

ZARZĄDZENIE Nr 1/2010
PREZESA
AGENCJI OCENY TECHNOLOGII MEDYCZNYCH
z dnia 4 stycznia 2010 r.
w sprawie wytycznych oceny świadczeń opieki zdrowotnej

Na podstawie art. 31o ust. 3 ustawy z dnia 27 sierpnia 2004 r. o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych (t.j. Dz. U. z 2008 r., Nr 164, poz. 1027 z późn. zm.) zarządza się, co następuje:

§ 1.

Ustala się wytyczne oceny świadczeń opieki zdrowotnej w Agencji Oceny Technologii Medycznych, stanowiące załącznik do zarządzenia.


§ 2.

Uchyla się Zarządzenie Dyrektora Agencji Oceny Technologii Medycznych Nr 21 z dnia 27 marca 2007 r. w sprawie wytycznych oceny technologii medycznych, utrzymane w mocy zarządzeniem Prezesa Agencji Oceny Technologii Medycznych Nr 3/2009 z dnia 12 sierpnia 2009 r. w sprawie utrzymania w mocy części zarządzeń wydanych przez Dyrektora Agencji Oceny Technologii Medycznych działającej na podstawie Zarządzenia Ministra Zdrowia z dnia 30 czerwca 2006 r. (Dz. Urz. Min. Zdr. Nr 10, poz. 53 z późn. zm.).

§ 3.

Zarządzenie wchodzi w życie z dniem podpisania.




Jacek Powalla
radca prawny

Wytyczne oceny technologii medycznych (HTA)

1. Informacje wstępne

1.1. Definicja oceny technologii medycznych

Ocena technologii medycznych (ang. *Health Technology Assessment*, HTA) to interdyscyplinarna dziedzina wiedzy, służąca podejmowaniu opartych na dowodach naukowych decyzji w zakresie polityki zdrowotnej i praktyki klinicznej. Dyscyplina ta łączy wiedzę z zakresu m.in.: medycyny, epidemiologii, biostatystyki, ekonomii, prawa i etyki. HTA dostarcza naukowych podstaw podejmowania racjonalnych decyzji dotyczących stosowania i finansowania świadczeń zdrowotnych.

1.2. Zakres oceny technologii medycznych

Pełna ocena technologii medycznej składa się z analiz:

- klinicznej,
- ekonomicznej,
- wpływu na system ochrony zdrowia.

1.3. Cel oceny technologii medycznych

Celem prowadzenia ocen technologii medycznych jest dostarczanie informacji, które są niezbędne do podejmowania decyzji z zakresu polityki zdrowotnej, opartych na racjonalnych podstawach. Mają one służyć dobru pacjentów i zmierzają do zapewnienia bezpieczeństwa zdrowotnego, uzyskania efektów o największej wartości oraz optymalnego wykorzystania dostępnych środków.

Mach

1.4. Cel wytycznych

Celem wytycznych jest wskazanie zasad i podstawowych metod przeprowadzania ocen technologii medycznych, zapewniających wysoką jakość analiz i wiarygodność ich wyników.

1.5. Informacja o autorach i konflikcie interesu

Ocena technologii medycznych wymaga informacji o zleceńodawcy opracowania oraz autorach i wkładzie każdego z nich w opracowanie analizy. Konieczne jest także załączenie przez nich informacji o konfliktach interesów.

12/2018

2. Problem decyzyjny

2.1. Definiowanie problemu

Pierwszym etapem przeprowadzanej analizy jest jasne sprecyzowanie badanej technologii, interwencji diagnostycznej, profilaktycznej lub terapeutycznej, stosowanych w określonej sytuacji klinicznej.

Wymagany jest pełny opis zagadnień kontekstu klinicznego według schematu PICO¹:

- populacji, w której dana interwencja ma być stosowana (P);
- proponowanej interwencji (I);
- komparatorów (C);
- efektów zdrowotnych, czyli punktów końcowych badań klinicznych (O).

W przypadku analiz dołączanych do wniosków o finansowanie technologii ze środków publicznych kontekst kliniczny analiz musi odpowiadać opisanemu we wniosku. Należy również wskazać, które technologie i w jakim stopniu mogą zostać zastąpione przez technologię ocenianą.

2.1.1. Populacja

Należy przedstawić charakterystykę docelowej populacji lub populacji, która będzie poddawana ocenianej interwencji. Opis ów powinien obejmować podstawowe informacje o chorobie lub problemie zdrowotnym, z uwzględnieniem historii naturalnej choroby, rokowania i stosowanych obecnie metod diagnostycznych lub terapeutycznych.

Trzeba określić potencjalną liczebność populacji i opisać metodę jej oszacowania oraz uzasadnić ją.

