



# GenerazioneD

Associazione culturale GenerazioneD  
Via privata del Gonfalone 3, 20123 Milano  
info@generazioneD.org

*Pubblichiamo la nostra traduzione degli estratti più significativi del documento [Supplement to Treatment for Pediatric Gender Dysphoria Report Peer Reviews and Responses](#) pubblicato nel novembre 2025 dall' U.S. Department of Health and Human Services.*

## Introduzione

“Il documento **“Trattamento della disforia di genere pediatrica: revisione delle prove e delle migliori pratiche (La Revisione)”** è stato pubblicato dal Department of Health and Human Services (HHS) degli Stati Uniti il 1° maggio 2025; le revisioni sono state apportate il 15 maggio. A seguito delle revisioni paritarie post-pubblicazione, la revisione e l'Appendice 4 separata sono state ulteriormente riviste. **Sono qui riportate le risposte individuali a sette revisioni paritarie sollecitate, insieme a una risposta a due revisioni paritarie non richieste** (Dowshen et al., 2025; Rider et al., 2025) ...

Nel tentativo di sollecitare e incorporare il feedback delle principali organizzazioni mediche che hanno espresso sostegno alla transizione medica pediatrica, l'HHS ha invitato la American Academy of Pediatrics (AAP), l'American Psychiatric Association (APA) e l'Endocrine Society a partecipare al processo di revisione paritaria. Tutti i tre gruppi hanno criticato la revisione, con l'AAP che l'ha condannata in una dichiarazione ufficiale poche ore dopo la sua pubblicazione. Sfortunatamente, l'AAP e l'Endocrine Society ha rifiutato l'offerta di partecipazione dell'HHS. Siamo grati all'APA per aver accettato l'invito”.<sup>1</sup> (pag. 3)

## Partecipanti

“Eugenia Abbruzzese, Society for Evidence-Based Gender Medicine

Alex Byrne, PhD, Massachusetts Institute of Technology

Farr Curlin, MD, Duke University

Moti Gorin, PhD, MBE, Colorado State University

Kristopher Kaliebe, MD, DFAACAP, University of South Florida

Michael K. Laidlaw, MD, Michael K. Laidlaw MD, Inc.

Kathleen McDeavitt, MD, Baylor College of Medicine

Leor Sapir, PhD, Manhattan Institute for Policy Research

Yuan Zhang, PhD, Evidence Bridge” (pag.4)

## Revisioni tra pari

### American Psychiatric Association

“Le nostre conclusioni sono che, sebbene il Report HHS si proponga di essere una valutazione approfondita e basata sulle prove dell'assistenza di genere per i giovani transgender, la sua metodologia di base manca di sufficiente trasparenza e chiarezza affinché i suoi risultati possano essere presi per oro colato. Elementi chiave, tra cui i criteri di selezione della letteratura, i quadri analitici e la giustificazione per l'esclusione di altri studi, nonché i risultati chiave degli studi su cui si basa il Report, sono poco spiegati o assenti. Di conseguenza, le affermazioni del Report non raggiungono lo standard di rigore metodologico che dovrebbe essere considerato un prerequisito per guidare le policy di assistenza. Di seguito alcuni commenti specifici sulla metodologia del Report:

- Con una sola eccezione, gli autori del rapporto non sono identificati. La trasparenza sulla paternità è essenziale per l'integrità dell'analisi scientifica e politica perché consente ai lettori di valutare l'esperienza dei collaboratori, valutare le loro qualifiche nei campi pertinenti e identificare potenziali conflitti di interesse o impegni ideologici.
- Il rapporto non riesce ad articolare chiaramente come sono stati selezionati gli studi, quali criteri regolavano la loro inclusione o esclusione, o come la loro qualità veniva valutata...
- Non vi è alcuna indicazione che le principali parti interessate, vale a dire le persone transgender, le loro famiglie e medici - siano stati consultati o che le loro prospettive siano state prese in considerazione...
- Sebbene il rapporto sia chiaro sui potenziali danni derivanti dall'intervento medico, non applica alcun tipo di analisi razionale ai potenziali danni associati alla mancata adozione di un intervento, tra cui tassi più elevati di depressione, ansia, tendenze suicide e isolamento sociale.”<sup>2</sup> (pag.7-8)

### Dott. Johan C. Bester

**Impressione Generale:** La Revisione è una revisione delle prove e dell'analisi etica degli interventi offerti a minori (adolescenti e bambini) affetti da disforia di genere. Si tratta di un lavoro importante e attuale. È ben scritto, metodologicamente rigoroso e apporta un contributo significativo alla discussione su questo argomento. Indicherò alcune aree in cui raccomanderei piccoli miglioramenti o ulteriori analisi. Queste servono come raccomandazioni per migliorare il lavoro e non influiscono sui risultati complessivi della Revisione. Il contenuto qui presente è completo, convincente e ben fatto. I principali risultati e le raccomandazioni della Revisione sono coerenti con i risultati e le raccomandazioni di altre revisioni e analisi di prove di alto livello pubblicate su questo argomento.

#### Principali risultati e conclusioni della revisione:

1. Non ci sono prove convincenti di benefici per gli interventi di transizione di genere offerti a minori con disforia di genere. Le evidenze scientifiche sono di scarsa qualità. Le evidenze disponibili non dimostrano alcun chiaro beneficio derivante dalla transizione sociale, dai bloccanti della pubertà, dagli ormoni del sesso opposto o dalla chirurgia di modificazione del genere.”<sup>3</sup> (pag. 11)
2. “Esistono notevoli preoccupazioni circa i potenziali danni derivanti dagli interventi di transizione di genere offerti ai minori con disforia di genere, in particolare ai bloccanti della pubertà, agli ormoni eterosessuali e agli interventi chirurgici di modificazione del genere. Le migliori prove disponibili sui potenziali danni derivano da quanto si sa su questi interventi, dall'uso in altri contesti e dalle

conoscenze scientifiche di base. Mancano studi solidi che abbiano indagato questi danni nel contesto della transizione di genere nei minori.

3. Le linee guida per la pratica clinica che hanno avuto maggiore influenza e hanno informato la pratica relativa alla disforia di genere nei minori negli Stati Uniti, sono di scarsa qualità. In particolare, gli Standard di Cura WPATH e le linee guida sviluppate dall'Endocrine Society presentano problemi che ne rendono l'utilizzo inadeguato. Tra questi rientrano conflitti di interesse, problemi metodologici e applicazione errata o travisamento delle prove disponibili.
4. Data la mancanza di benefici dimostrabili e la preoccupazione per i potenziali danni, l'uso di bloccanti della pubertà, ormoni eterosessuali e interventi chirurgici per la transizione di genere nei minori con disforia di genere non può essere giustificato eticamente...

**Discussione dei principali risultati e contributi di questa revisione:** ... Ecco cosa sappiamo. La prevalenza della disforia di genere è aumentata tra bambini e adolescenti negli ultimi 15-20 anni. Le ragioni di ciò non sono del tutto comprese, ma sono probabilmente multifattoriali. L'epidemiologia della disforia di genere è cambiata. La stragrande maggioranza dei casi si risolve senza alcun intervento o trattamento. Non vi sono prove convincenti di benefici derivanti dai trattamenti di transizione di genere per la risoluzione della disforia o per la gestione delle comorbidità. I trattamenti utilizzati per la transizione di genere non sono generalmente reversibili, hanno implicazioni a lungo termine e comportano potenziali effetti negativi.”<sup>4</sup> (pag. 12)

“Ciò che è venuto prima è stata la certezza che questi trattamenti fossero benefici e necessari dal punto di vista medico. Ciò che è venuto dopo è stata la ricerca di prove che giustificassero questa affermazione a posteriori. In assenza di prove, sono emerse giustificazioni variabili per il trattamento. In definitiva, l'impegno a fornire tali interventi è diventato una sorta di cartina di tornasole, in cui il sostegno ai bambini con disforia di genere e la credibilità come esperti in questo ambito dipendevano da un impegno solido negli interventi di transizione di genere. In secondo luogo, la questione è stata vista come una questione di diritti civili ed è diventata una questione di politica e attivismo, piuttosto che una questione medica. Qualsiasi tentativo di porre domande su questi approcci terapeutici è stato visto come un'azione di oppressione e ha portato alla denigrazione...”<sup>5</sup> (pag. 13)

**“Commenti sulla metodologia e sulla revisione delle prove: 1. Revisione delle prove a favore degli interventi di transizione di genere:** La revisione conduce una revisione sistematica delle revisioni sistematiche. Nei capitoli 5 e 6, viene descritta la metodologia che mostra i criteri di inclusione, i criteri di ricerca e un diagramma di flusso per tenere conto degli studi identificati, inclusi ed esclusi. La revisione si avvale di uno strumento appropriato per valutare i *bias* e analizza in modo appropriato gli studi inclusi. Inoltre, la revisione si occupa direttamente di un'analisi degli studi osservazionali che sono stati centrali nelle precedenti affermazioni di presunti benefici derivanti dai trattamenti di transizione di genere nei minori. Tutti questi metodi appaiono validi e si può essere certi che questa revisione sistematica sia riproducibile, coerente con i metodi delle revisioni sistematiche di questo tipo e abbia identificato le migliori prove disponibili relative ai quesiti clinici.”<sup>6</sup> (pag.14)

“2. Revisione delle prove di danno degli interventi di transizione di genere: La revisione analizza gli studi inclusi nella revisione sistematica e ne evidenzia le carenze nel monitoraggio dei potenziali danni. In linea con le migliori pratiche della medicina basata sulle prove di efficacia, la revisione ricerca quindi le migliori evidenze disponibili, ovvero quanto si conosce di questi interventi in generale quando utilizzati per altri scopi, e quanto si conosce di questi interventi da considerazioni scientifiche di base e cliniche. La revisione identifica una serie di danni che è probabile che si verifichino con elevata certezza e una serie di ulteriori potenziali danni che potrebbero verificarsi con l'uso di questi interventi. I metodi utilizzati sono appropriati e la conclusione che i medici dovrebbero essere cauti riguardo al potenziale danno associato a questi interventi è fondata.”<sup>7</sup> (pag. 15)

“4. Revisione delle prove relative alla psicoterapia come trattamento della disforia di genere: L'ultimo capitolo della Revisione è incentrato sulla psicoterapia. La revisione delle prove qui presentata si basa sulla revisione

sistematica del Capitolo 5 e conclude che mancano prove a favore della psicoterapia nei minori con disforia di genere, poiché non è stata adeguatamente studiata in questo contesto. Tuttavia, vi sono prove che la psicoterapia possa essere utile per gestire le comorbidità che spesso accompagnano la disforia di genere, come la depressione o l'ansia. Ciò sembra ragionevole, tuttavia è necessario procedere con cautela. Non possiamo semplicemente presumere che, poiché la psicoterapia apporta benefici ai minori con disturbi mentali ma senza disforia di genere, la psicoterapia abbia gli stessi benefici e rischi per il trattamento del disagio correlato alla disforia di genere...Una forte raccomandazione in questo caso è che siano necessari ulteriori studi sulla psicoterapia nel contesto della disforia di genere, data la mancanza di prove scientifiche.”<sup>8</sup> (pag. 15)

“La revisione riassume bene come la pratica della medicina di genere negli Stati Uniti sia scesa al di sotto degli standard richiesti per il consenso informato. Un processo di consenso valido richiede una serie di requisiti, tra cui la piena divulgazione delle informazioni rilevanti e la volontarietà. La revisione dimostra che la piena divulgazione non è avvenuta in molti contesti di pratica di trattamento di genere negli Stati Uniti, e la reale mancanza di benefici e i potenziali danni sono stati generalmente oscurati quando si è richiesto il consenso ai genitori e ai minori. Inoltre, è stato utilizzato sistematicamente un linguaggio che mina la volontarietà del consenso o l'autorizzazione al trattamento da parte dei genitori. La frase "Puoi avere una figlia morta o un figlio vivo" è stata utilizzata regolarmente dagli operatori della transizione di genere per spingere i genitori dubbiosi a fornire il consenso alla transizione di genere. Innanzitutto, questa frase oscura la verità. Non ci sono prove che la transizione di genere salvi la vita o che la disforia di genere porti inevitabilmente alla morte. Ma, cosa ancora più importante, questa è una frase coercitiva che esercita una pressione irresistibile sui genitori affinché acconsentano alla transizione di genere”<sup>9</sup> (pag. 17)

## **Professor Karleen Gribble**

“Nella sezione introduttiva del riassunto esecutivo della revisione si osserva che: Il comprensibile desiderio di evitare un linguaggio che possa causare disagio ai pazienti ha, in alcuni casi, dato origine a modalità di comunicazione prive di fondamento scientifico, che presuppongono risposte a controversie etiche irrisolte e che rischiano di fuorviare pazienti e famiglie. Questa revisione utilizza una terminologia scientificamente accurata e neutrale in ogni sua parte. Apprezzo questo approccio e la spiegazione fornita per questo ragionamento nel Capitolo 2. Le motivazioni fornite per il rifiuto di termini come "sesso assegnato alla nascita", l'uso di "identità di genere" anziché semplicemente "genere" e l'uso di un linguaggio sessuato in generale nella Revisione sono ben argomentate. Raccomando che sarebbe utile aggiungere del testo sui rischi dell'uso di una terminologia che suggerisce che le persone possono cambiare sesso.”<sup>10</sup> (pag. 20)

“Nel considerare i rischi e i benefici dei trattamenti per la disforia di genere pediatrica, la revisione osserva: Per assolvere ai propri doveri di non maleficenza e beneficenza, i medici devono garantire, per quanto ragionevolmente possibile, che tutti gli interventi offerti ai pazienti presentino profili rischio/ beneficio clinicamente favorevoli rispetto all'insieme delle alternative disponibili, tra cui anche il non fare nulla... Le affermazioni qui avanzate sulla probabilità e l'entità di danni e benefici si basano sulle migliori prove disponibili. Talvolta, le probabilità sono note con un elevato grado di certezza. Ad esempio, la probabilità che la mastectomia comporti l'incapacità di allattare al seno è pari a 1,0 o prossima a tale valore... Per quanto riguarda la natura dei benefici e dei danni medici e il loro peso relativo, i presupposti di lavoro della Revisione sono coerenti con l'intuizione morale comune, il giudizio medico standard come rivelato dai criteri diagnostici medici e i risultati di interesse per clinici e ricercatori, nonché con la legge. Ad esempio, l'analisi concluderebbe che un lieve miglioramento dei sintomi depressivi conta come un beneficio, ma che tale beneficio, anche se garantito, non supera i rischi moderati o addirittura bassi ma non trascurabili di infertilità o grave disfunzione sessuale, perdita della funzione di allattamento al seno o dipendenza medica permanente, che la Revisione considera danni...Mi congratulo con gli autori per aver adottato questo approccio. È stato frustrante vedere che le revisioni sistematiche delle prove di intervento non considerassero

i risultati noti semplicemente perché coloro che hanno condotto la ricerca primaria non li hanno inclusi...Tuttavia, nonostante le affermazioni sopra citate, né il corpo della Revisione né la panoramica delle revisioni sistematiche includono il danno in termini di incapacità di allattare al seno nell'analisi dei risultati. Suggestirei di affrontare questo aspetto. I contenuti aggiuntivi dovrebbero sottolineare che le implicazioni della chirurgia di mascolinizzazione del torace possono includere disagio psicologico.”<sup>11</sup> (pag.21)

“La revisione rileva che le persone che rinunciano alla propria identificazione transgender potrebbero provare rimpianto a seguito di interventi medici correlati all'identità di genere. Il rimpianto associato all'intervento chirurgico di mascolinizzazione del torace non viene menzionato, ma dovrebbe essere aggiunto poiché: 1) questa è la procedura chirurgica più comune per i minori con disforia di genere, 2) la WPATH e altre linee guida omettono di raccomandare di discutere l'impatto di questo intervento sull'allattamento al seno con coloro che lo stanno prendendo in considerazione, 3) i sostenitori di questo intervento chirurgico affermano regolarmente falsamente che questo intervento è reversibile, 4) l'insorgenza di un rimpianto acuto può verificarsi molti anni dopo l'intervento, poiché potrebbero trascorrere decenni tra l'intervento e il parto e l'incapacità della donna di allattare al seno.”<sup>12</sup> (pag. 22)

## Dr. Richard J. Santen

“Ho esaminato il documento del DHHS e ho scoperto che il riepilogo dei dati e le discussioni dettagliate riflettono ragionevolmente una panoramica delle informazioni attualmente disponibili e della loro interpretazione...Tuttavia, ritengo che un'area del documento DHHS non abbia ricevuto l'enfasi essenziale. Si tratta della questione se la terapia ormonale di affermazione di genere (ad esempio, bloccanti della pubertà e terapia ormonale cross-sex) è una pratica sperimentale o accettata. Ho esaminato i concetti alla base della definizione di terapia sperimentale in questa critica e suggerisco che venga aggiunta una sezione specifica per affrontare questo problema. A mio parere, se la terapia ormonale di affermazione di genere sia sperimentale o meno è la questione più importante che sottolinea tutta l'attuale controversia...Terapia sperimentale: definizione. Panoramica: Determinare se una terapia per un paziente è considerata sperimentale è un processo complesso. Un processo con implicazioni significative per l'assistenza ai pazienti, l'etica e la regolamentazione. Non esiste una definizione univoca e universale, ma generalmente si riferisce a trattamenti che non sono ancora riconosciuti dalla comunità medica professionale come efficaci, sicuri e comprovati per la specifica condizione per la quale vengono utilizzati. Di seguito è riportata una ripartizione della definizione e le considerazioni chiave.”<sup>13</sup> (pag.24)

“La terapia è tipicamente considerata sperimentale in caso di:

- **Mancanza di efficacia e sicurezza accertate:** non vi sono sufficienti prove scientifiche, ad esempio da studi clinici ben progettati, per dimostrare definitivamente la sua efficacia e sicurezza per gli scopi previsti. Ciò accade spesso perché si tratta di un intervento nuovo, sconosciuto o raramente utilizzato
- **Deviazione dallo standard di cura:** non è in linea con la pratica clinica usuale supportata da un consenso dei medici per la condizione specifica.
- **In fase di studio o in attesa di ricerca:** è attualmente oggetto di studi clinici oppure non è stato ancora sottoposto ai test rigorosi necessari per ottenere un'ampia accettazione.
- **Uso off-label in un contesto determinato:** sebbene l'uso off-label, ad esempio l'uso di un farmaco approvato per uno scopo non specificamente approvato dagli enti regolatori, possa talvolta essere considerato una pratica standard in base a prove emergenti, può anche essere considerato sperimentale se le prove per il suo nuovo uso sono limitate o speculative.
- **Non approvato dagli enti regolatori:** in molti paesi, le terapie sono considerate sperimentali finché non ricevono l'approvazione da agenzie regolatorie come la Food and Drug Administration statunitense per indicazioni specifiche.

In sintesi, la classificazione di una terapia come sperimentale si basa sul livello di solide prove scientifiche sulla sua sicurezza ed efficacia in un contesto specifico. La decisione di utilizzare una terapia richiede una discussione collaborativa e trasparente tra il paziente, la sua famiglia e il gruppo medico, garantendo un consenso informato completo, rigorose considerazioni etiche e l'aderenza ai relativi quadri normativi.”<sup>14</sup> (pag.25)

“Le linee guida finlandesi e svedesi e la revisione Cass considerano l'assistenza che afferma il genere come sperimentale...a seguito di una RS, le autorità finlandesi hanno concluso che il corpus di prove a sostegno dei bloccanti della pubertà e degli ormoni cross-sex per i giovani è inconcludente. È importante sottolineare che le linee guida affermano esplicitamente che "alla luce delle prove disponibili, la riassegnazione di genere a un minore è una pratica sperimentale" ... il NHS England ha introdotto importanti cambiamenti nelle sue policy. I bloccanti della pubertà non sono più prescritti di routine a causa di prove insufficienti sulla loro sicurezza ed efficacia a lungo termine...In netto contrasto, le linee guida di pratica clinica della Endocrine Society e WPATH considerano la cura di affermazione di genere, i bloccanti della pubertà e la terapia ormonale cross-sex come standard di cura e supportati da prove. La discussione in queste due linee guida evidenzia il fatto che la cura di affermazione di genere e la terapia ormonale cross-sex sono state utilizzate negli ultimi 30 anni e sono praticate in diversi paesi... Va notato che la questione sulla natura sperimentale dell'approccio di cura è stata sollevata solo negli ultimi 5 anni. I medici svedesi sono stati i primi a sollevare preoccupazioni riguardo a tale approccio, seguiti dai medici finlandesi. Il documento del DHHS, a mio parere, deve affrontare specificamente i componenti della definizione di medicina sperimentale e il modo in cui gli studi attuali si relazionano a questa definizione.”<sup>15</sup> (pag. 26)

“Il documento del DHHS è stato progettato per valutare le prove e non per stabilire linee guida. Il mio suggerimento, in quanto revisore, è che i pro e i contro di questo problema vengano discussi. La mia analisi indica che attualmente non esiste alcun accordo sul fatto che la terapia ormonale di affermazione di genere sia sperimentale o una pratica standard. I vantaggi di considerarlo sperimentale sono che l'inizio del trattamento richiederà tutti gli elevati standard applicati agli studi di ricerca, vale a dire: consenso informato, discussione dei rischi e dei benefici noti, monitoraggio rigoroso, supervisione del comitato di revisione della sicurezza, requisiti di formazione dei ricercatori e a lungo termine, follow up. Gli svantaggi sono che gli studi clinici in questo ambito sono difficili, in particolare gli RCT, e che considerare gli approcci come sperimentali comporterà una riduzione dei benefici per molti adolescenti con disforia di genere. Il secondo problema generale da discutere è l'accumulo di linee guida: l'Endocrine Society ha pubblicato un manoscritto sull'affidabilità delle linee guida nel 2022 (vedere JCEM107:129-2138,2022) mettendo in guardia sulla pratica dell'accumulo di linee guida per la pratica clinica (CPG). Il concetto di "stacking", la sua definizione e il suo ruolo nello sviluppo delle linee guida devono essere specificati nel documento DHHS. L'ES ha proposto quattro criteri per garantire un CPG affidabile. (1) per garantire un CPG multidisciplinare, che includa membri con competenze pertinenti all'argomento (2) per incoraggiare la diversità del panel con fattori quali internazionalità, genere, razza/etnia e fase di carriera (3) per evitare lo "stacking" e (4) per garantire l'aderenza alla politica del CGC sui conflitti di interesse/dualità di interessi. Lo "stacking" è stato definito come "limitare in modo inappropriato l'appartenenza al gruppo di sviluppo delle linee guida a coloro che hanno un punto di vista particolare".<sup>16</sup> (pag. 28)

## **Dr. Jilles Smids**

### **Revisione del capitolo 13 del rapporto HHS sulla disforia di genere**

La presente revisione analizza il capitolo 13 del rapporto HHS sulla disforia di genere, che tratta dell'etica della medicina di genere pediatrica. Vorrei iniziare con una precisazione: ho fornito un riscontro critico costruttivo su una prima versione del capitolo 13 e, in una fase molto successiva, su diversi altri capitoli del rapporto HHS. Di conseguenza, la domanda sul perché si dovrebbe collaborare alla stesura di un rapporto commissionato

dall'amministrazione Trump, che aveva appena definito l'assistenza medica pediatrica di genere come "mutilazione infantile", potrebbe essere rivolta, in misura minore, anche a me. Le mie considerazioni erano che una relazione sarebbe stata comunque prodotta per l'HHS e che era sempre meglio che venisse prodotta un'analisi di buona qualità anziché un documento scritto nello stesso stile del precedente ordine esecutivo redatto dall'amministrazione Trump. La composizione del team di autori, per quanto mi era noto in quel momento, mi dava sufficiente fiducia che molto probabilmente avrebbero scritto un'analisi equilibrata e basata su prove scientifiche. Ho quindi deciso di fornire un riscontro, nella speranza di contribuire alla realizzazione di una relazione con queste caratteristiche."<sup>17</sup> (pag. 30)

"Il capitolo 13 sostiene le seguenti tesi principali. In primo luogo, i principi etici medici comunemente accettati di beneficenza - agire per il bene del paziente - e non maleficenza - non causare danno - richiedono prove scientifiche sufficienti per un profilo rischio/beneficio favorevole che giustifichi la transizione medica pediatrica (PMT, come la definisce il rapporto). In secondo luogo, i recenti tentativi di giustificare la PMT sulla base del rispetto dell'autonomia del paziente travisano questo principio di etica medica e costituiscono un allontanamento radicale dalla comprensione della medicina di genere pediatrica, che ritiene la PMT giustificata dai suoi (presunti) benefici per la salute mentale. In terzo luogo, il capitolo si conclude con un'analisi etico-scientifica della potenziale ricerca sulla PMT che è scettica sulla giustificazione di offrirla anche nel contesto di studi clinici... quando non vi è alcuna prova ragionevole di un profilo rischio/beneficio positivo, i pazienti non hanno diritto a ricevere la PMT, e non offrire la PMT non costituisce una violazione della loro autonomia."<sup>18</sup> (pag.30)

"Per quanto riguarda la valutazione della PMT, il capitolo presenta un argomento cumulativo molto forte a sostegno della conclusione che, date le nostre attuali conoscenze, un approccio precauzionale è più che giustificato: la storia naturale della disforia di genere pediatrica è poco conosciuta e decenni di ricerca hanno dimostrato che la disforia di genere ad esordio precoce di solito si risolve senza intervento medico. Non ci sono prove convincenti che lo stesso non valga anche per i sintomi ad esordio adolescenziale, e le prove che suggeriscono il contrario sono limitate. E in ogni caso, è ampiamente riconosciuto che i medici non sono in grado di distinguere i pazienti la cui disforia persisterà da quelli la cui disforia si risolverà. Inoltre, vi sono preoccupazioni sul ruolo che la medicalizzazione stessa può svolgere nel contribuire alla persistenza delle condizioni trattate, e sono disponibili interventi meno invasivi e meno rischiosi. Infine, l'intervento medico presenta danni noti e plausibili, e decenni di ricerca condotta da importanti istituzioni accademiche non sono riusciti a produrre prove affidabili dei benefici medici."<sup>19</sup> (pag. 31)

"Non sappiamo quale percentuale di adolescenti avrebbe superato la disforia di genere senza la PMT, o la cui disforia sarebbe diminuita in misura sufficiente da non richiedere più la PMT. Potrebbe trattarsi del 5%, ma anche del 50% o addirittura dell'80%. Somministrare un trattamento così invasivo come la PMT, che comporta una dipendenza dall'assistenza medica per tutta la vita e presenta gravi rischi e danni per la salute, quando c'è una così profonda incertezza sul fatto che l'adolescente ne abbia davvero bisogno, è semplicemente inaccettabile. Questa mancanza di conoscenza della storia naturale si applica a coloro che soffrono di disforia di genere con esordio nell'infanzia (Baron & Dierckxsens, 2022; Byrne, 2024) e ancora di più a coloro che soffrono di disforia di genere con esordio nell'adolescenza (Kaltiala-Heino et al., 2018). Per entrambe le categorie, il potenziale di sovra diagnosi e di trattamento eccessivo e dannoso è molto elevato."<sup>20</sup>(pag. 31)

"Per quanto riguarda la preoccupazione che i bloccanti della pubertà blocchino gli adolescenti nella loro disforia di genere, l'altissima percentuale di loro che passa dai bloccanti della pubertà agli ormoni del sesso opposto, più del 95% (Brik et al., 2020; Carmichael et al., 2021) è motivo di grave preoccupazione a questo proposito. Ciò è particolarmente vero perché esistono meccanismi plausibili per tali effetti di blocco (Cass, 2022): la soppressione della pubertà arresta lo sviluppo fisico e psicosessuale, mentre l'esperienza sessuale e romantica può essere determinante per superare la disforia di genere (Steensma et al., 2011)."<sup>21</sup> (pag.31)

## Dr. Lane Strathearn

“Ringrazio per questa opportunità di recensire *“Trattamento della disforia di genere pediatrica: revisione delle prove e delle migliori pratiche”*. Sono professore ordinario di pediatria, psichiatria, scienze psicologiche e cerebrali, neuroscienze e farmacologia presso l'Università dell'Iowa. Sono anche direttore della Divisione di pediatria dello sviluppo e comportamentale e direttore medico del Centro per le disabilità e lo sviluppo (CDD).

