

Magdalena Władysiuk, Daria Szmurło, Piotr Wojciechowski

HTA Consulting, Kraków

Rola surogatów w ocenie skuteczności i bezpieczeństwa technologii medycznych w onkologii — analiza rekomendacji Agencji Oceny Technologii Medycznych w Polsce

Surrogate endpoints in assessment of clinical efficacy and safety of oncology drugs — analysis of AHTAPol recommendations

Adres do korespondencji:

Lek. med. Magdalena Władysiuk
HTA Consulting
ul. Starowiślna 17/3, 31-038 Kraków
Tel.: +48 (12) 421 88 32
Faks: +48 (12) 395 38 32
e-mail: m.władysiuk@hta.pl

STRESZCZENIE

Terapie stosowane w onkologii oceniane są pod względem skuteczności i bezpieczeństwa zarówno na etapie decyzji terapeutycznych podejmowanych przez lekarzy, jak i na etapie procesów refundacyjnych (m.in. przez Agencję Oceny Technologii Medycznych). Oceny te polegają głównie na analizie korzyści mierzonych za pomocą punktów końcowych dotyczących progresji (*progression free survival* — PFS, *time to progression* — TTP) oraz przeżywalności (*overall survival* — OS).

Decyzje kliniczne w onkologii w codziennej praktyce podejmuje się najczęściej na podstawie zastępczych punktów końcowych, tzw. surogatów (PFS, TTP, odpowiedź na leczenie, częściowa odpowiedź na leczenie itp.), w odróżnieniu od istotnych dla pacjenta punktów końcowych (OS). Wynika to z faktu, iż ocena przeżywalności często wiąże się z koniecznością przeprowadzenia wieloletnich badań w relatywnie dużych grupach pacjentów, co nie zawsze jest praktycznie możliwe. Wobec powyższego pojawia się pytanie, czy i jak bardzo surogaty są skorelowane z punktami takimi jak przeżywalność. Rozważane punkty końcowe mają różne znaczenie kliniczne, różną istotność dla pacjentów, a także w różnym stopniu przyczyniają się do decyzji o finansowaniu terapii ze środków publicznych.

Celem niniejszej pracy było przedstawienie podstawowych definicji związanych z najczęściej stosowanymi punktami końcowymi w onkologii, wskazanie podobieństw i różnic między nimi, które mogą potencjalnie wpływać na wnioskowanie, a także przeanalizowanie wpływu tych punktów końcowych na rekomendacje dotyczące finansowania terapii onkologicznych ze środków publicznych.

Słowa kluczowe: onkologia, pierwszorzędkowy punkt końcowy, zastępczy punkt końcowy, surogat, przeżycie całkowite, przeżycie wolne od progresji, czas do progresji, OS, PFS, TTP

ABSTRACT

Treatments in oncology are assessed in terms of clinical efficacy and safety at the stage of therapeutic decisions made by oncologists, but also during reimbursement processes (by Health Technology Assessment bodies). Those assessments are based on analysis of clinical benefits measured by progression-related endpoints (*progression free survival* — PFS, *time to progression* — TTP) or by survival (*overall survival* — OS).

Therapeutic decisions in clinical practice are usually based on so called surrogate endpoints (PFS, TTP, response rate, partial response rate), as opposed to „hard” endpoints like survival (OS). This results from a fact that it is much harder to prove survival benefits in clinical trials, because it usually requires long-term follow-up studies and large number of enrolled patients. This raises a question of whether surrogates are correlated with hard clinical endpoints,

and how strong the correlation is. All those endpoints have different clinical meaning, different significance for patients and also different impact on reimbursement decisions.

The aim of this paper was to present definitions of the most commonly used oncology endpoints, show differences and similarities between them and to analyze what is their impact on reimbursement recommendations made by Polish Agency for Health Technology Assessment.

Key words: oncology, primary endpoint, surrogate endpoint, overall survival, progression-free survival, time to progression, OS, PFS, TTP

Onkol. Prak. Klin. 2014; 10, 3: 138–145

Wstęp

W obecnych czasach w różnych bazach informacji medycznej codziennie publikowanych jest ponad tysiąc badań klinicznych dotyczących szerokiego spektrum problemów kliniczno-epidemiologicznych, reprezentujących mnogość założeń metodycznych, a także różne sposoby oceny i prezentacji uzyskanych wyników. Liczba parametrów stosowanych do oceny efektywności klinicznej technologii z dziedziny onkologii jest stosunkowo niewielka, a ich definicja posiada ugruntowane znaczenie, przez co są one szeroko wykorzystywane w procesach rejestracji, refundacji, jak również w ramach rzeczywistej praktyki klinicznej. Nie oznacza to jednak, iż sposób oceny skuteczności i bezpieczeństwa leków onkologicznych jest zunifikowany. Niestety — również i w tym obszarze klinicznym zwraca się uwagę na niepokojące zjawisko przyjmowania odmiennych definicji ugruntowanych już punktów końcowych. Sytuacja ta niejednokrotnie uniemożliwia wzajemne porównanie efektywności interwencji ocenianych w dwóch różnych pracach, a w skrajnych przypadkach może prowadzić do błędnych wniosków i nieoptymalnych wyborów klinicznych. Stąd zarówno jednostki odpowiedzialne za rejestrację produktów leczniczych (amerykańska *Food and Drug Administration* — FDA, czy europejska *European Medicines Agency* — EMA), jak i różnego rodzaju towarzystwa naukowe, stowarzyszenia eksperckie i grupy specjalistów podniosły kwestię konieczności ujednolicenia definicji i standaryzacji pomiarów wykorzystywanych w ocenie efektywności leków onkologicznych [1–5].

