



Agencja Oceny Technologii Medycznych i Taryfikacji



Fundusze Europejskie
Wiedza Edukacja Rozwój



**Rzeczpospolita
Polska**

Unia Europejska
Europejski Fundusz Społeczny



Porównania pośrednie w analizach klinicznych ocenianych przez AOTMiT – przegląd i opis zastosowanych metod

Natalia Szwarc¹, Anna Kaczorek-Juszkiewicz¹, Andrzej Kalbarczyk¹

¹Dział Refundacji (Reimbursement Division), Agencja Oceny Technologii Medycznych i Taryfikacji
(Agency for Health Technology Assessment and Tariff System)

Warszawa, sierpień 2019

Spis treści

1. Abstrakt	3
2. Wprowadzenie i cel publikacji	4
3. Porównanie pośrednie w świetle wytycznych HTA i rozporządzenia w sprawie wymagań minimalnych	5
3.1. Porównania pośrednie w stanowiskach europejskich agencji HTA.....	6
4. Porównania niebezpośrednie	7
4.1. Porównania pośrednie	7
4.2. Porównania mieszane	8
5. Wyniki przeprowadzonej analizy	13
5.1. Metodyka	13
5.2. Wyniki dotyczące AWA i analiz klinicznych załączonych do wniosków refundacyjnych	14
5.3. Wyniki dotyczące stanowisk Rady Przejrzystości i rekomendacji Prezesa AOTMiT.....	17
6. Dyskusja i wnioski	20
7. Bibliografia	22

1. Abstrakt

Porównania pośrednie stanowią coraz częściej wykorzystywane źródło wiedzy o skuteczności i bezpieczeństwie leków. Są one również coraz częściej wykorzystywane w analizach HTA. **Cel:** Identyfikacja ocenianych przez AOTMiT analiz klinicznych z porównaniami pośrednimi, analiza metod ich przeprowadzenia oraz analiza stanowisk Rady Przejrzystości i rekomendacji Prezesa Agencji wydanych na ich podstawie. **Metody:** W zasobach AOTMiT zidentyfikowano analizy kliniczne weryfikowane przez Agencję w latach 2016 – I połowa 2019 r., a następnie dokonano ekstrakcji danych dotyczących przeprowadzonych w nich porównań pośrednich. Ponadto dokonano analizy stanowisk Rady Przejrzystości i rekomendacji Prezesa AOTMiT. **Wyniki:** Zidentyfikowano 217 analiz weryfikacyjnych Agencji, spośród których w 76 (35%) przeprowadzono porównanie pośrednie. Do najczęściej wykorzystywanych metod należała metoda Buchera. Zaobserwowano również rosnące wykorzystanie metody porównania mieszanego z podejściem Bayesa (MTC). Najczęściej porównanie pośrednie stanowiło podstawę podejmowania decyzji w chorobach takich jak: nowotwory i choroby zapalne, a rzadziej choroby układu krążenia i choroby układu nerwowego. Analiza metodyki wykazała brak oceny heterogeniczności badań w ok. 15% dostarczonych wniosków. Jedynie w 50% MTC została przeprowadzona pełna ocena miary dopasowania modelu z zastosowaniem parametru DIC (z przedstawieniem wyników). Rekomendacje Prezesa częściej odwoływały się do porównań pośrednich niż stanowiska RP (68 vs. 55), a także częściej przedstawiano w nich wyniki porównań pośrednich (66 vs. 54), wskazywano na ograniczony charakter metod (45 vs. 16) oraz wymieniano poszczególne ograniczenia porównań (18 vs. 6). **Wnioski:** Analiza wskazuje na potrzebę opracowania jasnych zasad przeprowadzania i prezentacji porównań pośrednich oraz wnioskowania na ich podstawie o względnej efektywności technologii lekowych.

słowa kluczowe: porównania pośrednie, porównania mieszane, MTC, metaanaliza sieciowa, NMA.

Indirect comparisons are becoming an increasingly important source of knowledge about the efficacy and safety of medicinal products. They are also becoming more common in HTA analyses. **Aim:** Identification and analysis of indirect comparisons conducted within clinical analyses assessed by Agency for Health Technology Assessment and Tariff System, their methods and analysis of the standpoints of the Transparency Council and recommendations of the Agency's President based on indirect comparisons. **Methods:** We extracted data about indirect comparisons conducted within clinical analyses assessed by Agency for Health Technology Assessment and Tariff System from 2016 to June 2019. Then we conducted a review of standpoints of the Transparency Council and recommendations of the Agency's President. **Results:** We identified 217 verification analyses. In 76 (35%) there was an indirect comparison. The most frequently used method was Bucher's adjusted indirect comparison. We also observed an increasing use of mixed treatment comparison method (MTC). Indirect comparisons most frequently were used in oncology and inflammatory diseases. Less often they were used in cardiovascular and nervous system diseases. The analysis of the methodology showed lack of heterogeneity assessment in approx. 15% of submitted applications. Only 50% of MTC carried out a full assessment of the model fit measure using the DIC parameter. The President's recommendations more often referred to indirect comparisons than the standpoints of Transparency Council (68 vs. 55), more often presented the results of indirect comparisons (66 vs. 54), pointed limited nature of methods (45 vs. 16) and mentioned restrictions of indirect comparisons (18 vs. 6). **Conclusions:** The analysis indicated a need to develop clear rules for conducting, presenting results and drawing conclusions about the relative effectiveness of drug technologies from indirect comparisons.

key words: indirect comparison, mixed treatment comparison, MTC, network meta-analysis, NMA.

2. Wprowadzenie i cel publikacji

Obecnie na rynku dostępny jest coraz większy wybór technologii lekowych, mogących znaleźć zastosowanie w danej jednostce chorobowej. Rosnący wybór z jednej strony rozszerza możliwości terapeutyczne wielu chorób, z drugiej jednak stawia przed decydentami i klinicystami trudne zadanie oceny i wyboru najskuteczniejszej terapii w danym wskazaniu.

W celu podjęcia właściwej decyzji niezbędny jest dostęp do odpowiednich dowodów naukowych, z których najwyższą wiarygodnością charakteryzują się przeglądy systematyczne randomizowanych badań klinicznych z metaanalizą. Niemniej w wielu przypadkach brak jest badań RCT lub wnioskowanie na ich podstawie o względnej skuteczności ocenianej technologii jest ograniczone, gdyż nie zawierają porównań z aktywnym komparatorem, a jedynie z placebo. Brak badań bezpośrednio porównujących różne terapie wynika częściowo z przepisów dotyczących dopuszczania leków do obrotu, w których porównanie z placebo uznawane jest za wystarczające do wykazania skuteczności i bezpieczeństwa nowego leku. Ponadto zwraca się uwagę na znaczne koszty przeprowadzania badań typu „head to head” pomiędzy aktywnymi komparatorami, w związku z koniecznością włączenia do takich badań dużych grup pacjentów w celu zapewnienia odpowiedniej mocy testu [1]. Dodatkowo wadą tradycyjnie przeprowadzanej metaanalizy jest możliwość porównania jedynie dwóch interwencji jednocześnie [2].

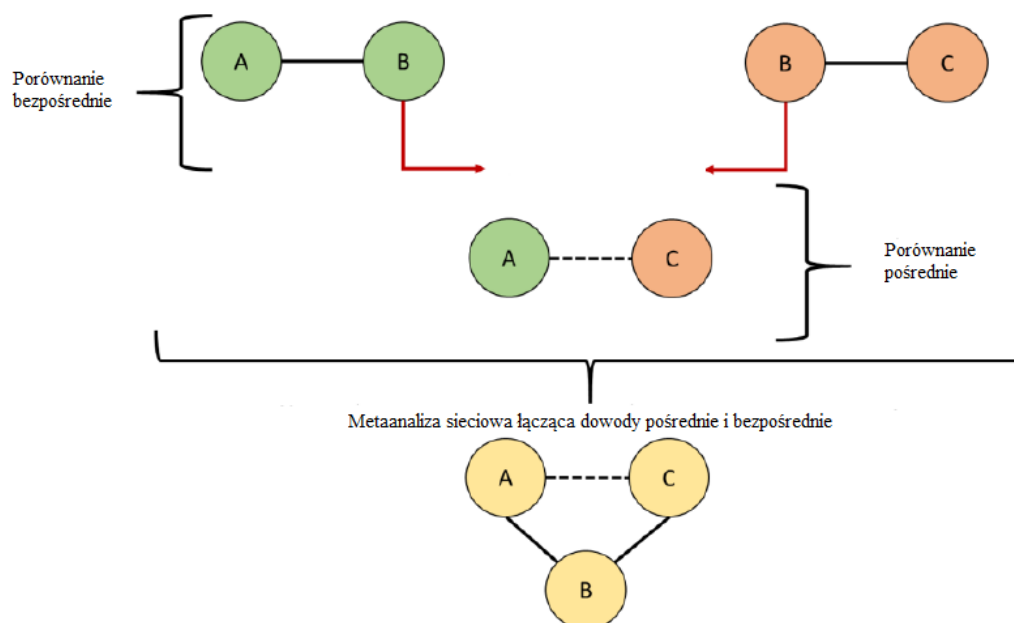
Z uwagi na powyższe zaistniała potrzeba stworzenia narzędzi umożliwiających pośrednie porównanie ze sobą interwencji przez wspólny komparator. Pierwsza metoda pozwalająca na pośrednie porównanie dwóch technologii lekowych została zaproponowana przez Buchera w 1997 r. Od tego czasu metoda ta doczekała się kilku modyfikacji. Jednocześnie rozwinęły się nowe metody porównań niebezpośrednich takie jak: metaanaliza sieciowa Lumley’a (2002 r.) i metaanaliza z podejściem Bayesa opracowana przez Lu i Ades (2004 r.). Możliwości stwarzane przez ww. metody, które pozwalają na wnioskowanie o skuteczności i bezpieczeństwie kilku terapii względem siebie na podstawie porównań niebezpośrednich, cieszą się coraz większym zainteresowaniem. Od 2008 r. wzrasta też liczba publikacji opartych na porównaniach pośrednich [3]. Należy się więc spodziewać, iż tego typu porównania coraz częściej będą wykorzystywane w analizach oceny technologii medycznych kierowanych do AOTMiT wraz z wnioskami o refundację technologii lekowych, co wiąże się z wieloma wyzwaniem metodologicznymi. Przede wszystkim podkreślić należy, iż pomimo licznych zalet porównań pośrednich ich stosowanie wiąże się z kontrowersjami i nie mają one wciąż ugruntowanej pozycji wiarygodnego źródła informacji o skuteczności i bezpieczeństwie terapii.

Biorąc pod uwagę powyższe, w niniejszej publikacji przybliżono główne założenia porównań przeprowadzanych w sposób niebezpośredni. Jednocześnie dokonano oceny częstości przeprowadzania w Agencji Oceny Technologii Medycznych i Taryfikacji raportów weryfikujących analizy kliniczne oparte na porównaniach pośrednich na podstawie danych z lat 2016 – I połowa 2019 r. Celem pracy było również sprawdzenie, jakie metody porównań wykorzystywane były w powyższych analizach. Podjęto się również uproszczonej oceny metodyki przeprowadzonych porównań oraz analizy decyzji wydanych przez Prezesa Agencji i Radę Przejrzystości z uwzględnieniem odwołań do wyników porównań pośrednich.

3. Porównanie pośrednie w świetle wytycznych HTA i rozporządzenia w sprawie wymagań minimalnych

Celem analizy klinicznej jest dostarczenie informacji o skuteczności i bezpieczeństwie ocenianej technologii medycznej w danej populacji w porównaniu z odpowiednimi komparatorami. Aby wykonać analizę kliniczną należy przeprowadzić przegląd systematyczny badań naukowych. W przypadku nieodnalezienia badań bezpośrednio porównujących ocenianą technologię medyczną z komparatorem (badań typu „head to head”) polskie wytyczne HTA zalecają przeprowadzenie porównania pośredniego z zastosowaniem metod skorygowanych o wynik grupy kontrolnej, np. metody Buchera, porównania mieszanego/sieciowego (ang. mixed treatment comparison) metodą Bayesa, metaanalizy sieciowej metodą Lumley’a (ang. Lumley network metaanalysis) lub metaregresji. Ostateczny wybór metody analitycznej powinien być uzależniony od typu dostępnych danych i celu porównania (porównanie ze sobą dwóch leków poprzez wspólną interwencję kontrolną lub porównanie więcej niż dwóch interwencji). Uzasadnienie wyboru metody powinno zostać opisane w raporcie.