La mia ricerca finanziata dal NIH comprende studi longitudinali su genitori e neonati, incentrati sugli effetti delle prime esperienze e dei maltrattamenti sullo sviluppo del bambino, nonché sulla neurobiologia dell'attaccamento madre-bambino. In qualità di co-direttore di un centro P50 finanziato dal NIH, l'Hawkeye Intellectual and Developmental Disabilities Research Center (Haw-IDDRC), nutro anche un ampio interesse per la cura dei bambini con disturbi intellettivi, dello sviluppo e comportamentali, compresi quelli con disforia di genere. Nell'agosto 2024, in qualità di membro del comitato editoriale del *Journal of Pediatrics*, ho ricevuto l'autorizzazione dal direttore a preparare un commento sulla valutazione e la gestione della disforia di genere pediatrica, riassumendo otto revisioni sistematiche collegate commissionate per la Cass Review britannica.

A quel tempo sembrava esserci un forte sostegno, sia a livello locale che in tutti gli Stati Uniti, per la “cura di affermazione di genere”, ma poca o nessuna consapevolezza della limitata base di prove scientifiche. Il commento mirava a evidenziare questa discrepanza e si intitolava “Cosa sappiamo e cosa non sappiamo: valutazione delle prove scientifiche a sostegno della cura di affermazione di genere in pediatria”, sollevando molte delle stesse preoccupazioni evidenziate in questa revisione. Purtroppo, ma non in maniera inaspettata, il commento non è stato accolto favorevolmente dai revisori, che hanno utilizzato molti degli argomenti efficacemente confutati in questa pubblicazione. Nonostante la presentazione di una confutazione completa di questi argomenti e la ripresentazione al “*Journal of Pediatrics: Clinical Practice*”, come raccomandato dalla rivista, il commento non è mai stato pubblicato.

Nel complesso la presente revisione fornisce una sintesi completa delle prove scientifiche relative a numerose pratiche terapeutiche nella medicina di genere pediatrica, tra cui la transizione sociale, i bloccanti della pubertà, gli ormoni cross sex, la chirurgia e la psicoterapia. Fornisce inoltre un interessante contesto storico per l'attuale panorama sanitario statunitense, compreso l'impatto delle linee guida internazionali, le risposte dell'associazione medica statunitense e le informazioni raccolte dai procedimenti legali. La revisione pone una forte enfasi sulla medicina basata sull'evidenza, delineandone sia i punti di forza che i limiti, integrata da prove indirette provenienti dalla scienza di base e dalla fisiologia per comprendere meglio i meccanismi e il probabile rapporto rischio/beneficio del trattamento. Ritengo che questa revisione fornisca un contributo prezioso e assolutamente necessario a questo importante campo di applicazione.”<sup>22</sup> (pag. 36)

## Prof. Patrik Vankrunkelsven - Dr. Trudy Bekkering

“Data la nostra esperienza nella medicina basata sulle evidenze (Evidence-Based Medicine, EBM), abbiamo focalizzato la revisione sul nucleo del rapporto, ovvero la “*umbrella review*” sui vari trattamenti [...]

**Metodi:** L'uso di una “*umbrella review*” è giustificato dal fatto che esistono molte revisioni sistematiche (SR), spesso basate sugli stessi studi. La revisione utilizza metodi solidi [...]

**Conclusioni generali:** Non abbiamo osservazioni importanti sul disegno dello studio né sulle conclusioni.

**Osservazioni minori:** La mancanza di una rigorosa dichiarazione dei conflitti di interesse (COI) degli autori è un problema rilevante, vista la delicatezza del tema; sarebbe stata utile una definizione formale di revisione sistematica ai fini dell'inclusione nella *umbrella review*; la registrazione del protocollo avrebbe aumentato la

trasparenza, così come dettagli aggiuntivi sulle modalità di sintesi dei risultati; non sono disponibili informazioni su supporto, dati, o accesso ad altri materiali [...]

## **Sintesi del Capitolo 5**

Tema della *umbrella* SR: Quali sono gli effetti di transizione sociale, bloccanti della pubertà (PB), ormoni cross-sex (CSH), interventi chirurgici e psicoterapia nei giovani con disforia di genere fino a 26 anni?

### **Risultati [generali]**

17 revisioni sistematiche incluse: 10 con basso rischio di bias, 7 con rischio elevato [...]

### **Risultati sulla transizione sociale**

2 SR, entrambe con basso rischio di bias; gli effetti a lungo termine su GD, salute psicologica e decisioni future sono poco compresi; evidenze molto limitate sul rimpianto; Qualità delle evidenze: molto bassa; studi prevalentemente trasversali, non comparativi.

### **Risultati sui bloccanti della pubertà (PB)**

9 SR totali (4 inglesi con basso rischio di bias); evidenza di certezza molto bassa sugli effetti su GD, salute mentale e sicurezza; evidenza di alta certezza sugli effetti fisiologici (soppressione ormonale) e rischio frequente di infertilità se seguiti da CSH; evidenza di bassa certezza su compromissione della salute ossea; molti giovani proseguono verso gli ormoni cross-sex

### **[Risultati su] Ormoni cross-sex (CSH)**

8 SR totali (4 inglesi a basso rischio di bias); Evidenza di certezza molto bassa su effetti clinici e sicurezza; Alta certezza sugli effetti fisiologici

### **[Risultati su] Chirurgia**

3 SR (2 a basso rischio di bias); Prevalentemente mastectomie; Alta certezza su complicanze chirurgiche prevedibili (necrosi, cicatrici); Evidenza molto bassa sugli effetti su GD, salute mentale, suicidalità, depressione, qualità di vita e rimpianto

### **[Risultati su] Psicoterapia**

5 SR (2 a basso rischio); evidenze limitate; Certezza molto bassa sugli esiti di salute mentale; nessun danno riportato

## **Discussione**

La certezza delle evidenze è molto bassa. Non solo per mancanza di RCT, ma anche perché mancano osservazionali ben progettati. Non ci sono studi nuovi o in corso che possano modificare sostanzialmente le conclusioni. Nuove SR difficilmente porteranno nuove evidenze: servono nuovi studi clinici”.<sup>23</sup>

## **Dr. Nadia Dowshen et al.**

Dowshen, N., Baker, K., Garofalo, R., Chen, D., Inwards-Breland, D. J., Sequeira, G., ... & McNamara, M. (2025). Una valutazione scientifica critica del rapporto del Dipartimento della Salute e dei Servizi Umani sulla disforia di genere pediatrica. *Journal of Adolescent Health*, 77(3), 342–345.

## **Professor G. Nic Rider et al.**

Rider, G. N., Weideman, B. C., Ehrensaft, D., Choudhary, K., Connor, J. J., Feldman, J., ... & Berg, D. (2025). Integrità scientifica e assistenza sanitaria pediatrica per la disforia di genere: contestazione della revisione dell’HHS. *Sexuality Research and Social Policy*.

## Repliche all'American Psychiatric Association (pag. 55)

### 1. L'APA afferma di:

- a) non poter valutare il rigore metodologico della Review a causa di una mancanza di “chiarezza metodologica” e “trasparenza”, sostenendo che ciò impedisce la verifica o la replicazione indipendente dei risultati della Review.
- b) In particolare, l'APA afferma che il Review “non fornisce la propria strategia di ricerca”, “non chiarisce come siano stati selezionati gli studi”, non spiega “quali criteri ne abbiano governato l'inclusione o l'esclusione”, non fornisce informazioni su “come sia stata valutata la loro qualità”, non fornisce informazioni sugli “schemi analitici” utilizzati per il Review e “non ... elenca gli studi revisionati con citazioni complete o identificatori digitali”.

### Replica:

- a) “La panoramica delle revisioni sistematiche (SR) della Review è stata sottoposta a peer review da due metodologi, la Dott.ssa Trudy Bekkering (Belgian Centre for Evidence-Based Medicine) e il Professor Patrik Vankrunkelsven (Direttore, Belgian Centre for Evidence-Based Medicine). Bekkering e Vankrunkelsven hanno utilizzato la checklist PRIOR (Preferred Reporting Items for Overviews of Reviews) per valutare la panoramica e ne hanno elogiato la metodologia solida. Non hanno identificato problemi significativi relativi alla progettazione o alle conclusioni, osservando che “i risultati finali sono descritti in modo trasparente e sono facili da seguire” e che “ci sono anche molte tabelle con informazioni necessarie e pertinenti”. La valutazione positiva di Bekkering e Vankrunkelsven riconosce il rigore metodologico dell'approccio della Review.
- b) Contrariamente a quanto affermato dall'APA, l'Appendice 4 fornisce una spiegazione chiara e trasparente della strategia di ricerca/criteri di selezione della letteratura (Sezione 1), dei criteri di esclusione (Sezione 2.2), dei principali risultati degli studi su cui si basa il Review (Sezioni 4–9) e include effettivamente l'elenco degli studi revisionati con citazioni complete e identificatori digitali (Sezione 11). Questo contraddice la peer review dell'APA al punto da suggerire che i revisori non abbiano notato i riferimenti nel Review (inclusi nell'indice) all'Appendice 4 di 174 pagine.”<sup>24</sup>

2. L'APA riconosce la chiarezza della Review riguardo ai “potenziali danni di un intervento medico”, ma critica la mancanza di “alcun tipo di analisi razionale ai potenziali danni associati alla mancata attuazione dell'intervento, inclusi tassi più elevati di depressione, ansia, suicidalità e isolamento sociale”.

### Replica:

“L'analisi della Review sui potenziali benefici e rischi della PMT consiste in:

(1) una panoramica delle revisioni sistematiche (Capitolo 5 e Appendice 4); e (2) evidenze provenienti dalle scienze di base e dalla fisiologia (Capitolo 7).

La panoramica delle revisioni sistematiche sugli interventi ha considerato tutta la letteratura pubblicata rilevante riguardo alla PMT, inclusi studi che confrontavano gli esiti per popolazioni che hanno ricevuto la PMT con quelle che non l'hanno ricevuta. La sintesi delle evidenze ha rilevato che non vi sono prove credibili dei benefici della PMT rispetto alla mancata PMT negli esiti citati dall'APA (depressione, ansia, suicidalità) — e, di conseguenza, non ha trovato prove credibili di danni derivanti dal non fornire la PMT. L'analisi delle scienze di base e della fisiologia ha presupposto che la pubertà endogena non sia patologica, ma un processo normale di sviluppo sessuale attraverso il quale un bambino matura in adulto. Interrompere questo processo può potenzialmente causare danni fisici. Pertanto, l'analisi delle scienze di base e della fisiologia poteva

solo fornire una valutazione dei danni derivanti dall'interruzione di un normale processo fisiologico.”<sup>25</sup>

**3. L'APA critica la Review per:**

- a) **“fare ampio riferimento alla Cass Review, che a sua volta è stata criticata dagli esperti per i suoi difetti metodologici e pregiudizi”.**
- b) **non aver “preso in considerazione le conclusioni della Cass Review che non supportano il risultato della [Review]”.**

**Replica:**

- a) “L'APA cita due fonti come “critiche di esperti” alla Cass Review. Una è un saggio online non sottoposto a peer review, la cui paternità è comunemente ma erroneamente attribuita alla Yale University (McNamara et al., 2024). L'altra è un articolo sottoposto a peer review (Noone et al., 2025) che critica principalmente le revisioni sistematiche dell'Università di York (una delle principali fonti di evidenza commissionate per la Cass Review) e commenta anche la Cass Review stessa. Ad oggi, almeno tre articoli hanno contestato le affermazioni centrali di McNamara et al. (2024) (Cheung et al., 2025; Kingdon et al., 2025; McDeavitt et al., 2025), e il primo e il terzo di questi hanno anche commentato Noone et al. (2025).

Come tutte le pubblicazioni scientifiche, la Cass Review presenta delle limitazioni. Inoltre, il disaccordo è comune nella scienza e il dibattito dovrebbe essere accolto positivamente. Tuttavia, i dibattiti attuali sulla Cass Review si basano in gran parte su dimostrabili travisamenti e semplici errori di fatto e sembrano far parte di una campagna di disinformazione scientifica (Kingdon et al., 2025). Cheung et al. (2025) criticano McNamara et al. (2024) per aver scambiato la Cass Review per una linea guida di pratica clinica (CPG). Le revisioni indipendenti sono un processo specifico del Regno Unito, utilizzato quando un'area della medicina inizia a operare in modo da mettere a rischio la sicurezza dei pazienti o compromettere la qualità delle cure. Le revisioni indipendenti rispettano i “termini di riferimento” stabiliti dall'ente commissionante, piuttosto che gli standard per lo sviluppo delle CPG. Concordiamo con Baxendale (2025) sul fatto che i dibattiti sull'efficacia degli interventi medici dovrebbero essere risolti dalle parti utilizzando evidenze provenienti da studi al vertice della piramide della medicina basata sulle evidenze (EBM) (vedi Appendice 3 del Review) e affrontando gli argomenti degli avversari in buona fede — piuttosto che fare affidamento sull'autorità, citare studi irrilevanti o leggere superficialmente.”<sup>26</sup>

- b) Passiamo ora alla seconda accusa dell'APA, secondo cui il Review selezionerebbe in modo parziale dalla Cass Review, omettendo la “conclusione” secondo cui “per alcuni, il miglior risultato sarà la transizione”. L'APA sostiene che questa “conclusione” sia incoerente con i risultati dell'HHS Review.

L'affermazione dell'APA secondo cui questa sarebbe una “conclusione” della Cass Review è un altro esempio della campagna di disinformazione in corso contro di essa. Quando viene compresa correttamente nel contesto, e alla luce di altre osservazioni critiche presenti nella Cass Review, la citazione è meglio interpretata come una considerazione all'interno di un ragionamento clinico più sfumato. Fondamentalmente, la citazione non supporta l'insinuazione dell'APA secondo cui la Cass Review sarebbe favorevole alla PMT in ambito clinico. Per riferimento, la citazione completa richiamata dall'APA è:

*Per alcuni, il miglior risultato sarà la transizione, mentre altri potrebbero risolvere il loro disagio in altri modi. Alcuni potrebbero intraprendere la transizione e poi de/transizionare e/o provare rimpianto. Il NHS deve prendersi cura di tutti coloro che cercano supporto. (Cass, 2025, p. 21)*

La Cass Review ha rilevato che le evidenze sui benefici sono “deboli”, che “i clinici ... non sono in grado di determinare con certezza quali bambini e giovani continueranno ad avere

un'identità trans duratura" e che "una diagnosi di disforia di genere ... non è un indicatore affidabile per prevedere se quel giovane avrà una incongruenza di genere persistente in futuro, o se l'intervento medico sarà la scelta migliore per lui/lei".<sup>27</sup>

4. Per quanto riguarda la mancata divulgazione iniziale dei contributori del Review, l'APA afferma che ciò impedisce agli altri di "valutare la competenza dei contributori, esaminare le loro qualifiche nei campi pertinenti e identificare potenziali conflitti di interesse o impegni ideologici".

**Replica:**

"Per ridurre al minimo i *bias*, l'HHS ha adottato i seguenti passaggi:

- **Primo**, riconoscendo pienamente il clima altamente politicizzato che circonda la cura dei giovani con disforia di genere, l'HHS ha deliberatamente cercato contributori esperti provenienti da un'ampia gamma di posizioni politiche, inclusi coloro che non erano politicamente allineati con l'amministrazione che ha commissionato la Review.
- **Secondo**, le sintesi delle evidenze della Review hanno seguito una metodologia consolidata, rigorosa e riproducibile (Pollock et al., 2024). Questo garantisce che, se lo stesso progetto di panoramica delle revisioni sistematiche fosse condotto nuovamente nelle stesse condizioni, un team indipendente arriverebbe a risultati comparabili e trarrebbe conclusioni simili riguardo alle evidenze di beneficio della PMT, classificate come di certezza molto bassa.
- **Terzo**, l'HHS ha condotto una revisione esterna dei risultati, cercando contributi da organizzazioni e singoli esperti con una varietà di prospettive e posizioni sulla PMT. Oltre all'APA, l'HHS ha richiesto contributi dall'American Academy of Pediatrics (AAP) e dalla Endocrine Society (ES), nonché da individui riconosciuti come esperti in questo campo (inclusi coloro dai quali ci si aspettava una critica al Review o ad alcuni suoi aspetti importanti)."<sup>28</sup>

## Repliche a Bester (pag. 69)

- c) Bester raccomanda di "suggerire con più forza ulteriori studi sulla psicoterapia come intervento per la disforia di genere pediatrica (GD)".

**Replica:**

"La revisione riporta che le prove a favore degli approcci psicoterapeutici per le condizioni di salute mentale che spesso accompagnano il disturbo bipolare (ad esempio, la depressione) sono più forti rispetto alle prove dei loro effetti sulla GD stessa e, quindi, che la psicoterapia è un trattamento promettente per le precedenti condizioni nei pazienti che presentano GDL. L'incidenza di condizioni di salute mentale concomitanti è molto alta in questa popolazione e non ci sono prove concrete che la transizione medica pediatrica (PMT) sia sicura o efficace per il trattamento di tali condizioni, così come non ci sono prove concrete che la PMT sia sicura o efficace nel trattamento della GD pediatrica stessa. La panoramica delle revisioni sistematiche della revisione (Appendice 4) conclude che le prove a beneficio degli interventi psicoterapeutici per il trattamento della GD pediatrica sono incerte. Ciò è probabilmente dovuto almeno in parte alla scarsità di studi primari che esaminano gli effetti della psicoterapia ... Siamo d'accordo con Bester sul fatto che una solida ricerca sulla psicoterapia sia necessaria. Incoraggiamo i ricercatori a condurre tali ricerche e a incorporare i loro risultati nello sviluppo di linee guida affidabili e basate sull'evidenza per la gestione della disforia di genere pediatrica."<sup>29</sup>

- d) **Bester raccomanda “che la revisione includa una discussione più dettagliata sul consenso informato e su come vengono prese le decisioni mediche per i minori, le ragioni etiche per cui ciò avviene ed infine come il processo decisionale e le procedure di consenso differiscano nel caso della transizione di genere.”**

**Replica:**

“Certamente, si potrebbe dire di più sul consenso informato. La Review si concentra sulla questione clinica ed etica preliminare se sia lecito offrire la PMT ai pazienti in primo luogo. Le questioni di autonomia e consenso diventano urgenti solo dopo che è stato stabilito che è clinicamente ed eticamente giustificato offrire un intervento ai pazienti (vedi Capitolo 13). Poiché un profilo rischio/beneficio favorevole è una condizione necessaria affinché qualsiasi intervento pediatrico sia eticamente giustificato, una discussione approfondita sulla possibilità che i minori (o i loro tutori legali) possano dare il consenso alla PMT sarebbe prematura. Sebbene possa essere utile esplorare il tema dell'autonomia e del consenso del paziente in un contesto ipotetico in cui la PMT sia nota per avere un profilo rischio/beneficio favorevole, l'ambito della Review è stato limitato alla valutazione delle migliori pratiche nel contesto delle evidenze esistenti sui rischi e benefici.”<sup>30</sup>

- e) **“Mi colpisce da tempo il fatto che la pressione per consentire ai minori di guidare il processo decisionale in questo ambito si discosti notevolmente da come normalmente avviene la presa di decisioni mediche per i minori. Di solito, i genitori sono i decisori per i minori e, insieme ai clinici, prendono decisioni che servono al miglior interesse del minore.”**

**Replica:**

“Certamente, si potrebbe dire di più sul consenso informato. La Review si concentra sulla questione clinica ed etica preliminare se sia lecito offrire la PMT ai pazienti in primo luogo. Le questioni di autonomia e consenso diventano urgenti solo dopo che è stato stabilito che è clinicamente ed eticamente giustificato offrire un intervento ai pazienti (vedi Capitolo 13). Poiché un profilo rischio/beneficio favorevole è una condizione necessaria affinché qualsiasi intervento pediatrico sia eticamente giustificato, una discussione approfondita sulla possibilità che i minori (o i loro tutori legali) possano dare il consenso alla PMT sarebbe prematura. Sebbene possa essere utile esplorare il tema dell'autonomia e del consenso del paziente in un contesto ipotetico in cui la PMT sia nota per avere un profilo rischio/beneficio favorevole, l'ambito della Review è stato limitato alla valutazione delle migliori pratiche nel contesto delle evidenze esistenti sui rischi e benefici.”<sup>31</sup>

## **Repliche a Gribble (pag. 72)**

**1. “Gribble raccomanda di:**

- a) **“menzionare i rischi dell'uso di una terminologia che suggerisce che le persone possono cambiare sesso”;**
- b) **“riconsiderare la terminologia “da maschio a femmina”, “da femmina a maschio” e “chirurgia di riassegnazione del sesso”;**
- c) **far notare “che non tutti applicano il concetto di identità di genere a sé stessi.”<sup>32</sup>**

**Replica:**

- a) **“...abbiamo aggiunto il seguente testo sottolineato alla nota 16 nella Sezione 2.1 (“Terminologia nella medicina di genere pediatrica”): American Psychological Association (2024a). L'APA ha capovolto l'“implicazione problematica”: la terminologia che suggerisce che il sesso di una persona sia una caratteristica mutevole è fuorviante per i pazienti e dovrebbe essere evitata.”<sup>33</sup>**

- b) "I termini "da maschio a femmina" e "da femmina a maschio" sono molto familiari e facilmente interpretabili come indicatori della direzione aspirazionale del percorso piuttosto che di un cambiamento letterale di sesso. "Chirurgia di riassegnazione sessuale" ha lo svantaggio di suggerire una precedente 'assegnazione', ma è meno problematico rispetto al vecchio "cambio di sesso" o all'attuale "chirurgia di conferma/affermazione di genere". Piuttosto che moltiplicare la terminologia, riteniamo opportuno mantenere il nostro uso originale".<sup>34</sup>
- c) "Per quanto riguarda il punto (c), citiamo Sullivan (2025) nella nota 41 (Sezione 2.2, "Terminologia utilizzata in questa revisione"): "Le domande sull'identità di genere dovrebbero tenere conto del fatto che il concetto di identità di genere in quanto tale può risultare sconosciuto, poco chiaro o irrilevante per alcuni intervistati e che molti intervistati potrebbero non percepire sé stessi come aventi un'identità di genere. Le domande non dovrebbero presupporre che gli intervistati concordino sul fatto di avere un'identità di genere". Abbiamo modificato l'inizio di tale nota a piè di pagina per evidenziare più chiaramente il punto sollevato da Gribble: Va sottolineato che non tutti accettano di avere un'identità di genere".<sup>35</sup>

