k, NB_l))), \quad (6)$$

zaś dla $n > 2$ rekurencyjnie¹⁰:

$$C^R(X) = \arg \max_{i \in X} (P(NB_i \geq NB_{C^R(-i)})). \quad (7)$$

Wtedy *rekurencyjna krzywa akceptowalności* jest dana jako

$$CEAC^R(i) = P(NB_i \geq NB_{C^R(-i)}). \quad (8)$$

Tak więc podejście rekurencyjne proponuje porównywanie każdego wariantu A do innego wariantu B, który z kolei zostałby wybrany, gdyby A nie był

¹⁰ W definicji nie uwzględniono możliwości wskazania przez $C^R(-i)$ kilku wariantów, co z praktycznego punktu widzenia jest nieistotne. Np. doprecyzowanie tego elementu przez wzięcie wówczas średniej wartości prawdopodobieństwa nie zmienia wyników opracowania.

dostępny, przy czym określenie *zostałby wybrany* zakłada wykorzystanie analogicznej reguły w mniejszym zbiorze (z pominięciem A).

Definicja 2. *Konserwatywna krzywa akceptowalności* jest dana jako

$$CEAC^C(i) = \min_{j \in -i} P(NB_i \geq NB_j). \quad (9)$$

Zdefiniujmy także *konserwatywną regułę wyboru* C^C :

$$C^C(X) = \arg \max_{i \in X} (CEAC^C(i)). \quad (10)$$

Podejście konserwatywne proponuje wskazywanie pesymistycznego wyniku, tj. porównania parami o najmniejszym CEAC. Oczywiście jednoczesne porównanie ze wszystkimi wariantami prowadzi do jeszcze mniejszych wartości CEAC i wydaje się bardziej konserwatywne, ale nawiązuje do filozofii, że zamiast danego wariantu można wybrać wszystkie inne.

2.2. Własności

Poniżej przedstawiono własności $CEAC^R$ i $CEAC^C$ w odniesieniu do wad klasycznych CEAC. Zauważmy, że analizę własności krzywych akceptowalności można prowadzić dla ustalonego WTP lub zmieniającego się WTP. W pierwszym przypadku badamy wpływ dodawania/modyfikacji wariantów, korelacji itd. W drugim można badać kwestie monotoniczności, ciągłości itp. [por. np. Jakubczyk i Kamiński, 2010]. W niniejszym opracowaniu badanie ograniczono do pierwszego obszaru zagadnień, drugi obszar wymagałby znacznego rozbudowania analizy.

Wpierw zajmijmy się niepewnością określoną wielowymiarowymi normalnymi rozkładami niepewności. Zachodzi poniższe twierdzenie.

Twierdzenie 2. Rozważamy NB_1, \dots, NB_n zadane (potencjalnie skorelowanymi) wielowymiarowymi rozkładami normalnymi. C^C i C^R (definicja 1, 2) mają własności: i) wskazują warianty maksymalizujące ENB¹¹; ii) klonowanie wariantów niewybranych nie prowadzi do zmiany decyzji; iii) zmiana korelacji nie zmienia decyzji; iv) spełniona jest własność α ; v) spełnione jest kryterium FOSD.

W tabelach 1 i 2 zapisano także wartości $CEAC^R$ i $CEAC^C$ pokazujące ich zachowanie dla przykładów, w których klasyczne CEAC zachowują się w nieintuicyjny sposób (por. punkt 1.2).

¹¹ Nie twierdzimy jednocześnie, że ranking generowany przez wartości $CEAC^R$ lub $CEAC^C$ jest zgodny z rankingiem według ENB.

Tabela 1. Przykład naruszenia własności α dla CEAC oraz wartości CEAC^C i CEAC^R

	T ₁	T ₂	T ₃
NB	N(9,8; 2)	N(10; 1)	N(10; 1)
CEAC	36,3%	31,8%	31,8%
CEAC ^C i CEAC ^R	46,44%	50%	50%

Tabela 2. Przykład naruszenia FOSD przez CEAC przy ujemnej i dodatniej korelacji

	T ₁	T ₂	T ₃
NB	N(9,8; 1)	N(9,8; 1)	N(10; 1)
korelacje	$\rho = -0,8$		
CEAC	34,3%	34,3%	31,8%
CEAC ^C , CEAC ^R	45%	45%	55%