2.1.2. Interwencja

Powinno się scharakteryzować ocenianą interwencję zdrowotną. W przypadku interwencji zarejestrowanej w Polsce należy podać datę rejestracji lub datę pierwszej deklaracji zgodności wyrobu medycznego i zarejestrowane wskazania, porównując je ze wskazaniami rozpatrywanymi w analizie. Dla technologii niezarejestrowanych w Polsce właściwe jest podanie wybranych dat i miejsc ich rejestracji w innych krajach oraz warunków określonych przez instytucje rejestrujące, w szczególności przez EMEA² i FDA³.

¹ PICO – z ang. *population, intervention, comparison, outcome*.

² *European Medicines Agency* – instytucja odpowiedzialna za sprawy dopuszczania do obrotu produktów leczniczych na poziomie centralnym w Unii Europejskiej.

³ *Food and Drug Administration* – instytucja odpowiedzialna za sprawy dopuszczania do obrotu m.in. produktów leczniczych i wyrobów medycznych w USA.

Mah

2.1.3. Komparatory

Analiza kliniczna polega na porównaniu skuteczności i bezpieczeństwa stosowania ocenianej interwencji (sposobu postępowania) z wynikami innych interwencji (opcjonalnych sposobów postępowania) stosowanych w docelowej populacji.

Komparatorem dla ocenianej interwencji w pierwszej kolejności musi być tzw. istniejąca praktyka. Jest to sposób postępowania, który w praktyce medycznej prawdopodobnie zostanie zastąpiony przez ocenianą technologię.

Zaleca się przeprowadzenie porównania również z innymi komparatorami, czyli z technologiami:

- najczęściej stosowaną,
- najtańszą,
- najskuteczniejszą
- zgodną ze standardami i wytycznymi postępowania klinicznego.

Istotne jest, aby wybrane komparatory odpowiadały warunkom polskim. Ich wybór powinien być rzetelnie uzasadniony oraz opatrzony źródłami danych.

2.1.4. Efekty zdrowotne

W analizie klinicznej powinny być oceniane efekty zdrowotne, które stanowią istotne klinicznie punkty końcowe⁴, odgrywające istotną rolę w danej jednostce chorobowej, tj.:

- zgony,
- zachorowania bądź wyleczenia,
- jakość życia,
- działania niepożądane (z podziałem na ciężkie i pozostałe) i/lub incydenty medyczne⁵.

Punkty końcowe w analizie klinicznej powinny:

- dotyczyć ocenianej jednostki chorobowej i jej przebiegu,
- odzwierciedlać medycznie istotne aspekty problemu zdrowotnego i jednocześnie umożliwiać wykrycie potencjalnych różnic między porównywanymi interwencjami,
- mieć zasadnicze znaczenie dla podejmowania racjonalnej decyzji (punkty krytyczne danego problemu zdrowotnego).

W przypadku, kiedy nie odnaleziono badań klinicznych z klinicznie istotnymi dla pacjenta punktami końcowymi, jako wyniki mogą być oceniane surogaty. Zalecane jest wtedy przedstawienie w analizie związku pomiędzy użytymi surogatami a klinicznie istotnymi punktami końcowymi.

⁴ Istotny klinicznie punkt końcowy mający znaczenie dla pacjenta (ang. *clinically important endpoint, clinically relevant endpoint, patient important outcome, patient-oriented endpoint* – parametr/wynik, którego zmiana pod wpływem leczenia sprawiłaby, że to leczenie będzie pożądanym przez chorych. Odzwierciedlają wpływ leczenia: przedłużający życie, poprawiający samopoczucie chorego bądź pozwalający żyć bez powikłań choroby lub jej leczenia.

⁵ Pojęcia precyzowane przez ustawę z 20.04.2004 o wyrobach medycznych (Dz. U. nr 93 poz. 896 z 2004 i nr. 64, poz. 565 z 2005), ustawę z 6.09.2001 prawo farmaceutyczne (tekst jednolity Dz. U. nr. 53, poz. 533 z 2004).

W przypadku, kiedy wyniki oceny klinicznej uzyskane są przy użyciu skal lub kwestionariuszy, należy przedstawić informacje o ich walidacji oraz istotności klinicznej wyników.



3. Analiza kliniczna

Analiza kliniczna dotyczy wyników zdrowotnych ocenianej technologii medycznej. Informuje także o jej skuteczności i bezpieczeństwie w określonej populacji w porównaniu z odpowiednimi komparatorami.

3.1. Dane

Poszukiwane w toku analizy klinicznej dane dotyczą zarówno efektywności eksperymentalnej (ang. *efficacy*), jak i efektywności praktycznej (ang. *effectiveness*). Wyszukiwanie i selekcja danych powinny odbywać się w oparciu o szczegółowy protokół, opracowany przed przystąpieniem do tego działania i zawierający określone kryteria włączania badań do analizy oraz kryteria ich wykluczania.

3.1.1. Źródła danych

We wstępnej części analizy należy przeprowadzić systematyczne poszukiwanie wszystkich badań klinicznych, które dotyczą rozpatrywanego problemu. Proces wyszukiwania danych i informacji trzeba szczegółowo opisać, aby można było ocenić jego poprawność oraz umożliwić jego powtórzenie w przypadku weryfikacji analizy HTA.