Na rysunku poniżej przedstawiono rodzaje dowodów wykorzystywanych w analizach klinicznych.



Rysunek 1. Rodzaje porównań - pośrednie (oznaczone na rysunku linią przerywaną) i bezpośrednie (oznaczone na rysunku linią ciągłą).

Źródło: opracowanie własne na podstawie Tonin 2017

Warunkiem przeprowadzenia porównania pośredniego jest dokonanie oceny heterogeniczności metodycznej i klinicznej badań włączonych do analizy. Wytyczne zalecają dokładną analizę populacji badanej, interwencji stanowiącej ramię referencyjne i badanych punktów końcowych, a następnie przedstawienie w formie tabelarycznej różnic zidentyfikowanych pomiędzy badaniami. Jeżeli heterogeniczność badań jest zbyt duża, należy poprzestać na jakościowym zestawieniu wyników. Jeżeli nie ma możliwości przeprowadzenia porównania przez ramię referencyjne można rozważyć zastosowanie jednej ze wskazanych przez wytyczne HTA metod, tj:

- proste zestawienie badań bez dostosowania (ang. naïve comparison);
- porównanie z danymi historycznymi (ang. benchmarking with historical controls);
- porównanie wyników badań po korekcji o różnicę w charakterystyce populacji – dopasowaniu populacji (ang. matching-adjusted indirect comparison, MAIC).

Wytyczne HTA podkreślają, że interpretacja wyników porównań pośrednich powinna być bardzo ostrożna. Należy również wskazać ograniczenia analizy oraz przeprowadzić analizę wrażliwości pozwalającą na ocenę wpływu poszczególnych badań (ze szczególnym uwzględnieniem badań odbiegających metodycznie) na wynik porównania.

Warto zwrócić uwagę, że Rozporządzenie Ministra Zdrowia z dnia 2 kwietnia 2012 r. w sprawie minimalnych wymagań, jakie muszą spełniać analizy uwzględnione we wnioskach o objęcie refundacją i ustalenie urzędowej ceny zbytu oraz o podwyższenie urzędowej ceny zbytu leku, środka spożywczego specjalnego przeznaczenia żywieniowego, wyrobu

medycznego, które nie mają odpowiednika refundowanego w danym wskazaniu nie odnosi się do kwestii porównań pośrednich.

3.1. Porównania pośrednie w stanowiskach europejskich agencji HTA

W związku z faktem, iż polskie wytyczne w ograniczonym stopniu odnoszą się do porównań pośrednich, zdecydowano o przeprowadzeniu wyszukiwania wytycznych dotyczących porównań pośrednich w innych europejskich agencjach HTA. Spośród europejskich agencji oceny technologii medycznych jedynie w przypadku 3 zidentyfikowano dokumenty odnoszące się do znaczenia i zasad przeprowadzania porównań pośrednich: brytyjskiej National Institute for Health Care Excellence (NICE), niemieckiej Institute for Quality and Efficiency in Health Care (IQWiG) oraz francuskiej Haute Autorité de Santé (HAS). Podejście do porównań pośrednich w dokumentach NICE i IQWiG jest w dużej mierze zbieżne. Dokument HAS prezentuje natomiast bardziej liberalne podejście do przeprowadzanych porównań pośrednich.

W dokumencie NICE „Guide to the methods of technology appraisal 2013” zarekomendowano przedstawianie wyników porównań z zastosowaniem metaanalizy sieciowej, w przypadku braku porównań bezpośrednich przeprowadzonych w ramach RCT. Wskazano również możliwość przedstawienia wyników metaanaliz sieciowych w przypadku istnienia RCT zawierającego porównanie bezpośrednie, jeśli porównanie to dostarczy dodatkowych informacji. W idealnej sytuacji metaanaliza sieciowa powinna uwzględniać wszystkie terapie, które zostały zidentyfikowane jako interwencje lub komparatory w danym wskazaniu [4].

W dokumencie przedstawiono wymagania dotyczące sposobu prezentacji przeprowadzonych porównań. Zaleca się w nim m.in.: przedstawienie jasnego opisu metod syntezy oraz uzasadnienia, w jaki sposób przeprowadzono identyfikację, wybór i wykluczenie RCT oraz przedstawienie metod i wyników poszczególnych badań, które zostały włączone do metaanalizy sieciowej. W przypadku wątpliwości dotyczących tego, czy dane badanie lub zestaw badań zostały włączone do analizy zasadnie, NICE zaleca przeprowadzenie analizy wrażliwości, raportowanie heterogeniczności pomiędzy wynikami porównań w parach i niespójności (ang. inconsistency), rozumianej przez NICE jako różnice pomiędzy dowodami bezpośrednimi i pośrednimi. W przypadku braku wystarczających danych do opracowania metaanalizy sieciowej, analiza może być ograniczona do przeglądu narracyjnego, który krytycznie ocenia poszczególne badania i przedstawia ich wyniki. W takim przypadku wskazano, iż komisja oceniająca NICE jest szczególnie ostrożna przy analizie wyników i wyciąganiu wniosków w zakresie względnej skuteczności klinicznej [4].

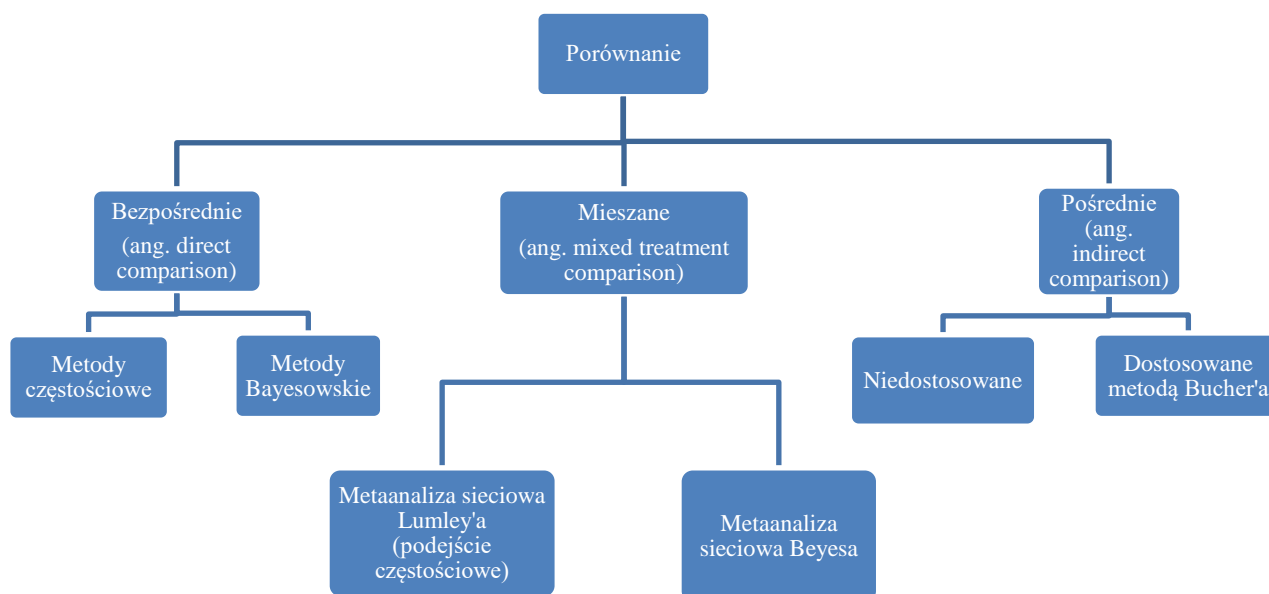
W dokumencie „General Methods 5.0” z 2017 r. IQWiG, podobnie jak NICE, dopuszcza przeprowadzenie porównań pośrednich w ramach oceny technologii medycznych. Instytucja wskazała, iż przeprowadzenie porównań pośrednich jest możliwe w przypadku istnienia tylko jednego badania o niepewnych wynikach, gdyż może pozwolić w takiej sytuacji na zmianę wnioskowania. Ponadto zauważono, iż w niektórych przypadkach porównania pośrednie mogą zwiększyć precyzję oszacowania efektu. Niemniej ze względu na wciąż istniejące problemy metodologiczne, IQWiG nie poleca wykonywania porównań pośrednich i wskazuje, że wnioski dotyczące oceny skuteczności powinny pochodzić z badań bezpośrednio porównujących analizowane technologie. Dodatkowo IQWiG nie dopuszcza przeprowadzania niedostosowanych porównań pośrednich. Jedynymi akceptowalnymi formami porównań są: dostosowane porównanie pośrednie metodą Buchera oraz metaanaliza sieciowa. W wytycznych, podobnie jak w dokumencie NICE, wymieniono konieczność przeprowadzenia oceny homogeniczności, a w przypadku metaanaliz sieciowych również oceny spójności (ang. consistency). Podkreślono, że w praktyce niezbędny jest dokładny opis modelu, wraz ze wszystkimi niepewnościami, które powinny być testowane w ramach analizy wrażliwości [5].

Odnaleziony na stronie francuskiej agencji dokument z 2009 r. „Indirect comparisons Methods and validity” nie stanowi wytycznych a raport. Jako kluczowe punkty raportu wskazano, iż w celu zrekompensowania braku badań porównawczych dla wielu jednostek chorobowych, wykorzystanie porównań pośrednich może zostać rozpatrzone w przypadku konieczności udzielenia odpowiedzi na pytanie dotyczące rankingu porównywanych technologii w zakresie bezpieczeństwa i skuteczności. Wykorzystanie porównań pośrednich zdaniem HAS pozwala na ocenę skuteczności względnej nowych terapii w momencie ich rejestracji, bez potrzeby oczekiwania na wyniki badań „head to head”. W raporcie zalecono, aby porównania bezpośrednie i pośrednie były porównane i jeśli to możliwe połączone w porównaniu mieszanym (ang. mixed treatment comparison). Idealnym podejściem według powyższej pracy jest przeprowadzanie dużych metaanaliz sieciowych, które pozwalałyby na hierarchizację terapii stosowanych w danym obszarze [7].

4. Porównania niebezpośrednie

Wśród metod porównań niebezpośrednich wyróżnia się metody pośrednie: niedostosowane (naiwne) i dostosowane metodą Bucher’a oraz porównania mieszane (ang. mixed treatment comparison, MTC), do których zalicza się metaanalizę sieciową Lumley’a oraz porównanie mieszane metodą Bayesa [6]. W metodach mieszanych wykorzystuje się dane z porównań pośrednich, jak i bezpośrednich.

Na rysunku poniżej przedstawiono rodzaje najczęściej stosowanych porównań, z wyszczególnieniem porównań pośrednich i mieszanych. Charakterystykę poszczególnych metod przedstawiono w kolejnych podrozdziałach.



Rysunek 2. Rodzaje porównań.

Źródło: opracowanie własne na podstawie EUnetHTA 2015

4.1. Porównania pośrednie

Niedostosowane

Wśród metod niedostosowanych wyróżnia się naiwne porównanie oszacowań punktowych i porównanie wyników dla pojedynczych ramion z niezależnych badań tak jakby pochodziły z jednego badania [7]. Metoda ta nie jest zalecana, ponieważ ignoruje randomizowaną naturę badań [6].

Jeśli ryzyko względne (ang. relative risk, RR) dla porównania A vs. placebo wynosi 0,76 (redukcja ryzyka względem placebo o 24%), a RR dla porównania B vs placebo wynosi 0,87 (redukcja ryzyka względem placebo o 13%) to naiwne porównanie oszacowań punktowych będzie polegało na wyciągnięciu wniosku, że A jest bardziej skuteczne niż B, ponieważ redukuje ryzyko względem placebo w większym stopniu niż B. Główną wadą takiego podejścia jest ignorowanie mocy statystycznej oszacowań i wykluczenie wnioskowania na podstawie przedziałów ufności.

