

NEUROLOGIA

Do analiz pośrednich należy podejść z ostrożnością

O porównaniach pośrednich – czym są, kiedy się je stosuje i czy można je uznać za wiarygodne w przypadku chorób rzadkich, w tym rdzeniowego zaniku mięśni (*spinal muscular atrophy* – SMA) – opowiedział „Kurierowi Medycznemu” prof. dr hab. n. med. Marcin Czech, prezes Polskiego Towarzystwa Farmakoekonomicznego, były wiceminister zdrowia.

Co należałoby zrobić, aby móc wiarygodnie porównać skuteczność terapii stosowanych w SMA? Czy badania pośrednie w przypadku tej choroby są dobrym rozwiązaniem?

Porównania pośrednie, których historia w ocenie technologii medycznych sięga zaledwie ostatnich 20 lat, wykonuje się wtedy, kiedy nie mamy bezpośrednio zestawionych ze sobą dwóch leków w jednym badaniu klinicznym (*head-to-head*). Oczywiście idealna sytuacja byłaby wtedy, kiedy mielibyśmy dwa leki, losowy dobór pacjentów do grup, czyli randomizację, zaślepienie procesu, czyli sytuację, w której ani badacz, ani chory nie wiedziałby, jakie badanie jest wykonywane. Poza tym leki w obu grupach pacjentów byłyby stosowane równolegle, protokół dla obu grup byłby identyczny. To jest świat idealny, ale w życiu tak nie jest. W związku z tym dopuszczamy porównania pośrednie, czyli takie, w których lek A i lek B nigdy nie były ze sobą porównane, ale były np. porównane z lekiem C albo z placebo. Ten lek C i placebo są wspólną referencją dla leków A i B, które są ze sobą zestawiane. W praktyce wygląda to tak, że mamy wykonywane osobne badanie dla leku A, potem dla leku B i oba te badania są porównywane z lekiem C lub z placebo. Problem tkwi w tym, że mogą się pojawić trudności wynikające np. z różnego doboru populacji, które biorą udział w tych badaniach. Bardzo często mamy też inny okres obserwacji, inną charakterystykę populacji nie tylko jeśli chodzi o wiek czy typ SMA, lecz także o stopień zaawansowania choroby. To oczywiście wpływa na rezultaty, ponieważ np. pacjenci z mniej zaawansowaną chorobą będą mieli wyjściowo lepsze wyniki, co może się przełożyć na lepszy wynik końcowy. Poza tym w niektórych badaniach pytamy samych pacjentów czy też ich opiekunów o jakość życia (np. w badaniu SUNFISH), a odpowiedzi, które otrzymujemy, mogą być niespójne z innymi punktami końcowymi lub z wynikami innych badań.

Z jakiej metodyki korzysta się najczęściej w przypadku porównań pośrednich przeprowadzanych na potrzeby oceny technologii medycznych?

Pierwszą taką metodą była metoda Buchera z 1997 r., która może być zastosowana do badań dwuramiennych. Rozwinięte zostały również inne metody porównań bezpośrednich, np. metaanaliza sieciowa, która jest też dopuszczalna w polskich wytycznych HTA w wersji trzeciej. Tutaj podwaliny położył Lumley w 2002 r. Są też metaanalizy z podejściem Bayesa opracowane



Fot. Paryk Bydyk

w 2004 r. W literaturze po 2008 r. obserwujemy wzrost liczby publikacji dotyczących porównań pośrednich, które cieszą się coraz większym zainteresowaniem. Znalazły się one w wytycznych najważniejszych światowych agencji. Takie wytyczne, w dość wąskiej formie, przedstawiły polska Agencja Oceny Technologii Medycznych i Taryfikacji (AOTMiT), brytyjski *National Institute for Health and Care Excellence* (NICE), niemiecki *German Institute for Quality and Efficiency in Health Care* (IQWiG) czy francuska *Haute Autorité de Santé* (HAS).

Wytyczne NICE z 2013 r. rekomendują przedstawianie wyników porównań z zastosowaniem metaanalizy sieciowej, a w przypadku braku

porównań bezpośrednich dopuszczają ich stosowanie w ramach kontrolowanych badań klinicznych z randomizacją (*randomized controlled trial* – RCT). W takiej sytuacji analiza sieciowa powinna uwzględniać wszystkie terapie, przy czym wszystkie komparatory muszą być zidentyfikowane. Ponadto NICE zaleca przeprowadzenie analizy wrażliwości, czyli sprawdzenie, jakie parametry wejściowe są podatne na zmiany i jaki mają wpływ na wyniki końcowe. Podkreśla się pełne raportowanie heterogeniczności, czyli wszystkich różnic między badaniami – pomiędzy wynikami, porównaniami w parach czy pewnej niespójności rozumianej jako różnice między badaniem bezpośrednim (jeśli takie istnieje) a badaniem pośrednim. Oczywiście NICE jest bardzo ostrożna przy wyciąganiu wniosków co do skuteczności i bezpieczeństwa na podstawie porównań pośrednich.