Właściwa interpretacja pojawiających się dowodów naukowych wymaga od odbiorcy znajomości podstawowych zagadnień z zakresu metodyki, używanych terminów oraz przyjętych definicji mających zastosowanie w projektowaniu, przeprowadzaniu i raportowaniu wyników badań klinicznych. Podstawowym terminem niezbędnym do właściwego zrozumienia wyników badania klinicznego jest punkt końcowy (*endpoint*), określający mierzalną zmianę stanu zdrowia uczestnika badania, której częstość występowania lub nasilenie może wiązać się z ekspozycją na określony czynnik lub badaną interwencją [6].

Punkty końcowe definiowane w badaniach klinicznych mogą być zarówno pozytywne, gdy wiążą się z poprawą stanu zdrowia na skutek stosowanej terapii, jak i negatywne, w sytuacji gdy ocenia się np. potencjalne szkodliwe działanie leku [6]. Właściwie raportowane badanie powinno zawierać informacje na temat tzw. pierwszorzędowego punktu końcowego (*primary endpoint*), czyli efektu klinicznego, zdefiniowanego na etapie projektowania badania, którego ocena będzie służyć jako podstawa wnioskowania o efektywności porównywanych technologii. To na podstawie informacji odnoszących się do pierwszorzędowego punktu końcowego przeprowadza się kalkulację minimalnej wielkości badanej próby, która z dużym prawdopodobieństwem pozwoli na statystyczne potwierdzenie założeń badania. Korzystna dla badania jest sytuacja, gdy pierwszorzędowy punkt końcowy zdefiniowany zostanie na podstawie zdarzeń klinicznie istotnych, czyli takich, które wpływają w sposób bezpośredni na stan zdrowia lub samopoczucie pacjenta, a ich wystąpienie lub utrata są odczuwalne i raportowane przez uczestnika bez konieczności przeprowadzania dodatkowych pomiarów. Przykładowo, klinicznie istotnymi punktami końcowymi są objawy wynikające z progresji choroby, zawał serca czy zgon pacjenta. Zmiana poziomu parametrów laboratoryjnych czy też bezobjawowe stany kliniczne pozostające bez wpływu na samopoczucie pacjenta określa się mianem zastępczych punktów końcowych (*surrogate endpoints*). Należy jednak zaznaczyć, że w momencie gdy znacząca zmiana nasilenia zastępczego punktu końcowego prowadzi do wystąpienia objawów, uzyskuje ona wówczas istotność kliniczną i staje się automatycznie klinicznym punktem końcowym. Sytuacja taka może dotyczyć np. spadku stężenia glukozy we krwi na skutek stosowania insuliny, który stanowi efekt pożądaný terapii przeciw cukrzycowej. Nieadekwatna dawka leku może jednak prowadzić do zbyt dużego obniżenia stężenia glukozy, co z kolei doprowadzi do łagodnych i nieobjawowych hipoglikemii (zastępczy punkt końcowy), ale również do stanów ciężkich, manifestujących się osłabieniem lub omdleniami, a nawet zdarzeń stanowiących bezpośrednie zagrożenie życia (klinicznie istotny punkt końcowy).

W niektórych obszarach terapeutycznych, w tym również w onkologii, wykorzystanie pierwszorzędowych punktów końcowych w ocenie skuteczności badanych interwencji nie jest jednak możliwe. Dotyczy to zwłaszcza sytuacji, w których do ocenianego zdarzenia dochodzi stosunkowo rzadko lub jego wystąpienie oczekiwane jest w dość odległym czasie, jak np. zgon pacjenta. Wykazanie różnic pomiędzy interwencjami wymaga w takim przypadku zebrania informacji o stosunkowo dużej liczbie zdarzeń, co w przypadku rzadkich punktów końcowych wiąże się z koniecznością przeprowadzenia długoterminowego badania klinicznego z dużą liczbą uczestników. W niektórych sytuacjach uzasadniona jest więc ocena efektywności klinicznej na podstawie zastępczych punktów końcowych (surogatów) — zdarzeń lub określonych wyników pomiarów laboratoryjnych, badań obrazowych, będących udokumentowanymi predyktorami wystąpienia klinicznie istotnych zdarzeń [7, 8].

Surogaty są stosowane w codziennej praktyce klinicznej lub w badaniach klinicznych w celu pomiaru parametrów fizjologicznych, bez równoczesnej oceny ich wpływu na jakość życia pacjenta, jego samopoczucie czy funkcjonowanie. Ze względu na liczne dowody wskazujące na związek wielu ocenianych w ramach testów laboratoryjnych badań diagnostycznych z klinicznie istotnymi punktami końcowymi są one podstawą podejmowania decyzji klinicznych. Jednocześnie jednak w wielu przypadkach dowody na istnienie zależności pomiędzy surogatami a klinicznie istotnymi zdarzeniami pochodzą z badań o niskiej wiarygodności, które nie podlegały późniejszej rewizji pod względem prawidłowości zastosowanych metod statystycznych.

Większość badań klinicznych, w tym również prób w dziedzinie onkologii, projektowana jest w celu wykazania przewagi (*superiority*) ocenianej interwencji nad obecnie stosowanymi terapiami lub brakiem leczenia. Badania te mają za zadanie wykazać, że nowa technologia pozwala na uzyskanie dodatkowego efektu zdrowotnego w ściśle zdefiniowanej populacji pacjentów, odpowiadającej często późniejszemu wskazaniu rejestracyjnemu. Potwierdzenie skuteczności technologii stanowi jeden z elementów procesu wprowadzania nowych terapii i zakłada przeprowadzenie szeregu badań porównawczych z wykorzystaniem jednej lub kilku grup referencyjnych, na podstawie których dokonuje się pomiaru częstości wystąpienia lub zmiany nasilenia danego punktu końcowego w czasie obserwacji. W ramach badań przeprowadzanych w obszarze onkologii zasadnicze znaczenie ma analiza przeżywalności, która pozwala na określenie prawdopodobieństwa wystąpienia ocenianego zdarzenia (punktu końcowego, np. zgonu, progresji) w trakcie procesu terapeutycznego. Analiza przeżywalności pozwala również na określenie przeciętnej przewidywanej długości życia lub czasu do wystąpienia zmiany stanu zdrowia.