W pierwszej kolejności należy poszukiwać już istniejących, niezależnych raportów oceny technologii (raporty HTA) oraz przeglądów systematycznych dostępnych w:

- Cochrane Library,
- bazie MEDLINE,
- bazie EMBASE
- bazie Centre for Reviews and Dissemination.

Kolejnym etapem analizy klinicznej powinno być przedstawienie wniosków płynących z odszukanych opracowań wtórnych. Opracowania można również wykorzystać jako źródło informacji o praktyce analitycznej w danym problemie decyzyjnym. Jeżeli nie dostarczą one dostatecznie aktualnej i wyczerpującej informacji, należy poszukiwać odpowiednich badań pierwotnych.

Istotnym warunkiem wykonania przeglądu systematycznego badań pierwotnych jest odnalezienie wszystkich doniesień naukowych, które dotyczą porównywanych interwencji i spełniają kryteria włączenia do analizy. W pierwszej kolejności poszukiwać należy badań, w których bezpośrednio porównywano badaną technologię z wybranym komparatorem⁶.

Podstawowymi bazami do poszukiwania badań pierwotnych są:

- Medline,
- EMBASE[®],
- Cochrane Library (CENTRAL).

⁶ Ang. *head to head trials*.

Zalecane jest przeszukanie także innych baz informacji medycznych, takich jak:

- BIOSIS Previews[®],
- CINAHL[®] Database,
- PsycINFO[®],
- European Public Assessment Report (EPAR)⁷,
- Health Canada⁸,
- Netherlands Pharmacovigilance Centre Lareb⁹,
- The Uppsala Monitoring Centre¹⁰,
- Thompson Micromedex^{®11}.

Konieczne jest również sprawdzenie doniesień w źródłach innych niż bazy informacji medycznej, czyli:

- wykorzystanie odniesień bibliograficznych zawartych w publikacjach dotyczących badań klinicznych,
- przegląd rejestrów badań klinicznych,
- konsultacje z ekspertami klinicznymi.

Należy także rozważyć potrzebę uzyskania dodatkowych informacji. Można to osiągnąć poprzez:

- wyszukiwanie danych publikowanych w czasopismach specjalistycznych zajmujących się ocenianą technologią, nieuwzględnionych w strategii wyszukiwania,
- kontaktowanie się z autorami badań klinicznych,
- wykorzystanie wyszukiwarek internetowych,
- konsultacje z producentami, w szczególności w zakresie informacji o działaniach niepożądanych (tzw. PSUR12).

Powinno się zwrócić uwagę, czy uwzględnienie jedynie takich badań, których wyniki zostały opublikowane, może prowadzić do nieprawidłowego odczytania wyników przeglądu, ze względu na błąd publikacji^{13,14}.

Dane dotyczące efektywności eksperymentalnej uzyskiwane są przede wszystkim w wyniku systematycznego przeglądu prób klinicznych z grupą kontrolną. Dane odnoszące się

⁷ www.emea.europa.eu/htms/human/epar/eparintro.htm

⁸ www.hc-sc.gc.ca

⁹ www.lareb.nl

¹⁰ www.WHO-umc.org

¹¹ www.micromedex.com

¹² Ang. *Periodic Safety Update Report*.

¹³ Błąd publikacji jest związany z częstszym publikowaniem w czasopismach naukowych doniesień z badań, w których uzyskano pozytywne wyniki, niż z takich, w których uzyskano wyniki negatywne albo wykazano brak różnic.

¹⁴ Wyniki badań niepublikowanych mogą dostarczyć ważnych danych, należy więc, we współpracy z producentem leku lub sprzętu, którego dotyczy analiza, poszukiwać również takich badań.

do efektywności praktycznej pochodzą z pragmatycznych prób klinicznych¹⁵. Można je również otrzymać z badań obserwacyjnych i baz danych (w tym rejestrów pacjentów), skupiających informacje dotyczące stosowania danej technologii. Dane te powinny być uzyskiwane również w trybie przeglądu systematycznego. Stopień zgodności efektywności eksperymentalnej i praktycznej powinien być opatrzony komentarzem.

3.1.2.Strategia wyszukiwania

Analitik winien opracować strategię wyszukiwania odpowiednią do zdefiniowanego problemu klinicznego. Zaleca się stosowanie strategii wyszukiwania o możliwie najwyższej czułości. Jedyne w przypadku dużej liczby trafień można zwiększać swoistość przeszukiwania. W przypadku stosowania strategii istotnie różniących się czułością w różnych wyszukiwarkach wymagane jest przedstawienie uzasadnienia takiego postępowania. Kryteria wyszukiwania mogą uwzględniać elementy wybrane spośród następujących (schemat PICOS¹⁶):

- (P) populacja,
- (I) interwencja,
- (C) komparatory,
- (O) punkty końcowe,
- (S) rodzaj badań.