Z kolei IQWiG dopuszcza przeprowadzenie analiz pośrednich w przypadku istnienia jednego badania o niepewnych wynikach, jeśli może to pozwolić na zmianę wnioskowania. Porównania pośrednie mogą oczywiście zwiększać precyzję szacowania efektu. Co do zasady IQWiG nie zaleca wykonania porównań pośrednich, kiedy można bezpośrednio porównać technologie. Metodycznie rekomenduje metodę Buchera i metaanalizę sieciową. Konieczne jest przeprowadzenie oceny homogeniczności czy też spójności. Z kolei francuska agencja HAS jest trochę bardziej liberalna – kiedy czas ma znaczenie, nie ma potrzeby oczekiwania na wyniki badań *head-to-head* i można zastosować porównanie mieszane, czyli *mixed treatment comparison*.

Na jakie aspekty należy zwrócić uwagę podczas wykonywania porównań pośrednich metodą MAIC?

Matching adjusted indirect comparison – MAIC – to metoda porównania pośredniego, której celem jest zniesienie heterogeniczności wynikającej z odmiennych charakterystyk populacji we włączonych do niego badaniach. Metoda MAIC wymaga dostępu do danych indywidualnych pacjentów (*individual patient data* – IPD), ponieważ w danych ogólnych mamy tylko grupy i statystyki przyporządkowane do tych grup, a potrzebujemy indywidualnych informacji. Na podstawie tych danych wykonuje się ważenia efektu interwencji w porównywalnych ramionach. Metoda MAIC może być stosowana do badań jednoramiennych lub badań bez wspólnego komparatora. Należy jednak zaznaczyć, że w polskiej agencji HTA w la-

tach 2016–2019 dominowała metoda Buchera czy *mixed treatment comparison*. Metoda MAIC jest stosunkowo rzadko stosowana i nie jest popularna w polskich warunkach.

Czy analizy pośrednie lub analizy MAIC powinny być przeprowadzane w chorobach rzadkich?

W chorobach rzadkich mamy trochę związane ręce. Jeżeli stosujemy tę metodę, powinniśmy połączyć jak najwięcej cech. Im więcej tych cech, im bardziej zharmonizujemy populację, im bardziej są one podobne w obu ramionach, tym uzyskujemy bardziej wiarygodne wyniki. W badaniach w SMA nie zawsze mamy taki komfort. Wtedy musimy się liczyć z tym, że heterogeniczność jest bardzo duża. Możemy indywidualnie przedstawić różne badania i powiedzieć, co wyszło w każdym z nich, ale nie jesteśmy w stanie zniwelować różnic pomiędzy populacjami. Wtedy wyciąganie wniosków jest ograniczone. Jeżeli próbujemy porównywać populację z medianą wieku 9 lat, jak w badaniu SUNFISH, z populacją osób starszych albo populację, gdzie *baseline*, czyli punkt startowy, jest prawie taki sam jak punkt, w którym kończy się inne badanie, to widzimy, że te popu-

”

Światowe agencje HTA wykazują dużą ostrożność, jeśli chodzi o wyciąganie wniosków z porównań pośrednich dotyczących leków stosowanych w SMA

lacje są heterogeniczne. Ta heterogeniczność jest zbyt duża, żeby zestawiać ze sobą różne analizy w sposób pośredni.

Jak inne kraje oceniły porównanie pośrednie metodą MAIC dotyczące terapii stosowanej w SMA?

W zasadzie konkluzja wszystkich agencji jest taka, że nie mamy porównania bezpośredniego w ogóle, a porównanie pośrednie jest dość słabe, ponieważ heterogeniczność pacjentów jest duża. W związku z tym trudno zestawiać wyniki i wysnuć jakieś ostateczne wnioski. Widzimy, że różne produkty inaczej działają na pacjentów. Według IQWiG nie ma dostępnych danych, które wskazują na wyższość jednej terapii nad drugą w badaniach pośrednich. Według NICE siła danych jest słaba. Argumentacje różnych światowych agencji HTA są publicznie dostępne i zachęcam do ich przesłania. Najważniejsze jest jednak to, że wszystkie one charakteryzują się dużą ostrożnością, jeśli chodzi o wyciąganie wniosków z porównań pośrednich dotyczących leków stosowanych w SMA.

Rozmawiała Agata Misiurewicz-Gabi



Fot. iStockphoto