Podstawowe definicje punktów końcowych w onkologii

Badania kliniczne dostarczają dowodów naukowych dotyczących efektywności klinicznej oraz bezpieczeństwa, które umożliwiają rejestrację oraz dopuszczenie technologii medycznej do obrotu, jak również — w późniejszym okresie — wykorzystywane są w procesie podejmowania decyzji refundacyjnych. W raportach oceny technologii medycznych (*Health Technology Assessment*) przeprowadza się analizę oraz statystyczną ocenę punktów końcowych raportowanych w opublikowanych badaniach klinicznych, przy uwzględnieniu specyfiki populacji docelowej.

W dziedzinie onkologii efektywność kliniczną nowych interwencji zarówno w badaniach eksperymentalnych, jak i obserwacyjnych ocenia się najczęściej, opierając się na takich punktach końcowych, jak przeżywalność całkowita (OS, *overall survival*) oraz przeżycie wolne od progresji (PFS, *progression-free survival*). Przeżywalność w badaniach klinicznych zazwyczaj zdefiniowana jest jako odsetek pacjentów pozostających przy życiu w trakcie trwania badania. Z kolei OS definiuje się jako czas, który upłynął od randomizacji do wystąpienia zgonu, i ocenia się ją zazwyczaj za pomocą tzw. krzywych przeżycia. Metody analizy przeżycia pozwalają na estymację prawdopodobieństwa przeżycia lub zgonu w poszczególnych punktach czasowych, przy uwzględnieniu danych dotyczących pacjentów, którzy z różnych przyczyn przedwcześnie zaprzestali udziału w badaniu. Kluczowym elementem w konstrukcji krzywych przeżycia jest więc możliwie precyzyjne określenie czasu, który upłynął od momentu randomizacji do momentu wystąpienia zgonu lub innego ocenianego punktu końcowego.

Przeżycie wolne od progresji (PFS) jest drugim najczęściej wykorzystywanym punktem końcowym w ocenie efektów zdrowotnych terapii onkologicznych i zdefiniowane jest zazwyczaj jako czas pomiędzy randomizacją oraz wystąpieniem złożonego punktu końcowego, obejmującego progresję choroby lub zgon. Dodatkowo w badaniach onkologicznych ocenia się niekiedy czas do wystąpienia progresji (TTP, *time to progression*), oceniany na podstawie prawdopodobieństwa wystąpienia progresji choroby w poszczególnych momentach okresu obserwacji.

Przeżycie całkowite (OS) bezwzględnie jest punktem końcowym o wysokiej istotności — zarówno z punktu widzenia pacjenta, jak i z perspektywy klinicysty. Z kolei definicja PFS obejmuje progresję nowotworu, która nie zawsze wpływa w sposób bezpośredni na stan zdrowia i samopoczucia chorego, przez co stanowi czasami jedynie zastępczy punkt końcowy.

Na podstawie powyższych informacji wydawać by się mogło, iż OS — jako punkt końcowy o wysokiej istotności klinicznej oraz istotny z perspektywy pacjentów

Tabela 1. Podział punktów końcowych

Istotne dla pacjenta punkty końcowe	Zastępcze punkty końcowe
Zgon	Progresja zgodna z definicją RECIST
Progresja objawowa	Próby wątrobowe
Jakość życia	Zmiana wielkości guza w badaniu obrazowym
Złamania patologiczne	

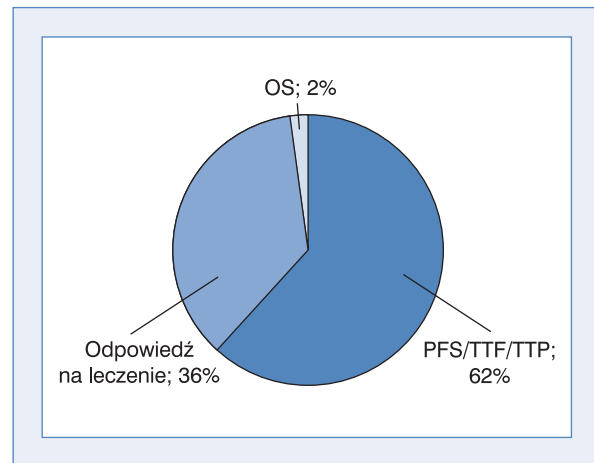
RECIST — *Response Evaluation Criteria In Solid Tumors*

— będzie stanowić „złoty standard” w ocenie skuteczności interwencji onkologicznych. Okazuje się jednak, że w większości badań eksperymentalnych w dziedzinie onkologii wnioskowanie opiera się na punktach końcowych o niższej istotności, tj. PFS, TTP i innych. Wybór pierwszorzędowego punktu końcowego różni się też w zależności od typu nowotworu i metodologii badania klinicznego (tab. 1).