Wyszukiwanie badań pierwotnych powinno dotyczyć publikacji w językach: angielskim, polskim, niemieckim i francuskim, a w uzasadnionych przypadkach – także w innych.

Efektem końcowym wyszukiwania powinno być zgromadzenie wszystkich dostępnych badań i danych odnoszących się do analizowanego problemu klinicznego.

Przedstawienie wyników wyszukiwania powinno opisywać użytą strategię z uwzględnieniem:

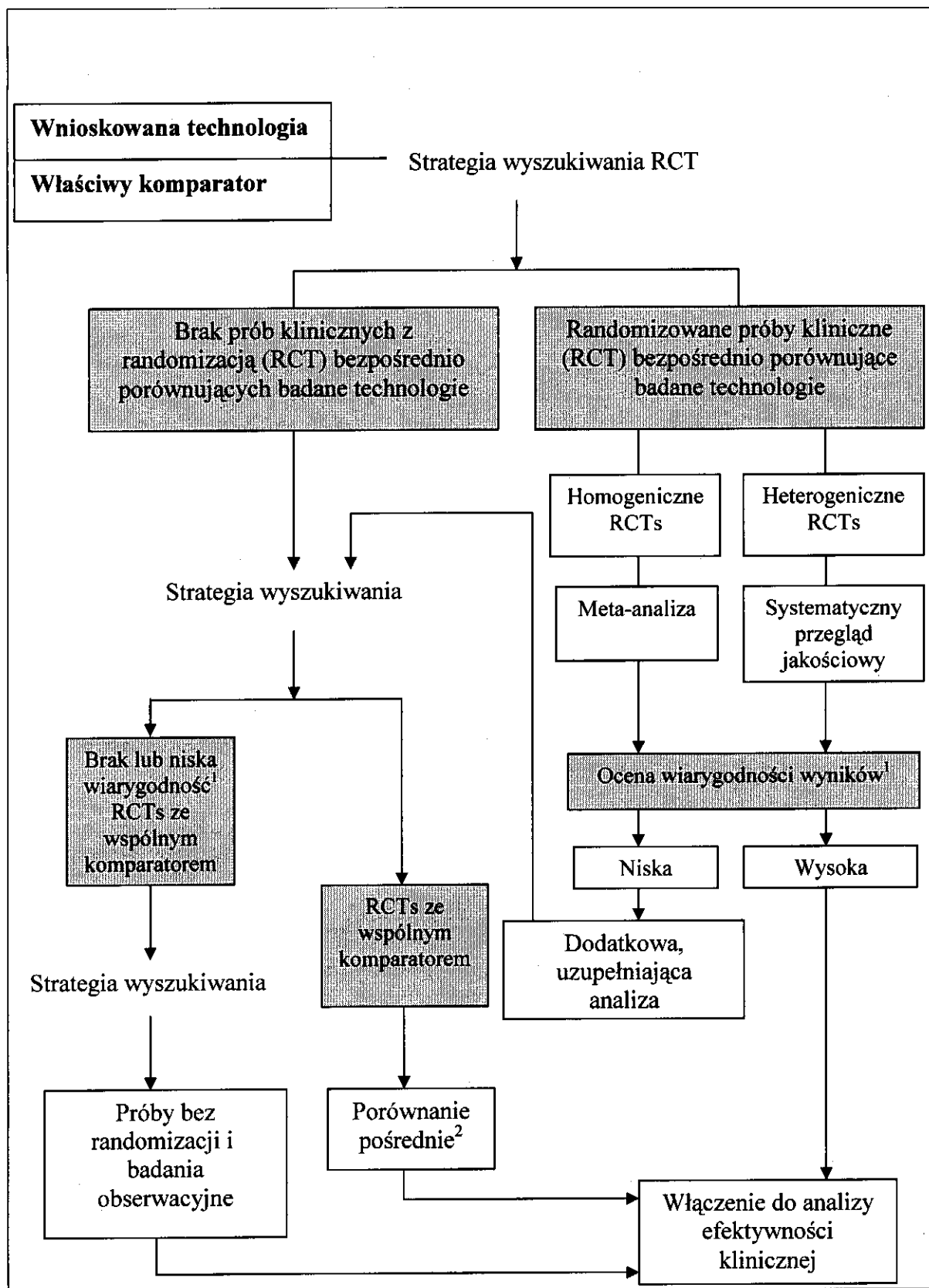
- słów kluczowych i deskryptorów zastosowanych w czasie wyszukiwania,
- użytych operatorów logiki Boole'a,
- użytych filtrów,
- przedziału czasowego objętego wyszukiwaniem.

