

Treatment for Pediatric Gender Dysphoria: Review of Evidence and Best Practices

Leczenie dziecięcej dysforii płciowej: przegląd dowodów i najlepszych praktyk

Tłumaczenie: Magda Lewandowska & WOKiEm

Rozdział 6

Ograniczenia przeglądów systematycznych

Jak zauważono w podsumowaniu przeglądów systematycznych (PS), jakość dowodów na korzyści płynące z interwencji stosowanych w pediatrycznej tranzycji medycznej (PTM) jest bardzo niska. Niemniej jednak praktyka ta nadal szybko się rozprzestrzenia. „Bardzo niska” jakość dowodów na korzyści wymaga dodatkowej analizy potencjalnych szkód, aby właściwie zrozumieć stosunek ryzyka do korzyści.

Powszechnie uznaje się, że przeglądy systematyczne są skuteczniejsze w wykrywaniu korzyści niż szkód[1]. W związku z tym brak dowodów na szkody w opublikowanych badaniach nie jest równoznaczny z dowodem na brak szkód. W pediatrycznej medycynie gender (PMG) nieodpowiednie wykrywanie szkód jest potęgowane przez następujące czynniki:

- Od czasu powszechnego przyjęcia PTM minęło zbyt mało czasu;
- Krótkoterminowe badania obserwacyjne przed i po interwencji, które stanowią większość dowodów, często charakteryzują się wysokim odsetkiem „utruty dla badań kontrolnych” i nie obejmują systematycznego szukania, śledzenia ani pełnego zgłaszania szkód;
- Błąd systematyczny publikacji utrudnia rozpoznanie szkód.

W pierwszych trzech sekcjach niniejszego rozdziału rozważono, w jaki sposób czynniki te wpływają na zdolność PS do wykrywania szkód, a podsumowanie przedstawiono w sekcji 6.4.

Kompleksowa analiza szkód związanych z PTM jest niezbędna do kształtowania praktyki klinicznej. Szkody związane z operacjami usuwania zdrowych części ciała są łatwe do

rozpoznania, podobnie jak niektóre niepożądane skutki interwencji hormonalnych. Te ostatnie są wyliczone zarówno w ulotkach leków wydawanych przez Amerykańską Agencję ds. Żywności i Leków (FDA [z ang. Food and Drug Administration, przyp. tł.])[\[2\]](#), jak i w formularzach świadomej zgody lub zgody rodziców i dzieci podpisywanych przed rozpoczęciem leczenia[\[3\]](#). Stosowanie tych leków poza wskazaniami rejestracyjnymi w PMG może zwiększać ryzyko ich stosowania. W przypadku osób dorosłych dobrze wiadomo, że stosowanie tego samego leku poza wskazaniami rejestracyjnymi wiąże się ze znacznie wyższym odsetkiem działań niepożądanych, zwłaszcza gdy brakuje mocnych dowodów naukowych[\[4\]](#). Nigdy nie przeprowadzono badań klinicznych dotyczących stosowania leków blokujących dojrzewanie (BP) w celu zatrzymania normalnego dojrzewania; takie stosowanie będzie miało inny profil ryzyka niż stosowanie BP w celu tymczasowego zatrzymania przedwczesnego dojrzewania.

Dodatkowa niepewność wynika z faktu, że w badaniach po stosowaniu BP w ponad 90% przypadków następuje podawanie hormonów płci przeciwnej (HPP)[\[5\]](#). Ta de facto terapia kojarzona wiąże się z nowymi i potencjalnie poważnymi zagrożeniami (np. dotyczącymi płodności) i nigdy nie została poddana regulowanym przez FDA badaniom klinicznym w żadnej populacji. Ponadto badania pokazują, że gdy dowody potwierdzające stosowanie danego leku poza wskazaniami rejestracyjnymi są bardzo niepewne – jak w przypadku PTM – już i tak podwyższone ryzyko działań niepożądanych związanych ze stosowaniem poza wskazaniami rejestracyjnymi ulega dalszemu zwiększeniu[\[6\]](#).

6.1 Niewystarczający upływ czasu

Chociaż protokół holenderski został opublikowany po raz pierwszy 20 lat temu[\[7\]](#), dopiero w ostatnich kilku latach zgromadzono wystarczającą liczbę pacjentów, aby przeprowadzić badania pozwalające oszacować wskaźniki szkód i korzyści[\[8\]](#). Ponadto, ponieważ pojawienie się szkód może być opóźnione o dziesięć lat lub więcej, przeprowadzone stosunkowo wcześnie przeglądy systematyczne tej dziedziny medycyny mogą zaniżać te wyniki[\[9\]](#). Przeciętny pacjent w momencie obserwacji jest nadal dość młody – zazwyczaj jest to nastolatek lub osoba w wieku około dwudziestu lat. Komplikuje to wykrywanie ryzyka, ponieważ wpływ hormonów na zdrowie może być kumulacyjny (np. wpływ hormonów na choroby sercowo-metaboliczne lub nowotwory złośliwe) i może w pełni ujawnić się dopiero w późniejszym okresie życia.