2. **"Né il corpo della revisione né la panoramica delle revisioni sistematiche includono il danno in termini di incapacità di allattare al seno nell'analisi dei risultati. Suggesterei di affrontare questo aspetto... Potrebbe essere utile fornire una citazione che spieghi la natura dell'intervento chirurgico di mascolinizzazione del torace per chiarire perché l'allattamento al seno è impedito."**<sup>36</sup>

**Replica:**

"Abbiamo aggiunto una frase dopo "Inoltre, gli interventi chirurgici per rimuovere organi sani e funzionanti introducono una serie unica di danni iatrogeni che non si riscontrano in altri settori della medicina" all'inizio della Sezione 7.5: Un esempio è la mastectomia eseguita come parte della PMT, che comporta l'impossibilità di allattare al seno e la potenziale perdita della sensibilità del capezzolo. Abbiamo anche aggiunto una nota a piè di pagina alla frase sopra riportata: La mastectomia rimuove le ghiandole mammarie insieme ai dotti che trasferiscono il latte da esse al capezzolo. La perdita di sensibilità al capezzolo è invariabilmente menzionata dai chirurghi che eseguono la "chirurgia superiore" come potenziale effetto collaterale, anche se le prove sulla portata del rischio negli adolescenti e nei giovani adulti sono molto incerte (Miroshnychenko et al., 2025)".<sup>37</sup>

3. **"La revisione non tratta il tema del bendaggio del seno... Invito gli autori a riconsiderare il bendaggio del seno e il tucking genitale nella revisione come un intervento fisico piuttosto che come parte della transizione sociale"**.<sup>38</sup>

**Replica:**

"Abbiamo aggiunto del testo alla nota a piè di pagina (26), che cita Cass (2024, p. 158): La transizione sociale può anche comportare il bendaggio del seno per le donne o il "tucking" (spostamento dei testicoli nei canali inguinali e posizionamento del pene e dello scroto nella regione perineale) per gli uomini. Si tratta di un intervento fisico (non medico) con potenziali effetti negativi sulla salute, a differenza dei tagli di capelli o dei cambiamenti di abbigliamento. Come afferma una recente revisione, "Per quanto riguarda il bendaggio del torace, sono state segnalate numerose implicazioni negative per la salute, con tassi che raggiungono il 97,2%" (Bumphenkatiikul, 2025, p. 5)".<sup>39</sup>

4. **"Il rimpianto associato alla chirurgia di mascolinizzazione del torace non viene menzionato, ma dovrebbe essere aggiunto"**.<sup>40</sup>

**Replica:**

"Abbiamo aggiunto il testo sottolineato al primo paragrafo all'inizio della Sezione 7.6.2 ("Detransizione e rimpianto"): I pazienti di qualsiasi età possono provare rimpianto per gli effetti fisici e fisiologici permanenti della CSH, indipendentemente da come si identificano. Ad esempio, i pazienti che si identificano stabilmente come transgender possono rimpiangere la perdita della fertilità. Lo sviluppo di calvizie o irritazioni/fastidi causati dalla clitoromegalia possono portare un paziente che si

identifica come uomo transgender a pentirsi di aver assunto testosterone. Tale paziente, o (in particolare) una donna che ha effettuato la detransizione, può pentirsi di aver subito una mastectomia con la conseguente perdita della capacità di allattare al seno".<sup>41</sup>

## Repliche a Santen (pag. 75)

1. **"Santen raccomanda che la revisione sia più esplicita riguardo al "concetto di 'stacking', alla sua definizione e al suo ruolo nello sviluppo delle linee guida".<sup>42</sup>**

### Replica:

Riteniamo che la revisione descriva adeguatamente i problemi relativi al conflitto di interessi nello sviluppo delle linee guida WPATH, ma abbiamo aggiunto una sintesi di tre paragrafi dei punti sopra indicati nella Sezione 9.2.3. Il termine "panel stacking" si riferisce alla pratica di popolare i gruppi di sviluppo delle linee guida cliniche con individui che condividono una posizione simile riguardo al trattamento in esame, spesso a causa di conflitti di interesse finanziari o non finanziari. A causa della sua tendenza a perpetuare il pensiero di gruppo, il panel stacking rappresenta una minaccia per l'affidabilità di qualsiasi CPG, soprattutto in assenza di revisioni sistematiche delle prove.<sup>39</sup> La gestione dei conflitti di interesse è essenziale nello sviluppo delle CPG perché queste ultime formulano raccomandazioni tenendo conto non solo delle prove, ma anche di una serie di altre considerazioni, come i "valori e le preferenze" (vedere la sezione 10.3.1). Le linee guida ES per questo settore della medicina sono state fortemente influenzate dai medici-ricercatori olandesi che hanno aperto la strada alla pratica della PMT. Tre degli otto autori delle linee guida ES del 2009, che hanno introdotto per la prima volta il PMT nella pratica clinica, erano i fondatori del Protocollo olandese: Peggy Cohen-Kettenis, Henriette A. Delemarre-van de Waal e Louis J. Gooren.<sup>40</sup> Il gruppo comprendeva anche Norman Spack, cofondatore della prima clinica pediatrica di genere degli Stati Uniti. La maggior parte degli autori erano anche membri e leader di spicco della WPATH. L'aggiornamento del 2017 delle linee guida ES ha continuato a mantenere un forte legame con il team clinico olandese e ha ulteriormente consolidato il rapporto con la WPATH attraverso la co-autorialità. L'impegno intellettuale dei membri del comitato delle linee guida ES trova espressione nel fatto che l'ES ha raccomandato la PMT nonostante non avesse condotto SR sulle prove dei benefici e della maggior parte dei rischi - un forte allontanamento dalle norme che regolano la stesura delle CPG - e nelle insolite dichiarazioni sui "valori e le preferenze" menzionate sopra.<sup>43</sup>

2. **"Santen afferma che è "essenziale" dichiarare "se la terapia ormonale di affermazione del genere (cioè i bloccanti della pubertà e la terapia ormonale cross-sex) sia una pratica sperimentale o accettata". Egli sottolinea che le autorità sanitarie di diversi paesi (ad esempio Svezia, Finlandia e Regno Unito) hanno ritenuto "sperimentali" alcuni o tutti gli aspetti del protocollo endocrino e ne hanno limitato l'uso alla ricerca. Santen raccomanda di includere una discussione separata sullo status "sperimentale" nella revisione dell'HHS."<sup>44</sup>**

### Replica:

"Santen ha ragione nel sostenere che la decisione di etichettare la PMT come "sperimentale" avrebbe conseguenze significative. Dal punto di vista dei pagatori, ciò giustificherebbe probabilmente il rifiuto della copertura. Dal punto di vista della ricerca, richiederebbe la sottoposizione della PMT a studi clinici adeguati e approvati dall'IRB, probabilmente dopo studi sugli animali per la sicurezza dei farmaci. Quest'ultimo aspetto è particolarmente importante, poiché è ormai ampiamente riconosciuto che i bloccanti della pubertà non sono un intervento autonomo, ma sono quasi sempre seguiti da ormoni del sesso opposto. Data la complessità della questione, riteniamo che la discussione su questo tema esuli dall'ambito della revisione dell'HHS."<sup>45</sup>

3. **“Santen contesta l'affermazione della revisione dell'HHS secondo cui “la storia naturale della disforia di genere pediatrica è poco conosciuta, anche se i dati esistenti suggeriscono che nella maggior parte dei casi si risolve senza intervento”. La revisione, sostiene, non distingue tra persistenza nell'infanzia e persistenza nell'adolescenza della disforia di genere (GD). Sebbene la GD “sia nota per risolversi comunemente” nei bambini, spiega Santen, l'evidenza che si risolva comunemente negli adolescenti senza una storia prepuberale di GD “non è considerata scientificamente valida”. (Un altro revisore, Strathearn, ha fatto un commento simile in una revisione pre-pubblicazione). Santen raccomanda di eliminare la frase dalla revisione”.**<sup>46</sup>

**Replica:**

“Concordiamo sul fatto che la revisione dovrebbe sottolineare la natura provvisoria delle ricerche emergenti sulla bassa stabilità diagnostica. Il termine “provvisoriamente” è stato aggiunto alla sezione 4.3.2.1 (“Nuove prove sulla storia naturale della disforia di genere”): Sebbene la storia naturale della DG, ovvero il suo decorso in assenza di interventi medici, sia attualmente impossibile da misurare data l'ampia disponibilità di interventi, prove più recenti suggeriscono provvisoriamente che la DG abbia una bassa stabilità diagnostica”.<sup>47</sup>

4. **“Santen suggerisce che gli intervalli specificati nella Revisione (Sezioni 7.4.3 e 7.4.4) per il testosterone nelle donne e l'estradiolo negli uomini sono troppo elevati e dovrebbero essere compresi tra 10 e 35 ng/100 mL per il testosterone (cfr. 2-45 ng/dL nella Revisione) e tra 10 e 40 pg/mL per l'estradiolo (cfr. 60-190 pg/mL nella Revisione)”.**<sup>48</sup>

**Replica:**

“Sebbene gli esperti possano non essere d'accordo sui valori di riferimento specifici, il punto fondamentale è che, indipendentemente dai valori di riferimento utilizzati, i regimi ormonali raccomandati dal WPATH e dalla Endocrine Society ai fini della transizione medica superano di gran lunga i valori normali di estrogeni/estradiolo nei maschi e di testosterone nelle femmine”.<sup>49</sup>

5. **“Santen non concorda con il commento della revisione dell'HHS (Sezione 7.6.1, “Effetti psichiatrici avversi”) sull'abuso di steroidi anabolizzanti in quanto “le quantità di steroidi anabolizzanti che causano i sintomi descritti [reazioni avverse cardiovascolari e psichiatriche] sono molto più elevate rispetto alle quantità utilizzate nella terapia ormonale cross-sex”.**<sup>50</sup>

**Replica:**

“Appreziamo il punto di vista di Santen e concordiamo sul fatto che ulteriori prove, in particolare sul fatto che sia il livello assoluto di testosterone o il livello relativo al range normale femminile ad aumentare il rischio di problemi psichiatrici, consentirebbero una valutazione più attendibile del fenomeno. Abbiamo apportato alcune modifiche alla Sezione 7.6.1. La frase “Negli uomini” è stata aggiunta al secondo paragrafo: Uno studio che ha valutato l'uso di steroidi anabolizzanti a dosi medie (300-1000 mg/settimana) e elevate (>1000 mg/settimana) negli uomini ha rilevato che il 23% degli utenti... Alla fine del terzo paragrafo è stato aggiunto quanto segue: Non è noto se questi pazienti avessero livelli di testosterone compresi tra 320 e 1000 ng/dL (il range raccomandato dalla Endocrine Society per le donne che si sottopongono a transizione medica) o livelli al di fuori di questo range. Ciò che si sa è che i pazienti erano di sesso femminile, erano stati classificati come aventi un'indicazione terapeutica correlata al “transgender”, presentavano problemi psichiatrici ed erano in terapia con testosterone.<sup>99</sup> Sebbene non sia possibile determinare la causalità dai dati FAERS, ciò sottolinea l'importanza di considerare gli eventi psichiatrici avversi come un potenziale rischio nelle pazienti di sesso femminile che iniziano la terapia con testosterone per la transizione medica. (Nota associata: 99 Vedi Gomez-Lumbreras & Villa-Zapata (2024), Tabella 2.)”.<sup>51</sup>

6. **“Santen rileva l'esistenza di una nuova ricerca pubblicata associata all'iniziativa Olson-Kennedy et al. finanziata dal NIH e incoraggia i contributori ad affrontare questa ricerca (“con l'avvertenza che non è stata sottoposta a revisione paritaria”)”.**<sup>52</sup>

### Replica:

“Nel 2025 sono stati pubblicati due studi associati all'iniziativa di Olson-Kennedy et al.: “Salute mentale ed emotiva dei giovani dopo 24 mesi di cure di affermazione del genere iniziate con la soppressione puberale” (Olson-Kennedy, Durazo-Arvizu et al. 2025) e “Salute emotiva dei giovani transgender 24 mesi dopo l'inizio della terapia ormonale di affermazione del genere” (Olson-Kennedy, Wang et al. 2025). Il primo è stato pubblicato come preprint (non sottoposto a revisione paritaria) e solo dopo la pubblicazione della revisione dell'HHS il 1° maggio. Per un'analisi critica, si veda Society for Evidence-Based Gender Medicine (2025). La revisione cita il secondo studio sugli ormoni del sesso opposto. Come spiegato nella sezione 5.7.3 (“Robustezza delle conclusioni di questa panoramica”), “anziché andare oltre ciò che le prove possono sostenere, questa panoramica si limita a riassumere le conclusioni delle SR [revisioni sistematiche]. Di conseguenza, potrebbe non includere alcuni degli studi pubblicati più di recente a causa della tempistica delle ricerche bibliografiche delle SR. Tuttavia, una ricerca mirata [la nota a piè di pagina cita Olson-Kennedy, Wang et al. (2025)] degli studi pubblicati di recente non ha rivelato alcuno studio pubblicato o in corso che possa modificare in modo significativo le conclusioni, in particolare quelle relative ai benefici. Ciò è dovuto a problemi persistenti quali l'assenza di gruppi di confronto, dimensioni inadeguate dei campioni e follow-up limitato”.<sup>53</sup>

7. **“Santen raccomanda che la revisione “evidenzi le differenze nei risultati tra i maschi e le femmine assegnati alla nascita [in Chen et al. (2023)] come complemento alla discussione del manoscritto di Olson-Kennedy”.**<sup>54</sup>

### Replica:

“La revisione evidenzia tali differenze in una sezione dedicata a Chen et al. (2023) (6.2.3): “L'unico miglioramento statisticamente significativo in entrambi i sessi è stato riscontrato nella ‘congruenza dell'aspetto’ misurata dalla ‘scala di congruenza transgender’, che non è stata convalidata nei minori. Gli autori hanno anche riportato miglioramenti statisticamente significativi nella depressione, nell'ansia e nella soddisfazione di vita. Tuttavia, questi miglioramenti erano modesti e di dubbia rilevanza clinica. I miglioramenti statisticamente significativi sono stati osservati solo nelle donne, mentre gli uomini non hanno registrato miglioramenti significativi in queste misure”.<sup>55</sup>

## Repliche a Smids (pag. 85)

1. **“Smids riconosce che il trattamento della terminologia pertinente nel capitolo 2 è “importante e fornisce approfondimenti essenziali”, ma teme che il suo “scetticismo riguardo al termine identità di genere possa essere facilmente interpretato come uno scetticismo totale riguardo all'esperienza dell'incongruenza di genere e possa sembrare sprezzante nei confronti dell'importanza che i sentimenti di genere hanno per le persone trans”.**<sup>56</sup>

### Replica:

“La Rivista riconosce che alcuni bambini e adolescenti provano disagio o angoscia riguardo al proprio corpo sessuato o ai ruoli e alle aspettative sociali ad esso associati. Il suo obiettivo è descrivere e valutare le migliori pratiche attuali per il trattamento di bambini e adolescenti che affrontano proprio questa sfida. La revisione rileva che le principali associazioni mediche professionali statunitensi e i medici che lavorano nelle cliniche di genere hanno adottato una terminologia che, nella migliore delle ipotesi, è inutile per descrivere le esperienze o i problemi dei loro pazienti, ma non nega né minimizza tali esperienze o problemi. Al contrario, la prefazione afferma che quando i pazienti cercano un aiuto professionale, “loro e le loro famiglie dovrebbero ricevere un'assistenza compassionevole, basata su prove scientifiche e adattata alle loro esigenze specifiche”. Per discernere quale assistenza sia adeguata alle esigenze specifiche dei pazienti sono necessari un linguaggio chiaro e termini

scientificamente accurati. A tal fine, sottolineiamo ancora una volta che “il comprensibile desiderio di evitare un linguaggio esclusivo o patologizzante, unito a convinzioni profondamente radicate nel settore, ha portato a un vocabolario e a una modalità di comunicazione scientificamente infondati, che presuppongono risposte a controversie etiche e che sono in altri modi fuorvianti” (Sezione 2.1).<sup>57</sup>

2. **“Sebbene Smids riconosca che l'analisi etica della ricerca condotta dalla Review si basi correttamente su “principi etici di ricerca consolidati” che richiedono una ragionevole previsione di un equilibrio positivo tra benefici e rischi, non è del tutto convinto dall'analisi. Questo perché non è certo che nel contesto della ricerca “siamo in grado di affermare che le previsioni di un possibile beneficio complessivo siano irragionevoli anche per un piccolo sottogruppo di adolescenti affetti da GD attualmente sottoposti a PMT”.**<sup>58</sup>

**Replica:**

“Come per gli studi su qualsiasi intervento clinico, il fatto che gli studi condotti finora non abbiano trovato prove evidenti che la PMT migliori i risultati sanitari non esclude, in linea di principio, la possibilità che alcune sottopopolazioni di soggetti traggano beneficio dagli interventi mentre altre ne siano danneggiate. Tuttavia, ad oggi non è stato dimostrato che alcuna sottopopolazione ne tragga beneficio. Inoltre, i medici non sono in grado di prevedere quali pazienti soffriranno di GD persistente in età adulta e quali vedranno una risoluzione dei sintomi. Né i medici che seguono il modello americano di “affermazione del genere” cercano di fare tali previsioni (vedi, ad esempio, il capitolo 11). Concordiamo con Smids sulla necessità di ulteriori ricerche, ma per i motivi esposti nella Sezione 13.5 riteniamo che “la somministrazione di PMT agli adolescenti, anche in un contesto di ricerca, sia in contrasto con le norme etiche consolidate per la ricerca sui soggetti umani”. Delineare aree specifiche di ricerca futura nella medicina di genere pediatrica è una sfida urgente che esula dall'ambito della presente revisione. Tuttavia, è importante che l'eziologia e la storia naturale della GD pediatrica continuino ad essere studiate e che vengano caratterizzati gli effetti a breve e lungo termine degli interventi ormonali sulla salute. Incoraggiamo i ricercatori a esplorare modi alternativi per studiare questa popolazione vulnerabile, ad esempio analizzando i dati esistenti, reclutando partecipanti alla ricerca tra la popolazione adulta che ha ricevuto una diagnosi di GD da bambini o adolescenti (che abbiano ricevuto cure mediche o meno) e conducendo studi clinici che utilizzano interventi psicosociali meno invasivi e rischiosi. Sottolineiamo inoltre che la ricerca clinica procede tipicamente con una descrizione ragionevolmente chiara della salute e della malattia nella popolazione interessata e con una buona comprensione degli obiettivi clinici degli interventi. Qui notiamo ancora una volta che nel campo della medicina di genere pediatrica le ragioni dell'intervento medico sono molto contestate, un problema rivelato dagli approcci nosologici fondamentalmente diversi adottati dal DSM-5 e dall'ICD-11, come discusso nella Sezione 13.3”.<sup>59</sup>

3. **“Secondo Smids, il capitolo 11 della Relazione è «molto più accusatorio di quanto sia appropriato per il tipo di relazione che l'analisi dell'HHS intende essere, accusando persino i medici che sono appena diventati oggetto di procedimenti legali». Pur riconoscendo al capitolo 11 il merito di «fornire preziose informazioni», egli sostiene che il «principio fondamentale» che avrebbe dovuto guidare il capitolo è quello secondo cui «si è innocenti fino a prova contraria».**<sup>60</sup>

**Replica:**

“Il capitolo 11 riguarda le mancanze nella medicina pediatrica di genere. Il capitolo descrive come medici e cliniche di primo piano si siano allontanati dagli standard etici della pediatria in modi che mettono a rischio di gravi danni i pazienti giovani e vulnerabili. Le prove presentate nel capitolo includono citazioni dirette di medici di primo piano, mentre ulteriori testimonianze sono fornite da informatori. Sarebbe irresponsabile ignorare queste realtà cliniche in una valutazione completa della transizione medica pediatrica negli Stati Uniti. Pur concordando con Smids sul fatto che “si è innocenti fino a prova contraria”, il capitolo cita prove sufficienti a sostegno delle sue conclusioni.

Naturalmente, il capitolo non formula alcuna affermazione in merito alla legge, poiché ciò eccederebbe il suo ambito di competenza e le competenze professionali dei suoi collaboratori.”<sup>61</sup>

## Repliche a Strathearn (pag. 88)

1. **“Strathearn suggerisce che di riconoscere nella Prefazione, “che non ci sono prove sufficienti per comprendere chiaramente il ‘rischio di potenziale danno’ per alcuni di questi trattamenti. Ad esempio, i risultati a lungo termine (sia i rischi che i benefici) sono incerti per tutte le modalità di trattamento... Tuttavia, la responsabilità dei medici di ‘non arrecare prima alcun danno’ implica che l’onere primario delle prove dovrebbe riguardare la probabilità del beneficio, soprattutto quando esiste anche un potenziale danno”.**<sup>62</sup>

### Replica:

“Il seguente testo sottolineato è stato aggiunto alla Prefazione (paragrafi quattro e sette): Avendo riconosciuto la natura sperimentale di questi interventi medici e il loro potenziale di danno (che è stato studiato in modo inadeguato, soprattutto per quanto riguarda i risultati a lungo termine), le autorità sanitarie di diversi paesi hanno imposto delle restrizioni. Ciononostante, il modello di cura “di affermazione del genere” include interventi endocrini e chirurgici irreversibili su minori senza patologie fisiche. Questi interventi comportano il rischio di danni significativi, tra cui infertilità/sterilità, disfunzioni sessuali, riduzione dell'accumulo di densità ossea, effetti cognitivi negativi, malattie cardiovascolari e disturbi metabolici, disturbi psichiatrici, complicanze chirurgiche e rimpianti, e la frequenza e la gravità di questi danni non sono state studiate in modo adeguato. Nel frattempo, revisioni sistematiche delle prove hanno rivelato una profonda incertezza sui presunti benefici di questi interventi”.

<sup>63</sup>

2. **“Strathearn suggerisce che la Parte 1 dell’executive summary Riepilogo Esecutivo dovrebbe menzionare che alcuni paesi hanno limitato i bloccanti della pubertà, gli ormoni eterosessuali e gli interventi chirurgici ai contesti di ricerca”**<sup>64</sup>

### Replica:

“Il seguente testo sottolineato è stato aggiunto alla Parte 1 della sintesi: ...le autorità sanitarie di un numero crescente di paesi hanno limitato l'accesso ai bloccanti della pubertà e agli ormoni del sesso opposto e, nei rari casi in cui sono stati offerti, agli interventi chirurgici per i minori. Queste autorità raccomandano ora approcci psicosociali, piuttosto che interventi ormonali o chirurgici, come trattamento primario e, in alcuni casi, hanno limitato questi ultimi a protocolli di ricerca supervisionati a livello nazionale.”<sup>65</sup>

3. **“Strathearn ritiene che il capitolo 3 sia “in parte basato su congetture e sentito dire” e rileva la che tale capitolo potrebbe essere esposto a pregiudizi”.**<sup>66</sup>

### Replica:

“Il capitolo 3... segue convenzioni accademiche consolidate, supportando le sue affermazioni con letteratura sottoposta a revisione paritaria e di fonti primarie, che i lettori possono consultare per la verifica”.