	T ₁	T ₂	T ₃
NB	N(10,2; 1)	N(10,2; 1)	N(10; 1)
korelacje	$\rho = 0,8$		
CEAC	31,5%	31,5%	36,9%
CEAC ^C , CEAC ^R	50%	50%	44,1%

Dla ogólnych (niekoniecznie normalnych) rozkładów prawdopodobieństwa występuje zarówno więcej problemów dla klasycznych CEAC, jak i trudniej wykazać korzystne własności rekurencyjnej i konserwatywnej modyfikacji. Ponieważ klasyczne CEAC mogą wskazywać inny wariant niż ENB już dla dwóch wariantów decyzyjnych, w ogólności CEAC, CEAC^R i CEAC^C nie są zgodne z kryterium ENB. W tym sensie analiza wrażliwości może być niespójna z wynikami wyboru podstawowego, jeśli ten następuje z wykorzystaniem wartości oczekiwanej. Pozostają natomiast zasadne pytania o wewnętrzną spójność wyników CEAC^R i CEAC^C. Twierdzenie 1 przytoczone powyżej obowiązuje dla CEAC także dla ogólnej klasy rozkładów (w przypadku niezależnych ocen wariantów wariant FOSD zdominowany nie może wygrać z wariantem go dominującym). Poniższe twierdzenie wskazuje, że zgodność z FOSD dla niezależnych rozkładów jest zachowana przez CEAC^C.

Twierdzenie 3. Dla danego n , X i niezależnych rozkładów NB_1, NB_2, \dots, NB_n , jeśli NB_1 FOSD NB_2 , to $CEAC^C(1) \geq CEAC^C(2)$.

Usunięcie założenia o niezależności nie pozwala na utrzymanie wniosku o zgodności z FOSD nawet dla $n = 2$. Rozważmy przykład: $P(NB_1 = 0) = P(NB_1 = 1)$

$= P(NB_1 = 2) = 1/3$, $NB_2 = (NB_1 + 1 \bmod 3) - \varepsilon$, gdzie np. $\varepsilon = 1/2$. Wtedy NB_1 FOSED NB_2 , ale $P(NB_2 > NB_1) > 50\%$.

W przypadku dowolnych rozkładów, nawet dla niezależnych zmiennych NB_i , może zostać naruszona własność α przy wyborze według C^R lub C^C . Rozważmy przykład z tabeli 3. Ze zbioru $X = \{1, 2, 3\}$ zarówno C^R , jak i C^C wskażą wybór T_2 (w przypadku C^R także T_1), tymczasem ze zbioru $\{2, 3\}$ wskazane zostanie przez C^R i C^C jedynie T_3 .

Tabela 3. Przykład naruszenia własności α dla niesymetrycznych rozkładów

	T_1	T_2	T_3
NB	$P(NB_1 = 2) = 50\%$	$P(NB_2 = 3) = 75\%$	$P(NB_3 = 1) = 25\%$
	$P(NB_1 = 5) = 50\%$	$P(NB_2 = 6) = 25\%$	$P(NB_3 = 4) = 75\%$
P	$P(NB_1 > NB_3) = 62,5\%$	$P(NB_2 > NB_1) = 62,5\%$	$P(NB_3 > NB_2) = 56,25\%$

Można wykazać natomiast następujące twierdzenie:

Twierdzenie 4. Rozważmy n wariantów opisanych NB_1, \dots, NB_n . Jeśli relacja R , określona $T_1 R T_2$ wtt. $P(NB_1 \geq NB_2) \geq 50\%$, jest przechodnia, to istnieje element największy (lub elementy największe) tej relacji i jeden z nich jest wybierany przez C^R (C^C) z $CEAC^R$ ($CEAC^C$) $\geq 50\%$.

Zatem, o ile nie ma miejsca naruszenie przechodniości w porównaniach parami, to sama wielość wariantów decyzyjnych nie powoduje istotnych wewnętrznych niespójności w zachowaniu reguł C^R i C^C (jakkolwiek są niezgodne z wyborem według ENB), np. ewentualne zmiany wyboru mogą tylko zachodzić w ramach podzbioru elementów największych relacji (często jednoelementowego).