6.2 Krótkoterminowe badania obserwacyjne

W przeglądzie Cass uznano dalsze poleganie na źle zaprojektowanych badaniach obserwacyjnych za „kluczowe ograniczenie w identyfikacji [...] bezpieczeństwa hormonów afirmujących tożsamość płciową u dzieci i młodzieży”[\[10\]](#). Kluczowym problemem, który utrudnia PS wykrywanie szkód, jest fakt, że całość dowodów w PMG składa się głównie z krótkoterminowych badań obserwacyjnych „przed i po”, które są szczególnie podatne na stronniczość. Badania takie koncentrują się na wykazaniu korzyści, nie informując w sposób spójny o szkodach[\[11\]](#). W przeciwieństwie do zamierzonych korzyści interwencji, które są elementem hipotezy przed rozpoczęciem badania i są badane jako wyniki pierwotne[\[12\]](#), szkody nie zawsze są przewidywane, monitorowane lub zgłaszane. W związku z tym przeglądy systematyczne są ograniczone, ponieważ mogą zidentyfikować tylko te wyniki, które zostały zgłoszone w badaniach pierwotnych[\[13\]](#).

Przykładem trudności w identyfikacji szkód jest kwestia czy BP stanowią nieszkodliwą „pauzę”, dającą „czas na zastanowienie się”[\[14\]](#), czy też znaczącą interwencję, sprzyjającą jatrogennemu utrwalaniu DP i kontynuacji HPP. Co najmniej pięć badań wskazuje na ponad 90% wskaźnik progresji od BP do HPP[\[15\]](#). Jak zauważono w przeglądzie Cass, „dane te sugerują, że blokery

dojrzewania nie dają czasu na zastanowienie się, biorąc pod uwagę, że zdecydowana większość pacjentów, którzy rozpoczynają hamowanie dojrzewania, kontynuuje przyjmowanie hormonów maskulinizujących/feminizujących, szczególnie jeśli przyjmowanie BP rozpoczyna się we wczesnym okresie dojrzewania”[16]. Ponieważ wysoki odsetek przyjmowania HPP po BP jest często wymieniany tylko w tekście badania (a nie przedstawiony w tabelach wyników lub streszczeniach), PS mogą przeoczyć tę informację na etapie selekcji. Tylko jeden PS o niskim ryzyku błędu systematycznego ocenił odsetek stosowania HPP po BP, identyfikując tylko dwa badania[17].

Istnieją inne przykłady niewłaściwego zgłaszania działań niepożądanych[18]. Poniżej przedstawiono trzy godne uwagi przykłady.

6.2.1 de Vries et al., 2011, 2014[19]

Para przełomowych holenderskich badań, które zapoczątkowały stosowanie PTM na całym świecie, może służyć jako przykład tego, jak szkody mogą nie zostać zgłoszone w badaniu. Wniosek z badania z 2014 r., że leczenie BP, HPP i chirurgiczne „prowadzi do poprawy funkcjonowania psychicznego”[20], nie skomentował tego, że co najmniej czterech z 70 młodych pacjentów doświadczyło działań niepożądanych podczas leczenia. U trzech nastolatków zdiagnozowano otyłość i cukrzycę, co wykluczyło ich z dalszego leczenia chirurgicznego, a jeden pacjent zmarł z powodu powikłań po operacji po wczesnym zahamowaniu dojrzewania[21]. Te poważne zdarzenia nie zostały zgłoszone jako formalne wyniki[22], co uniemożliwiło dyskusję na temat tego, czy były one spowodowane eksperymentalną interwencją, czy też wystąpiły przypadkowo. Niedawno opublikowane dowody sugerują, że zgon pooperacyjny był prawdopodobnie bezpośrednio związany z praktyką wczesnego hamowania dojrzewania płciowego[23]. Ponadto, ponieważ te działania niepożądane odnotowano jedynie w kontekście zmniejszenia wielkości próby pooperacyjnej, badacze mogli nie podjąć się wykrycia i zgłoszenia mniej poważnych działań niepożądanych, które nie dyskwalifikowały pacjentów z dalszego udziału w badaniu.

Niewystarczające zgłoszenie szkód w badaniu holenderskim zostało spotęgowane, gdy przegląd systematyczny wykluczył to badanie ze względu na jego niewiarygodny projekt (np. badanie łączyło fazę HPP i fazę operacyjną, co uniemożliwiło określenie, które wyniki były związane z którą interwencją)[24]. W związku z tym poważne szkody zgłoszone przypadkowo przez autorów badania zostały całkowicie pominięte w przeglądzie systematycznym ze względu na słaby projekt badania.