Dostosowane metodą Buchera

Metodą dostosowanego porównania pośredniego jest metoda Buchera. Pierwotnie została wprowadzona w celu obliczania OR (ilorazu szans, odds ratio) w sytuacji, gdy chcemy porównać A vs. C, podczas gdy mamy dostępne porównania bezpośrednie A vs. B i B vs. C [1]. Może być jednak łatwo przekształcona do obliczania innych parametrów EBM [6]. Na Rysunek 3 (A) przedstawiono schemat porównania metodą Bucher’a.

Przy tak skonstruowanej sieci ryzyko względne A vs. C wyniesie: $RR_{AC} = \frac{RR_{AB}}{RR_{CB}}$, a przedział ufności zostanie oszacowany na podstawie wariancji obliczonej z następującego wzoru: $WAR(\log RR_{AC}) = WAR(\log RR_{AB}) + WAR(\log RR_{CB})$ [7].

Metoda Bucher’a może być zastosowana tylko do badań dwuramiennych [1]. Zakłada ona również, że względna skuteczność ocenianej interwencji jest taka sama we wszystkich badaniach włączonych do analizy, dlatego istotne jest, żeby uwzględnione badania były zbliżone pod kątem metodyki, badanych punktów końcowych, stosowanych interwencji, włączonej populacji [6]. W celu określenia homogeniczności badań dokonuje się subiektywnego porównania charakterystyki badań pod kątem metodologicznym i klinicznym.

Jeśli możliwe jest przeprowadzenie porównania pośredniego metodą Bucher’a za pomocą więcej niż jednego wspólnego komparatora (w omawianym przykładzie będzie to interwencja inna niż B), wtedy możliwe jest obliczenie zgodności wyników uzyskanych za pomocą różnych komparatorów wspólnych przy wykorzystaniu testu na heterogeniczność [7].

Główna wada dostosowanego porównania pośredniego dotyczy niepewności wyniku związanej z kumulacją statystycznej niepewności (suma wariancji) z poszczególnych badań włączonych do takiego porównania [8]. W celu oszacowania efektu z taką samą precyzją jak w przypadku porównania bezpośredniego należy wykorzystać 4-krotnie więcej danych, co powoduje, że z matematycznego punktu widzenia jest to metoda mniej efektywna niż porównanie bezpośrednie [9].

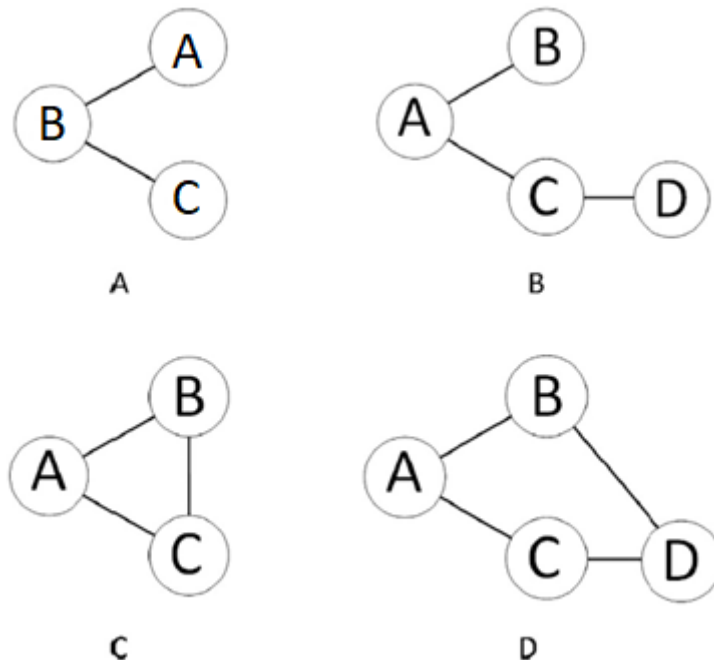
Porównanie pośrednie z dopasowaniem populacji

Analiza MAIC (ang. matching-adjusted indirect comparison) [10] jest metodą porównania pośredniego, której celem jest zniesienie heterogeniczności klinicznej wynikającej z odmiennej charakterystyki populacji w badaniach włączonych do porównania pośredniego. Zniesienie heterogeniczności badań ma na celu wiarygodniejsze porównanie interwencji.

W przeciwieństwie do klasycznego porównania pośredniego wykorzystującego zagregowane dane z publikacji badań, metoda MAIC wymaga dostępu do danych indywidualnych pacjentów przynajmniej jednej interwencji. Na podstawie tych danych wykonuje się ważenie efektu interwencji tak, aby określić jaki wynik zdrowotny powinien być uzyskany w populacji badania z ramienia komparatora.

4.2. Porównania mieszane

Stosowane są w celu oceny skuteczności względnej dwóch lub więcej interwencji wykorzystując jednocześnie dowody pośrednie i bezpośrednie. Na poniższym rysunku przedstawiono przykładowe sieci – bez pętli zamkniętej (B) i z pętlą zamkniętą (C, D).



Rysunek 3. Przykładowe schematy porównań pośrednich (A) i mieszanych z pętlą zamkniętą (C, D) i otwartą (B).

Źródło: Opracowanie własne na podstawie Jansen 2014

Im więcej porównań bezpośrednich znajduje się w sieci, tym wiarygodność wnioskowania jest większa [1]. Autorzy publikacji Tonin 2017, dotyczącej porównań mieszanych zwracają uwagę, że w celu pełniejszego zrozumienia i interpretacji wyników porównania mieszanego sieć powinna być szczegółowo opisana i przedstawiona za pomocą grafu [1].

Wśród porównań mieszanych wyróżnia się metaanalizy sieciowe Lumley’a (ang. network meta-analysis, NMA) oraz porównania mieszane (MTC) z podejściem Bayesa. MTC Bayesa stanowi wersję udoskonaloną NMA Lumley’a, opracowaną przez Lu i Ades w 2004 r. Przez swoje podobieństwo do metody Lumley’a często nazywana jest również NMA [1]. Przegląd raportów HTA weryfikowanych przez AOTMiT pokazuje, że MTC Bayesa nazywane jest metaanalizą sieciową z podejściem Bayesa.

Porównanie metodą Lumley’a

Zazwyczaj (w zależności od złożoności sieci) możliwe jest obliczenie skuteczności względnej danej interwencji przy wykorzystaniu różnych połączeń w sieci. Lumley wskazał, że jeśli wyniki uzyskane poprzez zastosowanie różnych komparatorów są zbieżne to istnieje duże prawdopodobieństwo, że porównanie pośrednie pokazuje rzeczywistą relację pomiędzy porównywanymi interwencjami [1]. Jeśli wyniki nie są zbieżne mówi się o tym, że występuje niespójność (ang. incoherence, w MTC nazywane również inconsistency) [1].

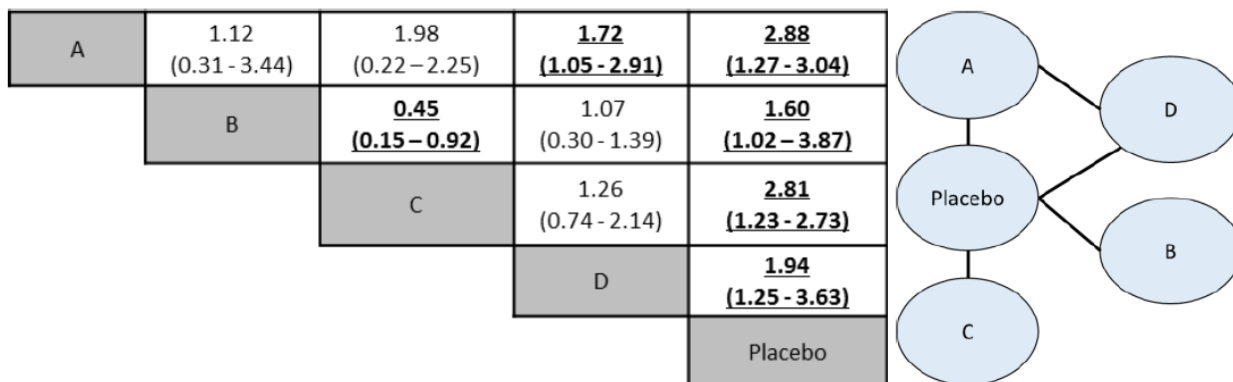
Podobnie jak metoda Bucher’a w NMA Lumley’a zakłada się, że względna skuteczność ocenianej interwencji jest taka sama we wszystkich badaniach włączonych do analizy. Dlatego jej ograniczeniem jest to, że umożliwia włączenie wyłącznie dwuramiennych badań. Problem ten został rozwiązany przez Lu i Ades w MTC, w którym stosuje się korektę do badań wieloramiennych [9].

Porównanie mieszane/metaanaliza metodą Bayesa

Porównanie MTC metodą Bayesa zostało wprowadzone przez Lu i Ades w 2004 roku [11]. Zastosowanie w metaanalizie sieciowej statystyki Bayesa miało na celu wzmocnienie siły wnioskowania dotyczącego względnej skuteczności jednej technologii nad drugą z wykorzystaniem dowodów pośrednich i bezpośrednich [1]. Zastosowanie statystyki Bayesa, a także uwzględnienie w analizie szerszej puli danych (tj. z porównań pośrednich i bezpośrednich) niż w przypadku zwykłego porównania pośredniego metodą Bucher’a wymieniane jest w literaturze jako argument na wyższą wiarygodność takich porównań [8]. Kluczowe jest jednak ściśle spełnienie założeń metodologicznych, ponieważ w MTC można wyróżnić dwa rodzaje heterogeniczności mogące zaburzać wnioskowanie: heterogeniczność badań wewnątrz poszczególnych porównań parami (co jest miarą zbliżoną do heterogeniczności w standardowej metaanalizie parami) i heterogeniczność pomiędzy poszczególnymi porównaniami parami (zwana niespójnością – inconsistency/incoherence) [9]. Istotny jest również transparentny opis przeprowadzonych czynności.

Prezentacja wyników MTC polega na przedstawieniu tabeli „ligowej” (ang. league table), która zawiera informację o relatywnej skuteczności każdej interwencji w porównaniu z pozostałymi wraz z określeniem stopnia niepewności w postaci przedziału wiarygodności (ang. credible interval) [13]. Poniżej przedstawiono przykładową tabelę z wynikami MTC.

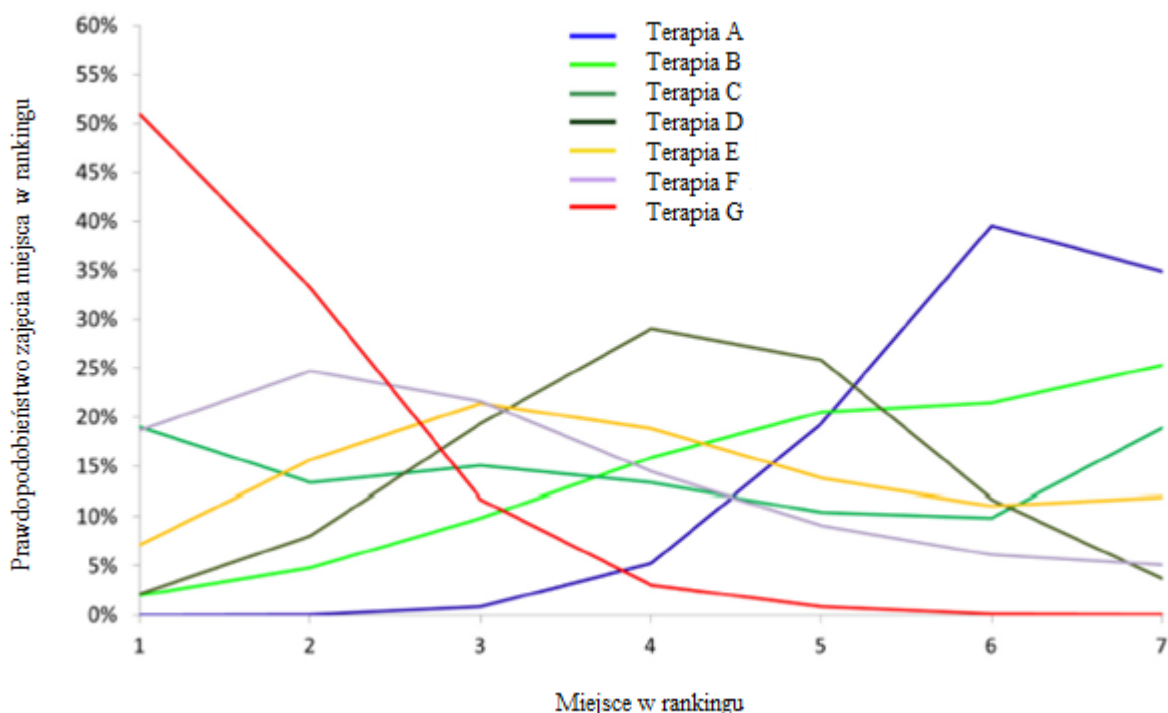
Rysunek 4. Przykład prezentacji wyników metaanalizy sieciowej w postaci tabeli. Literami oznaczono poszczególne interwencje



w sieci, której schemat narysowany jest po prawej stronie tabeli. Wyniki analizy zaprezentowane są w postaci ilorazu szans (OR) i 95% przedziału wiarygodności. Wyniki powinny być odczytywane w kolejności od lewej do prawej.