W 2009 r. Saad i wsp. opublikowali przegląd systematyczny, którego celem była ocena częstości stosowania różnych pierwszorzędowych punktów końcowych w badaniach dotyczących zaawansowanego raka piersi (ryc. 1) [9]. W pracy tej zidentyfikowano ogółem 58 randomizowanych badań klinicznych przeprowadzonych w latach 2000–2007. Tylko w jednym z nich pierwszorzędowym punktem końcowym, na podstawie którego skalkulowano wielkość próby badania, było OS. We wszystkich pozostałych badaniach wnioskowanie przeprowadzono, opierając się na punktach końcowych o niższej istotności oraz zastępczych punktach końcowych (m.in. PFS, TTP, TTF — *time to treatment failure*). Istotnym problemem w zidentyfikowanych badaniach były różnice definicji przyjętych punktów końcowych. W 21 pracach pierwszorzędowym punktem końcowym był TTP, przy czym w 5 nie podano jego definicji. Z kolei 13 badań oceniających TTP w definicji punktu końcowego uwzględniło również wystąpienie zgonu, przez co przyjęta definicja w istocie odpowiadała PFS (por. tab. 2). Wyniki przeglądu Saad i wsp. (2009) wskazują na dużą dowolność w definiowaniu TTP lub wręcz zamienne stosowanie go z PFS, co utrudnia interpretację wyników prac i znacząco obniża ich walor porównawczy. Przykładowo, w jednym badaniu raportowano zarówno PFS i TTP, przy czym w definicji obu punktów końcowych ujęto zarówno progresję choroby, jak i zgony, a jedyna różnica polegała na tym, że w przypadku TTP uwzględniano zgony z powodu raka piersi, a w przypadku PFS — zgony bez względu na przyczynę [9].

Wyniki przeglądu Saad i wsp. wskazują, że definicje punktów końcowych w badaniach klinicznych odbiegają od ustalonych standardów, w wyniku czego są one często stosowane i raportowane w sposób niewłaściwy.

Zgodnie z *Response Evaluation Criteria In Solid Tumors* (RECIST) 1.1 definicja progresji choroby nowotworowej obejmuje zwiększenie wymiarów guza lub sumy największych wymiarów guzów o przynajmniej



Rycina 1. Pierwszorzędowe punkty końcowe w badaniach randomizowanych dla rozlanego raka piersi (na podstawie [9]). OS (*overall survival*) — przeżycie całkowite; PFS (*progression-free survival*) — przeżycie wolne od progresji; TTF (*time to treatment failure*) — czas do niepowodzenia leczenia; TTP (*time to progression*) — czas do wystąpienia progresji

20% (bezwzględny wzrost sumy wymiarów o min. 5 mm) i/lub pojawienie się nowych zmian [10]. Pozostałe definicje dla określonych punktów końcowych z uwzględnieniem innych, często stosowanych zastępczych punktów końcowych, tj. TTP, TTF, ORR (*objective response rate*), podano w tabeli 2 [11].

W związku z oceną progresji pojawiają się liczne wątpliwości dotyczące kwestii związanych ze stanem zaawansowania choroby w momencie włączenia pacjenta do badania klinicznego, typu terapii lub momentu ich rozpoczęcia w zależności od rozwoju choroby. Z tego powodu czas do progresji (TTP) jest czasami przedstawiany obok PFS, gdyż ten pierwszy nie uwzględnia wszystkich zgonów w danym okresie obserwacji. Często w TTP uwzględniane są zgony z powodu raka, traktowane jako progresja choroby, nie uwzględnia się natomiast zgonów z innych przyczyn.

Relacja między przeżyciem (OS) a punktami mierzącymi progresję choroby (PFS/TTP) zależy od typu nowotworu, stanu zaawansowania nowotworu oraz linii leczenia. Przy próbie oceny takiej zależności pojawia się kilka zasadniczych pytań: po pierwsze — czy możliwe

Tabela 2. Punkty końcowe w onkologii — definicje, użyteczność (opracowanie własne na podstawie [4])

Typ punktu końcowego	Definicja	Surogat/istotny dla pacjenta punkt końcowy	Korzyści	Ograniczenia	Przykład rejestracji w Food and Drug Administration
Przeżycie całkowite (OS, overall survival)	Czas od momentu randomizacji do momentu zgonu pacjenta	istotny dla pacjenta punkt końcowy	<ul style="list-style-type: none"> • Łatwo i obiektywnie mierzalny punkt końcowy • Wysoka precyzja wyników 	<ul style="list-style-type: none"> • Wymaga relatywnie dużej próby w badaniu • Relatywnie długi okres obserwacji w zależności od typu nowotworu lub linii leczenia • Trudno mierzalny w przypadku kolejnych linii leczenia czy badania typu cross-over 	<ul style="list-style-type: none"> • Bevacizumab (rozsziany rak jelita grubego) • Capecytabine (rak jelita grubego) • Irinotecan (rak jelita grubego) • Leucovorin (rak jelita grubego)
Czas do progresji [TTP, time to (tumor) progression]	Punkt zastępczy, definiowany jako czas od momentu randomizacji do momentu obiektywnej progresji guza	Surogat	<ul style="list-style-type: none"> • Ocena jedynie progresji choroby, uważanej za lepiej prognozujący OS niż PFS • Mała wielkość populacji i krótszy okres obserwacji w porównaniu z OS • Nie jest zaburzony w badaniach typu cross-over • Obiektywna metoda oceny wielkości guza 	<ul style="list-style-type: none"> • Brak dowodów w przypadku większości nowotworów na związek pomiędzy PFS czy TTP z OS • Brak oceny stanu pacjenta • Błędy w definicjach w badaniach • Częstość oceny zmian wpływa na pomiary 	<ul style="list-style-type: none"> • Anastrozole (rak piersi, 2. linia) • Anastrozole (rak piersi, 1. linia) • Lapatinib (rozsziany rak piersi HER2-dodatni) • Topotecan (rak jajnika, 2. linia) • Fulvestrant (rak piersi, 2. linia) • Sunitinib (rak podścieliska przewodu pokarmowego) • Trastuzumab (rozsziany rak piersi, 1. linia)
Przeżycie wolne od progresji (PFS, progression-free survival)	Punkt zastępczy, definiowany jako czas od momentu randomizacji do momentu obiektywnej progresji guza lub zgonu pacjenta (z dowolnej przyczyny)	Surogat			<ul style="list-style-type: none"> • Alemtuzumab (przewlekła białaczka limfatyczna) • Bevacizumab (rozsziany rak piersi) • Bevacizumab (rozsziany rak nerki) • Gemcitabine (zaawansowany rak jajnika, po nawrocie) • Ixabepilone (rozsziany lub miejscowo zaawansowany rak piersi) • Panitumumab (zaawansowany rak jelita grubego) • Sorafenib (zaawansowany rak nerki) • Sunitinib (zaawansowany rak nerki)