6.2.2 Tordoff i in., 2023[25]

Niewystarczające zgłaszanie szkód jest również widoczne w nowszych badaniach. Bardzo wpływowe badanie z 2023 r. przeprowadzone przez Seattle Children's Gender Clinic przez Tordoff i in. skupiało się wyłącznie na wynikach psychologicznych, w przypadku których postawiono hipotezę o korzyściach interwencji. Nie zbadano ani nie zgłoszono wyników dotyczących zdrowia fizycznego, w przypadku których występowanie szkód byłoby bardziej prawdopodobne.

Badanie Tordoff i in.[26] zawierało dane dotyczące zdrowia psychicznego nastolatków i młodych dorosłych w wieku 13–20 lat, którzy byli leczeni za pomocą BP lub HPP przez okres do jednego roku. Badanie twierdziło, że dostarcza „ilościowych dowodów” poprawy zdrowia psychicznego; jednak wskaźniki depresji wśród młodzieży, która rozpoczęła stosowanie BP lub HPP, pozostały niezmienione – około 6 na 10 młodych ludzi cierpiało na umiarkowaną lub ciężką depresję

zarówno przed interwencją, jak i po niej[27]. Dane te zostały przedstawione w dodatku dostępnym wyłącznie online. Wyniki badania dotyczące poprawy zdrowia psychicznego zostały wyprowadzone z odsetków prawdopodobieństwa:

o 60% mniejsze prawdopodobieństwo wystąpienia depresji i o 73% mniejsze prawdopodobieństwo wystąpienia tendencji samobójczych u pacjentów leczonych BP lub HPP w porównaniu z pacjentami nieleczonymi tymi metodami. Jednak poważne problemy metodologiczne, w tym potencjalne błędne dobranie próby i czynniki zakłócające[28], niewłaściwe modelowanie statystyczne[29], brakujące dane uczestników[30] oraz wybiórcze raportowanie[31]— budzą poważne obawy co do wiarygodności przedstawionych wyników.

Imponująco brzmiące odsetki prawdopodobieństwa zostały zatem obliczone w oparciu o coraz mniejsze mianowniki w nieleczonej grupie „kontrolnej”. Dodatkowe dane dostępne wyłącznie w Internecie ujawniły, że do końca badania w grupie „kontrolnej” pacjentów nieleczonych BP lub HPP pozostało tylko siedem uczestników[32]. Nie było również jasne, czy i w jakim stopniu tę grupę „kontrolną” można uznać za reprezentatywną dla wszystkich nieleczonych pacjentów. Wydaje się, że grupa „nieleczonych” składała się z młodzieży, której problemy zdrowia psychicznego i tożsamości płciowej ustąpiły i która porzuciła klinikę gender oraz badanie (stanowiąc 80% całej „grupy nieleczonych”); pacjentów, którzy mieli przejść do grupy leczonej po krótkiej ewaluacji; oraz 8 pacjentów (z których 7 dostarczyło dane z obserwacji w badaniach kontrolnych), którzy nadal uczęszczali do kliniki gender 12 miesięcy później, ale z nieznanых przyczyn nie rozpoczęli terapii hormonalnej[33].

W dwóch ostatnich przeglądach systematycznych oceniono ryzyko błędu systematycznego w badaniu Tordoffa i wsp. przy użyciu dwóch różnych narzędzi (skala Newcastle Ottawa i ROBINS-I)[34]. Chociaż w obu analizach stwierdzono pewne słabości metodologiczne (np. krótki okres obserwacji, wysoki wskaźnik rezygnacji), włączenie do badania młodzieży nieleczonej jako grupy kontrolnej nieco podniosło ocenę jakości badania. Systematyczne przeglądy nie zawsze są w stanie uchwycić tego typu poważne problemy związane z grupami kontrolnymi.

Niewłaściwe przedstawienie wyników badania, poważne czynniki zakłócające oraz wyjątkowo wysoki odsetek pacjentów, którzy nie kontynuowali leczenia, zostały podkreślone przez dziennikarzy naukowych i omówione w recenzowanych artykułach[35]. Pomimo tych dobrze wykazanych problemów, badanie to, opublikowane w Journal of the American Medical Association (JAMA) Open Network, nadal jest bezkrytycznie opisywane przez autorytatywne źródła jako dowód, że BP i HPP są korzystne dla zdrowia psychicznego, medycznie konieczne i ratują życie[36].