<sup>67</sup>

4. **“Strathearn osserva che nella Sezione 4.1, la Figura 4.2 dovrebbe includere barre di errore per valutare la variabilità dei punteggi medi. Solleva anche una questione sulla distribuzione dei punteggi”.**<sup>68</sup>

### Replica:

“La Figura 4.2 è stata aggiornata per includere intervalli di confidenza al 95%, con “intervalli di confidenza al 95% aggiunti” inseriti in una nota a piè di pagina. I dati riportati nella ricerca olandese originale non sono sufficienti per rispondere alla ragionevole domanda di Strathearn sulla distribuzione”.<sup>69</sup>

5. **“Strathearn suggerisce che le prove incerte sui risultati della psicoterapia dovrebbero essere menzionate anche nella Sezione 5.7.5 (“Conclusione”)”.**<sup>70</sup>

**Replica:**

“Abbiamo aggiunto una frase alla fine di questa sezione: Questa panoramica sintetizza le migliori evidenze cliniche disponibili dai dati a livello di popolazione, evidenziando un modello coerente tra gli interventi per bambini e adolescenti con GD. I benefici e i rischi della transizione sociale rimangono sconosciuti; PB, CSH e interventi chirurgici producono costantemente determinati effetti fisici e fisiologici; e vi è una notevole incertezza riguardo ai loro esiti psicologici e di salute a lungo termine. Allo stesso modo, vi è incertezza riguardo agli effetti della psicoterapia per la GD”.<sup>71</sup>

6. **“Strathearn chiede chiarimenti in merito alla seguente affermazione contenuta nell'introduzione al capitolo 6: “È ampiamente dimostrato negli adulti che, per lo stesso farmaco, l'uso off-label è associato a tassi considerevolmente più elevati di effetti avversi, soprattutto quando mancano solide prove scientifiche”.**<sup>72</sup>

**Replica:**

Una citazione nell'introduzione al capitolo 5 (Egualte et al., 2016) riporta: Abbiamo riscontrato che l'uso off-label dei farmaci era associato a ADE dopo aver effettuato un aggiustamento per le caratteristiche importanti dei pazienti e dei farmaci. Inoltre, abbiamo notato un gradiente di rischio con tassi più elevati di ADE per gli usi off-label privi di solide prove scientifiche. Sebbene la natura intrinseca del farmaco di causare ADE sia la stessa per gli usi on-label e off-label, essa può essere modificata da una serie di fattori, tra cui la condizione patologica off-label. Inoltre, la mancanza di approvazione da parte di un organismo di regolamentazione implica una mancanza di intervalli di dosaggio sicuri e informazioni inadeguate sulle controindicazioni, che nel complesso rendono più probabili gli ADE. Abbiamo riscontrato che 4 prescrizioni off-label su 5 erano prive di solide prove scientifiche e che questo gruppo presentava tassi più elevati di ADE”.<sup>73</sup>

7. **“Strathearn suggerisce di migliorare e citare più approfonditamente i confronti sui tassi di suicidio riportati nella Sezione 6.2.3”.**<sup>74</sup>

**Replica:**

Il seguente testo sottolineato è stato aggiunto alla discussione sul suicidio: Tuttavia, due dei soggetti dello studio sono morti per suicidio entro un anno dall'inizio della terapia ormonale, con un tasso di suicidio annualizzato di 317 su 100.000 pazienti. Il tasso di suicidi riportato da Chen et al. era superiore a quelli segnalati dalle cliniche PGM nel Regno Unito e in Finlandia (13 ogni 100.000 e 51 ogni 100.000). Anche uno studio belga ha riportato un tasso di suicidi annuale relativamente elevato (1.126 su 100.000); come nello studio di Chen et al., tutti i suicidi dei pazienti in quello studio riguardavano pazienti che assumevano CSH<sup>75</sup>.

**Table 4.1. Suicide mortality in youth referred to pediatric gender medicine clinics (with estimated per 1,000 patient-years rates)**

Study; Country	Age range	Years	Referred youths	Suicides	%	Per 1K patient -years	Were patients who died by suicide taking PBs and/or CSH?
Van Cauwenberg et al. (2021); Belgium	12–18	2007–2016	148	5	3.38	11.26	All suicides among patients taking CSH
Chen et al. (2023); U.S.	12–20	2016–2021	315	2	0.63	3.17	All suicides among patients taking CSH
Ruuska et al. (2024); Finland	<23	1996–2019	2,083	7	0.34	0.51	Unknown (38% of cohort treated with PBs and/or CSH)
Biggs (2022); U.K.	<18	2010–2020	15,032	4	0.03	0.13	Unknown (59% of cohort treated with PBs and/or CSH)

Table adapted from Society for Evidence-Based Gender Medicine (2024a).

8. “Riferendosi alla discussione sulla psicoterapia nel capitolo 13, Strathearn osserva correttamente che, come nel caso degli interventi medici, "nessuna prova di danno non equivale a 'nessun danno potenziale'".<sup>76</sup>

**Replica:**

“Al paragrafo tre della Sezione 13.2.3 è stato aggiunto il seguente testo sottolineato:

Per quanto riguarda i potenziali danni della psicoterapia per gli adolescenti con GD, una revisione sistematica delle prove non ha trovato prove di effetti negativi o avversi in nessuno degli studi esaminati (sebbene l'assenza di prove di danno non implichi prove di nessun danno, la psicoterapia non comporta i rischi medici o chirurgici associati alla PMT)”.<sup>77</sup>

## Repliche A Bekkering & Vankrunkelsven (pag. 93)

1. “La mancanza di una rigorosa segnalazione dei conflitti di interesse (COI) da parte degli autori è la questione più importante in questo caso, dato l'argomento”.<sup>78</sup>

**Replica:**

“Data la natura fortemente polarizzata dell'argomento, i nomi dei collaboratori' sono stati omessi durante il processo di revisione paritaria, in modo che i revisori potessero concentrarsi sul contenuto della revisione anziché sui singoli collaboratori stessi. Questa è una pratica consolidata nella revisione scientifica”.<sup>79</sup>

2. “Sarebbe stata utile una definizione di SR (da includere nell'ombrello [revisione]), ma non abbiamo riscontrato problemi in merito all'inclusione o all'esclusione delle SR”.<sup>80</sup>

**Replica:**

“Abbiamo aggiunto una nota a piè di pagina alla Sezione 1 dell'Appendice 4: "Una SR deve : 1) formulare una domanda di ricerca definita in base agli elementi PICO: popolazione, intervento, controllo/comparatore, risultato; 2) prevedere criteri di ammissibilità per gli studi predefiniti ; 3) seguire adeguati metodi di ricerca sistematica che identifichino tutti gli studi che soddisferebbero i

criteri di ammissibilità; 4) proporre una valutazione della validità dei risultati degli studi inclusi, ad esempio attraverso una valutazione del rischio di bias; e 5) contenere una presentazione e sintesi sistematica delle caratteristiche e dei risultati degli studi inclusi, che può includere una meta-analisi. Le revisioni di ambito, le panoramiche delle revisioni sistematiche (revisioni ombrello) e le revisioni narrative non sono SR.”<sup>81</sup>

3. **“La registrazione del protocollo avrebbe aumentato la trasparenza, così come maggiori dettagli su come sono stati sintetizzati i risultati. Tuttavia, i risultati finali sono descritti in modo trasparente e sono facili da seguire. Sono presenti anche numerose tabelle con informazioni necessarie e pertinenti”.**<sup>82</sup>

**Replica:**

“Purtroppo, dati i limiti di tempo, in questo caso non è stato possibile effettuare la pre-registrazione. La pre-registrazione avrebbe inoltre rivelato i nomi dei collaboratori. Data la polarizzazione della questione, era importante per il processo di revisione tra pari che i nomi non fossero divulgati fino al completamento di tale processo. In breve, è stato necessario bilanciare considerazioni contrastanti e scendere a compromessi. Riteniamo che le informazioni metodologiche contenute nell'Appendice 4 siano sufficientemente dettagliate”.<sup>83</sup>

4. **“Non erano disponibili informazioni sul supporto, sull'autore e sulla disponibilità dei dati e altre informazioni”.**<sup>84</sup>

**Replica:**

“Per quanto riguarda il supporto, la revisione è stata commissionata dai collaboratori dell'appaltatore HHS per questo progetto. Non vi sono state altre fonti di supporto. Per quanto riguarda le informazioni sui collaboratori, si veda la risposta al punto 1 (sopra). Per quanto riguarda la disponibilità dei dati, tutti i dettagli rilevanti relativi alla metodologia e ai risultati sono riportati nel capitolo 5 e nell'appendice 4”.<sup>85</sup>

## **Repliche a Dowshen et al. (pag. 96)**

“Nell'agosto 2025 il Journal of Adolescent Health ha pubblicato un commento intitolato “Una valutazione scientifica critica del rapporto sulla salute e i servizi umani sulla disforia di genere pediatrica” (Dowshen et al., 2025).

Gli autori concludono che la Revisione comporta “numerose violazioni delle norme scientifiche, false dichiarazioni di prove scientifiche e false caratterizzazioni sia dell'identità di genere nei giovani che dello standard di cura.” Secondo loro, la Revisione “è un pericoloso esempio di incursione del governo nella fornitura di assistenza medica basata sulle prove” Le accuse contenute nel commento sono gravi; inoltre, i suoi autori sono leader nel campo della medicina di genere pediatrica. Abbiamo quindi deciso di trattare il commento come una revisione paritaria non richiesta. Apprezziamo l'opportunità di rispondere a questo feedback e alle preoccupazioni di medici e ricercatori del campo della salute di genere per i quali la transizione medica pediatrica (PMT) è vantaggiosa per i pazienti e in linea con gli standard esistenti in medicina”.<sup>86</sup>

### **A. Violazioni di norme scientifiche**

1. **“Dowshen et al. suggeriscono che la Revisione manchi di indipendenza e che “i risultati ... sono stati predeterminati dal potere esecutivo prima della stesura del rapporto stesso”.**<sup>87</sup>

**Replica:**

“...le conclusioni empiriche della Revisione sono state raggiunte tramite una metodologia trasparente e riproducibile e la Revisione ha seguito le norme accademiche standard di citazione e argomentazione. Le conclusioni principali della revisione si basano su una panoramica delle revisioni sistematiche (SR) o “revisione generale” (vedere Capitolo 5 e Appendice 4) e su un'analisi etica (vedere Capitolo 13). Queste hanno seguito principi consolidati nei rispettivi campi: medicina basata sull'evidenza ed etica biomedica. Ciascuna di esse è stata sottoposta a revisione paritaria e indipendente condotta da esperti in materia e, per ciascuna, questi esperti hanno concluso che le analisi sono solide e coerenti con elevati standard professionali. ... se i risultati della revisione avessero riflettuto i pregiudizi degli autori', “avrebbe dovuto essere possibile distinguere i punti in cui lo studio segue ragionamenti motivati, e quelli in cui non rende conto della letteratura esistente o mostra altri problemi” ... Dowshen et al. non hanno fatto nessuna di queste cose”.<sup>88</sup>

2. **“Dowshen et al. temono che il fatto che la Review “non faccia il nome degli autori, renda impossibile valutare l'esistenza di conflitti di interesse finanziari, intellettuali o altro tipo”<sup>89</sup>**

**Replica:**

“La scelta di non rivelare i nomi fino al completamento della revisione paritaria ha avuto lo scopo di contribuire a mantenere l'integrità del processo di revisione. Questa è una pratica standard nell'editoria scientifica”.<sup>90</sup>

3. **“Dowshen et al. sostengono che la Revisione manchi di credibilità in quanto “oltre il 20%” dei suoi riferimenti non provengono da “letteratura scientifica sottoposta a revisione paritaria”.<sup>91</sup>**

**Replica:**

“Concentrarsi sul rapporto tra fonti sottoposte a revisione paritaria e fonti non sottoposte a revisione paritaria è svitante. Ciò che conta è se tutte le prove rilevanti siano state adeguatamente rappresentate. La Revisione aveva il proposito di valutare sia gli elementi di prova che le migliori pratiche. Oltre a confrontarsi con la pertinente letteratura scientifica sottoposta a revisione paritaria, la Revisione esamina, quando opportuno, anche le pubblicazioni non sottoposte a revisione paritaria. Ad esempio, un saggio non sottoposto a revisione paritaria, prodotto dal “Integrity Project” e pubblicato sul sito web della Yale University Law School nel 2024, è rilevante per le prove della PMT ed è citato nella revisione. L'Integrity Project è stato co-fondato dal Dr. Meredith McNamara, autore principale di quel saggio e autore senior di Dowshen et al. La valutazione delle migliori pratiche condotta dalla Review esamina le realtà cliniche degli Stati Uniti, spesso documentate nei documenti depositati in tribunale e nei media quali il New York Times, il Washington Post, il Boston Globe, Reuters e The Economist. Il capitolo 10 cita documenti interni della World Professional Association for Transgender Health (WPATH) emersi nelle discovery afferenti a procedimenti giudiziari. Gli scambi di e-mail tra i membri senior di WPATH rivelano che il processo di sviluppo delle linee guida di WPATH ha violato standard ben riconosciuti. Dowshen e altri non contestano la veridicità delle affermazioni fatte nel capitolo 10 della revisione, che descrive l'accantonamento, da parte di WPATH, delle Revisioni Sistematiche (SR) e la rivisitazione delle proprie raccomandazioni cliniche a seguito ed in risposta di pressioni politiche”.<sup>92</sup>

**B. Errate rappresentazioni delle evidenze scientifiche**

4. **“Dowshen et al. sostengono che la Revisione abbia impropriamente applicato, alla PMT, i criteri di determinazione della qualità delle prove stabiliti dalla metodologia GRADE, che sono estremamente restrittivi, come giustificazione per “rigettare gli standard di cura per i giovani TGD [transgender e di genere diverso]”.<sup>93</sup>**

**Replica:**

“Le nostre conclusioni sul profilo rischi/benefici della PMT non si basano esclusivamente sul fatto che le prove a favore dei benefici siano di bassissima certezza (fatto peraltro dimostrato da diverse revisioni sistematiche (SR), e confermato dalla revisione generale della Revisione). Piuttosto, come spiegato nei capitoli 8 e 13, questa conclusione è supportata da un’analisi standard del rapporto rischi/benefici che incorpora sia i presunti benefici sia i rischi e i danni noti degli interventi rilevanti, paragonandoli al rapporto rischio/beneficio delle alternative”.<sup>94</sup>

5. **“Dowshen et al. sostengono che la Revisione “rappresenti erroneamente ... gli studi, ignorandone spesso le conclusioni principali.” Come esempio, citano gli scritti di Chen et al. del 2023. Dowshen ed altri valutano molto positivamente i risultati di questo studio: “la congruenza dell’aspetto, l’effetto positivo [sic], la soddisfazione di vita e i sintomi di depressione e ansia sono tutti migliorati significativamente dopo 2 anni di terapia ormonale.” Dowshen e altri sostengono che la revisione “ignori” questi risultati e “si concentra esclusivamente sui due decessi per suicidio tra i 315 partecipanti allo studio”.<sup>95</sup>**

**Replica:**

“I risultati dello studio di Chen et al. (oltre ad altri studi frequentemente citati) sono ampiamente discussi nel capitolo 6 della revisione. Sebbene i punteggi auto-dichiarati riferiti a congruenza di aspetto, depressione, ansia e qualità della vita riportino miglioramenti dopo i 24 mesi dall’inizio, 32 questi risultati non dovrebbero essere descritti senza precisare le circostanze del caso. Ad esempio, i maschi non hanno mostrato alcun miglioramento in nessun risultato, fatta eccezione per “la congruenza dell’aspetto”, che è stata misurata, come sottolineato nella Revisione, su una scala che non è mai stata convalidata nei minori. Originariamente la Revisione riportava che il miglioramento della congruenza dell’aspetto era “l’unico risultato statisticamente significativo” ma non precisava che questo riguardava i maschi”. La revisione ha riconosciuto i miglioramenti statisticamente significativi nelle donne. Tuttavia, un miglioramento statisticamente significativo non deve essere confuso con un miglioramento clinicamente significativo o rilevante. In Chen et al., il punteggio medio del Beck Depression Inventory è migliorato nell’arco di 24 mesi da 16,01 a 13,85 (scala a 63 punti); la media della Revised Children’s Manifest Anxiety Scale è migliorata da 59,84 a 57,32 (punteggio T, dove 50 è la media della popolazione e 10 è una deviazione standard); e il punteggio medio di soddisfazione di vita è migliorato da 40,03 a 44,68 (punteggio T) su una sottoscala della NIH Toolbox Emotion Battery. Si tratta di piccoli miglioramenti di dubbia importanza per medici e pazienti. Ad esempio, per il Beck Depression Inventory, i ricercatori hanno suggerito che una diminuzione del 17,5% rispetto al punteggio basale potrebbe rappresentare una differenza “minima clinicamente importante”. La diminuzione media di 2,16 punti su questa scala (63 punti) in Chen et al. (2023) non soddisfa, secondo questo criterio, la soglia minima di importanza clinica. Considerati i danni noti e plausibili di questi interventi, anche se tali benefici minori fossero stabiliti tramite studi ben condotti (ad esempio, studi randomizzati controllati), il profilo rischio/beneficio degli interventi ormonali rimarrebbe sfavorevole. Lo studio di Chen presenta molti altri problemi, tra cui la mancanza di dati, per il 31–34% dei partecipanti, di follow-up utili a misurare gli esiti sulla salute mentale, e la mancata segnalazione di molti esiti preregistrati, quali suicidio, autolesionismo e disforia di genere. Fatto ancora più importante, il disegno osservazionale incontrollato impedisce di trarre conclusioni circa l’eventuale miglioramento causato dal trattamento con ormoni eterosessuali (CSH). Tuttavia, ciò non ha impedito a Chen di fornire una descrizione errata dei risultati ... usando un linguaggio esplicitamente causale: CSH “ha migliorato la congruenza dell’aspetto e del funzionamento psicosociale””.<sup>96</sup>

6. **“Dowshen et al. sostengono che, tenuto conto del fatto che Chen et al. (2023) avrebbero raggiunto risultati molto positivi, ed “altri studi statistici [riportano] miglioramenti nel funzionamento psicosociale dopo il trattamento ...” la revisione stia selezionando immotivatamente i dati della ricerca”.<sup>97</sup>**

**Replica:**

“In proposito, Dowshen et al. citano altri cinque articoli a sostegno della loro affermazione sul miglioramento del funzionamento psicologico. Gli studi pre-post in questo campo hanno mostrato risultati incoerenti per quanto riguarda i miglioramenti psicologici (vedere Capitolo 4). I riferimenti selezionati da questi autori includono Achille et al. (2020), in cui (dopo l'analisi di regressione) la depressione è migliorata solo nei maschi (risultato opposto a quello segnalato da Chen et. al.); Chelliah et al., (2024), in cui il miglioramento riportato consiste in una piccola diminuzione del punteggio medio del Quick Inventory of Depressive Symptomatology (QIDS), da 10,7 a 8,2 (su una scala a 42 punti); e uno studio in cui si è verificato un piccolo miglioramento nella depressione auto-riferita ma non nella misura dell'esito della depressione riferita dal medico (Kuper et al., 2020).

L'ultima delle cinque citazioni (Olson-Kennedy, Wang et al., 2025) è un articolo che riporta dati derivati dalla stessa coorte di pazienti di Chen et al. (2023). È fuorviante insinuare che questo sia uno degli “altri” studi di coorte, quando in realtà si tratta della stessa coorte riportata in Chen et al. (Non è chiaro il motivo per cui i ricercatori di Trans Youth Care stiano diffondendo i dati sui risultati di questo progetto di ricerca su più pubblicazioni. I futuri revisori sistematici dovrebbero essere vigili sulla strategia di frammentazione di una ricerca in una molteplicità di pubblicazioni cd “salami slicing.” Questi problemi illustrano perché non è appropriato fare affidamento su studi osservazionali di bassa qualità quando sono disponibili Revisioni sistematiche rigorose ”.<sup>98</sup>

7. **“Dowshen et al. criticano la Revisione per aver omesso un riferimento a Nunes-Moreno et al. (2025).<sup>99</sup>**

**Replica:**

“Sebbene gli studi Nunes-Moreno e altri utilizzino un set di dati ampio e multicentrico, presentano limitazioni simili a quelle degli studi osservazionali precedentemente riportati. Il loro esame non avrebbe modificato la conclusione delle precedenti SR su PB e CSH, né la conclusione della revisione generale dell'Appendice 4'. Nella medicina basata sull'evidenza, la forza dell'evidenza è determinata dalla qualità, non dalla quantità, degli studi”.<sup>100</sup>

8. **“Dowshen et al. sostengono che la Revisione “non fornisce alcuna prova a sostegno della affermazione che i farmaci che interrompono la pubertà e la terapia ormonale siano dannosi per i giovani affetti da TGD” e che la Revisione afferma addirittura che “le prove dei danni sono ‘scarse’”<sup>101</sup>**

**Replica:**

“La citazione completa dalla Revisione è: 'Anche le prove dei danni associati alla transizione medica pediatrica nelle revisioni sistematiche sono scarse, ma questo risultato deve essere interpretato con cautela. (Riepilogo esecutivo, enfasi aggiunta)'. ... La valutazione dei danni associati alla PMT deve andare oltre le prove fornite dai SR. Come osservano Guyatt e colleghi, “Molte, se non la maggior parte, delle revisioni sistematiche non affrontano alcuni esiti rilevanti, in particolare i danni, associati a un intervento.” Ad esempio, il fatto che gli studi non abbiano riportato dati sull'infertilità (e quindi gli SR non siano stati in grado di acquisirli) non significa che l'infertilità possa essere ignorata in una valutazione completa delle prove. Lungi dal “[non fornire] alcuna prova” di danni, il Capitolo 7 della Revisione presenta prove indirette dettagliate derivate dalla scienza di base, dall'endocrinologia e dalla fisiologia dello sviluppo, dimostrando danni plausibili e biologicamente attesi”<sup>102</sup>

9. **“Dowshen et al. sostengono che i commenti della Revisione sulla mancanza di dati sui risultati a lungo termine siano “fuorvianti”. Menzionano a riguardo due studi olandesi che forniscono “oltre 20 anni di dati di follow-up” e uno studio americano che fornisce “fino a 10 anni” di dati di follow-up<sup>103</sup>.**

**Replica:**

“Il primo studio, van der Loos, Klink et al. (2023), ha preso in esame i percorsi di trattamento. È fuorviante descrivere questo studio come la rielaborazione di “20 anni di dati di follow-up” Qui, 20

anni si riferiscono al periodo di assunzione (dal 1997 al 2018). Per quanto riguarda il follow-up dopo l'inizio della terapia ormonale, il follow-up mediano dello studio è stato di 4,6 anni.

Il secondo studio olandese, van der Loos, Vlot et al. (2023), ha valutato la densità ossea di 25 maschi e 50 femmine trattati con PB seguiti da CSH. Come discusso nel capitolo 7 della revisione, questo studio ha rilevato che i punteggi Z sono tornati ai valori basali pre-trattamento al follow-up.<sup>42</sup> Una limitazione importante di questo studio è il tasso di non partecipazione del 40%.

Lo studio americano Olson et al. (2024) ha valutato la soddisfazione e il rimpianto dopo l'inizio della terapia ormonale. È fuorviante descrivere questo come un follow-up di 10 anni, poiché il follow-up medio è stato di 4,86 anni dopo l'inizio dei PB e di 3,4 anni dopo l'inizio del CSH (il follow-up mediano non è stato riportato e potrebbe essere stato considerevolmente inferiore). In particolare, lo studio potrebbe non essere rappresentativo, poiché i partecipanti a questa coorte hanno subito una transizione sociale completa prima della pubertà (alcuni avevano già due anni al momento della transizione sociale; l'età media alla transizione sociale era 6,4943). Inoltre, l'assenza di dati sugli esiti fisiologici o psichiatrici in questo studio rappresenta una limitazione critica.

Come spiega la Revisione (Sezione 13.4), la soddisfazione e il rimpianto, sebbene punti dati importanti, non sono indicatori validi per valutare la giustificazione della PMT. Una lettera all'editore in risposta a Olson et al. (2024) sottolinea che "la soddisfazione del paziente è generalmente considerata una misura complementare della qualità dell'assistenza sanitaria e viene generalmente valutata dopo aver stabilito la sicurezza e l'efficacia dell'intervento."<sup>104</sup>

### C. Sicurezza della transizione medica pediatrica (PMT)

10. "Dowshen et al. suggeriscono che i bloccanti della pubertà (PB) siano sicuri per "i giovani TGD" perché "sono stati utilizzati in modo sicuro ed efficace per decenni per trattare i giovani cisgender con condizioni mediche come la pubertà precoce"<sup>105</sup>

#### Replica:

"Come si vede al capitolo 7 della revisione esistono differenze tra uso dei bloccanti nel trattamento della pubertà precoce centrale (CPP) ed il loro utilizzo per la disforia di genere.

Innanzitutto, lo scopo: la pubertà precoce è una patologia fisica e i bloccanti vengono utilizzati per arrestare una pubertà insorta con tempistiche anomale. La disforia di genere (GD) non è una patologia fisica e in questo caso i bloccanti vengono utilizzati per interrompere una pubertà che si svolge nei tempi fisiologici.

In secondo luogo, la diagnosi: la pubertà precoce viene diagnosticata utilizzando test oggettivi come l'esame del sangue e la evoluzione della condizione è ben compresa, mentre la diagnosi di disforia si basa su criteri soggettivi e ha scarsa validità predittiva.

Terzo, prognosi: nei casi di pubertà precoce, i bloccanti vengono interrotti e la pubertà riprende. Nei casi pediatrici di disforia di genere, oltre il 90% dei giovani trattati con bloccanti procede con gli ormoni cross sex (CSH) e in questi pazienti la pubertà (correttamente intesa) non riprende.

L'autore principale degli studi Dowshen et al., McNamara, ha recentemente riconosciuto che i bloccanti non dovrebbero essere valutati come un intervento separato, ma piuttosto come una componente di un unico processo di trattamento, composto da bloccanti e ormoni cross sex. Per sua stessa ammissione, quindi, ha poco senso affermare che l'uso dei bloccanti sia sicuro rispetto alla disforia in quanto lo sono per il trattamento della pubertà precoce. Il processo di ricorso combinato (bloccanti seguiti da ormoni cross sex) presenta un profilo di rischio sostanzialmente diverso. Ad esempio, l'infertilità non è un risultato terapeutico atteso quando i bloccanti vengono utilizzati per la pubertà precoce, ma lo è quando detti bloccanti vengono utilizzati prima o assieme agli ormoni cross sex in caso di disforia di genere pediatrica.