Czas do niepowodzenia leczenia (TTF, <i>time to treatment failure</i>)	Złożony punkt końcowy, definiowany jako czas od momentu randomizacji do zaprzestania leczenia (zgon, progresja, działania niepożądane, inne przyczyny wycofania się pacjenta z badania)	Surogat	Uwzględnia rzeczywisty czas leczenia pacjenta	Niezalecany w ocenie skuteczności leczenia, gdyż uwzględnia także czynniki inne niż skuteczność terapii. Czynniki te mogą być zależne np. od protokołu badania
Obiektywna odpowiedź (ORR, <i>objective response rate</i>)	Suma odpowiedzi częściowej (<i>partial response</i>) oraz całkowitej (CRs) zgodnie z kryteriami RECIST (<i>Response Evaluation Criteria in Solid Tumors</i>)	Surogat	<ul style="list-style-type: none"> Może być oceniany w badaniach bez grupy kontrolnej Relatywnie mniejsza próba niż w badaniach dotyczących oceny OS Punkt końcowy bezpośrednio związany ze stosowaniem lekiem 	<ul style="list-style-type: none"> Liczne błędy w definicjach Ocena zależna od czasu obserwacji
				<ul style="list-style-type: none"> Altrektamine (oporny rak jajnika) Anastrozole (rak piersi, 1. i 2. linia) Capecytabine (oporny rak piersi) Exemestane (rak piersi, 2. linia) Imatinib (rak podścieliska przewodu pokarmowego) Irinotecan (rak jelita grubego, 2. linia) Letrozole (rak piersi, 1. i 2. linia)

jest wykazanie związku pomiędzy surogatami a istotnymi dla pacjenta punktami końcowymi dla tej konkretnej technologii; po drugie — czy inne terapie stosowane w danym wskazaniu wykazały ten związek i jak jest on silny; po trzecie zaś — czy jest możliwe zebranie dodatkowych danych dotyczących tych punktów końcowych w kontekście stawianego pytania. Odpowiedź na te pytania nie jest jednoznaczna — dla niektórych typów nowotworów wykazano korelację zastępczych punktów końcowych (m.in. PFS) z twardymi punktami końcowymi (OS), np. przerzutowy rak jelita grubego, niedrobnokomórkowy rak płuc [12, 13], w przypadku innych zaś nie stwierdzono takiej zależności, np. rak piersi (z ograniczeniami w zależności od typu metod analizy danych oraz poziomu wiarygodności danych) [13].

Analiza rekomendacji AOTM dotyczących technologii onkologicznych

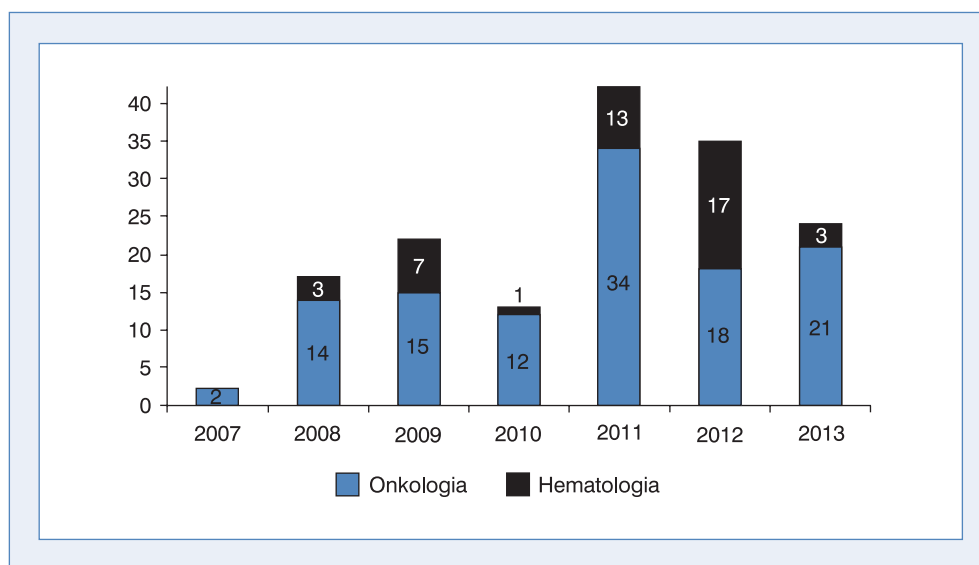
Agencja Oceny Technologii Medycznych (AOTM), działając na zlecenie Ministra Zdrowia, ocenia leki i technologie nielekowe i wydaje rekomendacje dotyczące finansowania tych technologii ze środków publicznych. Proces podejmowania decyzji przez AOTM obejmuje ocenę szeregu aspektów dotyczących nie tylko udokumentowanej badaniami klinicznymi efektywności terapii, ale także względów ekonomicznych, warunkujących opłacalność technologii.