6.2.3 Chen i in., 2023[37]

W 2016r, w czasie rejestracji finansowanego przez Narodowy Instytut Zdrowia (NIH) badania: „Wpływ wczesnego leczenia medycznego na młodzież trans”[38], protokół tego badania obejmował plany oceny bezpieczeństwa BP i HPP, a badacze zobowiązali się do krótkiej obserwacji wynoszącej zaledwie 24 miesiące[39]. Po przeanalizowaniu protokołu tego badania jego projekt wydaje się nieodpowiedni do kompleksowej oceny bezpieczeństwa PTM. Brak równoległej grupy kontrolnej utrudnia przypisanie zaobserwowanych wyników dotyczących bezpieczeństwa do konkretnej interwencji, ponieważ trudno jest odróżnić skutki leczenia od normalnych zmian rozwojowych. Ponadto wyłączne poleganie na markerach biochemicznych nie pozwala uchwycić innych krytycznych wyników dotyczących bezpieczeństwa, takich jak wzrost i płodność. Dodatkowo mała wielkość próby osłabia zdolność badania do oceny rzadkich, ale klinicznie istotnych szkód, takich jak zdarzenia zakrzepowo-zatorowe. Brak

długoterminowej obserwacji jeszcze bardziej ogranicza wykrywalność opóźnionych lub skumulowanych działań niepożądanych. Te ograniczenia metodologiczne osłabiają moc wszelkich wniosków, jakie można wyciągnąć na temat ogólnego bezpieczeństwa interwencji.

W badaniu przedstawiono dane dotyczące zdrowia psychicznego 315 nastolatków i młodych dorosłych w wieku 12–20 lat, którzy byli leczeni HPP przez okres 24 miesięcy. Badanie to, opublikowane w czasopiśmie „The New England Journal of Medicine” (NEJM), przedstawiło swoje wyniki w pozytywnym świetle: „nasze wyniki wykazały poprawę funkcjonowania psychospołecznego (...) co przemawia za stosowaniem [HPP] jako skutecznego leczenia młodzieży trans i niebinarnej”[40].

Jednak dwóch uczestników badania popełniło samobójstwo w ciągu roku od rozpoczęcia terapii hormonalnej, co stanowi wskaźnik 30–50 razy wyższy niż oczekiwany w przypadku nastolatków w podobnym wieku[41]. Jedynym statystycznie istotnym wynikiem była niewielka poprawa w „skali kongruencji trans”, która nie została walidowana u nieletnich[42]. Co istotne, zgłoszono tylko cztery z 19 wyników dotyczących zdrowia psychicznego opisanych we wstępie zarejestrowanym protokole badania; w szczególności nie zgłoszono wyników dotyczących DP, samookaleczeń i skłonności samobójczych, nie podano też żadnego wyjaśnienia pominięcia tych kluczowych danych[43]. Autorzy twierdzili, że objawy depresji zmniejszyły się w okresie badania[44]. Jednak dalsza analiza podgrup według płci wykazała, że wyniki zgłoszone u pacjentek płci żeńskiej miały wątpliwe znaczenie kliniczne, podczas gdy pacjenci płci męskiej nie odnotowali żadnej statystycznie istotnej poprawy w zgłoszonych pomiarach[45].

Zgodnie z wcześniej zarejestrowanym protokołem badania autorzy pierwotnie postawili hipotezę, że pacjenci będą wykazywać „zmniejszenie objawów (...) dysfornii płciowej, samookaleczeń, objawów traumy i skłonności samobójczych” (oraz inne zmiany) w miarę upływu czasu[46]. Zamiast tego, jak wspomniano, dwóch młodych pacjentów popełniło samobójstwo w ciągu roku od rozpoczęcia leczenia, a najczęściej występującym zdarzeniem niepożądanym podczas leczenia hormonalnego były „myśli samobójcze”[47]. Wyniki nie potwierdzają zatem pierwotnej hipotezy.

Chociaż w ostatnim przeglądzie systematycznym udało się uchwycić takie kwestie, jak spadek liczby uczestników badania w ocenie ryzyka błędu systematycznego dla Chen i in., analizy metodologiczne nie są w stanie uchwycić poważnych wątpliwości, które mogą być bardziej związane z etyką badań niż z metodologią (np. zmiana hipotez badawczych między protokołem a opublikowanym artykułem). Recenzowane krytyczne uwagi dotyczące kwestii poruszonych przez Chen i in. zostały ostatecznie opublikowane w czasopiśmie, ale dopiero rok po pojawieniu się artykułu[48].

Badanie to jest nadal szeroko cytowane i bezkrytycznie opisywane jako dostarczające dowodów na korzyści płynące z hormonalnej tranzycji nieletnich[49].