3.1.3.Selekcja informacji

Proces weryfikacji, czy odnalezione doniesienia naukowe nadają się do analizy, należy przeprowadzić etapowo. Pierwszy etap obejmuje selekcję na podstawie abstraktów, a dalsze – na podstawie pełnych tekstów publikacji. Selekcja badań powinna być dokonywana na podstawie przyjętych przed rozpoczęciem wyszukiwania kryteriów włączania i wyłączenia. W przypadku dostępności badań o bardzo wysokiej wiarygodności (klinicznej i statystycznej) analiza skuteczności ocenianej technologii może ograniczyć się wyłącznie do nich.

¹⁵ Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *Journal of Chronic Diseases* 1967;20:637-648. Armitage P. Attitudes in Clinical Trials. *Statistics In Medicine* 1998;17:2675-2683.

¹⁶ Ang. *population, intervention, comparison, outcome, study.*



¹ Może dotyczyć całych badań, jak i poszczególnych efektów korzystnych i niekorzystnych (zdarzeń, punktów końcowych ocenianych w badaniach).

² Zalecane metody wykonywania pośrednich porównań badań ze wspólnym komparatorem zależą od użytych miar efektu – w przypadku ilorazów szans zaleca się stosowanie regresji logicznej lub metaregresji, natomiast w przypadku miar, takich jak ryzyka względne, różnica ryzyk, różnica średnich czy stosunek hazardów, zalecanymi metodami są wyrównane pośrednie porównanie Buchera (*adjusted indirect comparison Bucher*) lub metaregresja. W uzasadnionych przypadkach może być stosowana meta-analiza sieciowa (*network meta-analysis*).

P. 166

Na wszystkich etapach selekcja badań klinicznych w ramach przeglądu systematycznego powinna być wykonywana przez co najmniej dwóch analityków, którzy pracują niezależnie. Należy podać stopień zgodności między analitykami dokonującymi selekcji na etapie analizy pełnych tekstów publikacji. Sposób prezentacji stopnia zgodności powinien uwzględniać wymienienie publikacji, co do których istnieje pomiędzy analitykami niezgodność oraz wykaz przyczyn tej niezgodności, a także sposób ostatecznego rozstrzygnięcia. Preferowaną metodą rozstrzygania niezgodności jest uzyskiwanie konsensusu. W odpowiednich miejscach opracowania należy podać inicjały analityków wykonujących poszczególne zadania.

Analiza powinna przejrzysto informować o liczbie dostępnych doniesień naukowych na poszczególnych etapach wyszukiwania i selekcji badań. Proces prowadzący do ostatecznej selekcji należy przedstawić w postaci diagramu zgodnego z zaleceniami QUOROM¹⁷. Powinno się także przedstawić przyczyny wykluczenia badań na poszczególnych etapach selekcji.

3.1.4. Ocena jakości informacji

Ocena jakości danych pozwala określić ich wiarygodność wewnętrzną¹⁸ i zewnętrzną¹⁹. Ocena informacji w wyszukanych i włączonych do analizy badaniach wymaga rozważenia kilku czynników:

- metodyki poszczególnych badań,
- zagrożenia wiarygodności wyników badań (niedociągnięć metodologicznych) – zalecane jest oszacowanie błędów systematycznych (selekcji, detekcji, przeprowadzenia, utraty), które mogą mieć wpływ na wiarygodność wyników,
- stałości obserwowanych wyników zdrowotnych w poszczególnych badaniach,
- stopnia, w jakim wyniki stwierdzone w badaniach naukowych można przenieść (generalizować) na populację, której ma dotyczyć analiza – należy ocenić podobieństwo próby z badań klinicznych do potencjalnej populacji, podobieństwo interwencji (np. problem efektu klasy w przypadku leków), zbieżność wyników obserwowanych w badaniach naukowych z wynikami oczekiwanymi (np. problem zastępczych punktów końcowych czyli surogatów).

Badania eksperymentalne, które dotyczą terapii, należy oceniać w skali Jadad²⁰, a w przypadku badań diagnostycznych – za pomocą QUADAS²¹. Badania obserwacyjne należy oceniać z pomocą kwestionariusza NOS²² rekomendowanego przez Cochrane Non-

¹⁷ Moher D, Cook DJ, Eastwood S i wsp. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet 1999;354(9193):1896-1900.

¹⁸ Wiarygodność wewnętrzna dotyczy tego, w jakim stopniu wnioski wyciągnięte z badania odpowiadają rzeczywistości związkowi między badanym postępowaniem a obserwowanym punktem końcowym badania.

¹⁹ Wiarygodność zewnętrzna oznacza możliwość uogólniania wniosków z badania na populację docelową dla danej technologii medycznej, tj. w jakim stopniu wnioski wyciągnięte na podstawie badanej próby można odnieść do populacji w warunkach rutynowej praktyki klinicznej.

²⁰ Jadad AR, Moore RA, Carroll D i wsp. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996;17(1):1-1.

²¹ Whiting P, Rutjes A, Reitsma J, Bossuyt P, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003;3:25.