Od stycznia 2007 do września 2013 r. AOTM wydała łącznie 160 rekomendacji dotyczących 116 leków i świadczeń nielekowych we wskazaniach onkologicznych oraz 44 we wskazaniach hematologicznych, 37% tych ocen odbyło się po wejściu w życie ustawy refundacyjnej (2012–2013) (ryc. 2) [14].

Sto sześćdziesiąt ocen dokonanych przez AOTM obejmowało łącznie 167 wskazań (kilka ocen zawierało więcej niż jedno wskazanie). W latach 2007–2011 AOTM najczęściej oceniała technologie wskazane w leczeniu raka nerki (20%), raka płuc i oskrzeli (17%) oraz raka piersi (13%), a w latach 2012–2013 — w raku gruczołu krokowego (12%) oraz raku piersi i raku skóry (po 7%).

Spośród 101 rekomendacji (lub ewentualnie stanowisk) wydanych w latach 2007–2011 w 46 odwoływano się do danych zarówno o przeżyciu (OS), jak i progresji choroby (PFS), w 17% wyłącznie do OS, w 8% wyłącznie do PFS, a w pozostałych 30% — do żadnego z nich. W latach 2012–2013 przeważały rekomendacje wydane na podstawie zarówno OS, jak i PFS (63%), dane tylko o OS podawano w 20% rekomendacji, tylko o PFS — w 2%, a pozostałe 20% wydano na podstawie innych punktów końcowych (tab. 3).

Negatywne rekomendacje argumentowano niską efektywnością terapii, niekorzystnym profilem bezpieczeństwa, ograniczoną wiarygodnością dowodów lub ich



Rycina 2. Liczba ocen dokonanych przez Agencję Oceny Technologii Medycznych (AOTM) dla technologii onkologicznych i hematologicznych według daty wydania rekomendacji (niektóre oceny kończą się wydaniem stanowiska, a nie rekomendacji)

Tabela 3. Rekomendacje Agencji Oceny Technologii Medycznych (AOTM) według dostępności danych klinicznych i typów rekomendacji (opracowanie własne)

Dostępne dane kliniczne	Finansować		Warunkowo finansować		Nie finansować		Razem
2007–2011							
OS i PFS	17	53%	17	44%	12	40%	46
Tylko OS	4	13%	6	15%	7	23%	17
Tylko PFS	1	3%	2	5%	5	17%	8
Brak OS i PFS	10	31%	14	36%	6	20%	30
Razem	32	100%	39	100%	30	100%	101
2012–2013*							
OS i PFS	12	80%	13	50%	11	65%	36
Tylko OS	0	0%	7	27%	2	12%	9
Tylko PFS	0	0%	1	4%	0	0%	1
Brak OS i PFS	3	20%	5	19%	4	24%	12
Razem	15	100%	26	100%	17	100%	58

*W jednej ocenie dokonanej w okresie 2012–2013 wydano częściowo pozytywną i częściowo negatywną rekomendację (dla wybranych subpopulacji), rekomendacja ta nie została uwzględniona w zestawieniu

OS (*overall survival*) — przeżycie całkowite; PFS (*progression-free survival*) — przeżycie wolne od progresji

brakiem, zbyt wysokimi kosztami oraz brakiem danych w zakresie efektywności kosztowej. W okresie zarówno przed, jak i po wejściu w życie ustawy refundacyjnej odpowiednio 73% i 76% rekomendacji negatywnych wydano ze względu m.in. na aspekty kliniczne (niska skuteczność, niska wiarygodność danych dotyczących skuteczności). Rekomendacje warunkowe zalecały najczęściej obniżenie ceny lub zmianę warunków realizacji ocenianego świadczenia.

Analiza rekomendacji wskazuje, że przede wszystkim pozytywne rekomendacje wydawano częściej w sytuacjach, w których technologia w sposób istotny statystycznie przedłużała OS lub PFS (tab. 4).

Zaobserwowano, że stopień, w jakim badana interwencja wydłuża OS, stanowi predyktor uzyskania odpowiedzi pozytywnej rekomendacji refundacyjnej przez AOTM. Wszystkie technologie przedłużające OS o więcej niż rok uzyskały rekomendację pozytywną, technologie przedłużające OS o 4–12 miesięcy otrzymały pozytywną rekomendację w 80% przypadków, a technologie przedłużające OS o nie więcej niż 4 miesiące — w 57% przypadków (tab. 5). Jednocześnie jednak pozytywne rekomendacje AOTM, w tym rekomendacje warunkowe, wydano również w przypadku 61% technologii, które nie prowadziły do istotnego przedłużenia OS, a także w przypadku 58% technologii nieprzedłużających PFS.

Tabela 4. Rekomendacje Agencji Oceny Technologii Medycznych (AOTM) według istotności wyników klinicznych

Typ rekomendacji	OS*		PFS*	
	Istotne przedłużenie	Nieistotna zmiana	Istotne przedłużenie	Nieistotna zmiana
Finansować (w tym: warunkowo)	27 (77%)	19 (61%)	41 (73%)	7 (58%)
Nie finansować	8 (23%)	12 (39%)	15 (27%)	5 (42%)
Razem	35 (100%)	31 (100%)	56 (100%)	12 (100%)

*Dane jakościowe dotyczące istotności OS podano w 66 na 108 rekomendacji, a istotności PFS — w 68 na 91 rekomendacji
OS (overall survival) — przeżycie całkowite; PFS (progression-free survival) — przeżycie wolne od progresji

Tabela 5. Liczba rekomendacji w zależności od długości istotnego przedłużenia przeżycia całkowitego (OS)

Typ rekomendacji	Istotne przedłużenie OS (miesiące)*		
	0–4	4–12	> 12
Finansować (w tym: warunkowo)	8 (57%)	8 (80%)	4 (100%)
Nie finansować	6 (43%)	2 (20%)	0 (0%)
Razem	14 (100%)	10 (100%)	4 (100%)