6.3 Stronniczość publikacji

Dobrze udokumentowanym problemem w badaniach biomedycznych jest to, że pozytywne wyniki potwierdzające pierwotną hipotezę są znacznie częściej publikowane lub publikowane w trybie przyspieszonym w porównaniu z badaniami, których wyniki są niejednoznaczne lub negatywne[50]. „Stronniczość publikacji” nie jest zjawiskiem charakterystycznym wyłącznie dla tej dziedziny. Krajowa klinika PMG w Wielkiej Brytanii, Gender Identity Development Service (GIDS), podjęła próbę powtórzenia przełomowego holenderskiego badania z 2011 r., rozpoczynając w 2011 r. własne badanie „wczesnej interwencji” w zakresie BP. Chociaż wyniki

były dostępne w latach 2016–2017, zostały opublikowane dopiero w 2021 r. w następstwie skargi złożonej do brytyjskiego organu ds. badań naukowych w dziedzinie zdrowia (UK Health Research Authority)[51]. Po opublikowaniu badania potwierdzono, że wyniki zgłoszone przez Holendrów nie zostały replikowane[52]. Badanie wykazało również, że blokada dojrzewania wiązała się z zaburzeniami gęstości masy kostnej. Późniejsza ponowna analiza tych samych danych wykazała, że aż jedna trzecia uczestników doświadczyła pogorszenia zdrowia psychicznego podczas stosowania BP[53].

Podobna sytuacja może mieć miejsce w Stanach Zjednoczonych. Główna badaczka finansowanego przez NIH badania nad BP przyznała się do opóźnienia publikacji wyników – najwyraźniej z obawy, że nieimponujące wyniki podważą wiarygodność tej dziedziny[54].

Podobną niechęć do informowania o rozczarowujących wynikach można zaobserwować również w innych publikacjach. Zgodnie z ujawnionymi informacjami sądowymi Światowe Stowarzyszenie Zawodowe na rzecz Zdrowia Osób Trans (WPATH) zablokowało publikację przeglądów systematycznych dotyczących co najmniej dziesięciu z trzynastu tematów związanych z leczeniem endokrynologicznym nastolatków z DP[55]. Nieopublikowane przeglądy systematyczne oceniały szkody, w tym wpływ BP na „wyniki kliniczne i szkody”; wpływ estrogenów na ryzyko zatorowości płucnej, zakrzepicy żył głębokich, udaru mózgu i zawału mięśnia sercowego; wpływ testosteronu na patologię macicy, jajników, szyjki macicy, pochwy i piersi; oraz wpływ terapii hormonalnej na płodność i na zespół metaboliczny[56].

Istnieją również przypadki przeglądów systematycznych, w których wyciągnięto wnioski sprzeczne z dowodami naukowymi analizowanymi w tych przeglądach. W jednym z przeglądów systematycznych zleconym przez WPATH stwierdzono jedynie „niską” i „niewystarczającą”[57] siłę dowodów dotyczących wpływu interwencji hormonalnych na zdrowie psychiczne. Niemniej jednak autorzy przeglądu systematycznego doszli do niepopartej dowodami konkluzji, że ta ścieżka leczenia „jest prawdopodobnie związana z poprawą” jakości życia, depresji i lęku. Autorzy stwierdzili również, że „korzyści te sprawiają, że terapia hormonalna jest niezbędnym elementem opieki, który promuje (...) zdrowie i dobre samopoczucie”. Co istotne, nie podjęto żadnej próby oceny wpływu hormonów na zdrowie fizyczne[58]. Na treść tego PS wpływ mieli kierownicy WPATH, którzy wymagali potwierdzenia [założonych z góry] wniosków jako warunek ostatecznej publikacji. Może to wyjaśniać rozbieżność między optymistycznymi wnioskami PS a bardziej problematycznymi ustaleniami. W sekcji 10.3.2 omówiono szczegółowo ukrywanie dowodów naukowych przez WPATH.

Kolejny przegląd systematyczny z 2021 r., przeprowadzony przez Bustos i in., zbadał wskaźniki żalu po operacji[59]. Przegląd ten jest często cytowany jako dowód, że wskaźnik żalu po operacji związanej z tranzycją jest bardzo niski i wynosi około 1%[60]. Jednak spośród 27 badań uwzględnionych w przeglądzie Bustos i in., 23 zostały ocenione jako „słabe” lub „przeciętne” pod względem jakości. Nieliczne badania o wyższej jakości charakteryzowały się znaczną utratą uczestników w badaniach kontrolnych, a zgłoszone wyniki obejmowały zgony z powodu samobójstwa i powikłań. Jedno z uwzględnionych badań stanowiło źródło prawie połowy wszystkich uwzględnionych w PS uczestników, ale w tym konkretnym badaniu zanotowano 36% utratę uczestników w badaniach kontrolnych. Autorzy badania przyznali, że krótki okres obserwacji był kluczowym ograniczeniem. Nie odnieśli się jednak do najpoważniejszej krytyki, a mianowicie tego, że ich główny wniosek o „niezwykle niskim poziomie żalu” nie był poparty dowodami[61].