W przypadku naruszenia przechodniości relacji zdefiniowanej powyżej, trudno oczekiwać spójności, ale co ważne z praktycznego punktu widzenia, zastosowanie $CEAC^R$ i $CEAC^C$ może automatycznie ostrzec decydenta o tym fakcie. Na przykład jeśli dla każdego wariantu i istnieje wariant j , że $P(NB_j \geq NB_i) > 50\%$, to: i) dla wszystkich wariantów $CEAC^C < 50\%$, ii) dla $n = 3$ dla wszystkich wariantów $CEAC^R > 50\%$.

3. Studium przypadku

Zaproponowane metody $CEAC^R$ i $CEAC^C$ można bezpośrednio stosować w rzeczywistych raportach oceny technologii medycznych. Mogą one być prezentowane w ramach analizy wrażliwości obok klasycznych krzywych CEAC, ilekroć porównywane są więcej niż dwie technologie medyczne. Aby zilustro-

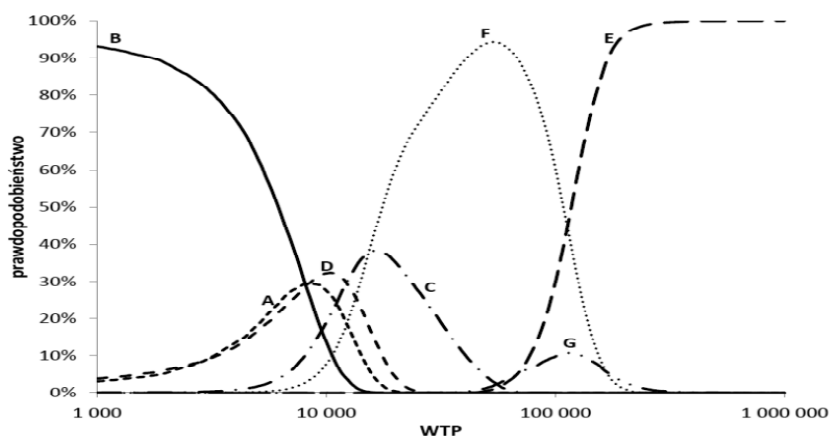
wać tę możliwość, poniżej zaprezentowano przykład wykorzystania krzywych CEAC^R i CEAC^C w rzeczywistym problemie decyzyjnym. Przykład pochodzi z badania Goeree i in. [2002] i był wykorzystany przez Barton, Briggs i Fenwick [2008] do badania własności CEAC i zdefiniowania tzw. granicy koszty-efektywność (*cost-effectiveness acceptability frontier*, CEAF)¹². W przykładzie tym niepewność oszacowań parametrów jest zadana poprzez wartości oczekiwane oraz macierze wariancji i kowariancji wielowymiarowego rozkładu normalnego dla zmiennych C_i i E_i . Porównywanych jest siedem wariantów. Wartości oczekiwane i odchylenia standardowe dla kosztów i efektów (wyrażonych odpowiednio w dolarach i QALY) przedstawiono w tabeli poniżej.

Tabela 4. Charakterystyka wariantów porównywanych w studium przypadku – wartości oczekiwane i odchylenia standardowe dla kosztów i efektów

Technologia	1	2	3	4	5	6	7
średni koszt (\$)	866,4	792,5	985,8	921	1314,1	1021,3	1257
SD kosztu (\$)	81,6	68,1	101,7	102	97,1	106,5	86
średni efekt (QALY)	0,892	0,883	0,902	0,897	0,908	0,905	0,907
SD efekty (QALY)	0,000717	0,000809	0,000644	0,000727	0,000568	0,000636	0,000576

Rys. 3 przedstawia klasyczne krzywe CEAC. Istotne jest, że technologie C i D są zdominowane w uogólnionym sensie ze względu na oczekiwane wartości kosztu i efektu, tzn. dla żadnej wartości WTP technologie te nie zostałyby wybrane przy wykorzystaniu ENB. Tymczasem krzywe akceptowalności rekomendują obie te technologie dla pewnych wartości WTP. Z drugiej strony technologia A maksymalizuje ENB dla wartości WTP około 10 tys. dolarów, lecz nie jest wskazywana przez CEAC dla żadnej wartości WTP. W tym sensie występuje w tym przykładzie istotna niezgodność kryterium opartego na wartości oczekiwanej i na prawdopodobieństwie. Niezgodność ta wynika z dużej niepewności oszacowań parametrów dla wariantów C i D, co prowadzi do dużego prawdopodobieństwa maksymalizacji NB przez te warianty w jednoczesnym porównaniu [por. Jakubczyk i Kamiński, 2010].