Randomized Studies Methods Working Group. Możliwa jest także osobna ocena badań za pomocą zmodyfikowanych skal²³; działanie to wymaga jednak uzasadnienia ich wyboru. Wszystkie skale i kwestionariusze powinny być przedstawione w załącznikach do przeglądu systematycznego.

3.1.5. Przedstawienie badań włączonych i ekstrakcja danych

W celu przedstawienia badań należy zestawić w tabelach wszystkie wyniki odnoszące się do określonego problemu klinicznego. Wykaz ten musi zawierać charakterystykę każdego z badań: okres obserwacji, liczbę ośrodków wykonujących, listę sponsorów, liczbę i rodzaj badań, wielkość próby badanej, charakterystykę pacjentów, szczegóły interwencji i uzyskane wyniki oraz inne informacje, które mają znaczenie dla oceny wiarygodności zewnętrznej.

Każde badanie włączone do analizy powinno posiadać zwięzłą ocenę krytyczną (*critical appraisal*), zgodną z zasadami Cochrane Collaboration²⁴.

Zestawienie powinno być wykonane w oparciu o klasyfikację doniesień naukowych wg. tab. 1 lub tab. 2 i z podaniem rodzaju włączonego badania.

Tabela 1. Klasyfikacja doniesień naukowych odnoszących się do terapii.²⁵

Typ badania	Rodzaj badania	Opis podtypu
Przegląd systematyczny RCT	IA	Metaanaliza na podstawie wyników przeglądu systematycznego RCT.
	IB	Systematyczny przegląd RCT bez metaanalizy.
Badanie eksperymentalne	IIA	Poprawnie zaprojektowana kontrolowana próba kliniczna z randomizacją ²⁶ .
	IIB	Poprawnie zaprojektowana kontrolowana próba kliniczna z pseudorandomizacją.
	IIC	Poprawnie zaprojektowana kontrolowana próba kliniczna bez randomizacji ²⁷ .
Badanie obserwacyjne z grupą kontrolną	IIIA	Przegląd systematyczny badań obserwacyjnych.
	IIIB	Poprawnie zaprojektowane prospektywne badanie kohortowe z równoczesną grupą kontrolną.

²² Wells GA, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. URL: <http://www.lri.ca/programs/ceu/oxford.htm>.

²³ Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovich C, Song F, et al. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003;7(27).

²⁴ Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

²⁵ Modyfikacja własna na podstawie: Undertaking systemic reviews of research on effectiveness: CRD guidelines for those carrying out or commissioning reviews. CRD report #4, University of York, York 1996.

²⁶ Kontrolowana próba kliniczna z randomizacją, ang. *randomised controlled trial*, RCT.

²⁷ Kontrolowana próba kliniczna, ang. *clinical controlled trial*, CCT.

	III C	Poprawnie zaprojektowane prospektywne badanie kohortowe z historyczną grupą kontrolną.
	III D	Poprawnie zaprojektowane retrospektywne badanie kohortowe z równoczesną grupą kontrolną.
	III E	Poprawnie zaprojektowane badanie kliniczno-kontrolne (retrospektywne).
Badanie opisowe	IV A	Seria przypadków – badanie pretest/posttest ²⁸ .
	IV B	Seria przypadków – badanie posttest ²⁹ .
	IV C	Inne badanie grupy pacjentów.
	IV D	Opis przypadku.
Opinia ekspertów	V	Opinia ekspertów w oparciu o doświadczenie kliniczne, badania opisowe oraz raporty paneli ekspertów.

W odniesieniu do badania eksperymentalnego należy określić przyjęte podejście do testowania hipotezy (*superiority, non-inferiority*).

Tabela 2. Klasyfikacja doniesień naukowych odnoszących się do diagnostyki.³⁰

Rodzaj badania	Opis
D I	Przegląd systematyczny badań poziomu D II.
D II	Badania kliniczne oceniające dokładność ³¹ metody diagnostycznej, w których zastosowano metodę ślepej próby oraz porównano ocenianą metodę diagnostyczną z testem referencyjnym (złotym standardem) w grupie pacjentów z określonym stanem klinicznym kolejno włączanych do badania.
D III-1	Badania oceniające dokładność metody diagnostycznej, w których zastosowano metodę ślepej próby oraz porównano ocenianą metodę diagnostyczną z testem referencyjnym (złotym standardem) w grupie pacjentów z określonym stanem klinicznym włączanych do badania nie w sposób kolejny.
D III-2	Badania porównujące ocenianą metodę diagnostyczną z testem referencyjnym, które nie spełniają kryteriów poziomu D II i D III-1.
D III-3	Diagnostyczne badania kliniczno-kontrolne.
D IV	Badania opisujące wyniki diagnostyczne, bez zastosowania testu referencyjnego.

²⁸ Badanie typu pretest/posttest – badanie opisowe ze zbieraniem danych przed zastosowaniem ocenianego postępowania i po nim.

²⁹ Badanie typu posttest – badanie opisowe ze zbieraniem danych tylko po zastosowaniu ocenianego postępowania.