6.4 Podsumowanie

Jakość przeglądów systematycznych w dziedzinie medycyny tożsamości płciowej jest bardzo zróżnicowana, ale nawet przeglądy najwyższej jakości zazwyczaj znajdują jedynie dowody o „niskim stopniu pewności” w odniesieniu do większości szkód. Istnieją jednak znane i prawdopodobne wyjaśnienia, dlaczego dowody na istnienie szkód mogły nie zostać znalezione, wykryte lub zgłoszone. Odnosząc się do stanu dowodów w PMG, w przeglądzie Cass zauważono, że „ewidentne jest, że nie udało się w sposób spójny i kompleksowy zebrać nawet najbardziej podstawowych danych i informacji; dane często nie były udostępniane lub były niedostępne”[62].

O ile przeglądy systematyczne nie są ukierunkowane konkretnie na poszukiwanie szkód (z zastrzeżeniami podanymi powyżej), nie nadają się one do wykrywania powikłań lub działań niepożądanych. Wynika to z faktu, że dziedzina PMG jest stosunkowo nowa i szybko się rozwija, badania obserwacyjne, które są przeważnie krótkoterminowe, skupiają się na korzyściach i nie zgłaszają wszystkich szkód, a także z powodu tendencji do publikowania wyników badań korzystnych dla danego leku lub terapii. EBM kładzie nacisk na rolę najlepszych dostępnych dowodów. Aby zaradzić ograniczonej zdolności przeglądów systematycznych do wykrywania szkód, decydenci muszą uwzględnić inne źródła informacji naukowej. Jest to temat następnego rozdziału.

Tłumaczenie: Magda Lewandowska

[1] Loke et al. (2007); Mayo-Wilson et al. (2023); Qureshi et al. (2022).

[2] Testosteron jest jednym z najczęściej stosowanych leków w medycynie gender i jest klasyfikowany jako substancja kontrolowana w wykazie III ze względu na ryzyko uzależnienia psychicznego i nadużywania.

[3] Np. Cornell Health (b.d.); Fenway Health (2024a, 2024b, 2025).

[4] Egualé i in. (2016).

[5] Np. Brik i in. (2020); Carmichael i in. (2021); Karakılıç Özturan i in. (2023); van der Loos i in. (2023); Wiepjes i in. (2018).

[6] Egualé i in. (2016).

[7] de Vries i in. (2006); Delemarre-van de Waal i Cohen-Kettenis (2006). Zob. rozdział 3.

[8] Respaut i Terhune (2022).

[9] Badanie medycyny gender z długoterminową obserwacją wykazało „gorsze wyniki [...], które można również wyjaśnić dłuższym okresem obserwacji (mediana 10 lat) w porównaniu z poprzednimi badaniami. Na poparcie tej tezy krzywa przeżywalności (rysunek 1) sugeruje wzrost śmiertelności od dziesiątego roku po zmianie płci” (Dhejne i in., 2011).

[10] Cass (2024a, s. 76).

[11] Loke i in. (2007); Mayo-Wilson i in. (2023); Qureshi i in. (2022).

[12] W przypadku badań prospektywnych.

[13] Badacze mogą chcieć, aby ich interwencja przyniosła korzystne wyniki i mogą być niechętni do poszukiwania lub skupiania się na szkodach.

[14] Zob. rozdział 4, w którym omówiono uzasadnienie „czasu na przemysłowanie”.

[15] Brik i in. (2020); Carmichael i in. (2021); Karakılıç Özturan i in. (2023); van der Loos i in. (2023); Wiepjes i in. (2018).

[16] Cass (2024a, s. 176).

[17] Miroschnychenko, Roldan i in. (2025) to PS, w którym zidentyfikowano badania Carmichael i in. (2021) oraz Karakılıç Özturan i in. (2023).

[18] Nie jest to niczym niezwykłym w medycynie. Badacze, którzy opracowują innowacyjne metody leczenia, często skupiają się na oczywistych korzyściach.

[19] de Vries i in. (2011); de Vries i in. (2014).

[20] de Vries i in. (2014, s. 703).

[21] Omówienie słabych punktów oryginalnych badań dotyczących protokołu holenderskiego znajduje się w rozdziale 4.

[22] Wyniki te zostały omówione jedynie w sekcji „metody” (de Vries i in., 2014).

[23] W badaniu de Vries i in. (2014) brakowało 15 uczestników z pierwotnej kohorty 70 uczestników. Spośród tych 15 uczestników trzech zostało zdyskwalifikowanych z badania, bo rozwinęła się u nich otyłość i/lub cukrzyca. Jeden zrezygnował z leczenia, a jeden zmarł z powodu powikłań chirurgicznych, które były prawdopodobnie konsekwencją zahamowania dojrzewania (Biggs, 2023b; Negenborn et al., 2017). Kilku innych nie zwróciło kwestionariuszy pooperacyjnych z nieznanymi przyczynami.