Źródło: Tonin 2017

Lu i Ades umożliwili także jednoczesne porównywanie więcej niż dwóch technologii poprzez ich rankingowanie, tj. określanie prawdopodobieństwa zajęcia określonego miejsca w sieci przez każdą z technologii. Sposoby raportowania rankogramu obejmują: medianę rangi wraz przedziałem ufności, skumulowane krzywe prawdopodobieństwa czy powierzchnię pod skumulowaną krzywą rankingową (ang. surface under the cumulative ranking curve, SUCRA). Nie zaleca się jednak rankingowania interwencji bazując na prawdopodobieństwie bycia najlepszą technologią. Lepszym podejściem jest mediana rangi lub skumulowane prawdopodobieństwo rangi [13][14].



Rysunek 5. Graficzny sposób prezentacji wyników w postaci rankogramu. Rankogram należy interpretować jako prawdopodobieństwo zajęcia określonego miejsca w rankingu przez poszczególne interwencje. Pole powierzchni pod krzywą dla danej interwencji (SUCRA) będzie wyrażało całkowitą pozycję interwencji w rankingu. SUCRA może przybierać wartości od 0% do 100%. Im wyższa wartość SUCRA, tym wyższe prawdopodobieństwo, że dana technologia będzie najbardziej skuteczna.

Źródło: Jansen 2014

Ze względu na zastosowane podejście Bayesowskie MTC oferuje bardziej wyrafinowane sposoby na uwzględnienie w sieci badań innych niż RTC (obserwacyjne, opinie ekspertów). Istnieją również metody pozwalające na uwzględnienie metaregresji w MTC, dzięki czemu możliwe jest uwzględnienie charakterystyk poszczególnych badań w celu zmniejszenia niespójności sieci (ang. inconsistency). Zastosowanie takiej metody zmniejsza jednak moc wnioskowania [6].

Główne ograniczenie związane z metodą MTC dotyczy stopnia jej złożoności, co wymaga od użytkowników zaawansowanej wiedzy statystycznej i dostępu do specjalistycznego oprogramowania [8]. Również Rouse 2017 podkreśla, że metoda ta jest skomplikowana i wymaga dużych nakładów pracy i wiedzy, żeby przeprowadzić ją w wiarygodny sposób.

Podejście częstościowe vs. bayesowskie

Główna różnica pomiędzy NMA Lumley'a a NMA z podejściem Bayesa polega na odmiennej definicji prawdopodobieństwa [1]. W statystyce częstościowej (NMA Lumley'a) wynikiem analizy jest estymator punktowy wraz z 95% przedziałem ufności (ang. confidence interval, CI), który mówi o prawdopodobieństwie pokrycia przez przedział ufności rzeczywistej wartości parametru. Analizy bayesowskie opierają się na rozkładach prawdopodobieństwa wszystkich parametrów modelu biorąc pod uwagę zaobserwowane dane oraz dodatkowo wcześniejsze przekonania (rozkład *a posteriori*). Wyniki NMA w ujęciu Bayesa zazwyczaj prezentowane są jako estymator punktowy wraz z 95% przedziałem wiarygodności (ang. Credible interval, CrI). Interpretacja wyniku NMA w ujęciu Bayesa dotyczy prawdopodobieństwa, że wynik przybierze daną wartość. [9].

Aby uzyskać rozkład parametrów *a posteriori*, stosuje się zazwyczaj metodę symulacji Monte Carlo przy użyciu łańcuchów Markowa (MCMC; ang. Markov Chain Monte Carlo) [9]. Proces próbkowania generuje dużą próbę parametrów z rozkładu aposteriorycznego, a im większa próba, tym lepiej reprezentuje ona rozkład aposterioryczny. Należy pamiętać, że próba parametrów otrzymana z MCMC nie jest tożsama z próbą danych empirycznych [19].

Ograniczeniem wynikającym wprost z bayesowskiej konstrukcji modelu metaanalizy sieciowej jest pewna wrażliwość ostatecznych wyników na dobór rozkładów *a priori*, dlatego dobór tych rozkładów jest istotny w kontekście wiarygodności MTC.

Model fixed i random

NMA może być przeprowadzona za pomocą modelu efektów stałych (ang. fixed effect approach) lub zmiennych (ang. random effect approach). W każdej sieci powinny być testowane oba modele [1]. W celu ustalenia, który z modeli jest lepiej dostosowany można zastosować parametr statystyczny DIC (ang. deviance information criteria), który informuje o jakości dopasowania modelu do danych przy uwzględnieniu stopnia złożoności modelu. Nie wskazano jednoznacznie na wartość różnicy, która byłaby ważna przy wyborze modelu, przy czym uznaje się, że niższa wartość DIC świadczy o lepszym dopasowaniu modelu [12].

Założenia NMA/MTC [1][15]:

Ze względu na złożoność metody, oprócz standardowych założeń występujących w klasycznej metaanalizie parami, NMA powinno spełniać założenie spójności. Z tego względu kolejne kroki przeprowadzenia NMA powinny obejmować:

- 1) Ocenę homogeniczności klinicznej i metodologicznej (w literaturze określane również jako ocenę podobieństwa – ang. similarity).

Wybór badań włączanych do sieci powinien opierać się na rygorystycznych kryteriach, tak aby zapewnić ich podobieństwo – [1].

Ocena ma na celu identyfikację zmiennych, które mogą wpływać na wynik analizy (ang. effect modifiers). Szczegółowej ocenie podlegać powinna charakterystyka badanej populacji, a także różnice metodologiczne w badaniach w kontekście czasu obserwacji i definicji punktów końcowych, a także czy wszystkie włączone do analizy badania oceniały określony punkt końcowy.

Nierównomierny rozkład modyfikatorów efektu pomiędzy badaniami zwiększa ryzyko popełnienia błędu wnioskowania i niespójności sieci. Założenie to dotyczy wszystkich porównań w sieci, tj. badania porównujące A z B powinny być podobne do badań porównujących B z C.

W literaturze zaleca się tabelaryczne przedstawienie porównania charakterystyki badań [13].

- 2) Ocenę homogeniczności statystycznej (ang. statistical homogeneity)

Dotyczy podobieństwa badań wewnątrz poszczególnych porównań parami.

Do badania heterogeniczności można również wykorzystać współczynnik I^2 , który określa jaki procent obserwowanej wariancji wynika z rzeczywistej różnicy w wielkości efektów (graficznie, odzwierciedla stopień nachodzenia na siebie przedziałów ufności poszczególnych badań). $I^2 > 50\%$ interpretuje się jako znaczną heterogeniczność. W przypadku zidentyfikowania znacznej heterogeniczności należy z analizy wykluczyć badania mogące ją powodować i przeprowadzić analizę wrażliwości pokazującą, w jaki sposób wyłączenie zidentyfikowanych badań wpłynie na wynik analizy.

- 3) Ocenę zwartości, spójności (ang. consistency) sieci

Dotyczy zgodności wnioskowania w sieci. Oznacza to, że niezależnie od drogi oszacowania efekt powinien być zbliżony.

Na Rysunek 6 przedstawiono przykład sieci porównującej skuteczność czterech interwencji (A, B, C i D), wraz z oszacowaniem efektu dla poszczególnych porównań bezpośrednich (wartości przy strzałkach). Oszacowany efekt może być przedstawiony zarówno w postaci różnicy ryzyk, jak i różnicy średnich. Kierunek strzałki oznacza kierunek przeprowadzonego porównania bezpośredniego. Na podstawie tak skonstruowanej sieci można porównać skuteczność A vs B (δ_{AB}) pośrednio, poprzez jeden z dwóch wspólnych komparatorów, tj. interwencję C lub D za pomocą poniższych działań:

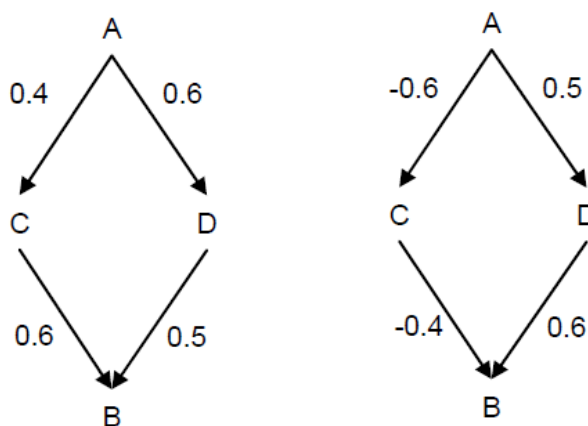
$$\delta_{AB} = \delta_{AC} + \delta_{BC} \text{ lub } \delta_{AB} = \delta_{AD} + \delta_{DB}.$$

Operując na konkretnych przykładach liczbowych z przykładu przedstawionego na rysunku poniżej, oszacowanie różnicy efektu w zależności od wybranej drogi wyniesie:

$$\delta_{AB} = 0,4 + 0,6 = 1 \text{ lub } \delta_{AB} = 0,6 + 0,5 = 1,1 \text{ w sieci z lewej strony rysunku, oraz}$$

$$\delta_{AB} = -0,6 - 0,4 = -1 \text{ lub } \delta_{AB} = 0,5 + 0,6 = 1,1. \text{ w sieci z prawej strony rysunku.}$$

Z powyższego wynika, że w sieci z lewej strony rysunku oszacowania efektu niezależnie od drogi oszacowania są ze sobą zbieżne, natomiast w sieci z prawej strony są rozbieżne, co świadczy o niespójności w sieci (ang. inconsistency) [7]. Na rysunku poniżej przedstawiono przykład sieci spójnej (po lewej stronie rysunku) i niespójnej (po prawej stronie).



Rysunek 6. Przykład sieci spójnej (consistency) i niespójnej (inconsistency).

Źródło: HAS 2011

W przypadku zdiagnozowania niespójności (ang. inconsistency) w sieci zaleca się następujące postępowanie [13, 15]:

- 1) upewnienie się, czy nie zostały popełnione błędy przy ekstrakcji danych,
- 2) wnikliwa analiza badań włączonych do sieci w celu odnalezienia badań o odmiennej charakterystyce lub badań tworzących niespójne węzły (ang. inconsistent loop),
- 3) przeprowadzenie metaregresji w celu sprawdzenia wpływu modyfikatorów efektu na wynik analizy,
- 4) przeprowadzenie analizy wrażliwości polegającej na usunięciu z sieci badań włączonych do porównań, dla których tę niezgodność stwierdzono.

W literaturze zwraca się uwagę, że brakuje obiektywnych i zwalidowanych metod służących do oceny podobieństwa badań (similarity), a akceptowany poziom rozbieżności między badaniami opiera się na subiektywnej ocenie badaczy [15, 16, 17]. Dodatkowo podkreśla się, że metody służące do oceny homogeniczności statystycznej i oceny spójności sieci są w MTC rzadko stosowane [15, 18].