³⁰ Według Medical Services Advisory Committee. Guidelines for the assessment of diagnostic technologies. August 2005.

³¹ Ang. *accuracy*.

W ostatecznej ocenie wykorzystuje się przede wszystkim odnalezione badania z najwyższego poziomu klasyfikacji. Najwyższy poziom wiarygodności zajmują przeglądy systematyczne (z metaanalizą lub bez niej), odzwierciedlające problem kliniczny pod względem badanego punktu końcowego, populacji, komparatora, pod warunkiem, że są one aktualne i zgodne z wytycznymi przeprowadzania takich badań. Gdy dane dotyczące skuteczności pochodzące z kontrolowanych prób klinicznych są ograniczone do wąsko zdefiniowanej populacji lub krótkiego horyzontu czasowego, powinny być uzupełnione badaniami obserwacyjnymi o dobrej jakości. Wartość materiałów dowodowych na każdym poziomie hierarchii zależy przede wszystkim od jakości metodycznej tych badań oraz spełnienia warunków oceny technologii medycznych.

Plan postępowania w procesie ekstrahowania danych z wyselekcjonowanych badań winien określać:

- rodzaje informacji wypisywanych z publikacji,
- liczbę osób dokonujących ekstrakcji i ich identyfikatory,
- formularz ekstrakcji danych.

Należy wykonać ilościowe zestawienie danych dotyczących skuteczności (korzystnych wyników) oraz bezpieczeństwa (szkodliwości, czyli niekorzystnych wyników) rozpatrywanej technologii, umieszczając je w jednolitym formularzu tabelarycznym. Zestawienie to powinno uwzględniać dokonaną uprzednio ocenę wiarygodności źródła i jakości danych. W zestawieniu należy ująć klinicznie istotne punkty końcowe (korzystne i niepożądane). Wyniki należy zestawić na podstawie wszystkich odnalezionych w toku przeglądu systematycznego badań poświęconych ocenianym technologiom lub wybranemu problemowi klinicznemu.

3.2. Synteza danych

Synteza wyników służy uzyskaniu informacji oraz określeniu poziomu związanej z nią niepewności oszacowań. Obejmuje ona systematyczny przegląd literatury (z metaanalizą lub bez niej) oraz podsumowanie.

Zalecane jest opracowanie wyników za pomocą metaanalizy. Jeśli nie jest możliwe jej wykonanie, analiza może być ograniczona do przeglądu jakościowego. Krytycznie ocenia się w nim poszczególne badania i przedstawia się ich wyniki w tabelach. Na podstawie omówienia wyników wyciągane są wnioski.

3.2.1. Synteza jakościowa

Zalecane jest przedstawienie lub oszacowanie wyników dla wszystkich analizowanych punktów końcowych każdego badania, które będą uwzględniały przedziały ufności i/lub istotność statystyczną.

Uzyskane wyniki dla punktu końcowego każdego badania powinny być omówione. W przypadku stwierdzenia heterogeniczności uzyskanych wyników konieczne jest prześledzenie i scharakteryzowanie różnic.

Zestawienie należy przedstawić w formie umożliwiającej porównanie wyników poszczególnych badań dla określonego punktu końcowego. Taki sposób prezentowania służy

wykrywaniu potencjalnych podobieństw lub różnic między włączonymi badaniami, a także między porównywanymi technologiami medycznymi.

Dane liczbowe należy przedstawić w formie tabeli zawierającej:

- wielkość próby dla każdej interwencji,
- wynik w postaci miar centralnych i miar rozrzutu dla każdego punktu końcowego w przypadku zmiennych ciągłych, a w przypadku zmiennych dychotomicznych, liczby i odsetki pacjentów, u których stwierdzono osiągnięcie punktu końcowego,
- różnice między średnimi wynikami porównywanych interwencji dla każdego punktu końcowego w poszczególnych badaniach wraz z podaniem przedziałów ufności i/lub istotności statystycznej dla parametrów ciągłych, zaś parametry względne i bezwzględne – z przedziałami ufności i/lub istotnością statystyczną w przypadku parametrów dychotomicznych.

3.2.2. Metaanaliza (synteza ilościowa)

Przed zastosowaniem statystycznych metod syntezy należy określić stopień i przyczyny niejednorodności (heterogeniczności) wyników badań. Sposób jej oceny i dalszego postępowania powinny być zgodne z wytycznymi Cochrane Collaboration³².

Jeśli istnieją wątpliwości dotyczące jakości badań lub związku poszczególnych badań z tematem analizy, należy oddzielnie przedstawić wyniki metaanaliz wykonanych z wyłączeniem wątpliwych badań. Konieczne jest wówczas osobne przedstawienie wyników badań o najwyższej wiarygodności. Wymagany jest dokładny opis kryteriów włączenia badań do metaanalizy oraz kryteriów wykluczenia.