[24] de Vries i in. (2014) zostało wykluczone z systematycznego przeglądu HPP przeprowadzonego przez PS przeprowadzonego przez brytyjski Narodowy Instytut Zdrowia i Doskonałości Opieki (NICE). Jako powód wykluczenia podano: „Wykluczenie ze względu na interwencję – wszyscy uczestnicy przeszli operację po podaniu hormonów afirmujących tożsamość płciową. Nie można ustalić, czy zmiany były spowodowane hormonami, czy operacją. Pełne dane dostępne tylko dla 40 pacjentów. Szczegóły dotyczące hormonów afirmujących tożsamość płciową są słabo opisane” (National Institute for Health and Care Excellence, 2020, s. 72). Badanie to zostało również wykluczone z przeglądu systematycznego przeprowadzonego przez Miroschnychenko i in. (2025b) na temat HPP. Chociaż zostało ono uwzględnione w przeglądzie systematycznym przeprowadzonym przez Taylora (2024b) na temat HPP, zostało sklasyfikowane jako badanie niskiej jakości i nie zostało uwzględnione w syntezie dowodów.

[25] Tordoff i in. (2022).

[26] Tordoff i in. (2022). Badania te zostały przeprowadzone w Stanach Zjednoczonych w Seattle Children's Hospital/Washington State University.

[27] Nie podano surowych danych z kwestionariusza zdrowia pacjentów (PHQ-9) i kwestionariusza zaburzeń lękowych GAD-7, nawet w dodatku dostępnym wyłącznie online.

[28] Spośród 104 młodych osób w wieku 13–20 lat 62 rozpoczęło leczenie za pomocą PB lub HPP w ciągu 12 miesięcy trwania badania, a 7 było już w trakcie leczenia w momencie rozpoczęcia badania. Uczestnicy, którzy rozpoczęli leczenie w trakcie badania, w różnych momentach należeli zarówno do grupy „leczonej”, jak i „nieleczonej”, co utrudnia interpretację wyników.

[29] Zamiast analizować indywidualne trajektorie zdrowia psychicznego przed i po leczeniu, w badaniu przeprowadzono porównania przekrojowe w różnych punktach czasowych, co mogło spowodować potencjalne odchylenia.

[30] Po 12 miesiącach brakuje danych dotyczących 17% młodzieży leczonej farmakologicznie (12 z 69) i aż 80% młodzieży nieleczonej farmakologicznie (28 z 35). Jest prawdopodobne, że uczestnicy niepoddani leczeniu medycznemu, bez zaburzeń zdrowia psychicznego, w tym osoby, u których ustąpiły zaburzenia tożsamości płciowej, byłyby najmniej skłonne do powrotu do kliniki.

[31] Autorzy przeprowadzili analizy stosując wiele strategii, ale w podsumowaniu badania podali tylko wersję analizy, która dała najbardziej dramatyczne wyniki (najsilniejszy aOR i najmniejsze wartości p), co budzi obawy dotyczące manipulacji danymi.

[32] Na podstawie eTabeli 3 w materiałach uzupełniających szacowana częstość występowania depresji w grupie kontrolnej wynosiła 59% na początku badania, 76% w 3. miesiącu, 58% w 6. miesiącu i 86% w 12. miesiącu. Znaczne wahania w poszczególnych punktach czasowych wskazują, że szacunki dla grupy kontrolnej były niestabilne i prawdopodobnie miały na nie wpływ błędy losowe.

[33] Dane dotyczące 80% pacjentów z grupy nieleczonej były niekompletne. Możliwe, że pacjenci, u których DP ustąpiła przestali zgłaszać się na wizyty. Ponadto 7 nieleczonych pacjentów, dla których dostępne były dane z obserwacji po roku badania, mogło nie uzyskać zgody na tranzycję medyczną, ponieważ nie mieli wystarczającej zdolności do wyrażenia zgody – jest to jedyny scenariusz, w którym ta pediatryczna klinika gender nie zaleciłaby tranzycji medycznej, u pacjenta, który tranzycji pragnie, według głównego autora badania ("W praktyce jedynymi przypadkami, w których opóźnienie rozpoczęcia BP/GAH byłoby właściwe, są sytuacje, w których istnieje obawa, że pacjent nie jest zdolny do wyrażenia świadomej zgody (...) co jest niezwykle rzadkie w okresie dojrzewania" (Tordoff, 2022). Bariery w wyrażeniu świadomej zgody mogą obejmować aktywną psychozę lub manię oraz głęboką niepełnosprawność intelektualną.

[34] Miroshnichenko i in. (2025); Taylor i in. (2024). ROBIS: Ryzyko błędu systematycznego w badaniach nierandomizowanych – interwencji.

[35] Np. Abbruzzese i in. (2023); Biggs (2022); Singal (2022a). Uniwersytet, który przeprowadził badania, został poinformowany o problemach związanych z badaniem, ale postanowił nie zakłócać bardzo pozytywnego obrazu medycznego, który fałszywie przedstawiał wyniki autorów Singal (2022b).