Va inoltre osservato che la descrizione fatta da Dowshen et al. dei pazienti affetti da pubertà precoce come “giovani cisgender” è imprecisa. Ogni bambino a cui viene diagnosticata la pubertà precoce sarà candidato ai bloccanti, indipendentemente da come si identifica. Pubertà precoce e disforia di genere rappresentano due scenari clinici distinti e l’implicazione che le preoccupazioni sull’uso dei bloccanti in uno scenario ma non nell’altro siano dovute a discriminazione a causa della propria identità è seriamente fuorviante”<sup>106</sup>

11. **“Dowshen et al. citano “una recente revisione completa commissionata dalla legislatura dello stato dello Utah e completata da esperti dell’Università dello Utah,” che “ha concluso “che i bloccanti della pubertà e la terapia ormonale possono essere utilizzati in modo sicuro anche nei giovani con TGD”<sup>107</sup>**

**Replica:**

“La “Utah Review” citata da Dowshen et al. è stata pubblicata successivamente alla HHS Review e non era stata inclusa per questo. Trattiamo di tale pubblicazione nella risposta all’APA”<sup>108</sup>

#### **D. Identità di genere e accuse di “terapia di conversione”**

12. **“Secondo Dowshen et al. una “premessa centrale” della Revisione è l’affermazione, non supportata, secondo la quale l’identità di genere tra gli adolescenti sia intrinsecamente instabile”<sup>109</sup>**

**Replica:**

“In primo luogo, la Revisione si occupa di disforia di genere, non di “identità di genere,” e non dice nulla sul fatto che la stabilità di quest’ultima rappresenti una “premessa centrale.” In secondo luogo, come discusso nella Revisione, la presunta permanenza della disforia di genere adolescenziale (a differenza di quella infantile) è servita come base per il Protocollo olandese, ma tale ipotesi non si è mai basata su prove convincenti. Nella Revisione abbiamo esaminato prove più recenti che suggeriscono che, per un numero significativo di bambini e adolescenti, la disforia di genere possa essere un fenomeno transitorio.

L’onere della prova sulla utilità di bloccanti, ormoni cross sex e interventi chirurgici nel trattamento della disforia di genere adolescenziale —che è una condizione di salute mentale— incombe a coloro che sostengono l’utilità di tali interventi”<sup>110</sup>.

13. **“Dowshen et al. contestano la valutazione molto positiva data alla psicoterapia (“esplorativa”) da parte della Revisione, sostenendo che si tratta di “una pratica mal definita che mira a cambiare l’identità di un giovane, cosa che è analoga alla terapia di conversione” “Decenni di prove”, dicono “dimostrano che le pratiche di conversione sono inefficaci e pericolose per la salute psicologica dei transgender [giovani].” Dowshen et al. sostengono che la Revisione include una “raccomandazione” per “pratiche di conversione non basate sull’evidenza”<sup>111</sup>**

**Replica:**

“Dowshen et al. non hanno preso in esame la Sezione 14.5.2.1, che anticipa e confuta questa accusa... Il noto sostenitore della transizione medica pediatrica (PMT), il dott. Jack Turban, direttore del programma di psichiatria di genere presso l’UCSF, ha affermato che “tentativi di conversione e psicoterapia esplorativa sono pratiche distinte e reciprocamente esclusive.” E anche WPATH —la principale organizzazione che sostiene la transizione medico pediatrica (PMT)— ha affermato, in una dichiarazione a condanna della revisione dell’HHS, che “[loro] si oppongono inequivocabilmente” “[equiparando] la terapia di conversione alla psicoterapia” per “i giovani che stanno esplorando la loro identità di genere”<sup>112</sup>

## E. Linee guida e prassi cliniche

14. **“Dowshen et al. criticano la Revisione riguardo al tema del deterioramento degli standard e del crollo delle garanzie di salvaguardia medica negli Stati Uniti. Tali autori si oppongono all'inclusione dei resoconti di "whistleblower" "non verificati", qualificando tali informatori come "individui non direttamente coinvolti nel processo decisionale clinico per i pazienti.”<sup>113</sup>**

### Replica:

“Contrariamente a quanto descritto da Dowshen et al., la maggior parte dei whistleblower (Capitolo 11) sono medici praticanti che hanno curato questa popolazione di pazienti. I loro resoconti sono estremamente rilevanti per comprendere la realtà clinica della medicina pediatrica di genere negli Stati Uniti”<sup>114</sup>

15. **“Dowshen et al. implicano che i resoconti dei whistleblower non dovrebbero essere presi in considerazione perché le linee guida WPATH ed ES prevedono la presenza di “un team interdisciplinare che esegua una valutazione biopsicosociale completa” prima dell'inizio dei bloccanti e ormoni cross-sex, o di intervento chirurgico”<sup>115</sup>**

### Replica:

“Il punto è che i resoconti degli informatori descrivono medici che offrono trattamenti inappropriati nel contesto di team “multidisciplinari” (o “interdisciplinari”) e “valutazioni” (vedere Capitolo 11 e Sezione 14.3).”<sup>116</sup>

16. **“Dowshen et al. affermano che le linee guida che raccomandano la transizione medica pediatrica (PMT) come standard di cura (ad esempio, le linee guida di WPATH ed Endocrine Society) si fondano sulle “prove della migliore qualità disponibile, che attestano il conseguimento migliori risultati riguardo a salute mentale, benessere psicologico e suicidalità”. Inoltre, Dowshen et al. sostengono che l'approccio terapeutico raccomandato nelle linee guida di WPATH sia “basato sull'evidenza” e “su oltre 70 revisioni sistematiche”<sup>117</sup>**

### Replica:

“Le migliori prove disponibili (vale a dire quelle tratte dalle Revisioni sistematiche) non supportano l'affermazione di Dowshen et al. che i risultati psicologici migliorino con la transizione medica pediatrica. La revisione i cui all'Appendice 4' svela che gli effetti sui risultati psicologici sono sconosciuti.

I capitoli 9–11 della revisione individuano gravi problemi nelle linee guida WPATH ed Endocrine Society (ES). Per quanto riguarda la cura dei bambini e degli adolescenti (capitoli 6 e 7), le linee guida del WPATH sono basate sul consenso, non sull'evidenza, in quanto non prendono in considerazione gli elementi di prova desumibili dalle revisioni sistematiche. Il documento "Standards of Care, Versione 8" (SOC-8), infatti, afferma —falsamente— che una revisione sistematica in tema di interventi ormonali nei minori sia “non possibile.”

La sezione 10.3.2 descrive in dettaglio come WPATH abbia soppresso la pubblicazione di alcune revisioni sistematiche commissionate come base per il suddetto documento SOC-8, comprese le revisioni riguardanti il trattamento dei minori. Ciò solleva serie preoccupazioni circa l'integrità scientifica della WPATH in quanto organizzazione sanitaria autoproclamatasi.”<sup>118</sup>

17. **“Dowshen et al. definiscono le linee guida per la pratica clinica WPATH/ES (e altre linee guida che vi fanno riferimento) come “lo standard di cura esistente”, implicando che le linee guida WPATH dovrebbero essere considerate affidabili in quanto “hanno ricevuto ampia approvazione”, “risultano osservate dal 1979”, “sono attualmente alla loro ottava edizione”, “sono state elaborate nell'arco di quasi un decennio” hanno richiesto quasi un decennio per essere elaborate ,” e**

**rappresentano “raccomandazioni formulate con l'accordo di oltre 100 esperti in salute transgender.”<sup>119</sup>**

**Replica:**

“Vorremmo chiarire che non esiste uno standard di cura accettato “esistente” per il trattamento dei pazienti pediatrici affetti da disforia di genere, e che linee guida e policy di tutto il mondo raccomandano approcci terapeutici molto diversi.

Riguardo all'elenco delle caratteristiche di Dowshen et al., a nessuno di questi viene riconosciuta rilevanza ai fini dell'affidabilità di una linea guida. AGREE II, strumento ampiamente utilizzato per valutare l'affidabilità delle linee guida nella pratica clinica, individua sei aspetti rilevanti: ambito e scopo, coinvolgimento delle parti interessate, rigore nel metodo di definizione, chiarezza della presentazione, applicabilità e indipendenza editoriale. Il rigore metodologico nella definizione di linee guida —in particolare, se si basano su revisioni sistematiche invece che sul consenso degli esperti— è considerato l'aspetto più importante. La medicina dovrebbe essere basata sull'evidenza, non su rango ed eminenza.”<sup>120</sup>

## **Repliche a Rider (pag. 112)**

“Come nel caso di Dowshen et al. (2025), abbiamo deciso di considerare Rider et al. (2025) come una revisione tra pari non richiesta. La parte A della presente risposta riassume le nostre risposte alle critiche mosse da Rider et al. che compaiono in altre revisioni tra pari e pubblicazioni, vale a dire Dowshen et al. e la revisione tra pari dell'American Psychiatric Association (APA). La parte B risponde ai pochi punti che non compaiono nelle altre revisioni tra pari. La parte C discute come il commento di Rider et al. esemplifichi i problemi comuni in questo campo: rappresentazione errata della ricerca, pratiche di citazione inadeguate e scarsa comprensione dei principi della medicina basata sull'evidenza (EBM). La continua disponibilità delle riviste sottoposte a revisione tra pari a pubblicare affermazioni palesemente false o fuorvianti sulla medicina basata sull'evidenza e sulle pratiche mediche relative alla salute dei bambini e degli adolescenti è profondamente deplorabile”.<sup>121</sup>

### **A. “Critiche già affrontate nelle risposte a Dowshen et al. e/o all'APA”.<sup>122</sup>**

**1. “Rider et al. obiettano che “gli autori [della recensione] non sono stati citati”.<sup>123</sup>**

**Replica:**

“Come inizialmente spiegato dall'HHS, l'identità dei collaboratori è stata temporaneamente omessa per “contribuire a mantenere l'integrità di questo processo [di revisione tra pari]”. L'omissione dei nomi nella revisione tra pari è una pratica standard nell'editoria accademica. Si veda il punto 2 della nostra risposta a Dowshen et al., nonché la sezione 4 della nostra risposta all'APA”.<sup>124</sup>

**2. “Citando Dowshen et al. (2025), Rider et al. obiettano che “più di un quinto” dei riferimenti della revisione “provengono da articoli di media popolari, blog o social media”.<sup>125</sup>**

**Replica:**

“Per ripetere parte della nostra risposta a Dowshen et al. (punto 4), concentrarsi sul rapporto tra fonti sottoposte a revisione paritaria e fonti non sottoposte a revisione paritaria è fuorviante. La questione importante è se tutte le prove rilevanti siano rappresentate in modo adeguato. La revisione aveva il compito di valutare sia le prove che le migliori pratiche. Oltre a prendere in considerazione la letteratura scientifica pertinente sottoposta a revisione paritaria, la revisione discute, ove opportuno,

anche le pubblicazioni non sottoposte a revisione paritaria. La valutazione delle migliori pratiche effettuata dalla revisione esamina le realtà cliniche negli Stati Uniti, che sono spesso documentate in atti giudiziari e mezzi di comunicazione come il New York Times, il Washington Post, il Boston Globe, Reuters e The Economist”.<sup>126</sup>

- 3. “Rider et al. criticano la provenienza della Revisione (ovvero il decreto esecutivo del gennaio 2025 che incarica il Segretario della Salute e dei Servizi Umani di commissionare la Revisione), descrivendo la Revisione come una “[entità] con un'agenda politica che prende di mira [pazienti, famiglie e fornitori]” e sostenendo che la Revisione “presenta l'agenda politica della Casa Bianca come scienza oggettiva”.<sup>127</sup>**

**Replica:**

“L'ordine esecutivo ha incaricato l'HHS di «pubblicare una revisione della letteratura esistente sulle migliori pratiche per promuovere la salute dei bambini che affermano di soffrire di disforia di genere». Se i risultati della revisione fossero stati dettati da agende politiche preesistenti, dovrebbe essere possibile individuarne gli errori. Come spiegato nella nostra risposta a Dowshen et al. e all'APA, e come ulteriormente dimostrato di seguito, non sono stati forniti esempi di questo tipo. Si veda il punto 1 della nostra risposta a Dowshen et al., nonché la sezione 4 della nostra risposta all'APA.”.<sup>128</sup>

- 4. “Rider et al. definiscono la transizione medica pediatrica (PMT) (tre volte) come “medicamento necessaria”, sostenendo che “le prove scientifiche dimostrano la sua sicurezza ed efficacia nel migliorare i risultati di salute a breve e lungo termine per gli adolescenti TGNB [transgender e non binari]”.<sup>129</sup>**

**Replica:**

“Rider et al. dimostrano ripetutamente di non comprendere i principi fondamentali dell'EBM relativi alla qualità (o certezza) delle prove. È semplicemente errato affermare che le prove scientifiche “dimostrano” la “sicurezza ed efficacia” della PMT. La revisione generale della Review (Appendice 4) lo dimostra chiaramente e Rider et al. non dicono nulla che metta in dubbio tali conclusioni. Si veda il punto 6 della nostra risposta a Dowshen et al. riguardo all'inversione della gerarchia delle prove e la sezione 1 della nostra risposta all'APA riguardo all'elenco dei singoli studi forniti nella loro revisione tra pari. Altri esempi dell'errata interpretazione dell'EBM da parte di Rider et al. sono discussi nei punti 8 e 9 di seguito”.<sup>130</sup>

- 5. “Rider et al. criticano l'impegno della Review nei confronti della Cass Review. Essi sostengono che la Review non abbia affrontato la presunta “omissione di risultati chiave dalla letteratura più ampia” della Cass Review, né il fatto che essa sia stata “criticata negativamente e contestata ripetutamente da organizzazioni professionali ed esperti individuali nel campo dell'assistenza pediatrica di genere”. Essi sostengono inoltre che la Review dell'HHS citi in modo selettivo le conclusioni della Cass Review”.<sup>131</sup>**

**Replica:**

“I risultati della Cass Review sono stati accettati dai due principali partiti politici del Regno Unito e le sue raccomandazioni sono state implementate dal Servizio Sanitario Nazionale britannico. Non sorprende che i medici specializzati in questioni di genere e le associazioni professionali che li rappresentano denigrino una revisione che ha stravolto il loro modello di trattamento preferito nel Regno Unito. Rimandiamo i lettori alle confutazioni complete delle critiche alla Cass Review: si veda la nota 77 nella Parte C qui sotto e la Sezione 3 della nostra risposta all'APA, che affronta anche l'affermazione di Rider et al. secondo cui la Review cita in modo selettivo dalla Cass Review. In breve, le critiche sono piene di falsità dimostrabili e alcune sembrano motivate da obiettivi legali piuttosto che scientifici. Vale anche la pena notare che uno dei riferimenti forniti da Rider et al. (Horton, 2024) a sostegno della loro affermazione secondo cui “approfonditi studi scientifici e giuridici, nonché le critiche degli esperti del settore ... confutano direttamente la valutazione delle prove...” nella Cass Review è stato pubblicato un mese prima della pubblicazione della Cass Review definitiva”.<sup>132</sup>

6. **“Rider et al. sostengono che la Rivista “promuove una pratica dannosa nota come ‘terapia esplorativa’... che è stata definita una forma di terapia di conversione che incoraggia un bambino o un adolescente ad accettare il genere associato al sesso loro assegnato alla nascita”.**<sup>133</sup>

**Replica:**

“In questo caso, Rider et al. si basano sul lavoro dell'avvocato Florence Ashley, che si oppone ai requisiti di valutazione della salute mentale prima dell'inizio della PMT e sostiene l'ampia disponibilità della PMT perché facilita gli “obiettivi di incarnazione del genere” dei minori. Ashley ha ripetutamente confuso la psicoterapia per la GD pediatrica con la terapia di conversione. Inoltre, Rider et al. tentano di rafforzare la loro tesi con citazioni fuorvianti. Un riferimento, la Dichiarazione congiunta degli Stati Uniti, afferma esplicitamente che “l'esplorazione di questioni relative all'identità di genere e all'orientamento sessuale in un modo che non favorisca o presuma una particolare identità o esperienza, non sarebbe considerata terapia di conversione”. Un altro riferimento è un rapporto delle Nazioni Unite che conclude che la terapia di conversione “può costituire tortura”. Tuttavia, gli esempi di terapia di conversione citati nel rapporto includono individui gay “bendati e picchiati con palloni da basket, legati con nastro adesivo, avvolti in coperte e sottoposti a insulti omofobi”. Queste e altre pratiche simili sono certamente ripugnanti, ma non hanno alcuna attinenza con la terapia verbale per i minori affetti da disforia di genere (GD)”.<sup>134</sup>

7. **“Rider et al. mettono a confronto gli interventi ormonali per il GD pediatrico con l'uso di interventi ormonali per altre condizioni mediche pediatriche, ad esempio la pubertà precoce centrale (CPP), nei “giovani cisgender in età prepuberale e puberale”, suggerendo che gli “adolescenti TGNB” vengono ingiustamente discriminati”.**<sup>135</sup>

**Replica:**

“Questa impostazione è estremamente fuorviante. Per riassumere il punto 10 della nostra risposta a Dowshen et al., il capitolo 7 della revisione contrappone l'uso dei bloccanti della pubertà (PB) nel trattamento della CPP al loro uso per la GD. Quando vengono utilizzati per il CPP, i PB arrestano la pubertà precoce (al contrario della pubertà normale) e non sono seguiti dalla somministrazione di CSH, che può causare infertilità permanente e disfunzioni sessuali, oltre ad altri rischi per la salute.<sup>62</sup> Inoltre, come Dowshen et al., anche Rider et al. descrivono in modo impreciso i pazienti affetti da CPP come “giovani cisgender”. Qualsiasi bambino con diagnosi di CPP sarà candidato all'uso dei PB, indipendentemente da come si identifichi. Il CPP e il GD sono due scenari clinici distinti ed è del tutto errato suggerire che le preoccupazioni relative all'uso dei PB in uno scenario ma non nell'altro siano dovute a una discriminazione basata sull'identità”.<sup>136</sup>

## **B. Punti innovativi in Rider et al. (2025)**

8. **“Rider et al. criticano la revisione per aver “omesso” il contesto relativo al “clima politico” e ai “divieti legislativi proposti o esistenti in materia di GAMC [assistenza medica per l'affermazione di genere] per gli adolescenti TGNB e i loro caregiver”. Inoltre, sostengono che ciò “distorce l'applicazione della medicina basata sull'evidenza”.**<sup>137</sup>

**Replica:**

“Questa critica sembra volta a sviare l'attenzione dal contenuto della revisione, concentrandola sul “clima politico” circostante. L'EBM si occupa del processo decisionale clinico basato sulle migliori prove disponibili in materia di sicurezza ed efficacia dei trattamenti, non del clima politico più ampio o della legislazione. Concordiamo sul fatto che il clima politico abbia reso molto difficile il dibattito scientifico, ma sottolineiamo ancora una volta che, se nella revisione sono presenti errori, dovrebbe essere possibile identificarli chiaramente. Rider et al. non individuano alcun errore”.<sup>138</sup>

9. **“Rider et al. sostengono che la revisione “trascuri il pregiudizio nelle revisioni sistematiche [SR] che cita e sottovaluti altri livelli nella ‘gerarchia delle prove’”. Secondo Rider et al., “i dati e gli studi multifattoriali su più livelli nella ‘gerarchia delle prove’ costituiscono il solido corpus di prove a sostegno della [PMT]”.<sup>139</sup>**

**Replica:**

“Per quanto riguarda i presunti problemi di qualità nelle revisioni sistematiche della PMT che hanno trovato prove di certezza molto bassa, Rider et al. fanno riferimento a un'analisi che critica la revisione di Cass, Noone et al. (2025), che sosteneva che due specifiche SR fossero parziali/imperfette.<sup>63</sup> Utilizzando lo strumento di valutazione del rischio di parzialità per le revisioni sistematiche (ROBIS), la revisione è giunta a una conclusione diversa da quella di Noone et al., riscontrando che queste due SR sui bloccanti della pubertà e sugli ormoni del sesso opposto presentavano generalmente un basso rischio di parzialità. Inoltre, anche se queste due SR fossero escluse, la conclusione che la qualità delle prove a favore dei benefici della PMT è di certezza molto bassa rimarrebbe invariata.<sup>64</sup> Rimandiamo Rider et al. alla revisione tra pari inclusa in questo supplemento dai metodologi Dr. Trudy Bekkering e Professor Patrik Vankrunkelsven, che ha concluso che la revisione ombrello della Review è stata condotta in modo appropriato. Se Rider et al. non sono d'accordo con l'analisi della revisione, sarebbe stato utile che spiegassero il perché. La semplice citazione di Noone et al. non fa progredire la comprensione scientifica. L'affermazione di Rider et al. secondo cui la Review “sottovaluta altri livelli nella ‘gerarchia delle prove’” sembra suggerire che avrebbe dovuto ignorare o minimizzare i risultati di revisioni sistematiche di qualità a favore delle conclusioni raggiunte da alcuni studi individuali (“altri livelli”). Tuttavia, ciò costituirebbe un'inversione della gerarchia delle prove e una violazione di un principio fondamentale dell'EBM. Le prove di bassa qualità preferite da Rider et al. non possono essere definite “robuste”.<sup>140</sup>

10. **“Rider et al. affermano che la revisione “non riconosce [...] che la maggior parte dell'assistenza sanitaria pediatrica è guidata da prove di qualità e forza simili a quelle a sostegno della [PMT]”.<sup>141</sup>**

**Replica:**

“Non è vero che «la maggior parte dell'assistenza sanitaria pediatrica sia guidata da prove di qualità e forza simili». La dottoressa Hilary Cass, autrice della Cass Review ed ex presidente del Royal College of Paediatrics and Child Health, ha osservato che le prove dell'efficacia della PMT sono molto deboli, anche rispetto ad altre aree della medicina pediatrica. Ma non è necessario fare appello all'autorità della Cass: il suo giudizio è supportato dalla citazione stessa di Rider et al. (Matheny Antommara et al., 2025). Matheny Antommara et al. hanno analizzato 14 attuali linee guida di pratica clinica pediatrica, scoprendo che il 58% era basato su prove di livello A o B: “Le prove di livello A includono studi randomizzati controllati ben progettati e condotti; le prove di livello B includono studi randomizzati controllati con limitazioni minori o prove coerenti provenienti da più studi osservazionali”. La PMT non è supportata da evidenze di questi livelli: non sono stati condotti studi randomizzati e gli studi osservazionali esistenti sono generalmente di bassa qualità. Pertanto, l'affermazione di Rider et al. secondo cui “la maggior parte” dell'assistenza sanitaria pediatrica è supportata da evidenze di qualità simile a quella che supporta la PMT è falsa. C'è un punto più importante. La qualità delle prove, valutata attraverso una rigorosa revisione sistematica, può informare le parti interessate su ciò che si conosce dell'efficacia di un intervento. Ma non è l'unica considerazione da tenere in conto nel processo decisionale clinico per la PMT o per qualsiasi altro intervento. Occorre considerare anche i danni, la storia naturale della condizione e i profili di rischio/beneficio delle opzioni terapeutiche alternative. Si veda anche la sezione 2 della nostra risposta all'APA. Inoltre, la discussione di Rider et al. sulle prove scientifiche rasenta l'incoerenza. In un passaggio si fa riferimento a “prove sostanziali dei benefici” della PMT, sottintendendo che esistono studi che forniscono prove scientifiche di alta qualità. Tuttavia, due paragrafi prima, gli autori sembrano ammettere che “la maggior parte dell'assistenza sanitaria pediatrica” e la PMT “[non sono

all'altezza della 'gerarchia delle prove scientifiche'". Rider et al. non possono avere entrambe le cose."<sup>142</sup>

11. **“La revisione discute la “disforia di genere a insorgenza rapida” (ROGD); secondo Rider et al., si tratta di una “diagnosi ampiamente screditata”.**<sup>143</sup>

**Replica:**

“La revisione affronta il recente aumento delle adolescenti affette da GD e i vari tentativi di spiegare il nuovo sviluppo della disforia di genere in questa popolazione clinica. “ROGD” è semplicemente un'etichetta per un nuovo fenomeno clinico; contrariamente a quanto affermato da Rider et al., ROGD non è mai stato presentato come una “diagnosi”. Per una discussione, rimandiamo Rider et al. alla sezione 4.3.1.4 della revisione. Rider et al. citano una fonte a sostegno della tesi dello “scredittamento”; questo argomento è discusso nella Parte C di seguito”.<sup>144</sup>

12. **“Rider et al. sostengono che “la revisione dell'HHS paragona il campo della GAC [“cura di affermazione del genere”] allo studio sulla sifilide di Tuskegee” e affermano inoltre che questo non è un paragone legittimo perché i genitori danno il consenso per interventi che causano infertilità, ecc. nei loro figli consenzienti, mentre i partecipanti di Tuskegee non hanno fornito il consenso informato”.**<sup>145</sup>