5. Wyniki przeprowadzonej analizy

5.1. Metodyka

W niniejszym opracowaniu dokonano analizy metod przeprowadzenia porównań pośrednich w analizach klinicznych załączonych do wniosków refundacyjnych dotyczących kategorii dostępności refundacyjnych: lek dostępny w aptece na receptę oraz program lekowy. W dalszej kolejności dokonano analizy stanowisk Rady Przejrzystości i rekomendacji Prezesa AOTMiT odnoszących się do powyższych wniosków.

W celu identyfikacji analiz klinicznych, w ramach których przeprowadzono porównania pośrednie, w zasobach Agencji Oceny Technologii Medycznych i Taryfikacji dokonano przeglądu wszystkich analiz weryfikacyjnych Agencji (AWA) opracowanych w latach 2016 – czerwiec 2019, a następnie wyszukano analizy kliniczne (oraz ich uzupełnienia przekazane w ramach odpowiedzi na pisma w sprawie niespełnienia wymagań minimalnych) przekazane wraz z wnioskami refundacyjnymi, które zostały zweryfikowane w odnalezionych AWA. Podkreślić należy, że przegląd dotyczył analiz weryfikacyjnych, a nie liczby zleceń. Liczba opracowań AWA Agencji jest mniejsza niż liczba Zleceń MZ, gdyż w ramach wspólnych AWA rozpatrywane były przypadki: złożenia odrębnych zleceń dla różnych prezentacji leków OT.4351.11.2017 oraz zleceń, które dotyczyły różnej odpłatności za lek w zależności od wskazań – OT.4330.16.2018.

Analizy weryfikacyjne, w których porównania pośrednie zostały opracowane lub przedstawione przez analityków Agencji, a nie zostały przedstawione w analizach wnioskodawców, nie były brane pod uwagę.

Z analiz włączonych do niniejszego opracowania ekstrahowano informacje dotyczące proponowanej kategorii dostępności refundacyjnej, problemów zdrowotnych, których dotyczyły wnioski oraz metody przeprowadzenia porównania pośredniego. Dokonano następującego podziału metod przeprowadzenia porównań pośrednich:

- zestawienie wyników – porównanie jakościowe,
- metoda Buchera,
- MAIC – skorygowane porównanie pośrednie (ang. matching-adjusted indirect comparison),
- MTC – mieszane porównanie pośrednie (ang. mixed treatment comparison) – jako MTC przyjęto wszystkie analizy zdefiniowane w analizach wnioskodawców jako MTC oraz analizy określone jako metaanalizy sieciowe z podejściem Bayesa,
- model regresji.

W ramach analizy sprawdzono również informacje dotyczące liczby komparatorów oraz ewentualnego przeprowadzenia w ramach analiz klinicznych, oprócz porównań pośrednich, porównań bezpośrednich.

W przypadku istnienia komparatora, względem którego dokonano porównania bezpośredniego, analizy kwalifikowano następująco:

- porównanie bezpośrednie dla głównego komparatora,
- porównanie bezpośrednie dla części komparatorów, w przypadku istnienia porównań z kilkoma równorzędnymi komparatorami,
- porównanie bezpośrednie dla komparatorów dodatkowych, w przypadku gdy zostały one tak określone przez wnioskodawcę lub zostały uznane w AWA za komparatory o ograniczonym znaczeniu klinicznym.

Następnie dokumenty analizowano pod kątem metodyki przeprowadzonych porównań pośrednich, w tym przeprowadzenia oceny heterogeniczności. W analizie uznawano istnienie oceny heterogeniczności, bez względu na to czy została ona szczegółowo przeprowadzona na łamach analizy, czy jedynie wskazano wnioski z niej płynące lub odniesiono się skrótowo do homogeniczności badań w podsumowaniach.

W przypadku analiz przeprowadzonych metodą metaanalizy sieciowej ekstrahowano dane dotyczące przeprowadzenia oceny dopasowania modelu. Dla przypadków wskazania w analizach klinicznych, iż ocena adekwatności dopasowania została przeprowadzona, jednak nie przedstawiono jej wyników, w niniejszej publikacji przyjęto określenie „DIC bez przedstawienia wyników”.

Analizy przeprowadzone metodą MTC i regresji zweryfikowano również w zakresie wykonania rankogramu.

Jednocześnie na stronie AOTMiT dokonano wyszukiwania stanowisk Rady Przejrzystości oraz rekomendacji Prezesa AOTMiT odnoszących się do analizowanych wniosków. Analizowano w nich rodzaj decyzji (pozytywna/negatywna) oraz sprawdzano czy i w jaki sposób odniesiono się do porównań pośrednich w uzasadnieniach decyzji.

Na potrzeby analizy jako pozytywne określono wszystkie rekomendacje i stanowiska Rady Przejrzystości, w których decyzja była pozytywna, warunkowo pozytywna, negatywna dla wskazania określonego we wniosku, jednak pozytywna dla wskazania zaproponowanego przez Radę Przejrzystości/Prezesa AOTMiT oraz w przypadku gdy rekomendacja była pozytywna dla jednego rodzaju odpłatności, jednak negatywna dla innego rodzaju odpłatności w zależności od wskazania.

W odniesieniu do analizy uzasadnień powyższych decyzji ekstrahowano informacje odnośnie przedstawienia wyników porównań pośrednich, wskazania charakteru metody jako ograniczenia analizy oraz przedstawienia szczegółowych ograniczeń samych porównań. Jako ograniczenie charakteru metody przyjęto zarówno bezpośrednie stwierdzenia o ograniczonym charakterze metody, jak i stwierdzenie, iż brak jest badań bezpośrednio porównujących analizowane technologie.

Analizę statystyczną uwzględniającą powyższe założenia wykonano przy pomocy aplikacji MS Excel.

Wyniki analizy przedstawiono zarówno zbiorczo dla wszystkich lat, jak i w podziale na lata 2016 – I połowa 2019 r. Dane zaprezentowano w postaci tabel i wykresów.

5.2. Wyniki dotyczące AWA i analiz klinicznych załączonych do wniosków refundacyjnych

Porównania pośrednie w analizach weryfikacyjnych Agencji (AWA)

W latach 2016 – I połowa 2019 r. AOTMiT przeprowadziło 217 analiz weryfikacyjnych dla produktów leczniczych, stosowanych w ramach programów lekowych i refundacji aptecznej. W 76 (35%) z nich rozpatrywano skuteczność kliniczną leków z zastosowaniem porównań pośrednich. W 2016 r. wykonano 18 analiz weryfikacyjnych, w których analizy kliniczne oparte były na porównaniu pośrednim, natomiast w 2017 i 2018 r. liczba ta wyniosła 25.

Porównania pośrednie częściej wykonywane były w analizach dotyczących leków stosowanych w programach lekowych (38% wszystkich analiz dotyczących stosowania leków w ramach programów lekowych) niż w refundacji aptecznej (28,4% wszystkich analiz dotyczących leków stosowanych w ramach refundacji aptecznej). Szczegóły prezentuje tabela poniżej.

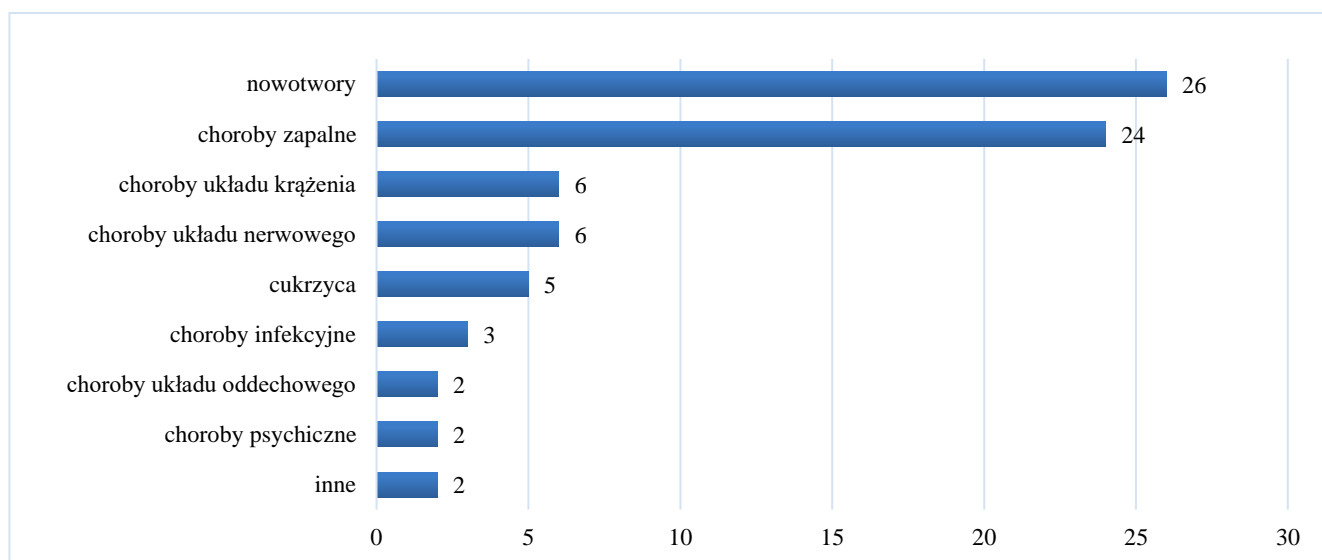
Tabela 1. Liczba wszystkich analiz weryfikacyjnych oraz analiz, w ramach których przeprowadzono porównanie pośrednie wykonanych w latach 2016 – I połowa 2019 roku.

Rok	AWA z porównaniem pośrednim/ Wszystkie AWA (%)		
	Program lekowy	Apteka	Razem
2016	14/42 (33,3)	4/19 (21,0)	18/61 (29,5)
2017	20/52 (38,5)	5/22 (22,7)	25/74 (33,8)
2018	18/43 (41,9)	7/20 (35,0)	25/63 (39,7)
I połowa 2019	5/13 (38,5)	3/6 (50,0)	8/19 (42,1)
Razem	57/150 (38,0)	19/67 (28,4)	76/217 (35,0)

Źródło: opracowanie własne

Wskazania, dla których przeprowadzono porównania pośrednie

Do wskazań, dla których najczęściej przeprowadzano porównania pośrednie we wnioskach analizowanych przez AOTMiT w latach 2016 – I połowa 2019 r. należały: nowotwory (26 z 76 analiz (34,5%)) i choroby zapalne (m.in. łuszczyca plackowata i łuszczykowe zapalenie stawów) (24 z 76 analiz (31,5%)). Rzadziej porównania pośrednie przeprowadzono dla wskazań: choroby układu krążenia i choroby układu nerwowego (m.in. stwardnienie rozsiane, padaczka) – po 6 analiz. Szczegółowe zestawienie wskazań pochodzących z analizowanych wniosków prezentuje wykres poniżej.



Rysunek 7. Wskazania, dla których przeprowadzono porównania pośrednie w latach 2016 – I połowa 2019 r.

Źródło: opracowanie własne

Porównania bezpośrednie w analizach zawierających porównania pośrednie

Spośród 76 analiz, w ramach których dokonano porównania pośredniego, do 55 włączono więcej niż 1 komparator. W 34 spośród 55 analiz zaprezentowano wyniki porównania bezpośredniego. W 6 analizach porównanie bezpośrednie było przeprowadzone dla głównego komparatora, natomiast w 28 dla części równorzędnych komparatorów i komparatorów dodatkowych. Najczęściej porównanie bezpośrednie było przeprowadzone w przypadku istnienia równorzędnych komparatorów (24 z 34). Szczegółowe dane przedstawiono w tabeli poniżej.