*Dane ilościowe dotyczące OS podano w 28 rekomendacjach z 35, w których stwierdzono istotne przedłużenie OS

Podsumowanie

W ramach przeprowadzonego przeglądu literatury odnaleziono tylko kilka publikacji, które nie oddają całego obrazu badań przeprowadzanych w onkologii oraz nie pozwalają na jednoznaczne wskazanie roli surogatów w ocenie skuteczności i bezpieczeństwa technologii medycznych. Istotny jest fakt, że zarówno klinicyści, jak i decydenci — w tym instytucje publiczne, takie jak urzędy rejestracji oraz agencje oceny technologii medycznych (HTA, *Health Technology Assessment*) — często muszą podejmować decyzje na podstawie zastępczych punktów końcowych. Analiza rekomendacji AOTM wskazuje, że na 159 rekomendacji z lat 2007–2013 ponad 30% nie pozwalało wnioskować, opierając się na OS, lecz na innych punktach końcowych — głównie zastępczych. Jednocześnie porównanie tych statystyk z danymi z FDA z lat 1999–2002 wskazuje, że w 68% przypadków decyzja o dopuszczeniu do obrotu została wydana na podstawie innych punktów końcowych niż OS [15], co pozwala sądzić, iż regulacje i kryteria dotyczące refundacji są bardziej restrykcyjne niż kryteria rejestracyjne. Jednocześnie niektóre opracowania na świecie wskazują na brak spójności w kwestii akceptowalności surogatów pomiędzy urzędami rejestracji czy agencjami HTA [16]. Wyniki analizy rekomendacji HTA w Polsce wskazują na konieczność wypracowania wytycznych czy konsensusu w aspekcie stosowania, walidacji oraz akceptowalności zastępczych punktów końcowych w kontekście raportów oceny technologii medycznych.

Źródła finansowania — brak konfliktu interesów

Piśmiennictwo

- Allegra C., Blanke C., Buyse M. i wsp. End points in advanced colon cancer clinical trials: a review and proposal. *J. Clin. Oncol.* 2007; 25: 3572–3575.
- Hudis C.A., Barlow W.E., Costantino J.P. i wsp. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J. Clin. Oncol.* 2007; 25: 2127–2132.
- Punt C.J., Buyse M., Kohne C.H. i wsp. Endpoints in adjuvant treatment trials: a systematic review of the literature in colon cancer and proposed definitions for future trials. *J. Natl. Cancer Inst.* 2007; 99: 998–1003.
- McCain J.A. The ongoing evolution of endpoints in oncology. *Managed Care* 2010; 19 (supl. 1).
- European Medicines Agency. ICH Topic E9 — Statistical Principles for Clinical Trials. Note for Guidance on Statistical Principles for Clinical Trials. www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002928.pdf.
- Jaeschke R., Niżankowski R. Ocena technologii medycznych dla ich finansowania ze środków publicznych zgodna z zasadami EBM. Propozycja polskich rozwiązań. *Med. Prakt.* 2004; 10: 160–166.
- Frank R., Hargreaves R. Clinical biomarkers in drug discovery and development. *Nat. Rev. Drug Discov.* 2003; 2: 566–580.
- Katz R. Biomarkers and surrogate markers: an FDA perspective. *NeuroRx* 2004; 1: 189–195.
- Saad E.D., Katz A. Progression-free survival and time to progression as primary end points in advanced breast cancer: often used, sometimes loosely defined. *Ann. Oncol.* 2009; 20: 460–464.
- Eisenhauer E.A., Therasse P., Bogaerts J. i wsp. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* 2009; 45: 228–247.
- Fleming T.R., DeMets D.L. Surrogate end points in clinical trials: Are we being misled? *Ann Intern Med.* 1996; 125: 605–613.
- Fleming T.R., Rothmann D.R. Issues in Using Progression-Free Survival When Evaluating Oncology Products. *J. Clin. Oncol.* 2009; 27: 2874–2880.
- Davis S., Tappenden P., Cantrell A. A review of studies examining the relationship between progression-free survival and overall survival in advanced or metastatic cancer. School of Health and Related Research, University of Sheffield 2012.
- Ustawa z dnia 12 maja 2011 r. o refundacji leków, środków spożywczych specjalnego przeznaczenia żywieniowego oraz wyrobów medycznych. *Dz.U.* 2011 nr 122 poz. 696.
- Johnson J.R., Williams G., Pazdur R. End points and United States Food and Drug Administration approval of oncology drugs. *J. Clin. Oncol.* 2003; 21: 1404–1411.
- Rocchi A., Khoudigian S., Hopkins R., Goeree R. Surrogate outcomes: experiences at the Common drug Review. *Cost effectiveness and Resource Allocation* 2013; 11: 31.

prof. dr. hab. n. med. Piotr J. Wysocki

Centrum Onkologii — Instytut im. Marii Skłodowskiej-Curie w Warszawie

Komentarz

do pracy *Rola surogatów w ocenie skuteczności i bezpieczeństwa technologii medycznych w onkologii — analiza rekomendacji Agencji Oceny Technologii Medycznych w Polsce*