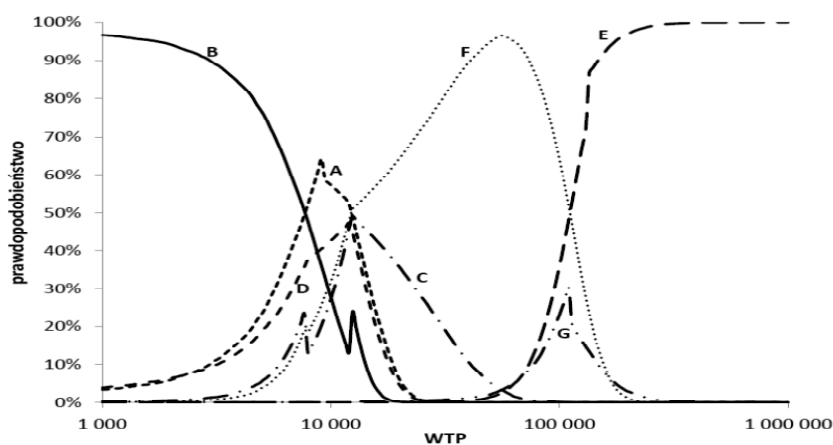
¹² Chciałbym w tym miejscu wyrazić podziękowania dla A. Briggsa i pozostałych autorów pracy [Barton i in., 2008] za udostępnienie szczegółowych danych.



Rys. 3. CEAC dla przykładu Barton i in. [2008], WTP na skali logarytmicznej

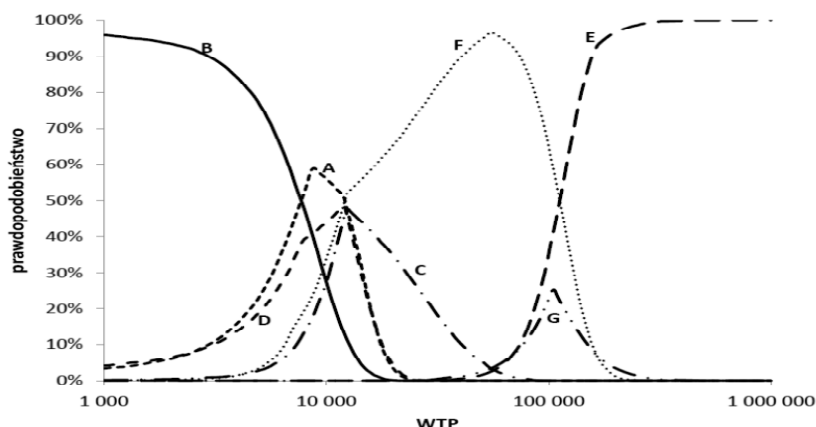
Źródło: Na podstawie danych [Barton i in., 2008].

Rys. 4 przedstawia krzywe CEAC^R obliczone dla tego problemu decyzyjnego. Kolejno, dla rosnących wartości WTP, wybierane są technologie B, A, F i E. Wybór ten jest zgodny z kryterium maksymalizacji ENB, ale zastosowanie CEAC^R pozwala decydentowi dodatkowo na określenie stopnia niepewności związanego z poszczególnymi wyborami i wartości WTP, dla których nie ma zdecydowanego zwycięzcy. Krzywe CEAC^C przedstawia rys. 5. Przebiegają podobnie do CEAC^R i wskazują na te same warianty. Różnice jakościowe polegają na braku nieciągłości.



Rys. 4. CEAC^R dla przykładu Barton i in. [2008], WTP na skali logarytmicznej

Źródło: Na podstawie danych [Barton i in., 2008].



Rys. 5. CEAC^C dla przykładu Barton i in. [2008], WTP na skali logarytmicznej

Źródło: Na podstawie danych [Barton i in., 2008].

Podsumowanie

Zaproponowano modyfikację sposobu wykorzystania CEAC – narzędzia standardowo wykorzystywanego w HTA – aby usunąć nieintuicyjne z punktu widzenia teorii własności, a jednocześnie zachować przydatność w analizie wrażliwości, poprzez zdolność do mierzenia niepewności związanej z podjętą decyzją i wpływu skłonności do płacenia. Proponowane modyfikacje – podejście rekurencyjne i konserwatywne – wydają się lepiej intuicyjnie nawiązywać do logiki procesu decyzyjnego i pozbawione są wszystkich nieintuicyjnych własności w ważnym przypadku niepewności określonej przez rozkład normalny. W przypadku ogólnych rozkładów podejście konserwatywne wydaje się lepsze (zachowuje zgodność z kryterium FOSD, gwarantuje wygenerowanie sygnału ostrzegawczego w przypadku istotnego naruszenia przechodności w porównaniach parami).