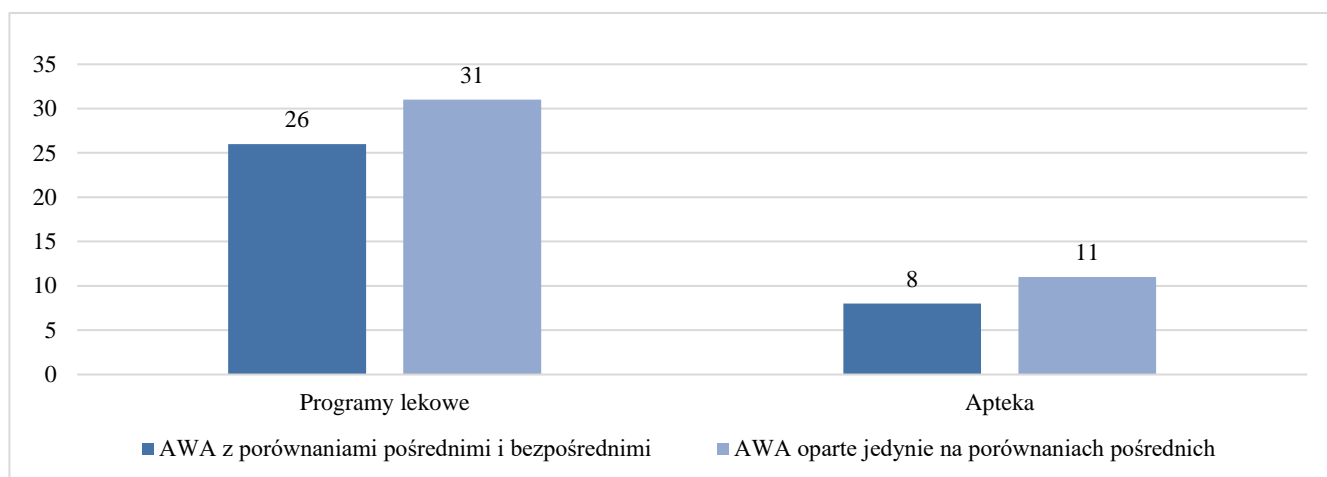
W dwóch przypadkach przeprowadzenia porównania bezpośredniego dla części komparatorów, wybrany przez wnioskodawcę komparator nie był odpowiedni, ze względu na brak jego refundacji w Polsce (dane nie przedstawione w tabeli).

Tabela 2. Liczba analiz, w których obok porównania pośredniego przeprowadzono porównanie bezpośrednie z podziałem na lata 2016 – I połowa 2019 r.

Rok	Więcej niż 1 komparator	Porównanie bezpośrednie	Porównanie bezpośrednie		
			Z komparatorem głównym	Z częścią komparatorów	Z komparatorem dodatkowym
2016	17	12	2	9	1
2017	17	11	1	10	0
2018	17	9	3	5	1
I połowa 2019	4	2	0	0	2
Razem	55	34	6	24	4

Źródło: opracowanie własne

W 31 z 57 (54,4%) analiz dotyczących refundacji leków w ramach programach lekowych i 11 z 19 (57,9%) analiz dotyczących wniosków o refundację leków w refundacji aptecznej analizy kliniczne oparte były jedynie na porównaniu pośrednim. Odsetek analiz, w których przeprowadzono jedynie porównania pośrednie, był w obydwóch przypadkach zbliżony.

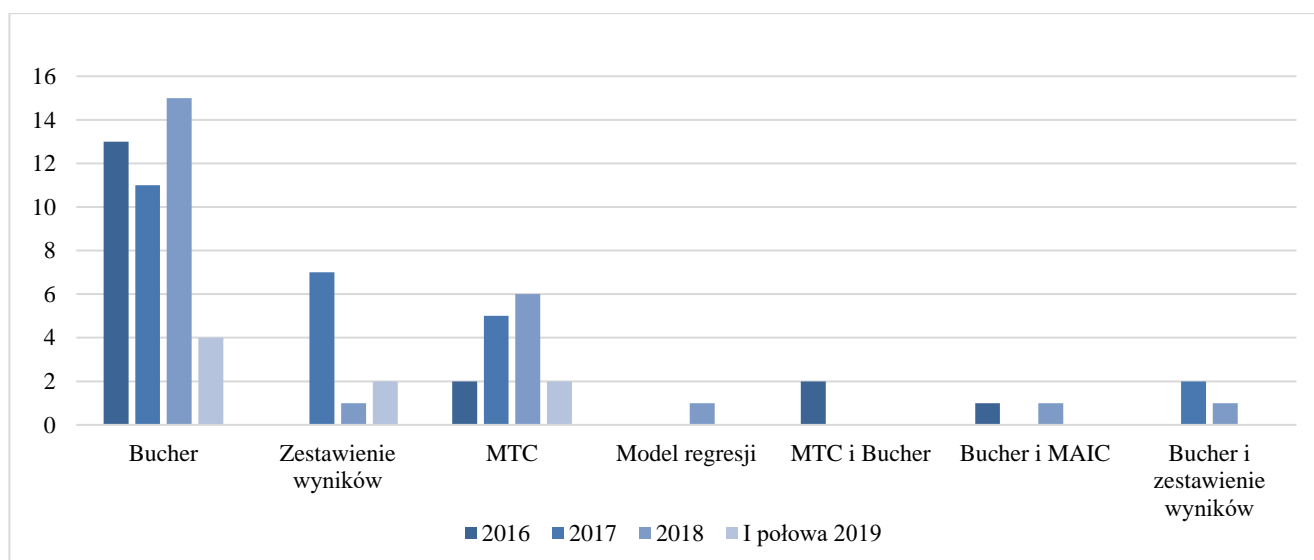


Rysunek 8. Porównania pośrednie i bezpośrednie w analizach klinicznych dotyczących refundacji leków w ramach programów lekowych i refundacji aptecznej w latach 2016 – I połowa 2019 r.

Źródło: opracowanie własne

Metody przeprowadzonych analiz pośrednich we wnioskach refundacyjnych wpływających do AOTMiT

Najczęściej wykorzystywaną metodą przeprowadzania porównań pośrednich we wnioskach refundacyjnych we wszystkich analizowanych latach była metoda Buchera (od 13 analiz w 2017 r. do 17 analiz w 2018 r.). Zestawienie wyników, czyli analiza jakościowa była przeprowadzana stosunkowo często w 2017 r. (7 analiz), jednak jej wykorzystanie uległo zmniejszeniu w kolejnych latach. W 2016 r. nie analizowano tego typu porównania w żadnym z wniosków. Zastosowanie metody MTC wzrastało stopniowo od 2016 r. W pierwszym analizowanym roku liczba analiz przeprowadzonych tą metodą wynosiła 2, w 2017 r. przeprowadzono 5 takich analiz, natomiast w 2018 r. – 6 analiz. Model regresji został zastosowany tylko w 1 analizie na przestrzeni 3,5 roku. W niewielkim odsetku przedłożonych analiz (9%) porównanie pośrednie przeprowadzono z zastosowaniem więcej niż jednej metody. Porównania te były przeprowadzane dla tych samych komparatorów, a zastosowanie więcej niż jednej metody wynikało z ograniczonej dostępności danych niezbędnych do przeprowadzenia poszczególnych typów analiz. Nadmienić należy, że w każdym z powyższych przypadków jedną z wykorzystanych metod analitycznych była metoda Buchera.



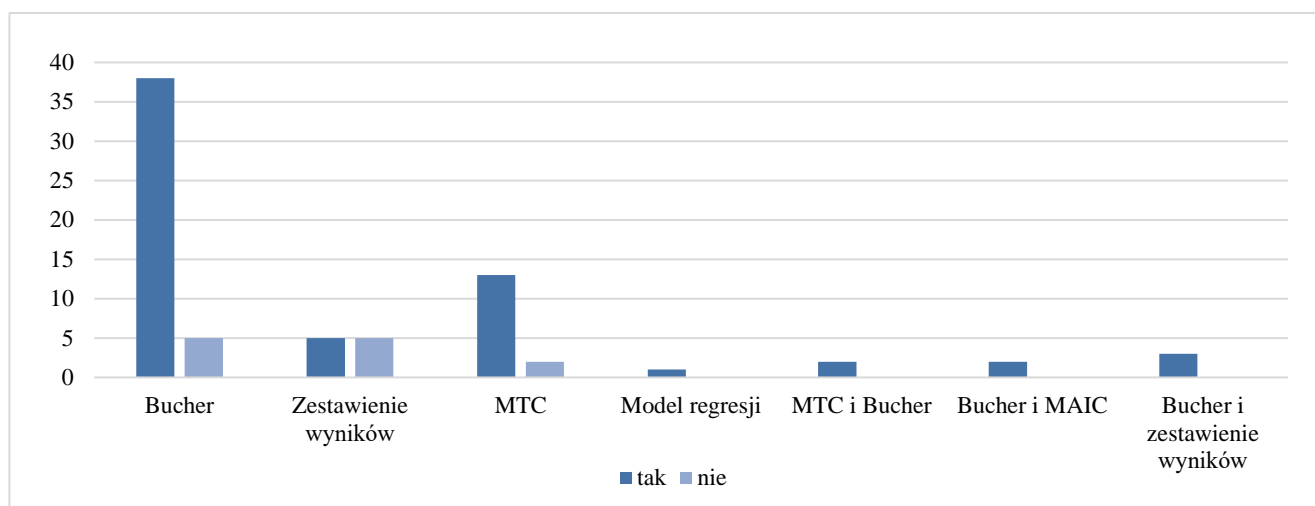
Rysunek 9. Rodzaje i liczba porównań pośrednich przeprowadzonych w latach 2016 – I połowa 2019 r.

Źródło: opracowanie własne

Ocena heterogeniczności badań włączonych do analiz

Spośród 76 analiz zawierających porównanie pośrednie, ocenę heterogeniczności przeprowadzono w 64 (84,2%). Analizę heterogeniczności najrzadziej wykonywano w przypadku przeprowadzenia porównań jakościowych w postaci zestawienia wyników – jedynie w 50% takich analiz. Szczegóły przedstawiono na wykresie poniżej.

W ramach niniejszej analizy nie przeprowadzono oceny sposobu i jakości przeprowadzonych ocen heterogeniczności.



Rysunek 10. Ocena heterogeniczności badań w analizach klinicznych wnioskodawców, w których przeprowadzono porównania pośrednie w latach 2016 – I połowa 2019 r. – dane dla całego analizowanego okresu.

Źródło: opracowanie własne

Ocena dopasowania modelu

W ramach prac nad publikacją sprawdzono wykorzystanie w porównaniach pośrednich przeprowadzonych metodą MTC i z zastosowaniem modelu regresji oceny dopasowania modelu. We wszystkich przypadkach oceny miary dopasowania estymowanych modeli statystycznych do rzeczywistości dokonano z zastosowaniem parametru DIC (ang. *Deviance Information Criterion*).

Wyniki analizy wskazują, iż w 9 z 18 analiz przeprowadzonych ww. metodami przedstawiono szczegółowo wartości osiągnięte przez parametr DIC, a tym samym wskazano uzasadnienie dla przyjętego w analizach modelu. W 7 analizach wskazano, iż sprawdzono dopasowanie modelu z zastosowaniem parametru DIC, nie wskazano natomiast szczegółowych wyników przeprowadzonej analizy. W dwóch analizach nie zamieszczono informacji dotyczących oceny dopasowania zastosowanych modeli.

Tabela 3. Zastosowanie parametru DIC w ocenie dopasowania modelu w porównaniach pośrednich analizowanych przez AOTMiT w latach 2016 – I połowa 2019 r. – dane dla całego analizowanego okresu.

Metoda	Liczba analiz łącznie	DIC	DIC bez przedstawienia wyników	Brak
MTC	15	8	6	1
MTC i Bucher	2	0	1	1
Model regresji	1	1	0	0
Razem	18	9	7	2

Źródło: opracowanie własne

Rankogram

Jedynie w 4 z 18 badanych analiz, w ramach których przeprowadzono porównanie pośrednie metodą MTC lub regresji przedstawiono rankogram.