[36] Np. American Psychiatric Association (2023a); Boerner (2022); Stringer (2023); U.S. v. Skrmetti (nr 23-477), American Academy of Pediatrics oraz dodatkowe krajowe i stanowe organizacje medyczne i zajmujące się zdrowiem psychicznym, opis amicus curiae (2023). Do tej pory badanie to zostało zacytowane ponad 300 razy w innych recenzowanych publikacjach i pobrane prawie pół miliona razy.

[37] Chen et al. (2023). Było to finansowane przez NIH wieloośrodkowe badanie przeprowadzone w Stanach Zjednoczonych w Szpitalu Dziecięcym w Chicago, Szpitalu Dziecięcym w Bostonie i Szpitalu Dziecięcym w Los Angeles.

[38] Protokół badania można znaleźć w materiałach uzupełniających w Chen et al. (2023).

[39] Do badania rekrutowano pacjentów z czterech pediatrycznych klinik zajmujących się tożsamością płciową: Benioff Children's Hospital, Boston Children's Hospital, Children's Hospital Los Angeles oraz Lurie's Children's Hospital of Chicago. Dwóch młodych pacjentów popełniło samobójstwo podczas przyjmowania HPP, a „myśli samobójcze” zostały odnotowane jako najczęstsze zdarzenie niepożądane (Chen i in., 2023). Protokół można znaleźć w materiałach uzupełniających.

[40] Chen i in. (2023, s. 249).

[41] Porównanie wskaźnika samobójstw w badaniu Chen i in. z innymi populacjami (np. innymi kohortami skierowanymi do pediatrycznych klinik medycyny gender; pacjentami przyjmującymi leki przeciwdepresyjne) można znaleźć w McDeavitt (2025).

[42] Strauss i in. (2024).

[43] W badaniu Chen i in. (2023) satysfakcję z życia/pozytywne afekty zmierzono za pomocą NIH Toolbox Emotion Battery. W najnowszym artykule opisującym tę samą kohortę pacjentów opublikowano więcej informacji na temat wyników dotyczących zadowolenia z życia/pozytywnych afektów wraz z informacjami na temat dodatkowych pomiarów z NIH Toolbox Emotion Battery Olson-Kennedy i in. (2025), ale nie podano wyników dotyczących dysforii płciowej, samookaleczeń ani skłonności samobójczych, ani nie odniesiono się do dwóch przypadków samobójstw pacjentów.

[44] Średni wynik w skali depresji Becka poprawił się z 15,46 do 12,92 w 63-punktowej skali (Chen i in. 2023).

[45] Średni wynik w skali depresji Becka poprawił się z 15,46 do 12,92 w 63-punktowej skali (Chen i in. 2023).

[46] Zob. str. 25–26 wstępnie zarejestrowanego protokołu badania, do którego link znajduje się w artykule Chen i in. (2023) na stronie internetowej New England Journal of Medicine.

[47] Chen i in. (2023, s. 241).

[48] Np. Biggs (2023a); Hare (2023); Jorgensen (2023c), holenderscy badacze opieki afirmującej tożsamość płciową opublikowali również komentarz do Chen i in., w którym uznali wiele problemów związanych z badaniem, de Vries & Hannema (2023).

[49] Np. U.S. v. Skrametti (nr 23-477), Eksperci, badacze i lekarze, opis amicus curiae (2023).

[50] Loke i in. (2007).

[51] Biggs (2023b).

[52] Carmichael i in. (2021).

[53] McPherson i Freedman (2023).

[54] Jak podano w New York Times (Ghorayshi, 2024). Należy zauważyć, że śledcza Johanna Olson-Kennedy zaprzeczyła, jakoby wypowiedziała te słowa w późniejszym zeznaniu (Zeznanie J. Olson-Kennedy w sprawie LB przeciwko Premiera Blue Cross nr 2:23-cv-00953-TSZ -11/18/2024. b.d.).

[55] Zob. rozdział 10.

[56] Sharma i in. (2018). Zob. również Society for Evidence-Based Gender Medicine (2024).

[57] Baker i in. (2021, s. 13).

[58] Załącznik 4 zawiera ocenę ROBIS badania Baker et al. (2021).

[59] Bustos et al. (2021) podali odsetek żalu wynoszący około 1%. Ten PS nie został uwzględniony w niniejszym przeglądzie, ponieważ większość uwzględnionych badań nie koncentrowała się na populacji nastolatków.

[60] Zob. przykłady takich stwierdzeń w Cohn (2023).

[61] Krytykę można znaleźć w Expósito-Campos i D'Angelo (2021). Autorzy udzielili odpowiedzi na krytykę (Escandón i in., 2021), w której odnieśli się do niektórych, ale nie wszystkich kluczowych kwestii.

[62] Cass (2024a, s. 39).