Przedstawione rozważania można odnosić do trzech typów problemów. Po pierwsze, uzyskane wyniki oczywiście bezpośrednio stosują się do problemów decyzyjnych w obszarze HTA. Po drugie, przedstawione techniki można bezpośrednio przenieść do problemów decyzyjnych, w których porównywane są zmienne losowe, niekoniecznie związane z oceną technologii medycznych (także bez wielokryterialności). Po trzecie, możliwe wydaje się zastosowanie analogicznego podejścia do innych problemów decyzyjnych, także bez niepewności, w których reguły wyboru naruszają intuicyjnie pożądane własności. Takim obszarem jest wybór grupowy, w którym wiele proponowanych reguł narusza czy to własność przechodności, czy też niezależność od nieistotnych wariantów.

Aneks

Dowód twierdzenia 1. Możemy zapisać:

$$\begin{aligned}
 P(NB_1 \geq \max(NB_2, \dots)) &= \int_{-\infty}^{+\infty} P(NB_1 \geq y) dF_{-1}(y) = \\
 &= \int_{-\infty}^{+\infty} (1 - F_{NB_1}(y)) dF_{-1}(y) \geq \int_{-\infty}^{+\infty} (1 - F_{NB_2}(y)) dF_{-1}(y),
 \end{aligned} \tag{11}$$

gdzie F_{-1} to dystrybuanta $\max(NB_2, NB_3, \dots, NB_n)$. Dalej mamy:

$$\begin{aligned}
 P(NB_2 \geq \max(NB_1, NB_3, \dots)) &= \int_{-\infty}^{+\infty} F_{NB_1}(y) F_{NB_3}(y) \dots dF_{NB_2}(y) \\
 &\leq \int_{-\infty}^{+\infty} F_{NB_2}(y) F_{NB_3}(y) \dots dF_{NB_2}(y).
 \end{aligned} \tag{12}$$

Równość ostatnich wyrażen w obu wyrażeniach daje żądany wynik.

Dowód twierdzenia 2. $P(NB_i > NB_j) = P(NB_i - NB_j > 0)$ i różnice NB mają rozkład normalny, więc $P(NB_i - NB_j > 0) > 1/2$ wtt. $E(NB_i) > E(NB_j)$. Ad. i) Zatem tylko jeden wariant (pomijając graniczny przypadek, że kilka wariantów ma takie samo ENB) wygra wszystkie porównania i zostanie wskazany przez CEAC^R i CEAC^C. Ad. ii) warianty niewybrane miały niemaksymalne ENB i ich powielenie nie zmieni rekomendacji. Ad. iii) korelacja nie wpływa na prawdziwość $P(NB_i - NB_j > 0) > 50\%$ wtt. $E(NB_i) > E(NB_j)$. Ad. iv) usunięcie wariantu nieistotne (o niemaksymalnym ENB) nie powoduje zmiany. Ad. v) NB_1 FOSD NB_2 implikuje $E(NB_1) > E(NB_2)$.

Dowód twierdzenia 3. Analogicznie jak w dowodzie twierdzenia 2 otrzymujemy, że $P(NB_1 \geq NB_i) \geq P(NB_2 \geq NB_i)$ dla dowolnego $i > 2$. Natychmiast też otrzymujemy, że $P(NB_1 \geq NB_2) \geq P(NB_2 \geq NB_1)$. Oznacza to, że CEAC^C dla opcji drugiej będzie wybierany ze zbioru elementów o odpowiednio mniejszych wartościach niż dla opcji pierwszej.