**Replica:**

“Si tratta di un'interpretazione errata. Rider et al. citano la sezione 13.2.4 della Relazione, che si limita a osservare che il Rapporto Belmont è stato “pubblicato nel 1978 sulla scia dello studio sulla sifilide non trattata condotto dal Servizio sanitario pubblico statunitense a Tuskegee”.<sup>68</sup> Non viene fatto né implicito alcun paragone tra Tuskegee e “il campo della GAC”. La Relazione fa riferimento a Tuskegee un'altra volta, ma in un capitolo diverso. I sostenitori della PMT a volte si oppongono al controllo di questa pratica con la motivazione che il numero di minori sottoposti a questi interventi è relativamente esiguo. Tuskegee è citato nella Sezione 11.2 come esempio di un esperimento medico ampiamente riconosciuto come profondamente immorale, nonostante coinvolgesse un numero relativamente esiguo di persone (e molto inferiore al numero di giovani sottoposti a PMT)”.<sup>146</sup>

### **C. Rider et al. (2025) come esempio di problemi generali**

“Innanzitutto, il commento utilizza una retorica provocatoria. Rider et al. sostengono che la Rivista abbia «scarsa considerazione» per i «diritti civili» dei giovani vulnerabili e confonde la psicoterapia per il GD pediatrico con la terapia di conversione, che descrive come simile alla «tortura». È inappropriato che una rivista sottoposta a revisione paritaria pubblici accuse così gravi, che mettono in discussione non solo la Rivista ma anche la moralità dei suoi collaboratori, senza prove a sostegno... In secondo luogo, come discusso sopra, Rider et al. basano la loro argomentazione contro la Review su una serie di errori informali e ragionamenti sospetti. Gli esempi includono errori genetici (ad esempio, giudicare il contenuto e le conclusioni della Review sulla base dell'Ordine Esecutivo che l'ha commissionata), diversivi (ad esempio, fingere di criticare la Review concentrandosi invece su questioni politiche o legislative) e appelli all'autorità (ad esempio, affermare che la PMT è benefica perché alcune organizzazioni mediche lo dicono). Inoltre, Rider et al. lanciano attacchi *ad hominem*, sembrando insinuare che i collaboratori della Review – presumibilmente sconosciuti a Rider et al. al momento della stesura – siano “individui non qualificati senza competenze nel campo dell'assistenza pediatrica di genere... La decisione di Rider et al. di sostituire l'impegno accademico con appelli all'autorità delle organizzazioni mediche con sede negli Stati Uniti è particolarmente spiacevole, dato che la Rivista dedica intere sezioni alla presentazione di prove di come queste organizzazioni abbiano fuorviato i propri membri, i pazienti e il pubblico. Rider et al. non contestano mai nessuna di queste prove. Un'altra tattica utilizzata da Rider et al. è la citazione acritica di una dichiarazione di denuncia come prova che qualcosa, in questo caso il ROGD, è “scredittato”. Ciò fa parte di un modello diffuso di

problemi di citazione nella letteratura sulla medicina di genere pediatrica. Una dichiarazione che di per sé non contiene bibliografia o collegamenti ipertestuali alla letteratura pertinente viene citata come un dato di fatto; questo è un esempio di “riferimento senza sbocco” ... Rider et al. danno per scontato, ad esempio, che i bambini “di appena cinque anni” possano essere “transgender o non binari”, come se queste categorie potessero essere applicate senza problemi a una fascia d'età che ha una comprensione rudimentale delle differenze di sesso in termini di stereotipi. Presumibilmente Rider et al. seguono l'esempio di una delle loro coautrici, la dottoressa Diane Ehrensaft, che ha insegnato che le azioni dei bambini piccoli, come togliere le mollette dai capelli o sbottonare le tutine, possono essere «comunicazioni preverbalì sul genere» ... L'effetto Woozle consiste nel citare frequentemente una fonte inadeguata a sostegno di una particolare affermazione; ciò può creare l'illusione che la fonte sia autorevole. L'effetto Woozle affligge la ricerca e la difesa della medicina di genere e fa la sua comparsa in Rider et al. Ad esempio, Rider et al. citano Chen et al. (2023) come parte delle “prove scientifiche che dimostrano ... miglioramenti nel benessere e nella qualità della vita”. Qui Rider et al. riproducono le parole di Budge et al. (2024), che cita anche Chen et al. come parte della “ricerca esistente [che] dimostra l'efficacia della [PMT]”, attraverso il “miglioramento del benessere e della qualità della vita”. Allo stesso modo, Dowshen et al. (2025) affermano che Chen et al. hanno riscontrato miglioramenti nella “congruenza dell'aspetto, effetto positivo [sic], soddisfazione di vita e sintomi di depressione e ansia”. Nessuno di questi tre articoli menziona l'elevato tasso di suicidi riportato da Chen et al.; la mancanza di misure di risultato; i miglioramenti di dubbia importanza clinica nelle pazienti di sesso femminile; il fatto che i pazienti di sesso maschile non hanno mostrato miglioramenti nella salute mentale, ma solo nella “congruenza dell'aspetto” (su una scala che non è stata convalidata nei minori); o l'alterazione delle ipotesi centrali dello studio tra il protocollo scritto e l'articolo pubblicato. Si veda anche il punto 5 della nostra risposta a Dowshen et al. e le sezioni 4.3.4 e 6.2.3 della revisione. Rider et al. citano due volte anche Tordoff et al. (2022), un altro studio influente che esemplifica l'effetto Woozle. Tordoff et al. pretendono di dimostrare una riduzione della morbilità psichiatrica a seguito della somministrazione di interventi ormonali per la TMP. Secondo Google Scholar, questo studio è stato citato oltre 700 volte, con una media di oltre 200 citazioni all'anno dalla sua pubblicazione, nonostante una tabella supplementare online dell'articolo non riveli alcun miglioramento statisticamente significativo nei pazienti che hanno ricevuto gli interventi. La revisione discute Tordoff et al. in dettaglio nella sezione 6.2.2. In qualità di ricercatori che hanno osservato attentamente la medicina di genere pediatrica per anni, ci aspettiamo che l'effetto Woozle si applichi a Rider et al. (2025) e Dowshen et al. (2025), che quasi certamente saranno citati acriticamente in futuri articoli sottoposti a revisione paritaria come prova che la revisione è stata “smentita”, nonostante i gravi problemi presenti in questi articoli. Esortiamo vivamente i revisori e gli editori delle riviste scientifiche a prestare maggiore attenzione alle norme accademiche poco rigorose in questo campo e a lavorare per rafforzarle”.<sup>147</sup>

## Appendice

*N.B. Considerato che l'appendice contiene dati analitici e grafici non direttamente trasponibili nel testo, proponiamo di seguito una sintesi elaborata da GenerAzioneD, redatta nel modo più fedele possibile al contenuto originale.*

### **Quando la ricerca scientifica viene amplificata dai rotocalchi**

*Sempre più spesso la ricerca scientifica e i risultati trovano spazio su riviste non specializzate e su rotocalchi, e le informazioni vengono poi amplificate grazie al web, arrivando, sotto forma di articoli divulgativi, sullo smartphone di ognuno e da lì sui social. Una grande eco e un'amplificazione fuori controllo, quindi. Quanti dei fruitori delle notizie vogliono o sono in grado di andare a ritroso e cercare le fonti di ciò che è scritto e di valutare in maniera critica (oggettiva) il messaggio che è arrivato a tante persone?*

L'appendice del « Supplement to Treatment for Pediatric Gender Dysphoria. Review of Evidence and Best Practices. Peer Reviews and Replies», descrive lo strumento ROBINS-I, acronimo di *Risk Of Bias In Non-randomized Studies - of Interventions*, uno strumento messo a punto per valutare ed eventualmente validare le pubblicazioni in campo scientifico, evidenziando la possibilità di distorsioni in esso presenti (*bias*), che rappresentano falle nell'architettura dello studio; si tratta di un protocollo rigoroso, nato per sottoporre a controllo tutti i punti dello studio, evidenziandone le criticità. Nel mondo della ricerca scientifica, infatti, non tutti gli studi sono condotti con una metodica sovrapponibile e la valutazione della qualità metodologica è alla base delle conclusioni raggiunte dallo studio stesso. Lo strumento ROBINS - I è stato studiato e messo a disposizione per offrire una possibilità oggettiva e applicabile in maniera sistematica per evidenziare i livelli di distorsione. Detto in termini semplici, è importante non solo cosa si trova o si "dimostra", ma anche, e soprattutto, "come" si fa la "scoperta".

Questo strumento è particolarmente utile per tutti quegli studi non randomizzati, cioè per quelli in cui il campione inserito nello studio, è composto da partecipanti non assegnati casualmente ai gruppi di trattamento, aumentando la possibilità che fattori esterni influenzino i risultati.

***Un risultato distorto, o affetto da bias, può portare a conclusioni errate e a decisioni cliniche o politiche inadeguate.***

### **Studi solidi?**

Lo strumento ROBINS-I V2 è costruito con 7 domini, cioè sette punti da analizzare e ai quali si attribuisce un valore in base al risultato ottenuto:

- rischio basso → preoccupazione per la distorsione scarsa o nulla
- rischio moderato → esiste una certa preoccupazione riguardo alla distorsione
- rischio grave → lo studio presenta problemi importanti che danno origine a grave rischio ed infine
- rischio critico → lo studio è molto problematico in questo ambito: le caratteristiche dello studio danno origine a un rischio critico di distorsione, tale che il risultato dovrebbe essere generalmente escluso dalle sintesi delle prove.

I sette domini di distorsione individuati in ROBIN I sono

- Dominio 1: Rischio di distorsione dovuto a fattori confondenti (fattori che variano in contemporanea, ma non sono dipendenti tra loro; valga l'esempio classico: con il caldo si mangiano più gelati e si fanno più bagni con conseguente aumento degli annegamenti, ma non sono i gelati a causare, ovviamente, gli annegamenti)

- Dominio 2: Rischio di distorsione nella classificazione degli interventi (necessita di avere la certezza di “chi” ha ricevuto “cosa”)
- Dominio 3: Rischio di distorsione nella selezione dei partecipanti allo studio (o all'analisi) (certezza che il reclutamento dei partecipanti possa essere stato già orientato verso un particolare target)
- Dominio 4: Rischio di distorsione dovuto a deviazioni dagli interventi previsti (i soggetti facenti parte dello studio hanno seguito il programma o hanno fatto altrimenti)
- Dominio 5: Rischio di distorsione dovuto a dati mancanti (non solo relativo all'abbandono di individui precedentemente reclutati, ma anche alle motivazioni che hanno condotto a tale abbandono; le persone che abbandonano lo studio potrebbero infatti essere con caratteristiche “diverse” da quelle che invece ne fanno parte fino alla conclusione, cioè, quanti mancano, quali mancano e perché mancano)
- Dominio 6: Rischio di distorsione derivante dalla misurazione del risultato (variabile che dipende per esempio dal fatto se il risultato è stato misurato con un esame oggettivo o con la somministrazione di un test soggettivo)
- Dominio 7: Rischio di distorsione nella selezione del risultato riportato (oggetto della pubblicazione sono stati tutti i risultati o solo una parte, quella più confacente ai risultati attesi, atti a confermare l'ipotesi iniziale).

Con il metodo ROBINS-I sono stati sottoposti ad esame due importanti e molto diffusi studi e l'analisi dettagliata ed i risultati ottenuti, sono stati inseriti nell'appendice del supplemento “Supplement to Treatment for Pediatric Gender Dysphoria. Review of Evidence and Best Practices. Peer Reviews and Replies”.

I due studi analizzati sono:

- 1) *Changes in Gender Dysphoria, Interpersonal Minority Stress, and Mental Health Among Transgender Youth After One Year of Hormone Therapy*, Chelliah et al. (2024) <https://pubmed.ncbi.nlm.nih.gov/38340124/>
- 2) *Mental Health Diagnoses and Suicidality Among Transgender Youth in Hospital Settings* Nunes – Moreno et al. (2024) <https://pubmed.ncbi.nlm.nih.gov/39016468/>

*In questo breve resoconto vengono illustrati, per semplicità espositiva, solo alcuni dei bias che sono stati messi in evidenza.*

## **Analisi studio N. 1**

Studio di coorte prospettica: il campione era costituito da un gruppo di adolescenti con disforia di genere seguiti per un anno, dall'inizio della terapia ormonale, valutando i cambiamenti nella loro salute mentale e nel loro benessere. I risultati riportavano riduzioni significative di insoddisfazione per il proprio corpo, riduzione di sintomi di depressione ed ansia. Analizzando lo studio, già nel dominio 1, quello sui **fattori confondenti** è stato messo in luce che lo studio aveva una valutazione critica in quanto accanto al gruppo analizzato non era presente il gruppo di controllo. Non si può conoscere cosa sarebbe occorso ad un simile gruppo di giovani ai quali non fosse somministrato il protocollo farmacologico di affermazione di genere. Per esempio, quanti giovani, maturando di un anno, avrebbero ridotto la loro disforia di genere, quanti con un supporto psicologico adeguato, o con una presenza più incisiva del clan familiare?

Altra valutazione di livello critico si è vista nel dominio 5, **quello sui dati mancanti**. 156 partecipanti ad inizio dello studio e 115 al termine. Indicano la perdita di circa il 27% del campione. Il sistema di valutazione ROBINS-I mette in evidenza che non si è a conoscenza delle caratteristiche di coloro che hanno abbandonato lo studio: non avendo dati su chi e perché lo studio è stato abbandonato, potrebbe essere lecito supporre che la parte di campione che non ha tratto giovamento dai farmaci o che ha avuto addirittura effetti collaterali abbia abbandonato la partecipazione allo studio.

Analizzando il dominio 4, quello relativo alle **deviazioni dagli interventi** è stato riscontrato ancora un rischio critico, in quanto non si era a conoscenza di altri eventuali supporti ricevuti dai partecipanti (psicoterapia? Antidepressivi ed ansiolitici?).

Un altro rischio, questa volta grave, è stato riscontrato nel dominio 6, relativamente alla **misurazione dei risultati**. I risultati hanno infatti tratto origine da risposte ottenute mediante somministrazione di questionari. Si è trattato quindi di dati autoriferiti. In uno studio di tale tipo, sia i medici che i pazienti erano a conoscenza del tipo di protocollo farmacologico adottato, e, pertanto, non si può escludere che il pregiudizio abbia influenzato le risposte o che il maggior benessere fosse il risultato di un effetto placebo, dovuto alla consapevolezza di aver ricevuto un trattamento farmacologico.

Alla luce di tali distorsioni è evidente che i risultati ottenuti non possono essere giudicati affidabili e non può essere costruito un legame certo di causa - effetto. ROBINS-I non stabilisce se i giovani siano migliorati o meno, non ne ha la possibilità, ma stabilisce che il nesso tra assunzione di farmaci e miglioramenti è troppo debole per orientare scelte mediche e politiche.

L'utilità di tale studio può essere solo quella di porre quesiti, non certo di fornire risposte affidabili.

## Analisi studio N. 2

Il secondo studio analizzato in appendice voleva dimostrare una correlazione tra ideazione suicidaria e giovani transgender, e, in particolare la sua minore incidenza in caso di prescrizione di farmaci per l'affermazione del genere percepito, o bloccanti della pubertà. A differenza del precedente, lo studio pubblicato partiva da una mole di dati considerevole, in quanto venivano considerati i dati raccolti in banca dati in un arco di oltre 10 anni (dal 2009 al 2019), riferiti a 3414 individui. Applicando lo strumento ROBINS-I, anche in questo caso il dominio 1, sui **fattori confondenti**, ha dato i risultati negativi, raggiungendo la valutazione di rischio critico. Mancavano infatti importanti informazioni sulla salute mentale dei soggetti, sull'atteggiamento delle famiglie, su eventuale psicoterapia associata a meno a trattamenti ormonali.

Altri domini a rischio grave erano: dominio 5 **dati mancanti**, poiché dati confondenti, come potrebbero essere la salute mentale di base e il supporto familiare erano incompleti; dominio 6, **misurazione dei risultati**, in quanto le diagnosi di suicidalità erano raccolte dalla presentazione al pronto soccorso o al ricovero e potrebbero essere sottostimati o registrati in modo differenziale tra i gruppi.

La suicidalità, infatti veniva intesa in base alle diagnosi emesse dai pronto soccorso o a seguito di ricovero ospedaliero, senza considerare la possibile ideazione suicidaria in chi non ha avuto accesso ai servizi sanitari; dominio 7, **selezione dei risultati**, in quanto per i giovani trattati con GAHT, gli autori hanno riportato solo i risultati principali. Ma per i GnRHa, (bloccanti della pubertà) dopo risultati non significativi, in quanto non vi è stata riduzione del rischio suicidalità, gli autori hanno effettuato analisi aggiuntive. (pag. 221)

## Conclusioni

*Con questa sintesi semplificata di ciò che è stato pubblicato in Appendice, si vuole mettere in luce quanto il fattore comunicazione, quando usato in maniera non corretta o superficiale, possa influenzare l'opinione pubblica e le scelte politiche.*

*L'enfasi ricevuta dagli studi analizzati non è stata infatti controbilanciata da opportune osservazioni critiche, che avrebbero dovuto metterne in luce i punti deboli. A monte, in realtà, vi è la questione di ciò che viene effettivamente accettato da riviste scientifiche, che, seppur sottoposte a valutazione di pari, presentano sovente bias che non ne dovrebbero consentire la divulgazione.*

---

<sup>1</sup> "Treatment for Pediatric Gender Dysphoria: Review of Evidence and Best Practices ("the Review") was published by the U.S. Department of Health and Human Services (HHS) on May 1, 2025; revisions were made on May 15.

Following post-publication peer reviews, the Review and the separate Appendix 4 were revised further.

Individual replies to seven solicited peer reviews, together with a reply to two unsolicited peer reviews (Dowshen et al., 2025; Rider et al., 2025), follow. ... In an effort to solicit and incorporate feedback from major medical organizations

---

that have expressed support for pediatric medical transition, HHS invited the American Academy of Pediatrics (AAP), the American Psychiatric Association (APA), and the Endocrine Society to participate in the peer review process. All

three groups have criticized the Review, with the AAP condemning it in an official statement within hours of its release. Unfortunately, the AAP and the Endocrine Society refused HHS's offer to participate. We are grateful to the APA for accepting the invitation".

<sup>2</sup> "American Psychiatric Association ...Our conclusions are that while the HHS Report purports to be a thorough, evidence-based assessment of gender-affirming care for transgender youth, its underlying methodology lacks sufficient transparency and clarity for its findings to be taken at face value. Key elements including literature selection criteria, analytical frameworks, and justification for excluding other studies, and key findings in studies on which the Report relies, are either underexplained or absent. As a result, the Report's claims fall short of the standard of methodological rigor that should be considered a prerequisite for policy guidance in clinical care.

Below are some specific comments on the Report's methodology:

- With one exception, the authors of the report are not identified. Transparency regarding authorship is essential to the integrity of scientific and policy analysis because it allows readers to assess the expertise of contributors, evaluate their qualifications in relevant fields, and identify potential conflicts of interest or ideological commitments.
- The Report fails to clearly articulate how the studies were selected, what criteria governed their inclusion or exclusion, or how their quality was assessed.
- There is no indication that key stakeholders - namely, transgender individuals, their families, and clinicians - were consulted or that their perspectives were considered.
- While the Report is clear about the potential harms of intervening medically, it does not apply any kind of rational scrutiny to potential harms that have been associated with withholding intervention, including higher rates of depression, anxiety, suicidality, and social withdrawal".

<sup>3</sup> Overall impression: The Review is a review of the evidence and ethical analysis of interventions offered to minors (adolescents and children) who have gender dysphoria. This is an important and timely work. It is well written, methodologically rigorous, and makes a significant contribution to the discussion on this topic. I will point out some areas where I would recommend minor improvements or further analysis. These serve as recommendations for improving the work, and do not affect the overall findings of the Review. What is here is thorough, compelling, and well done. The main findings and recommendations of the Review is consistent with the findings and recommendations of other high-level evidence reviews and analyses already published on this topic. Main findings and conclusions of the review: 1. There is no compelling evidence of benefit for gender transition interventions offered to minors with gender dysphoria. The evidence base is of low quality. What evidence there is does not demonstrate any clear benefit resulting from social transition, puberty blockers, cross-sex hormones, or gender modification surgery."

<sup>4</sup> "2. There are significant concerns about the potential harms from gender transition interventions offered to minors with gender dysphoria, and in particular puberty blockers, cross-sex hormones, and gender modification surgery. The best available evidence for potential harms comes from what is known about these interventions from use in other settings and from basic science knowledge. There is a lack of robust studies that have investigated these harms in the setting of gender transition in minors. 3. The Clinical Practice Guidelines that have been most influential and have informed practice related to gender dysphoria in minors in the United States are of low quality. In particular, the WPATH Standards of Care and the guidelines developed by the Endocrine Society are beset by problems that make them unsuitable for use. These include conflicts of interest, methodological problems, and misapplication or misrepresentation of available evidence. 4. Given the lack of demonstrable benefit and concern about potential harms, the use of puberty blockers, cross-sex hormones, and gender transition surgery in minors with gender dysphoria cannot be ethically justified. Discussion of the main findings and contributions of this review: ...Here is what we know. The prevalence of gender dysphoria has increased among children and adolescents in the last 15 to 20 years. Reasons for this are not quite understood but are likely multifactorial. The epidemiology of gender dysphoria has changed. By far the majority of cases will resolve without any intervention or treatment. There is no compelling evidence of benefit from gender transition treatments for resolution of dysphoria or for management of co-morbidities. The treatments used for gender transition are generally not reversible, have long-term implications, and carry the potential for harms".

<sup>5</sup> "What came first was certainty that these treatments are beneficial and medically necessary. What came second was seeking for evidence that would justify this proposition after the fact. When no evidence was forthcoming, shifting justifications for treatment emerged. Ultimately, a commitment to the provision of such interventions became a sort of litmus test, where one's support for children with gender dysphoria and one's credibility as an expert in this arena was

---

contingent on a rock-solid commitment to gender transition interventions. Second, the issue became seen as a civil rights issue and became a matter of politics and activism, rather than a medical issue. Any attempt to ask questions about these treatment approaches was seen as actions of oppression and led to vilification”.

<sup>6</sup> “Comments on methodology and evidence review: 1. The review of evidence for benefit of gender transition interventions: The Review conducts a Systematic Review of Systematic Reviews. In chapters 5 and 6, there appears a description of methodology that shows inclusion criteria, search criteria, a flow diagram to account for studies identified, included, and excluded. The Review makes use of an appropriate tool to assess for bias and analyzes the included observational studies that have been central to previous claims of purported benefit from gender transition treatments in minors. These methods all appear sound, and one can be confident that this systematic review is reproducible, consistent with the methods of systematic reviews of this kind and has identified the best available evidence related to the clinical questions”.

<sup>7</sup> The review of evidence for harm of gender transition interventions: The Review analyzes studies included in the systematic review and shows how they fall short in monitoring for potential harms. Consistent with best practice in evidence-based medicine, the Review then seeks the best available evidence, which is what is known about these interventions in general when used for other purposes, and what is known about these interventions from basic and clinical science considerations. The Review identifies a set of harms that are likely to occur with great certainty, and a set of additional potential harms that may occur with use of these interventions. The methods here are appropriate, and the conclusion that clinicians should be weary of the potential for harm with these interventions is sound.

<sup>8</sup> 4. The review of evidence related to psychotherapy as treatment for gender dysphoria: The final chapter of the Review is focused on psychotherapy. The evidence review here draws on the systematic review in Chapter 5, and it concludes that there is a lack of evidence for psychotherapy in minors with gender dysphoria because it has not been

adequately studied in this context. However, there is evidence that psychotherapy can be useful for managing comorbidities that often accompany gender dysphoria, such as depression or anxiety. This seems reasonable, however one would need to proceed with caution. We cannot just assume that because psychotherapy benefits minors with mental illness but don’t have gender dysphoria, that psychotherapy would have the same benefits and risks for the treatment of distress related to gender dysphoria...A strong recommendation here should be that further studies of psychotherapy in the context of gender dysphoria are needed, given the lack of evidence.

<sup>9</sup> The Review does summarize well how the practice of gender medicine in the US has fallen below the standards of what is required of informed consent. A valid consent process requires a number of things, including full disclosure of relevant information, and voluntariness. The Review demonstrates that full disclosure has not happened in many gender treatment practice settings in the United States, and the true state of lack of benefit and potential harms have generally been obscured when consent is sought from parents and minors. Further, language has been employed on a routine basis that undermines the voluntariness of consent or permission for treatment from parents. The phrase “You can either have a dead daughter or a live son” has been used on routine basis by practitioners of gender transition to push doubting parents to provide consent for gender transition. For one thing, this phrase obscures the truth – there is no evidence that gender transition is lifesaving or that gender dysphoria inevitably leads to death. But more importantly, this is a coercive phrase that places irresistible pressure on parents to acquiesce to gender transition.

<sup>10</sup> It is noted in the introductory section of the executive summary of the review that, the understandable desire to avoid language that may cause discomfort to patients has, in some cases, given rise to modes of communication that lack scientific grounding, that presuppose answers to unresolved ethical controversies, and that risk misleading patients and families. This Review uses scientifically accurate and neutral terminology throughout. I commend this approach and the explanation provided for this reasoning in Chapter 2. Reasoning provided for rejection of terms such as ‘sex assigned at birth,’ use of ‘gender identity’ rather than just ‘gender’ and use of sexed language generally in the Review is well argued. I recommend that it would be worthwhile to add text on the risks of using terminology suggesting that people can change their sex.