5.3. Wyniki dotyczące stanowisk Rady Przejrzystości i rekomendacji Prezesa AOTMiT

Stanowiska Rady Przejrzystości (RP)

Analiza stanowisk RP wydanych dla wniosków rozpatrywanych w latach 2016 – I połowa 2019 r., w ramach których przeprowadzono porównania pośrednie wykazała, iż spośród 76 stanowisk, 49 stanowiły stanowiska pozytywne, natomiast 27 negatywne. W 55 (72,4%) stanowiskach odwołano się do przeprowadzonego porównania pośredniego. Rada Przejrzystości odnosiła się również często do porównań pośrednich w przypadku analiz, w ramach których

przeprowadzono wyłącznie porównanie pośrednie, jak i analiz, w których oprócz porównania pośredniego wykonano porównanie bezpośrednie (73,8% vs. 70,6%). Różnice zauważalne są jednak w przypadku rozróżnienia pozytywnych i negatywnych decyzji. W przypadku decyzji odnoszących się do analiz klinicznych, w których przeprowadzono wyłącznie porównanie pośrednie RP częściej odnosiła się do porównań pośrednich w przypadku wydania pozytywnej decyzji niż negatywnej (76,6% vs. 66,7%). W przypadku analiz, w których oprócz porównań pośrednich dostępne były porównania typu „head to head” do porównań pośrednich odwołano się w 68,4% stanowisk pozytywnych i 73,3% stanowisk negatywnych. Szczegóły zaprezentowano w tabeli poniżej.

Tabela 4. Stanowiska RP wydane dla zleceń weryfikowanych przez AOTMiT w latach 2016 – I połowa 2019 r. dla wniosków, w ramach których wykonano porównanie pośrednie.

	Stanowisko pozytywne	Stanowisko negatywne	Razem
AWA, w których przeprowadzono porównanie pośrednie	49	27	76
Odwołanie do porównania pośredniego/ AWA z porównaniami pośrednimi (%)	36/49 (73,5%)	19/27 (70,4%)	55/76 (72,4%)
Odwołanie w stanowiskach RP do porównania pośredniego w analizach bez porównań bezpośrednich/ liczba analiz bez porównań bezpośrednich (%)	23/30 (76,6%)	8/12 (66,7%)	31/42 (73,8%)
Odwołanie w stanowiskach RP do porównań pośrednich w analizach z porównaniami bezpośrednimi/ liczba analiz z porównaniami bezpośrednimi (%)	13/19 (68,4%)	11/15 (73,3%)	24/34 (70,6%)

Źródło: opracowanie własne

W prawie wszystkich stanowiskach, w których odwołano się do porównań pośrednich, przedstawiono wyniki przeprowadzonych porównań. Na ograniczony charakter metody wskazano w 16 stanowiskach; w zbliżonym odsetku stanowisk pozytywnych i negatywnych (27,8% vs. 31,6% stanowisk, w których odniesiono się do porównań pośrednich). Jedynie w 6 stanowiskach RP odniosła się do poszczególnych ograniczeń analizowanych porównań pośrednich. Jako ograniczenia porównań pośrednich wymieniano: heterogeniczność badań w zakresie populacji, metodyki, punktów końcowych i okresów ich oceny, leczenia towarzyszącego dozwolonego w badaniach oraz różnic w wynikach uzyskiwanych w badaniach włączanych do porównań w grupach otrzymujących placebo. Poszczególne ograniczenia porównań wymieniano częściej w stanowiskach negatywnych niż pozytywnych (21,1% vs. 5,6% stanowisk, w których odniesiono się do porównań pośrednich). Szczegóły przedstawiono poniżej.

Tabela 5. Sposób odniesienia się do porównań pośrednich w stanowiskach RP wydanych dla analiz weryfikowanych przez Agencję w latach 2016 – I połowa 2019 r.

Sposób odniesienia się do porównania pośredniego	Stanowisko pozytywne (%)	Stanowisko negatywne (%)	Razem (%)
Odwołanie do porównania pośredniego:	36 (100,0)	19 (100,0)	55 (100,0)
– Przedstawienie wyników	36 (94,7)	18 (98,2)	54 (98,2)
– Wskazanie charakteru metody jako ograniczenia	10 (27,8)	6 (31,6)	16 (29,1)
– Wskazanie poszczególnych ograniczeń porównania	2 (5,6)	4 (21,1)	6 (10,9)

Źródło: opracowanie własne

Rekomendacje Prezesa AOTMiT

Spośród 76 rekomendacji Prezesa 46 było pozytywnych, natomiast 30 negatywnych. W 68 (89,5%) dokumentach odwołano się do porównania pośredniego. W rekomendacjach Prezesa AOTMiT częściej odwoływano się do porównań pośrednich w przypadku analiz, w których stanowiły one jedyne źródło informacji o skuteczności klinicznej w porównaniu do analiz, w których przeprowadzono porównanie bezpośrednie (97,6% vs. 79,4%). Ponadto niezależnie od obecności porównania bezpośredniego do porównań pośrednich częściej odwoływano się rekomendacjach pozytywnych. Szczegółowe zestawienie uzyskanych wyników zamieszczono w tabeli poniżej.

Tabela 6. Rekomendacje Prezesa AOTMiT wydane dla zleceń weryfikowanych w latach 2016 – I połowa 2019 r. dla wniosków, w ramach których wykonano porównanie pośrednie.

	Rekomendacja pozytywna	Rekomendacja negatywna	Razem
AWA, w których przeprowadzono porównanie pośrednie	46	30	76
Odwołanie do porównania pośredniego/ AWA z porównaniami pośrednimi (%)	43/46 (93,5%)	25/30 (83,3%)	68/76 (89,5%)
Odwołanie w rekomendacji do porównania pośredniego w analizach bez porównań bezpośrednich/ liczba analiz bez porównań bezpośrednich (%)	28/28 (100%)	13/14 (92,9%)	41/42 (97,6%)
Odwołanie w rekomendacji do porównania pośredniego w analizach z porównaniami bezpośrednimi/ liczba analiz z porównaniami bezpośrednimi (%)	15/18 (83,3%)	12/16 (75,0%)	27/34 (79,4%)

Źródło: opracowanie własne

Większość odwołań do porównań pośrednich w uzasadnieniach rekomendacji Prezesa polegało na przedstawieniu ich wyników. Na ograniczony charakter porównania pośredniego wskazano w 66,2% rekomendacji, przy czym stosunkowo częściej ograniczenie to wymieniano w rekomendacjach pozytywnych niż w negatywnych (79,1% vs. 44,0% rekomendacji, w których odniesiono się do porównań pośrednich). Ograniczenia przeprowadzonych porównań pośrednich częściej wymieniano natomiast w rekomendacjach negatywnych niż pozytywnych (36,0% vs. 20,9% rekomendacji, w których odniesiono się do porównań pośrednich). Ograniczenia porównań pośrednich wymieniane w uzasadnieniach rekomendacji Prezesa AOTMiT obejmowały: ogólną heterogeniczność badań, heterogeniczność kliniczną, heterogeniczność w zakresie populacji, metodyki, definicji punktów końcowych, w czasie oceny punktów końcowych oraz brak dostępności danych dla części analizowanych punktów końcowych.

Tabela 7. Sposób odniesienia się do porównań pośrednich w rekomendacjach Prezesa Agencji wydanych dla analiz weryfikowanych w latach 2016 – I połowa 2019 r.

Sposób odniesienia się do porównania pośredniego	Rekomendacja pozytywna	Rekomendacja negatywna	Razem
Odwołanie do porównania pośredniego:	43 (100,0)	25 (100,0)	68 (100,0)
– Przedstawienie wyników	42 (97,7)	24 (96,0)	66 (97,1)
– Wskazanie charakteru metody jako ograniczenia	34 (79,1)	11(44,0)	45 (66,2)
– Wskazanie poszczególnych ograniczeń porównania	9 (20,9)	9 (36,0)	18 (26,5)

Źródło: opracowanie własne

Rekomendacje Prezesa AOTMiT vs Stanowiska RP

Analizując powyższe dane należy zauważyć, iż rekomendacje Prezesa częściej odwoływały się do porównań pośrednich niż Stanowiska RP (68 vs. 55). W rekomendacjach Prezesa częściej przedstawiano wyniki porównań (66 vs. 54), wskazano na ograniczony charakter porównań pośrednich (45 vs. 16) oraz wymieniano ograniczenia porównań pośrednich przeprowadzanych w analizach (18 vs. 6).

Na uwagę zasługuje również fakt, iż na ograniczony charakter porównań pośrednich częściej wskazywano w pozytywnych rekomendacjach Prezesa AOTMiT niż w negatywnych (79,1% vs. 66,2% decyzji). Natomiast w przypadku Stanowisk RP na ograniczony charakter metody nieznacznie częściej wskazywano w przypadku decyzji negatywnych (31,6% vs. 27,8%).

Zarówno w pozytywnych, jak i negatywnych rekomendacjach Prezesa w uzasadnieniu wyszczególniano ograniczenia porównań pośrednich (26,5% vs. 20,9%). Tymczasem w przypadku Stanowisk RP znacznie częściej ograniczenia wymieniano w rekomendacjach negatywnych niż pozytywnych (21,1% vs. 5,6%).

6. Dyskusja i wnioski

W ciągu ostatnich 20 lat dokonał się znaczny postęp w możliwościach przeprowadzania porównawczej oceny skuteczności i bezpieczeństwa terapii lekowych. Znana od 1997 r. metoda Buchera pozwala na porównanie skuteczności technologii, dla których nie ma dostępnych porównań bezpośrednich. Natomiast najnowsze techniki MTC umożliwiają jednoczesne przeprowadzenie analiz na podstawie danych z porównań bezpośrednich, jak i pośrednich. Rozwój powyższych metod wpływa na uproszczenie i przyspieszenie procesu refundacji leków, dzięki uniknięciu konieczności przeprowadzania kosztownych RCT. Warto jednak pamiętać, że nie są to techniki pozbawione wad. Przede wszystkim wyniki porównań pośrednich mogą być zaburzone poprzez wysoką heterogeniczność kliniczną, metodologiczną i statystyczną, a dodatkowo przez brak spójności (ang. inconsistency). W publikacji Reken 2016 wskazano, że nie ma jasnych wytycznych dotyczących przesłanek jakie muszą być spełnione, aby można było odstąpić od wykonania MTC oraz dotyczących postępowania w przypadku braku możliwości identyfikacji przyczyny niespójności sieci. Dodatkowo jako ograniczenie metaanaliz sieciowych należy podkreślić ryzyko braku transparentności związane ze złożonością metodologiczną [15].

Problemem ze stosowaniem powyższych metod w warunkach polskich jest również brak zaleceń dotyczących sposobu przeprowadzania i prezentowania wyników porównań pośrednich. Polskie Wytyczne HTA z 2016 r. wskazują jedynie preferowane metody ich przeprowadzenia. Wobec faktu rosnącego zainteresowania i wykorzystania porównań pośrednich niezbędnym wydaje się więc rozszerzenie rekomendacji w tym zakresie.

Pomimo licznych niepewności związanych ze stosowaniem porównań niebezpośrednich porównania takie są przeprowadzane w analizach klinicznych załączanych do wniosków o refundację leków zarówno w ramach programów lekowych, jak i w ramach refundacji aptecznej. Przeprowadzona w ramach niniejszej publikacji analiza metod przeprowadzenia porównań w analizach klinicznych załączonych do wniosków refundacyjnych w latach 2016 – I połowa 2019 r. wykazała, iż 35% (76 z 217) z nich zostało przeprowadzonych z zastosowaniem porównań pośrednich. Do najczęstszych metod wykorzystywanych w powyższych analizach należała metoda Buchera. Zaobserwowano również rosnące wykorzystanie metody MTC i brak metaanaliz sieciowych Lumley'a. Wyniki te nie są zbieżne z analizą przeprowadzoną przez EUnetHTA w 2015 r., w ramach której zidentyfikowano 40 badań, w których zastosowano MTC Bayesa, co według autorów analizy czyniło go najpopularniejszą metodologią. W ramach publikacji EUnetHTA zidentyfikowano 27 i 24 badania z zastosowaniem odpowiednio metaanalizy sieciowej Lumley'a i skorygowanego porównania pośredniego Buchera. Zidentyfikowano również jeden artykuł, w którym zastosowano naiwne podejście do porównania pośredniego. W publikacji EUnetHTA stwierdzono, iż metaanaliza sieciowa Lumleya i Bayesian MTC wydają się zyskiwać na popularności, podczas gdy skorygowana metoda porównywania pośredniego Buchera wydaje się już nie być powszechnie stosowana [6]. Inne niż w raporcie EUnetHTA wnioski płynące z niniejszej analizy mogą wynikać z faktu, iż metoda Buchera pozostaje metodą preferowaną przez agencje oceny technologii medycznych w stosunku do innych typów porównań niebezpośrednich. Również w wytycznych EUnetHTA podkreślono, że chociaż wyniki analiz uzyskane za pomocą metody Buchera i MTC są zbieżne, to przewagą metody Buchera jest jej prostota i związana z nią transparentność analizy.