3.2.3. Porównanie pośrednie

W przypadku braku badań, które bezpośrednio porównują technologię ocenianą i komparator (badań typu „head to head”), zaleca się przeprowadzenie porównania pośredniego.

Warunkiem porównania pośredniego jest dostępność wiarygodnych badań klinicznych, w których zestawiane technologie badano w odniesieniu do tej samej, trzeciej interwencji (placebo lub aktywna interwencja)³³. Identyfikacja badań w celu przeprowadzenia porównania pośredniego powinna być oparta na przeglądzie systematycznym. Zaleca się dokładną analizę metodyki badań oraz ocenę różnic w zakresie stosowanej trzeciej interwencji, otrzymujących ją populacji i badanych punktów końcowych. Różnice te należy przedstawić w formie tabelarycznego zestawienia. Jeżeli różnice zostaną uznane za zbyt duże, należy odstąpić od obliczania wyników, gdyż takie porównanie byłoby mało wiarygodne. Interpretacja wyników jakichkolwiek porównań pośrednich powinna być bardzo ostrożna. W każdym przypadku przeprowadzania porównania pośredniego jego wyniki powinny być poddane wszechstronnej interpretacji wraz z opisem ograniczeń oraz analizą wrażliwości.

Porównania pośrednie mogą być wykonywane i prezentowane niezależnie od porównań bezpośrednich. W przypadku porównań mieszanych, obejmujących zarówno porównania

³² Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

³³ To znaczy badań porównujących bezpośrednio zarówno ocenianą interwencję z trzecią interwencją, jak i badań komparatora z trzecią interwencją.

bezpośrednie, jak i pośrednie, wyniki samych porównań bezpośrednich należy przedstawić osobno i niezależnie od wyników porównania mieszanego.

3.3. Ocena bezpieczeństwa

3.3.1. Cel

Analizę bezpieczeństwa wykonuje się w celu oceny ryzyka stosowania technologii medycznej. Uwzględnia się tu działania niepożądane i incydenty medyczne, także te, które ujawniają się w długich okresach obserwacji i występują rzadko. Wyniki analizy bezpieczeństwa powinny zostać uwzględnione w ocenie technologii medycznych.

Raporty HTA mogą korzystać z danych, którymi dysponują urzędy zajmujące się nadzorem i monitorowaniem bezpieczeństwa produktów leczniczych (np. EMEA, FDA, URPL, WHO Uppsala Monitoring Centre).

3.3.2. Zakres analizy bezpieczeństwa

Zakres oceny bezpieczeństwa należy dostosować do problemu decyzyjnego oraz specyfiki ocenianej technologii medycznej. Inny zakres należy przyjąć przy ocenie leków, urządzeń medycznych, procedur medycznych i chirurgicznych, a inny w przypadku oceny testów diagnostycznych. W niektórych przypadkach zakres ten może być zbliżony do stosowanego przy ocenie skuteczności, często jednak wymaga rozszerzenia. Ocena bezpieczeństwa powinna być poszerzona szczególnie w przypadku technologii innowacyjnych, leków o nowym mechanizmie działania lub występowania działań niepożądanych generujących duże koszty.

Dane dotyczące oceny bezpieczeństwa pochodzące z badań RCT, włączonych wcześniej do analizy skuteczności, często nie są wystarczające z powodu zbyt krótkiego okresu obserwacji lub zbyt małej liczby pacjentów do nich włączonych. W celu oceny różnorodnych działań niepożądanych, zarówno zidentyfikowanych w badaniach RCT, jak i innych, należy wykonać możliwie jak najszerszy przegląd systematyczny (zarówno jeśli chodzi o strategię wyszukiwania, jak i rodzaj włączanych badań). Tego typu przegląd może być niezwykle pracochłonny i wymagać m.in. analizy serii przypadków lub danych pochodzących z rejestrów pacjentów. Obejmuje on także dane z raportów o działaniach niepożądanych, zbieranych zarówno przez firmy farmaceutyczne w postaci *Periodic Safety Update Reports* (PSUR), jak i urzędy zajmujące się nadzorem i monitorowaniem bezpieczeństwa produktów leczniczych (np. EMEA, FDA, URPL, WHO Uppsala Monitoring Centre). Dopuszcza się zawężenie oceny bezpieczeństwa poprzez:

- Identyfikację możliwych działań niepożądanych, na podstawie:
 - EPAR (EMEA), w szczególności ChPL,
 - analiz FDA.
- Ograniczenie zakresu badań w ocenie bezpieczeństwa, jeżeli podstawą przeglądu są badania typu RCT włączone do analizy skuteczności klinicznej, gdy oceniano w nich wszystkie wybrane do oceny działania niepożądane, okres obserwacji był wystarczająco długi, aby się one ujawniły, i uczestniczyło w nich wystarczająco dużo pacjentów, lub badania RCT były zaprojektowane pod kątem oceny działań niepożądanych (tj. działania niepożądane są w nich istotnym klinicznie punktem końcowym).

Mot