<sup>11</sup> In considering the risks and benefits of treatments for pediatric gender dysphoria the Review notes: To discharge their duties of non-maleficence and beneficence, clinicians must ensure, insofar as reasonably possible, that any interventions they offer to patients have clinically favorable risk/benefit profiles relative to the set of available alternatives, which includes doing nothing...The claims made here about the probability and magnitude of harms and benefits are grounded in the best available evidence. Sometimes, the probabilities are known with a high degree of certainty. For example, the probability that mastectomy will lead to an inability to breastfeed is 1.0 or close to

---

it... As for the nature of medical benefits and harms and their relative weights, the Review's working assumptions cohere with common moral intuition, standard medical judgment as revealed in medical diagnostic criteria, and the outcomes of interest to clinicians and researchers, as well as the law. For example, the analysis would conclude that a minor improvement in depressive symptoms does count as a benefit but that such a benefit, even if assured, does not outweigh moderate or even low but nonnegligible risks of infertility or serious sexual dysfunction, loss of breastfeeding function, or lifelong medical dependency, which the Review considers harms. ... I commend the authors for taking this approach. It has been frustrating to see systematic reviews of the evidence of interventions not consider known outcomes simply because those undertaking the primary research have not included them. However, despite the statements quoted above, neither the body of the Review nor the overview of the systematic reviews includes harm in terms of inability to breastfeed in the analysis of findings. I would suggest that this be addressed.

<sup>12</sup> The Review notes that individuals who desist in their transgender identification may experience regret as a result of gender identity-related medical interventions. Regret associated with chest masculinisation surgery is not mentioned but should be added since: 1) this is the most common surgical procedure for minors with gender dysphoria, 2) WPATH and other guidelines omit to recommend that the impact of this surgery on breastfeeding be discussed with those considering it, 3) proponents of this surgery regularly falsely state that this surgery can be reversed, 4) emergence of acute regret may occur many years after surgery as there may be decades between surgery and a woman giving birth and being unable to breastfeed.

<sup>13</sup> I have reviewed the DHHS document and find that the summary of data and detailed discussions reasonably reflect an overview of the information currently available and its interpretation. However, I believe that one area of the DHHS document has not received the emphasis that is essential. That is the issue whether gender-affirming hormone therapy (i.e., puberty blockers and cross-sex hormone therapy) is experimental or accepted practice. I have reviewed the concepts underlying the definition of experimental therapy in this critique and suggest that a specific section be added to address this. In my opinion, whether or not gender affirming hormone therapy is experimental or not is the most important issue underlining all of the current controversy...Experimental therapy: Definition Overview: Determining whether a therapy for a patient is considered experimental is a complex process with significant implications for patient care, ethics, and regulation. There is not a single universal definition, but it generally refers to treatments that are not yet recognized by the professional medical community as effective, safe, and proven for this specific condition

for which they are being used. Considered below is a breakdown of the definition and key considerations.

<sup>14</sup> Therapy is to be typically considered experimental if:

Lack of established efficacy and safety: There is insufficient scientific evidence, for example from well-designed clinical trials, to definitively prove its effectiveness and safety for the intended use. This is often because it is a new, unknown or a rarely used intervention.

Deviation from standard of care: It does not align with the usual clinical practice supported by a consensus of medical practitioners for the specific condition.

Undergoing or awaiting research: It is currently being studied in clinical trials, or it has not yet undergone the necessary rigorous testing or to gain widespread acceptance.

Off-label use in certain context: While off label use, for example using an approved drug for a purpose not specifically approved by regulatory bodies, can sometimes be considered standard practice based on emerging evidence, it can also be considered experimental if the evidence for its new use is limited or speculative.

Not approved by regulatory bodies: In many countries, therapies are considered experimental until they receive approval from regulatory agencies like the U.S. Food and Drug Administration for specific indication.

...In summary, classifying a therapy as experimental hinges on the level of robust scientific evidence for its safety and efficacy in the particular context. The decision to use a therapy requires a collaborative and transparent discussion between the patient, their family, and the medical team, ensuring comprehensive informed consent, rigorous ethical considerations, and adherence to relative regulatory frameworks.

<sup>15</sup> The Finnish and Swedish guidelines and the Cass Review consider gender affirming care to be experimental...Following an SR, Finnish authorities concluded that the body of evidence supporting puberty blockers and cross -sex hormones for youth is inconclusive. Importantly, the guidelines explicitly state that "in the light of available evidence, gender reassignment of a minor is an experimental practice." NHS England introduced major policy changes . "Puberty blockers are no longer routinely commissioned due to insufficient evidence regarding their long term safety and effectiveness. ...In marked contrast, the Clinical Practice Guidelines of the Endocrine Society and WPATH consider gender-affirming

---

care, puberty blockers and cross-sex hormone therapy to be standards of care and supported by evidence. The discussion in these two guidelines highlights the fact that gender affirming care and cross-sex hormone therapy have been used over the past 30 years and are practiced in multiple countries...The DHHS document, in my opinion, needs to specifically address the components of the definition of experimental medicine and how current studies relate to this definition.

<sup>16</sup> The DHHS document was designed to evaluate evidence and not establish guidelines. My suggestion as a reviewer is that the pros and cons of this issue be discussed. My analysis indicates that there is currently no agreement whether gender-affirming hormonal therapy is experimental or standard practice. ...The pros of considering it experimental are that initiating treatment will necessitate all of the high standards applied to research studies, namely: informed consent, discussion of known risks and benefits, rigorous monitoring, safety review board oversight, training requirements of the researchers, and long-term follow-up. The cons are that clinical research trials in this area are difficult, particularly RCTs, and that considering the approaches experimental will result in withholding benefit from many adolescents with gender dysphoria. The second general issue to be discussed is Guideline Stacking: the Endocrine Society published a manuscript on the trustworthiness of guidelines in 2022 (see *JCEM* 107:129- 2138, 2022) cautioning about the practice of “stacking” of clinical practice guideline (CPG) writing committees. The concept of “stacking”, its definition, and its role in guideline development needs to be stated in the DHHS document. Four criteria were proposed by the ES to ensure a trustworthy CPG. (1) to ensure a multidisciplinary CPG, including members with expertise relevant to the topic (2) to encourage panel diversity with factors such as internationality, gender, race/ethnic, and career stage (3) to avoid “stacking” and (4) to ensure adherence to the CGC’s conflict of interest/duality of interest policy. “Stacking” was defined as “inappropriately restricting guideline development group membership to those with a particular point of view.”

<sup>17</sup> **Review Chapter 13 of the HHS report on Gender Dysphoria** This review analyses chapter 13 from the HHS report on gender dysphoria (GD), which deals with the ethics of pediatric gender medicine. I would like to start with a disclosure: I provided constructive critical feedback on an early version of chapter 13, and at a much later stage on several other chapters of the HHS report. Accordingly, the question of why one would cooperate with the production of a report commissioned by the Trump administration that had just characterized pediatric medical gender care as ‘child mutilation’ may, to a lesser extent, also be asked to me. My considerations were that a report would be produced for the HHS anyway, and that it was always better if a good quality analysis would be produced instead of a document written in the same style as the earlier executive order written by the Trump administration. The composition of the team of authors, as far as known by me at that point in time, gave me sufficient confidence that most likely they would write a balanced and evidence-based analysis. So I decided to provide feedback, hoping to help achieving the production of a report with these characteristics.

<sup>18</sup> Chapter 13 argues for the following main theses. First, the commonly held medical ethical principles of beneficence and non-maleficence require sufficient scientific evidence for a favorable risk/benefit profile to justify pediatric medical transition (PMT, as the report calls it). Second, recent attempts to justify PMT on the basis of respect for patient autonomy misconstrue this medical ethical principle, and constitute a radical departure from standard understandings in pediatric gender medicine which take PMT to be justified by its (purported) resulting mental health benefits. Third, the chapter ends with a research-ethical analysis of potential research into PMT that is skeptical of the justification for offering it even in the context of clinical trials...when there is no reasonable evidence of a positive risk/benefit profile, patients do not have a claim on receiving PMT, and not offering PMT is not an infringement of their autonomy.

<sup>19</sup> When it comes to the evaluation of PMT, the chapter makes a very strong cumulative argument for the conclusion that, given our current knowledge, a precautionary approach is most warranted: The natural history of pediatric GD is poorly understood and decades of research has shown that early onset GD usually resolves without medical intervention. There is no compelling evidence that the same will not prove true in the case of adolescent-onset symptoms, and limited evidence suggesting it will. And in any case, it is widely acknowledged that clinicians are unable to distinguish patients whose GD will persist from those whose GD will resolve. Further, there are concerns about the role medicalization itself may play in contributing to the persistence of the conditions being treated, and less invasive and less risky interventions are available. Lastly, medical intervention has known and plausible harms, and decades of research conducted by leading academic institutions have failed to produce reliable evidence of medical benefit.

<sup>20</sup> Yet, we do not know which percentage of adolescents would have outgrown their GD without PMT, or whose GD would sufficiently have decreased for adolescents no longer to desire PMT. It can be 5 %, but also 50 % or even 80%. Administering such invasive treatment as PMT that results in life-long dependency of medical care and has serious medical risks and harms, when there is such profound uncertainty whether the adolescent even needs it, is simply unacceptable. This lack of knowledge of the natural history applies to those with childhood onset GD (Baron & Dierckxsens, 2022; Byrne, 2024) and even more for those with adolescent onset GD (Kaltiala-Heino et al., 2018). For both categories, the potential for overdiagnosis and harmful overtreatment is very high.

---

<sup>21</sup> Regarding the worry that puberty blockers lock adolescents into their GD, the very high percentage of them continuing from puberty blockers to cross-sex hormones, more than 95% (Brik et al., 2020; Carmichael et al., 2021) is a reason for grave concern in this respect. This is especially the case because there are plausible mechanisms for such lock-in effects (Cass, 2022): puberty suppression halts bodily and psychosexual development, while sexual and romantic experience may be instrumental in outgrowing GD (Steensma et al., 2011).

<sup>22</sup> Thank you for this opportunity to review “Treatment for Pediatric Gender Dysphoria: Review of Evidence and Best Practice”. I am a tenured Professor of Pediatrics, Psychiatry, Psychological and Brain Sciences, and Neuroscience and Pharmacology, at the University of Iowa. I am also the Director of the Division of Developmental and Behavioral Pediatrics, and Physician Director of the Center for Disabilities and Development (CDD). My NIH-funded research includes longitudinal studies of parents and infants, focusing on the effects of early experience and maltreatment on child development, as well as the neurobiology of mother-infant attachment. As co-director of an NIH-funded P50 Center, the Hawkeye Intellectual and Developmental Disabilities Research Center (Hawk-IDDRC), I also have a broad interest in the care of children with intellectual, developmental and behavioral conditions, including those with gender dysphoria. In August 2024, as an editorial board member for the Journal of Pediatrics, I received approval from the editor to prepare a Commentary on the assessment and management of pediatric gender dysphoria, summarizing 8 linked systematic reviews commissioned for the U.K. Cass Review. At that time, there appeared to be strong support, both locally and across the U.S., for “gender affirming care”, but little if any acknowledgement of the limited evidence base. The Commentary aimed to highlight this discrepancy and was titled “What We Know and What We Don’t: Evaluating the Evidence for Gender-Affirming Care in Pediatrics”, raising many of the same concerns highlighted in this Review. Unfortunately, but not unexpectedly, the Commentary was not received favorably by peer reviewers, who used many of the arguments effectively countered in this publication. Despite submitting a comprehensive rebuttal of these arguments (attached) and resubmitting to the “Journal of Pediatrics: Clinical Practice”, as recommended by the Journal, the commentary was never published. Overall, the current Review provides a comprehensive summary of the evidence base for many treatment practices in pediatric gender medicine, including social transition, puberty blockers, cross-sex hormones, surgery, and psychotherapy. It also provides a compelling historical context for the current U.S. medical care environment, including the impact of international guidelines, U.S. medical association responses, and information garnered from legal proceedings. The Review provides a strong focus on evidence-based medicine, outlining both the strengths and limitations, supplemented by indirect evidence from basic science and physiology to better understand mechanisms and the likely risk/benefit ratio of treatment. I believe that this Review provides a valuable and much needed contribution to this important field of practice.

<sup>23</sup> Professor Patrik Vankrunkelsven & Dr. Trudy Bekkering

Given our expertise in evidence-based medicine (EBM), we focused the review on the core of the report, namely the umbrella systematic review (SR) about the various treatments....

#### 1. Methods:

The use of an umbrella review is justified by the fact there are many SRs, most using the same studies.

The review used robust methods:

#### ...2. General conclusions

We have no major remarks on the study design, nor on the conclusions. Minor remarks:

- The lack of rigorous reporting of conflict of interest (COI) by authors is the most important issue here, given the topic.
- A definition of an SR (to be included in the umbrella SR) would have been useful, but we found no issues on inclusion or exclusion of SRs.
- The registration of the protocol would have increased transparency, as would more details about how the results were summarized. ... No information was available on support, author information, and availability of data and other information.

#### III. Summary of Chapter 5

Subject of the umbrella SR: What are the effects of social transition, PBs, CSHs, surgeries, and psychotherapy for youth with GD up to 26 years of age? ...

#### 1. Results:

- 17 SR were included: 10 had low risk of bias overall, 7 had high risk of bias overall –

##### (1) 5.2 Outcomes of social transition

- 2 SRs, both low risk of bias - The results suggest that the impact of social transition on long-term GD, psychological outcomes and well-being, and future treatment decisions such as hormones or surgeries remains poorly understood.

##### ...(2) 5.3 Outcomes of puberty blockers (PB)

- 9 SRs, among which are 4 English reviews with low risk of bias - The certainty of evidence is very low regarding the effect of PBs on GD (or gender incongruence), improvement in mental health, and safety. There is high certainty evidence that PBs exert physiological effects (such as sex hormone suppression) and often cause infertility when

---

followed by CSHs, depending on the patient's pubertal stage and sex. Low certainty evidence suggests that PBs may compromise bone health. A high proportion of youth proceed to CSHs.

(3) 5.4 Outcomes of cross-sex hormones

- 8 SRs, among which are 4 English reviews with low risk of bias - The certainty of evidence is very low regarding the effect on GD or incongruence, improvement in mental health, and safety metrics including fertility and bone health. There is high certainty evidence that CSH exert physiological effects.

(4) 5.5 Outcomes of surgery

- 3 SRs, of which 2 are with low risk of bias - Most studies considered mastectomy only. - There is high certainty evidence that mastectomy is associated with predictable surgical complications such as necrosis and scarring. The certainty of evidence is very low regarding the effect of surgery on GD or incongruence, improvement in mental health including suicidality and depression, and long-term outcomes such as sexual function, quality of life, and regret. -

(5) 5.6 Outcomes of psychotherapy - 5 SRs, of which 2 are with low risk of bias. The evidence on the effects of psychotherapy is limited. For mental health outcomes, the certainty of evidence was very low. However, no harms were reported.

2. Discussion Certainty of evidence is very low. Not just because there are no randomized controlled trials (RCTs), as well designed observational studies would also be very helpful. There are no new or ongoing studies that would have an important impact. New studies are needed. New SRs are unlikely to yield novel insights.

<sup>24</sup> The Review's overview of systematic reviews (SRs) was peer-reviewed by two methodologists, Dr. Trudy Bekkering (Belgian Centre for Evidence-Based Medicine) and Professor Patrik Vankrunkelsven (Director, Belgian Centre for Evidence-Based Medicine). Bekkering and Vankrunkelsven used the PRIOR (Preferred Reporting Items for Overviews of Reviews) checklist to assess the overview, and commended its robust methodology. They identified no major issues related to its design or conclusions, noting that "the final results are described transparently and are easy to follow" and that "there are also many tables with necessary and relevant information." Bekkering and Vankrunkelsven's favorable peer review recognizes the methodological rigor of the Review's approach. Contrary to the APA's assertion, Appendix 4 provides a clear, transparent explanation of the Review's search strategy/literature selection criteria (Section 1), exclusion criteria (Section 2.2), the key findings in the studies on which the Review relies (Sections 4–9) and does indeed list the reviewed studies with full citations and digital object identifiers (Section 11). This contradicts the APA's peer review to such an extent that it suggests the reviewers failed to notice the references in the Review (including in the table of contents) to the 174-page Appendix 4.

<sup>25</sup> The Review's analysis of the potential benefits and harms of PMT consists of (1) an overview of systematic reviews (Chapter 5 and Appendix 4); and (2) evidence from basic science and physiology (Chapter 7). The overview of systematic reviews of interventions considered all relevant published literature regarding PMT, including studies that compared the outcomes for populations that received PMT with those that did not. The evidence synthesis found there was no credible evidence of benefits of PMT compared with no PMT in the outcomes referenced by the APA (depression, anxiety, suicidality)—and, by extension, found no credible evidence of harms from not providing PMT. The basic science and physiology analysis assumed that endogenous puberty is not pathological, but a normal process of sexual development through which a child matures into an adult. Disrupting this process has the potential to result in physical harms. Therefore, the basic science and physiology analysis could only yield an assessment of the harms of interrupting a normal physiological process.

<sup>26</sup> The APA cites two sources as "expert criticism" of the Cass Review. One is a non-peerreviewed online essay whose authorship is commonly but erroneously attributed to Yale University (McNamara et al., 2024). The other is a peer-reviewed article (Noone et al., 2025) that primarily critiques the University of York systematic reviews (one of the main sources of evidence commissioned for the Cass Review) and also comments on the Cass Review itself. At least three papers to date have contested the central claims made by McNamara et al. (2024) (Cheung et al., 2025; Kingdon et al., 2025; McDeavitt et al., 2025), with the first and third of these papers also having commented on Noone et al. (2025).<sup>5</sup> Like all scientific publications, the Cass Review has limitations. Further, disagreement is common in science, and debate should be welcomed. However, current debates surrounding the Cass Review are based largely on demonstrable mischaracterizations<sup>5</sup> and simple errors of fact and appear to be part of a scientific misinformation campaign (Kingdon et al., 2025). Cheung et al. (2025) criticize McNamara et al. (2024) for mistaking the Cass Review for a clinical practice guideline (CPG). Independent reviews are a U.K.-specific process deployed when an area of medicine begins to operate in a way that jeopardizes patient safety or compromises care quality. Independent reviews adhere to the "terms of reference" set out by the commissioning body rather than the standards for CPG development. We agree with Baxendale (2025) that debates about the efficacy of medical interventions should be settled by parties using evidence from studies at the top of the evidence-based medicine (EBM) pyramid (see Appendix 3 of the Review), and addressing

---

their opponents' arguments in good faith—rather than relying on authority, citing irrelevant studies, or reading perfunctorily.

<sup>27</sup> We now turn to the APA's second allegation, that the Review selectively picks from the Cass Review, omitting the "conclusion" that "for some, the best outcome will be transition." The APA asserts that this "conclusion" is inconsistent with the findings of the HHS Review. The APA's claim that this was a "conclusion" of the Cass Review is another example of the ongoing misinformation campaign against it.<sup>7</sup> When properly understood in context, and given other critical observations made in the Cass Review, the quote is best understood as one consideration in a more nuanced line of clinical reasoning. Crucially, the quote does not support the APA's insinuation that the Cass Review is supportive of PMT in clinical settings. For reference, the full quotation alluded to by the APA is: For some, the best outcome will be transition, whereas others may resolve their distress in other ways. Some may transition and then de/retransition and/or experience regret. The NHS needs to care for all those seeking support. (Cass, 2025, p. 21) The Cass Review found the evidence for benefit to be "weak," that "clinicians ... are unable to determine with any certainty which children and young people will go on to have an enduring trans identity,"<sup>8</sup> and that "a diagnosis of gender dysphoria ... is not reliably predictive of whether that young person will have longstanding gender incongruence in the future, or whether medical intervention will be the best option for them."

<sup>28</sup> To minimize bias, HHS took the following steps: • First, with full recognition of the highly politicized climate that surrounds the care for gender-dysphoric youth, HHS deliberately sought expert contributors from a wide range of political positions, including those not politically aligned with the administration commissioning the Review. • Second, the Review's evidence syntheses followed a well-established, rigorous, and reproducible methodology (Pollock et al, 2024). This ensures that if the same project of overviewing systematic reviews was conducted again under the same conditions, an independent team would arrive at comparable results and draw similar conclusions of very low certainty evidence for the benefits of PMT. • Third, HHS conducted an external peer review of the findings, seeking input from organizations and individual experts with a diverse set of perspectives and positions on PMT. In addition to the APA, HHS sought input from the American Academy of Pediatrics (AAP) and the Endocrine Society (ES), as well as from individuals who are recognized as experts in this field (including those who would be expected to be critical of the Review or some important aspects of it).

<sup>29</sup> The Review reports that the evidence for benefit of psychotherapeutic approaches for mental health conditions that often accompany GD (for instance, depression) is stronger than the evidence for their effects on GD itself and, therefore, that psychotherapy is a promising treatment for the former conditions in patients presenting with GD. The incidence of co-occurring mental health conditions is very high in this population and there is no good evidence that pediatric medical transition (PMT) is a safe or effective intervention for these indications, just as there is no good evidence that PMT is safe or effective in treating pediatric GD itself. However, crucially, the Review also points out that psychotherapy carries lower risks than PMT. Research indicating that psychotherapy is an effective treatment for a wide range of psychosocial problems, combined with its carrying a lower risk than PMT, suggests that the risk/benefit profile of psychotherapy for treatment of pediatric GD is favorable when compared to more medically aggressive alternatives such as puberty blockers, cross-sex hormones, and surgeries. The Review's overview of systematic reviews (Appendix 4) concludes that the evidence for benefit of psychotherapeutic interventions for the treatment of pediatric GD is uncertain...We agree with Bester that robust research on psychotherapeutic approaches is needed. We encourage researchers to conduct such research and to incorporate their findings when developing trustworthy, evidence-based clinical practice guidelines for the management of pediatric GD.

<sup>30</sup> Certainly, there is more that could be said about informed consent. The Review focuses on the clinically and ethically prior question of whether it is permissible to offer PMT to patients in the first place. Issues of autonomy and consent become pressing only after it has first been established that it is clinically and ethically justified to offer some intervention to patients (see Chapter 13). Because a favorable risk/benefit profile is a necessary condition of any pediatric intervention being ethically justified, a robust discussion of the question of whether minors (or their legal guardians) can consent to PMT would be premature. While it may be valuable to explore the issue of patient autonomy and consent within a hypothetical context in which PMT were known to have a favorable risk/benefit profile, the scope of the Review was limited to an assessment of best practices within the context of existing evidence of risks and benefits.

<sup>31</sup> We agree with Bester that the child-led "affirming" approach that has come to dominate some U.S. gender clinics departs markedly from how pediatric medicine is generally practiced, where medical decision-making is grounded in the health-related best interests of the patient. As quoted in Section 13.2.2 of the Review, the AAP's Committee on Bioethics emphasizes that "parental authority regarding medical decision-making for their minor child or young adult who lacks the capacity for medical decision-making is constrained compared with the more robust autonomy in medical decision-making enjoyed by competent adults" and, moreover, that clinicians' fiduciary duties "to protect and promote the

---

health-related interests of the child and adolescent ... may conflict with the parent's or patient's wishes ..." (Committee on Bioethics et al., 2016, pp. e5, e2). The centrality of the best interest standard to ethical clinical practice in pediatrics is not controversial.

<sup>32</sup> "Gribble recommends (a) mentioning "the risks of using terminology suggesting that people can change their sex"; (b) reconsidering the terminology of "male-to-female," "female-to-male," and "sex reassignment surgery"; and (c) noting that "not everyone applies the concept of gender identity to themselves."

<sup>33</sup> Regarding (a), we have added the following underlined text to footnote 16 in Section 2.1 ("Terminology in pediatric gender medicine"): American Psychological Association (2024a). The APA has the "problematic implication" backwards: terminology that suggests a person's sex is a mutable characteristic is misleading to patients and should be avoided.

<sup>34</sup> "Regarding (b), there is a tradeoff between coining new terminology which may tax the reader and using familiar terminology that is less-than-ideal. "Male-to-female" and "female-to-male" are very familiar and readily interpretable as indicating the aspirational direction of travel rather than a literal change of sex. "Sex reassignment surgery" has the disadvantage that it suggests a prior "assignment" but is less problematic than the older "sex change" or the current "gender confirmation/affirmation surgery." Rather than multiplying terminology, we think it best to keep to our original usage".

<sup>35</sup> "Regarding (c), we do quote from Sullivan (2025) in footnote 41 (Section 2.2, "Terminology in this Review"): "Questions on gender identity should recognize that the concept of gender identity as such will be unfamiliar, unclear or irrelevant to some respondents, and that many respondents may not perceive themselves as having a gender identity. Questions should not assume that respondents will agree that they have a gender identity." We have altered the start of that footnote to bring out Gribble's point more clearly: It should be emphasized that not everyone accepts that they have a gender identity".