Obserwowany od końca XX wieku gwałtowny wzrost liczby nowych leków onkologicznych, o potwierdzonej w badaniach klinicznych aktywności, stanowi ogromne wyzwanie dla instytucji zajmujących się ich rejestracją i finansowaniem. W latach 1995–2004 amerykańska Agencja ds. Żywności i Leków (FDA, *Food and Drug Administration*) zarejestrowała 57 leków onkologicznych, opierając się przede wszystkim na takich punktach końcowych, jak odsetek odpowiedzi klinicznych (26) oraz czas przeżycia całkowitego (18). Czas wolny od progresji (PFS, *progression-free survival*) lub czas wolny od choroby były podstawą rejestracji tylko w przypadku 3 leków [1]. Od połowy ubiegłej dekady obserwujemy coraz więcej rejestracji nowych leków (przede wszystkim ukierunkowanych molekularnie), które dokonały się na podstawie badań klinicznych niewykazujących znamiennych korzyści w zakresie czasu przeżycia całkowitego (OS, *overall survival*). W przypadku raka nerkowokomórkowego od 2006 roku zarejestrowano 7 leków ukierunkowanych molekularnie, spośród których tylko w przypadku jednego (temsrolimus) wykazano znamienne poprawę rokowania chorych w zakresie OS. Wszystkie pozostałe leki zostały zarejestrowane wyłącznie na podstawie badań klinicznych wykazujących znamienne korzyść z ich zastosowania jedynie w odniesieniu do PFS.

Kilka lat temu rozgorzała intensywna dyskusja dotycząca zasadności wykorzystania parametru PFS jako pierwszorzędowego punktu końcowego badań klinicznych nad nowymi lekami onkologicznymi. Nie podlega dyskusji, że OS jest najbardziej istotnym punktem końcowym, który jest obiektywny i łatwy do ustalenia. Jednak jego wiarygodna ocena wymaga zaangażowania dużych populacji chorych, jest bardzo czasochłonna oraz zaburzana przez procedurę *cross-over* (stosownie leku badanego w populacji chorych z ramienia kontrolnego) i kolejne linie leczenia po zakończeniu udziału chorego w badaniu. Z kolei PFS, jako pierwszorzędowy punkt końcowy, umożliwia wcześniejszą weryfikację skutecz-

ności terapii, co związane jest z krótszym czasem trwania badania oraz możliwością kwalifikowania mniejszych populacji chorych. Niestety, ocena PFS nie jest tak obiektywna jak ocena OS, bowiem w wielu przypadkach trudno jednoznacznie zdefiniować progresję (zmiany przerzutowe w kościach, inne zmiany niemierzalne), a poza tym stwierdzenie momentu progresji uzależnione jest od częstotliwości wykonywania badań kontrolnych. Dodatkowo, w przypadku większości nowotworów, nie potwierdzono istnienia korelacji PFS z OS. Pomimo wszystkich tych wątpliwości w 2007 roku FDA w swoim stanowisku „*Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics*” uznała, że PFS może być podstawą rejestracji leku, jeżeli korzyść w tym zakresie została wykazana w randomizowanym badaniu klinicznym, poddanym niezależnej ocenie. W 2009 roku przedstawiciele FDA, Narodowego Instytutu Raka (NCI, *National Cancer Institute*) oraz organizacji skupiającej firmy farmaceutyczne — PhARMA podczas *Progression Free Survival Oncology Workshop* wydali opinię, że w pewnych sytuacjach PFS może stanowić pierwszorzędowy punkt końcowy. Ekspertsi uznali jednak, że ewentualna decyzja rejestracyjna w stosunku do określonego leku musi być uzasadniona nie tylko znamiennej korzyścią w zakresie PFS, ale również istotnymi korzyściami klinicznymi. Ekspertsi uznali jednocześnie, że OS pozostaje bardzo ważnym drugorzędowym punktem końcowym, niezbędnym w absolutnej większości badań klinicznych, ponieważ świadczy nie tylko o aktywności nowego leku, ale i o jego bezpieczeństwie.

Dyskutując o znaczeniu PFS jako punktu końcowego, warunkującego rejestrację/refundację nowych technologii medycznych w onkologii, należy pamiętać, że wykazanie znamiennego wydłużenia czasu przeżycia całkowitego w badaniach klinicznych nad lekami stosowanymi w leczeniu paliatywnym u chorych onkologicznych staje się coraz mniej prawdopodobne. W badaniach klinicznych zdecydowaną większość nowych strategii

terapeutycznych ocenianych u chorych poddawanych leczeniu paliatywnemu stosuje się w ramach początkowych linii leczenia. Tym samym, po zakończeniu udziału w badaniu chorzy przez wiele miesięcy, a czasem i lat, otrzymują kolejne linie leczenia systemowego, które pozwalają na długotrwałą kontrolę choroby. Kilka lat temu, na łamach „*Journal of National Cancer Institute*”, zaprezentowano teoretyczną analizę statystyczną wykazującą, jak czas przeżycia po progresji wpływa na prawdopodobieństwo stwierdzenia znamiennej korzyści w zakresie OS u chorych, u których wykazano znamienne korzyści w zakresie PFS [2]. W analizie tej dowiedziono, iż w przypadku spodziewanego przeżycia po progresji wynoszącego około 2 miesięcy istnieje 90% szans, że różnice w zakresie OS będą znamienne statystycznie. W przypadku przeżycia po progresji wyno-

szącego około 2 lata to prawdopodobieństwo zmniejsza się do 20%. Biorąc na przykład pod uwagę rokowanie chorych na rozlanego raka piersi, jelita grubego czy nerki, których przeżycie po niepowodzeniu pierwszej linii leczenia systemowego wynosi aktualnie co najmniej kilkanaście miesięcy i z roku na rok się wydłuża, można łatwo zrozumieć, dlaczego coraz trudniej będzie wykazać w badaniach klinicznych znamienne poprawę w zakresie OS.

Piśmiennictwo

1. Johnson J.R., Williams G., Pazdur R. End points and United States Food and Drug Administration approval of oncology drugs. *J. Clin. Oncol.* 2003; 21: 1404–1411.
2. Broglio K.R., Berry D.A. Detecting an overall survival benefit that is derived from progression-free survival. *J. Natl. Cancer Inst.* 2009; 101: 1642–1649.