Należy również zauważyć, iż w ok. 45% analiz, opartych na porównaniu pośrednim przeprowadzono również porównanie bezpośrednie, co jest zgodne z wytycznymi zagranicznymi, wskazującymi, iż porównanie pośrednie może mieć charakter uzupełniający w przypadku dostępu do RCT.

Analiza wskazań, dla których przeprowadzono porównanie pośrednie w ciągu ostatnich 3,5 roku jednoznacznie wskazała, iż najczęściej porównanie pośrednie stanowi podstawę podejmowania decyzji w chorobach takich jak: nowotwory i choroby zapalne, a rzadziej dla chorób układu krążenia i chorób układu nerwowego. Wyniki te są częściowo zbieżne z publikacją Tonin 2017, w której wskazano, iż spośród 360 metaanaliz sieciowych opublikowanych w ponad 30 krajach najczęściej analizowanymi wskazaniem w ramach porównań pośrednich były: choroby serca, nowotwory, choroby psychiczne i choroby infekcyjne. Należy zauważyć, iż według danych AOTMiT choroby infekcyjne i psychiczne nie są częstymi problemami zdrowotnymi rozpatrywanymi w ramach porównań pośrednich. Może to jednak wynikać z faktu mniejszej ilości wniosków refundacyjnych ocenianych w AOTMiT w tych obszarach terapeutycznych.

W publikacji Song 2009 jako błędy metodologiczne w porównaniach pośrednich oceniających skuteczność technologii medycznych wymieniono: brak dyskusji dotyczącej podobieństwa badań włączonych do analiz (metody sprawdzenia tego założenia i ewentualnej poprawy obejmowały najczęściej subiektywną ocenę charakterystyk badań, rzadziej analizę w subgroupach lub metaregresję). Błędy w metodyce zidentyfikowane przez Song 2009 obejmowały:

- niepełne zrozumienie założeń poszczególnych metod porównań pośrednich,
- zastosowanie nieodpowiednich metod porównań pośrednich,
- niepoprawnie przeprowadzony przegląd systematyczny,
- brak obiektywnych i zwalidowanych metod służących do oceny podobieństwa badań i postępowania w celu zidentyfikowania braku podobieństwa,

- nieodpowiednie wykorzystanie dowodów pośrednich i bezpośrednich.

Uproszczona analiza metodyki porównań pośrednich przeprowadzona w niniejszej publikacji wykazała, iż ocena heterogeniczności badań nie była przeprowadzona w ok. 15% dostarczonych wniosków. Przede wszystkim ocena ta była pomijana w analizach zawierających naiwne porównanie pośrednie. W przypadku MTC dokonano również oceny sprawdzenia dopasowania modelu. W większości analiz (16 z 18) wskazano na jej przeprowadzenie z zastosowaniem parametru DIC, niemniej jedynie w 9 przedstawiono szczegółowo jej wyniki. Tym samym w przypadku siedmiu analiz zidentyfikowano niewystarczające raportowanie w tym zakresie, a dodatkowo w dwóch całkowity jego brak. Ponadto zaobserwowano, iż w wielu analizach zbyt mała waga przykładana była do omówienia założeń wykonanych porównań pośrednich. Heterogeniczność badań przeprowadzona była jedynie w sposób subiektywny i w wielu przypadkach ograniczała się do kilku zdań podsumowania. W analizach nie przeprowadzano oceny heterogeniczności statystycznej w porównaniach parami. W analizach MTC nie zidentyfikowano również ocen niespójności, których znaczenie podkreślane jest w literaturze przedmiotu. Według wytycznych NICE zidentyfikowane niespójności w sieci powinny być wyjaśniane, a wszelkie wątpliwości powinny zostać przetestowane w ramach analizy wrażliwości. Powyżej wymienione niedociągnięcia metodyki mogą wynikać z braku jasnych zasad prezentacji porównań pośrednich w ramach analiz HTA.

Przegląd stanowisk RP wydanych dla wniosków rozpatrywanych w latach 2016 – I połowa 2019 r., w ramach których przeprowadzono porównania pośrednie wykazała, iż spośród 76 stanowisk, 49 stanowiły stanowiska pozytywne, natomiast 27 negatywne. W tym samym czasie wydano 46 pozytywnych i 30 negatywnych rekomendacji Prezesa AOTMiT. rekomendacje Prezesa częściej odwoływały się do porównań pośrednich niż Stanowiska RP (68 vs. 55), a także częściej przedstawiano w nich wyniki porównań pośrednich (66 vs. 54), wskazywano na ograniczony charakter metod (45 vs. 16) oraz wymieniano poszczególne ograniczenia porównań (18 vs. 6). Na ograniczony charakter porównań pośrednich częściej wskazywano w pozytywnych rekomendacjach Prezesa AOTMiT niż w negatywnych (79,1% vs. 66,2% decyzji). Natomiast w przypadku Stanowisk RP na ograniczony charakter metody nieznacznie częściej wskazywano w przypadku decyzji negatywnych (31,6% vs. 27,8%). Należy również zauważyć, iż Rada odwoływała się do porównań pośrednich równie często w przypadku obecności w analizach badań typu „head to head”, jak i ich braku. Natomiast rekomendacje Prezesa częściej odwoływały się w uzasadnieniu do porównań pośrednich w przypadku braku w analizie klinicznej badań bezpośrednio porównujących analizowane technologie.

Wyniki niniejszej analizy wskazują na potrzebę opracowania jasnych zasad przeprowadzania i prezentacji porównań pośrednich, szczególnie z zastosowaniem najnowszych metod MTC oraz wnioskowania na ich podstawie o efektywności technologii lekowych.

7. Bibliografia

1. Tonin 2017 - Tonin FS, Rotta I, Mendes AM, Pontarolo R. Network meta-analysis: a technique to gather evidence from direct and indirect comparisons. *Pharmacy Practice* 2017 Jan-Mar;15(1):943. <https://doi.org/10.18549/PharmPract.2017.01.943>
2. Sobieraj 2013 - Sobieraj DM, Cappelleri JC, Baker WL, et al. Methods used to conduct and report Bayesian mixed treatment comparisons published in the medical literature: a systematic review. *BMJ Open* 2013;3:e003111. doi:10.1136/bmjopen-2013-003111.
3. Shim 2017 - Sungryul Shim, Byung-Ho Yoon, In-Soo Shin, Jong-Myon Bae. Network meta-analysis: application and practice using Stata. Received: Sep 6, 2017 / Accepted: Oct 27, 2017 / Published: Oct 27, 2017. This article is available from: <http://e-epih.org/>
4. NICE 2013 - Guide to the methods of technology appraisal. National Institute for Health and Care Excellence appraisal 2013. nice.org.uk/process/pmg9
5. IQWiG "General Methods 5.0" - Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. 10 July 2017
6. EUnetHTA 2015 - European Network for Health Technology Assessment. Guideline "Direct and indirect comparisons". Adapted version (2015)
7. HAS 2011 - Indirect comparisons Methods and validity. Haute Autorité de Santé. July 2009
8. Kim 2013 - Kim H, Gurrin L, Ademi Z, Liew D. Overview of methods for comparing the efficacies of drugs in the absence of head-to-head clinical trial data. *Br J Clin Pharmacol*. 2013; DOI:10.1111/bcp.12150
9. Edwards 2009 - S. J. Edwards, M. J. Clarke, S. Wordsworth, J. Borrill. Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. 2009 Blackwell Publishing Ltd *Int J Clin Pract*, June 2009, 63, 6, 841–854. doi: 10.1111/j.1742-1241.2009.02072.x
10. Phillippo 2018 - David M. Phillippo, Anthony E. Ades, Sofia Dias, Stephen Palmer, Keith R. Abrams, and Nicky J. Welton. Methods for Population-Adjusted Indirect Comparisons in Health Technology Appraisal. *Medical Decision Making* 2018, Vol. 38(2) 200–211. DOI: 10.1177/0272989X17725740.
11. Lu 2004 - G. Lu, A. E. Ades. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. 2004;23(20):3105-3124.
12. Jansen 2014 - Jeroen P. Jansen, et al. Indirect Treatment Comparison/Network Meta-Analysis Study Questionnaire to Assess Relevance and Credibility to Inform Health Care Decision Making: An ISPOR-AMCP-NPC GoodPractice Task Force Report, *Value in health* 17 (2014) 157–173
13. Rouse 2017 - Benjamin Rouse, MHS, Anna Chaimani, PhD, Tianjing Li, MD, MHS, PhD. Network Meta-Analysis: An Introduction for Clinicians. *Intern Emerg Med*. 2017 February ; 12(1): 103–111. doi:10.1007/s11739-016-1583-7.
14. Mbuagbaw 2017 - Mbuagbaw et al. Approaches to interpreting and choosing the best treatments in network meta-analyses. *Systematic Reviews* (2017) 6:79. DOI 10.1186/s13643-017-0473-z
15. Reken 2016 - Reken S, Sturtz S, Kiefer C, Böhler Y-B, Wieseler B (2016) Assumptions of Mixed Treatment Comparisons in Health Technology Assessments - Challenges and Possible Steps for Practical Application. *PLoS ONE* 11(8): e0160712. doi:10.1371/journal.pone.0160712
16. Song 2009 - Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ*. 2009; 338:b1147. doi: 10.1136/bmj.b1147 PMID: 19346285; PubMed Central PMCID: PMC2665205.
17. Fu 2010 - Fu R, Gartlehner G, Grant M, Shamliyan T, Sedrakyan A, Wilt T, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program: AHRQ; 2010 [11.03.2015]. 55: Available: <http://effectivehealthcare.ahrq.gov/>.
18. Donegan 2013 - Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. *Res Synth Methods*. 2013; 4(4):291–323. doi: 10.1002/jrsm.1085 PMID: 26053945.
19. <https://www.statystyka.az.pl/estymacja-bayesowska.php>, dostęp 26.08.2019

Inne:

- Ustawa z dnia 12 maja 2011 r. o refundacji leków, środków spożywczych specjalnego przeznaczenia żywieniowego oraz wyrobów medycznych (Dz. U. z 2017 r., poz. 1844 z późn. zm.).
- Ustawa z dnia 27 sierpnia 2004 r. o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych (Dz.U. z 2018 r., poz. 1510, z późn. zm.).
- Rozporządzenie Ministra Zdrowia z dnia 18 grudnia 2013 r. w sprawie sposobu i procedur przygotowania analizy weryfikacyjnej AOTMiT Oceny Technologii Medycznych oraz wysokości opłaty za tę analizę (Dz. U. z 2014 r., poz. 4).
- Rozporządzenie Ministra Zdrowia z dnia 2 kwietnia 2012 r. w sprawie minimalnych wymagań, jakie muszą spełniać analizy uwzględnione we wnioskach o objęcie refundacją i ustalenie urzędowej ceny zbytu oraz podwyższenie urzędowej ceny zbytu leku, środka spożywczego specjalnego przeznaczenia żywieniowego, wyrobu medycznego, które nie mają odpowiednika refundowanego w danym wskazaniu (Dz. U. z 2012 r., poz. 388).
- Wytyczne przeprowadzania Oceny Technologii Medycznych (HTA). Załącznik do Zarządzenia nr 40/2016 Prezesa AOTMiT Oceny Technologii Medycznych i Taryfikacji z dnia 13 września 2016 r. w sprawie wytycznych oceny świadczeń opieki zdrowotnej http://www.aotm.gov.pl/www/wp-content/uploads/wytyczne_hta/2016/20160913_Wytyczne_AOTMiT.pdf
- Dokumenty własne AOTMiT.