Dowód twierdzenia 4. Oczywiście dla relacji przechodniej w skończonym zbiorze istnieje co najmniej jeden element największy. Załóżmy najpierw, że istnieje dokładnie jeden, i , bez utraty ogólności $i = 1$. Zatem $P(NB_j \geq NB_1) < 50\%$ i $P(NB_1 \geq NB_j) > 50\%$, dla $j > 1$. Oznacza to, że jako jedyny będzie miał CEAC^C $> 50\%$ i zostanie wybrany przez C^C . Na mocy rozumowania indukcyjnego (po

rozmiarze X) widać, że zostanie wybrany przez C^R i jako jedyny będzie miał $CEAC^R > 50\%$. Dla przypadku więcej niż jednego elementu największego, b.u.o. założmy, że to elementy $1, \dots, k, k \leq n$. Dla $i \leq k < j$ $P(NB_j \geq NB_i) < 50\%$ i $P(NB_i \geq NB_j) > 50\%$, a dla $i, j \leq k$, $P(NB_i \geq NB_j) \geq 50\%$. Oznacza to, że C^C może wybrać jedynie któryś z $1, \dots, k$ i oczywiście $CEAC^C \geq 50\%$. Na mocy indukcji (po rozmiarze zbioru) C^R może wybrać jedynie któryś z elementów $1, \dots, k$ i $CEAC^R \geq 50\%$.

Literatura

- 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”, Vol. 11, Iss. 5.
- Bleichrodt H., Wakker P., Johannesson M. (1997), *Characterizing QALYs by Risk Neutrality*, „Journal of Risk and Uncertainty”, Vol. 15.
- Briggs A., Fenn P. (1998), *Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane*, „Health Economics”, Vol. 7, Iss. 8.
- 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”, Vol. 13.
- Eckermann S., Willan A.R. (2011), *Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies*, „Pharmacoeconomics”, Vol. 29, Iss. 7.
- Fenwick E., O'Brien B., Briggs A. (2004), *Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions*, „Health Economics”, Vol. 13.
- Fenwick E., Claxton K., Sculpher M. (2001), *Representing uncertainty: the role of cost-effectiveness acceptability curves*, „Health Economics”, Vol. 10, Iss. 8.
- Goeree R., O'Brien B.J., Blackhouse G., Marshall J., Briggs A., Lad R. (2002), *Cost effectiveness and cost-utility of long-term management strategies for heartburn*, „Value in Health”, Vol. 5, Iss. 4.
- Gold M.R., Siegel J.E., Russel L.B. i in. (1996), *Cost-Effectiveness in Health and Medicine*, Oxford University Press, New York.
- Hout B. van, Al M., Gordon G., Rutten F. (1994), *Costs, effects and C:E-ratios alongside a clinical trial*, „Health Economics”, Vol. 3.
- Jakubczyk M., Kamiński B. (2010), *Cost-Effectiveness Acceptability Curves – Caveats Quantified*, „Health Economics”, Vol. 19.
- Löthgren M., Zethraeus N. (2000), *Definition, interpretation and calculation of cost-effectiveness acceptability curves*, „Health Economics”, Vol. 9.

- Sadatsafavi M., Najafzadeh M., Marra C. (2008), *Acceptability curves could be misleading when correlated strategies are compared*, „Medical Decision Making”, Vol. 28, Iss. 3.
- Sen A. (1971), *Choice Functions and Revealed Preference*, „The Review of Economic Studies”, Vol. 38, Iss. 3.
- Ustawa z dnia 12 maja 2011 r. o refundacji leków, środków spożywczych specjalnego przeznaczenia żywieniowego oraz wyrobów medycznych, <http://isap.sejm.gov.pl/DetailsServlet?id=WDU20111220696> (dostęp: 1.10.2014).

RECURSIVE AND CONSERVATIVE DECISION RULES IN CHOICE OF THE HEALTH TECHNOLOGY UNDER RISK

Summary: In real-life decision problems evaluations of decision alternatives are estimated. Decision makers typically want the uncertainty to be visualised in an intuitive way. Presence of multiple criteria complicates the situation further. In health technology assessment cost-effectiveness acceptability curves (CEAC) are often used for sensitivity analysis. The goal of this study is to modify the way CEAC are used, so as to remove non-intuitive properties reported in the literature (while preserving their usefulness in sensitivity analysis). I suggest restricting CEAC to specifically selected pair-wise comparisons, in a recursive and a conservative approach. I show that both approaches remove problems typical for regular CEAC when uncertainty is normally distributed (e.g., CEACs become robust against alternative cloning and modifying correlations; guarantee α property; agree with first-order stochastic dominance). For general distributions the properties of CEACs are also somewhat improved.

Keywords: health technology assessment, cost-effectiveness acceptability curves, choice function, α property, first-order stochastic dominance.