<sup>36</sup> "[N]either the body of the Review nor the overview of the systematic reviews includes harm in terms of inability to breastfeed in the analysis of findings. I would suggest that this be addressed ... It may be helpful to provide a citation explaining the nature of chest masculinisation surgery to make it clear why breastfeeding is prevented."

<sup>37</sup> "We have added a sentence after "Further, surgeries to remove healthy and functioning organs introduce a unique set of iatrogenic harms not encountered in other areas of medicine" at the beginning of Section 7.5: An example is mastectomy performed as part of PMT, which results in an inability to breastfeed and potential loss of nipple sensation. We have also added a footnote to the above sentence: A mastectomy removes the mammary glands together with the ducts that transfer milk from them to the nipple. Loss of nipple sensation is invariably mentioned by surgeons performing "top surgery" as a potential side effect, although there is very low certainty evidence about the magnitude of the risk in adolescents and young adults (Miroshnychenko et al., 2025)".

<sup>38</sup> "The Review does not discuss breast binding ... I would encourage the authors to consider reconceptualising breast binding and genital tucking in the Review as a physical intervention rather than as part of social transition."

<sup>39</sup> "We have added text to the accompanying footnote (26), which cites Cass (2024, p. 158): Social transition may also involve breast binding for females or "tucking" (moving the testes into the inguinal canals and positioning the penis and scrotum in the perineal region) for males. This is a (non-medical) physical intervention with potentially adverse health effects, unlike haircuts or clothing changes. As a recent review puts it, "For chest binding, a significant number of negative health implications have been reported, with rates as high as 97.2%" (Bumphenkiatikul, 2025, p. 5)".

<sup>40</sup> "Regret associated with chest masculinisation surgery is not mentioned but should be added."

<sup>41</sup> "We have added the underlined text to the first paragraph at the start of Section 7.6.2 ("Detransition and regret"): Patients of any age may experience regret regarding the permanent physical and physiologic effects of CSH, regardless of how they identify. For example, stably transgender-identified patients may regret loss of fertility. Developing baldness, or chafing/discomfort caused by clitoromegaly, may lead a patient who identifies as a transgender man to regret taking testosterone. Such a patient, or (especially) a detransitioned female, may regret having had a mastectomy with the consequent loss of the ability to breastfeed."

<sup>42</sup> "Santen recommends that the Review be more explicit about "the concept of 'stacking', its definition, and its role in guideline development."

<sup>43</sup> "We believe the Review adequately describes COI problems in the development of WPATH guidelines, but we have added a three-paragraph summary of the points above in Section 9.2.3. "Panel stacking" refers to the practice of populating clinical practice guideline development groups with individuals who share a similar position regarding the treatment under consideration, often due to having financial or non-financial conflicts of interest. Because of its

---

tendency to perpetuate groupthink, panel stacking represents a threat to the trustworthiness of any CPG, especially in the absence of systematic reviews of evidence.<sup>39</sup> Managing conflicts of interest is essential in CPG development because CPGs make recommendations factoring in not only the evidence, but a range of other considerations, such as “values and preferences” (see Section 10.3.1). ES guidelines for this area of medicine have been heavily influenced by the Dutch clinician-researchers who pioneered the practice of PMT. Three of the eight authors of the 2009 ES guidelines, which first introduced PMT into clinical practice, were the founders of the Dutch Protocol: Peggy Cohen-Kettenis, Henriette A. Delemarre-van de Waal, and Louis J. Gooren.<sup>40</sup> The group also included Norman Spack, who co-founded the first U.S. pediatric gender clinic. Most of the authors were also prominent WPATH members and leaders. The 2017 update to the ES guidelines continued to maintain a strong link with the Dutch clinical team and further cemented the relationship with WPATH through common authorship. The intellectual commitments of ES guideline panel members find expression in the fact ES recommended PMT despite not having conducted SRs of evidence for benefits and most risks—a strong departure from norms governing how CPGs should be drafted—and the unusual “values and preference” statements mentioned above”.

<sup>44</sup> “Santen asserts that it is “essential” to state “whether gender-affirming hormone therapy (i.e., puberty blockers and cross-sex hormone therapy) is experimental or accepted practice.” He points out that health authorities in several countries (e.g., Sweden, Finland, and the U.K.) deemed some or all aspects of the endocrine protocol “experimental” and restricted it to research. Santen recommends including a separate discussion of “experimental” status in the HHS Review.”

<sup>45</sup> “Santen is right that a decision to label PMT “experimental” would have significant consequences. From the payor perspective, it likely would justify denial of coverage. From a research perspective, it would require submitting PMT to proper, IRB-approved clinical trials, likely following animal studies for drug safety. The latter is particularly important as it is now widely recognized that puberty blockers are not a standalone intervention but nearly always followed by cross-sex hormones. Given the complexities involved, we believe discussion of this issue is beyond the scope of the HHS Review.”

<sup>46</sup> “Santen takes issue with the HHS Review’s claim that “the natural history of pediatric gender dysphoria is poorly understood, though existing data suggests it will remit without intervention in most cases.” The Review, he argues, fails to distinguish persistence in childhood versus persistence in adolescent gender dysphoria (GD). Although GD “is known to commonly resolve” in children, Santen explains, the evidence that it commonly resolves in adolescents without a prepubertal history of GD is “not considered scientifically sound.” (Another reviewer—Strathearn—made a similar comment in a prepublication review.) Santen recommends deleting the sentence from the Review.”

<sup>47</sup> “We agree that the Review should note the tentative nature of emerging research on low diagnostic stability. “Tentatively” was added to Section 4.3.2.1 (“New evidence about the natural history of gender dysphoria”): Although the natural history of GD—i.e., its course absent medical interventions—is currently impossible to measure given the wide availability of interventions, newer evidence tentatively suggests that GD has a low diagnostic stability”.

<sup>48</sup> “Santen suggests that the ranges specified in the Review (Sections 7.4.3 and 7.4.4) for testosterone in females and estradiol in males are too high and should be 10 to 35 ng/100 mL for testosterone (cf. 2-45 ng/dL in the Review) and 10 to 40 pg/mL for estradiol (cf. 60–190 pg/mL in the Review).”

<sup>49</sup> “While experts may disagree about the specific reference ranges, the critical point is that whatever reference range is used, the hormonal regimens recommended by WPATH and the Endocrine Society for purposes of medical transition far exceed the normal ranges of estrogen/estradiol in males and testosterone in females.”

<sup>50</sup> “Santen disagrees with the HHS Review’s comment (Section 7.6.1, “Adverse psychiatric effects”) about anabolic steroid abuse as “the amounts of anabolic steroid that cause the symptoms described [cardiovascular and psychiatric adverse reactions] are very much higher than the amounts used as cross-sex hormone therapy.”

<sup>51</sup> “We appreciate Santen’s perspective and agree that more evidence—specifically, on whether it is the absolute level of testosterone or the level relative to the normal female range that increases risk for psychiatric problems—would allow for a more confident assessment of the phenomenon. We have made some changes to Section 7.6.1. “In men” was added to a sentence in the second paragraph: One study assessing medium (300–1000 mg/week) and high (>1000 mg/week) anabolic steroid use in men found that 23% of users ... The following has been added to the end of the third paragraph: It is unknown whether these patients had testosterone levels between 320 to 1000 ng/dL (the range recommended by the Endocrine Society for females undergoing medical transition), or levels outside of this range. What is known is that the patients were female, were categorized as having a “transgender” related treatment indication, presented with psychiatric problems, and were on testosterone.<sup>99</sup> Although it is not possible to determine causation from FAERS data, this underscores the importance of considering adverse psychiatric events as a potential risk in female

---

patients initiating testosterone for medical transition. (Associated footnote: 99 See Gomez-Lumbreras & Villa-Zapata (2024), Table 2.)”

<sup>52</sup> “Santen notes the existence of new published research associated with the NIH funded Olson-Kennedy et al. initiative and encourages the contributors to address this research (“with the caveat that it is not peer-reviewed”).”

<sup>53</sup> “Two studies associated with the Olson-Kennedy et al. initiative have been published in 2025: “Mental and emotional health of youth after 24 months of gender-affirming care initiated with pubertal suppression” (Olson-Kennedy, Durazo-Arvizu et al. 2025) and “Emotional health of transgender youth 24 months after initiating gender-affirming hormone therapy” (Olson-Kennedy, Wang et al. 2025). The first was published as a 83 preprint (not peer-reviewed), and only after the HHS Review was published on May 1. For a critical analysis, see Society for Evidence-Based Gender Medicine (2025). The Review cites the second study on cross-sex hormones. As explained in Section 5.7.3 (“Robustness of this overview’s conclusion”), “rather than extending beyond what the evidence can support, this overview is confined to summarizing the conclusions of SRs [systematic reviews]. As a result, it may not include some of the most recently published studies due to the timing of the SRs’ literature searches. However, a targeted search [the footnote cites Olson-Kennedy, Wang et al. (2025)] of recently published studies did not reveal any published or ongoing studies that would significantly change the conclusions, especially those pertaining to benefits. This is due to ongoing problems such as an absence of comparison groups, inadequate sample sizes, and limited follow up.”

<sup>54</sup> “Santen recommends that the Review “highlight the differences in results between birth assigned males and females [in Chen et al. (2023)] as an adjunct to the discussion of the Olson-Kennedy manuscript.”

<sup>55</sup> “The Review does highlight those differences in a section devoted to Chen et al. (2023) (6.2.3): “The only statistically significant improvement in both sexes was in ‘appearance congruence’ as measured by the ‘transgender congruence scale,’ which has not been validated in minors. The authors also reported that there were statistically significant improvements in depression, anxiety and life satisfaction. However, these improvements were small and of questionable clinical significance. The statistically significant improvements were observed only in females, whereas males experienced no significant improvement in these measures”

<sup>56</sup> “Smids acknowledges that Chapter 2’s treatment of relevant terminology is “important and provides essential insights” but worries that its “skepticism regarding the term gender identity may easily be taken for a wholesale skepticism regarding the experience of gender incongruence and may come across as dismissive to the importance that gendered feelings have for trans persons”.

<sup>57</sup> “The Review recognizes that some children and adolescents experience discomfort or distress regarding their sexed bodies or associated social roles and expectations. It aims to describe and assess current best practices for the treatment of children and adolescents facing precisely this challenge. The Review notes that leading U.S. professional medical societies and clinicians working in gender clinics have adopted terminology that is unhelpful at best for describing their patients’ experiences or problems, but it does not deny or minimize these experiences or problems. On the contrary, the Foreword states that when patients seek professional help, “they and their families should receive compassionate, evidence-based care tailored to their specific needs.” Discerning what care is tailored to patients’ specific needs requires clear language and scientifically accurate terms. Toward that end, we emphasize again that “the understandable desire to avoid exclusionary or pathologizing language—combined with beliefs firmly embedded in the field—has led to a vocabulary and a mode of communicating that is scientifically ungrounded, that presupposes answers to ethical controversies, and that is in other ways misleading” (Section 2.1).”

<sup>58</sup> “While Smids acknowledges that the Review’s research ethics analysis appropriately relies on “established research ethical principles” requiring a reasonable anticipation of a positive balance of benefits over risks, he is not wholly convinced by the analysis. This is because he is not certain that in the research context “we are in a position to claim that predictions of possible overall benefit are unreasonable even for a small subset of GD adolescents currently undergoing PMT.”

<sup>59</sup> “As with studies of any clinical intervention, the fact that studies to date do not find strong evidence that PMT improves health outcomes does not, in principle, rule out the possibility that some subpopulation of subjects benefits from the interventions while others are harmed by them. However, to date no subpopulation has been shown to benefit. Moreover, clinicians are unable to predict which patients will experience persistent GD into adulthood and which will experience a resolution of symptoms. Nor do those clinicians who follow the American “gender-affirming” model try to make such predictions (see, e.g., Chapter 11). We agree with Smids on the need for further research, but for reasons set out in Section 13.5 we find that “administering PMT to adolescents, even in a research context, is in tension with well-established ethical norms for human subjects research.” Delineating specific areas of future research in pediatric gender medicine is an urgent challenge that is outside the scope of the Review. However, it is important that the etiology and

---

natural history of pediatric GD continue to be studied and that the short- and long-term health effects of hormonal interventions be characterized. We encourage researchers to explore alternative ways to study this vulnerable population, for example by analyzing existing data, recruiting research participants from the adult population who received a diagnosis of GD as children or adolescents (whether they medicalized or not), and conducting trials using less invasive and risky psychosocial interventions. We also emphasize that clinical research typically proceeds with a reasonably clear account of health and disease in the relevant population and with a good understanding of the clinical aims of the interventions. Here we note once again that in the field of pediatric gender medicine the rationales for medical intervention are much contested, a problem revealed in the fundamentally different nosological approaches adopted by the DSM-5 and the ICD-11, as discussed in Section 13.3”.

<sup>60</sup> “According to Smids, Chapter 11 of the Review is “far more accusative than fitting for the type of report the HHS analysis aims to be, accusing even clinicians who have just become the target of legal procedures.” While he credits Chapter 11 with “providing valuable insights,” he claims the “fundamental principle” that ought to have guided the chapter is the principle that “one is innocent until proven guilty”.

<sup>61</sup> “We agree that the Review should note the tentative nature of emerging research on low diagnostic stability. “Tentatively” was added to Section 4.3.2.1 (“New evidence about the natural history of gender dysphoria”): Although the natural history of GD—i.e., its course absent medical interventions—is currently impossible to measure given the wide availability of interventions, newer evidence tentatively suggests that GD has a low diagnostic stability”.

<sup>62</sup> “Strathearn suggests that in the Foreword, “it is important to acknowledge that there is also insufficient evidence to clearly understand the ‘risk of potential harm’ for some of these treatments. For example, the long-term outcomes (both risks and benefits) are uncertain for all treatment modalities ... Nevertheless, the responsibility for medical practitioners to ‘first do no harm’ means that the primary burden of evidence should be for the likelihood of benefit, especially when there is even a potential for harm.”

<sup>63</sup> “The following underlined text was added to the Foreword (paragraphs four and seven): Having recognized the experimental nature of these medical interventions and their potential for harm (which has been inadequately studied, especially with respect to long-term outcomes), health authorities in a number of countries have imposed restrictions. Nevertheless, the “gender-affirming” model of care includes irreversible endocrine and surgical interventions on minors with no physical pathology. These interventions carry risk of significant harms including infertility/sterility, sexual dysfunction, impaired bone density accrual, adverse cognitive impacts, cardiovascular disease and metabolic disorders, psychiatric disorders, surgical complications, and regret, and there has been inadequate research into the frequency and severity of these harms. Meanwhile, systematic reviews of the evidence have revealed deep uncertainty about the purported benefits of these interventions.)”.

<sup>64</sup> “Strathearn suggests that Part 1 of the Executive Summary should mention that some countries have restricted puberty blockers, cross-sex hormones and surgeries to research settings”

<sup>65</sup> “The following underlined text was added to Part 1 of the Executive Summary: ... health authorities in an increasing number of countries have restricted access to puberty blockers and cross-sex hormones, and, in the rare cases where they were offered, surgeries for minors. These authorities now recommend psychosocial approaches, rather than hormonal or surgical interventions, as the primary treatment, and in some cases have restricted the latter to nationally overseen research protocols”.

<sup>66</sup> “Strathearn found Chapter 3 to be “somewhat based on conjecture and hearsay” and noted it could be vulnerable to bias.”

<sup>67</sup> “Chapter 3... follows well- established scholarly conventions, supporting its claims with peer-reviewed and primary source literature, which readers can consult for verification”.

<sup>68</sup> “Strathearn notes that in Section 4.1, Figure 4.2 should include error bars to assess the variability of the mean scores. He also raises a question about the distribution of the scores”

<sup>69</sup> “Figure 4.2 was updated to include 95% confidence intervals, with “95% confidence intervals added” placed in a footnote. The data reported in the original Dutch research are insufficient to answer Strathearn’s reasonable question about distribution”

<sup>70</sup> “Strathearn suggests that the uncertain evidence for psychotherapy outcomes should also be mentioned in Section 5.7.5 (“Conclusion”)”

<sup>71</sup> “We have added a sentence at the end of this section: This overview synthesizes the best available clinical evidence from population level data, highlighting a consistent pattern across interventions for children and adolescents with GD.

---

The benefits and harms of social transition remain unknown; PBs, CSH, and surgeries consistently produce certain physical and physiological effects; and there is considerable uncertainty regarding their psychological and long-term health outcomes. Likewise, there is uncertainty regarding the effects of psychotherapy for GD.”

<sup>72</sup> “Strathearn requests clarification regarding the following statement in the introduction to Chapter 6: “It is well-established in adults that for the same drug, off-label uses are associated with considerably higher rates of adverse effects, especially when strong scientific evidence is lacking.”

<sup>73</sup> “A citation in the introduction to Chapter 5 (Egualo et al., 2016) reports: We found that off-label use of drugs was associated with ADEs after adjusting for important patient and drug characteristics. Moreover, we noted a risk gradient with higher rates of ADEs for off-label uses lacking strong scientific evidence. Although the intrinsic nature of the drug to cause ADEs is the same for on-label and off-label uses, it may be modified by a number of factors, including the off label disease condition. In addition, the lack of approval from a regulatory body implies a lack of safe dose ranges and inadequate information on contraindications, which in aggregate make ADEs more likely. We found that 4 in 5 off-label prescriptions lacked strong scientific evidence, and this group had higher rates of ADEs.”

<sup>74</sup> “Strathearn suggests that the suicide rate comparisons in Section 6.2.3 be improved and more thoroughly cited.”

<sup>75</sup> “The following underlined text was added to the discussion of suicide: However, two of the study subjects died by suicide within one year of initiating hormones, representing an annualized suicide rate of 317 per 100,000 patients. The suicide rate in Chen et al. was higher than rates that have been reported by PGM clinics in the U.K. and Finland (13 per 100,000 and 51 per 100,000).<sup>40</sup> One Belgian study<sup>41</sup> has also reported a comparatively high annual suicide rate (1,126 per 100,000); like in Chen et al., all patient suicides in that study were among patients taking CSH”.

<sup>76</sup> “Referring to Chapter 13’s discussion of psychotherapy, Strathearn correctly notes that, as in the case of medical interventions, “no evidence for harm does not equate with ‘no potential harm.”

<sup>77</sup> “The following underlined text was added to paragraph three of Section 13.2.3:

Regarding the potential harms of psychotherapy for adolescents with GD, a systematic review of the evidence found no evidence of negative or adverse effects in any of the studies examined (although absence of evidence for harm does not imply evidence of no harm, psychotherapy does not carry the medical or surgical risks associated with PMT).”

<sup>78</sup> “The lack of rigorous reporting of conflicts of interest (COI) by authors is the most important issue here, given the topic.”

<sup>79</sup> “Given the highly polarized nature of the topic, contributors’ names were withheld during the peer-review process so that reviewers could focus on the content of the review, rather than on the individual contributors themselves. This is an established practice in scientific review”.

<sup>80</sup> “A definition of an SR (to be included in the umbrella [review]) would have been useful, but we found no issues on inclusion or exclusion of SRs.”

<sup>81</sup> “We have added a footnote to Section 1 of Appendix 4: “An SR needs to have: 1) a defined research question according to PICO elements: Population, Intervention, Control/Comparator, Outcome; 2) pre-defined eligibility criteria for studies; 3) adequate systematic search methods that identify all studies that would meet the eligibility criteria; 4) an assessment of the validity of the findings of the included studies, for example through a risk-of-bias assessment; and 5) a systematic presentation and synthesis of the characteristics and findings of the included studies, which may include a meta-analysis. Scoping reviews, overviews of systematic reviews (umbrella reviews), and narrative reviews, are not SRs”.

<sup>82</sup> “The registration of the protocol would have increased transparency, as would more details about how the results were summarized. However, the final results are described transparently and are easy to follow. There are also many tables with necessary and relevant information.”

<sup>83</sup> “Unfortunately, given the time constraints, preregistration was not possible in this case. Preregistration also would have revealed contributor names. Given the polarization of this issue, it was important for the peer-review process that names not be disclosed until that process was completed.<sup>27</sup> In short, competing considerations had to be balanced, and trade-offs had to be made. We think the methodological information in Appendix 4 is sufficiently detailed”.

<sup>84</sup> “No information was available on support, author information and availability of data and other information”.

<sup>85</sup> “With respect to support, the Review was commissioned from the contributors by the HHS contractor for this project. There were no other sources of support. Regarding contributor information, see reply to point 1 (above). Regarding data availability, all relevant details regarding the methodology and results are in Chapter 5 and Appendix 4.”.

---

<sup>86</sup> “In August 2025, the Journal of Adolescent Health published a commentary titled “A critical scientific appraisal of the Health and Human Services Report on pediatric gender dysphoria” (Dowshen et al., 2025). The authors conclude that the Review engages in “numerous violations of scientific norms, misrepresentation of scientific evidence, and mischaracterizations of both gender identity in youth and the standard of care.” The Review, they suggest, “is a dangerous example of government incursion into the provision of evidence-based medical care.” The commentary’s allegations are serious; moreover, its authors are leaders in the field of pediatric gender medicine. We have therefore decided to treat the commentary as an unsolicited peer review. We are grateful for the opportunity to address the collated feedback and concerns of gender clinicians and researchers who believe pediatric medical transition (PMT) is beneficial to patients and in line with existing standards in pediatric medicine.”

<sup>87</sup> “A Dowshen et al. suggest the Review lacks independence and that “the findings ... were predetermined by the EO [executive order] that predated the writing of the report itself ...”

<sup>88</sup> “We note that no members of the contributor team are employed by the commissioning administration, that the empirical conclusions of the Review were arrived at via a transparent, reproducible methodology, and that the Review followed standard, scholarly norms of citation and argumentation. The Review’s central conclusions are based on an overview of systematic reviews (SRs) or “umbrella review” (see Chapter 5 and Appendix 4) and an ethics analysis (see Chapter 13). Each of these followed well-established principles in the relevant fields: evidence-based medicine and biomedical ethics, respectively. Each underwent independent peer-review by subject matter experts, and in both cases these experts concluded that the analyses are robust and consistent with high professional standards. ... if the Review’s findings reflect the authors’ bias, “it should be possible to point out where the report engages in motivated reasoning, fails to do justice to the extant literature, or shows other problems.” ... Dowshen et al. have done none of these things”.

<sup>89</sup> “Dowshen et al. are concerned that the Review “declines to name its authors, making assessment of their financial, intellectual, or other conflicts of interest impossible”

<sup>90</sup> “The decision to withhold names until completion of peer-review was intended to help maintain the integrity of the review process. This is standard practice in scientific publishing”.

<sup>91</sup> “Dowshen et al. claim that the Review lacks credibility because “over 20%” of its references are not from the “peer-reviewed scientific literature”

<sup>92</sup> “Focusing on the ratio of peer-reviewed to non-peer-reviewed sources is misguided. The important question is whether all relevant evidence is appropriately represented. The Review was tasked with evaluating both evidence and best practices. In addition to engaging with the relevant peer-reviewed scientific literature, the Review discusses non-peer-reviewed publications where appropriate. For example, a non-peer-reviewed essay, produced by the “Integrity Project” and posted on the Yale University Law School website in 2024, is relevant to the evidence for PMT and is cited in the Review.<sup>31</sup> The Integrity Project was co-founded by Dr. Meredith McNamara, lead author of that essay and senior author of Dowshen et al. The Review’s appraisal of best practices examines clinical realities in the U.S., which are frequently documented in court filings and media outlets such as The New York Times, The Washington Post, The Boston Globe, Reuters, and The Economist. Chapter 10 cites internal World Professional Association for Transgender Health (WPATH) documents which were obtained via the discovery process in a lawsuit. The email exchanges between senior WPATH members reveal that WPATH’s guideline development process flouted well-recognized standards.

Dowshen et al. do not dispute the veracity of claims made in Chapter 10 of the Review, which describes WPATH’s suppression of SRs and its revision of clinical recommendations in response to political pressures”.

<sup>93</sup> “Dowshen et al. claim the Review misused the very low GRADE designation of evidence quality for PMT as justification for “rejecting the standard of care for TGD [transgender and gender diverse] youth.”

<sup>94</sup> “Our conclusion regarding the risk/benefit profile of PMT does not rely exclusively on the fact (now demonstrated by many SRs and confirmed by the Review’s umbrella review) that the evidence for benefit is of very low certainty. Rather, as explained in Chapter 8 and Chapter 13, this conclusion is supported by a standard risk/benefit analysis that incorporates both the purported benefits and the known risks and harms of the relevant interventions, as compared to the risk/benefit profile of the alternatives.”

<sup>95</sup> “Dowshen et al. claim the Review “misrepresents ... studies, often ignoring their primary conclusions.” They give Chen et al. (2023) as an example. Dowshen et al. describe the findings of this study very positively: “appearance congruence, positive effect [sic], life satisfaction, and depression and anxiety symptoms all improved significantly following 2 years of hormone therapy.” Dowshen et al. allege that the Review “ignores” these findings and “focuses solely on the two deaths by suicide among the study’s 315 participants”.

---

<sup>96</sup> “The findings of Chen et al. (and other commonly cited studies) are extensively discussed in Chapter 6 of the Review.

Although self-report scores for appearance congruence, depression, anxiety, and life satisfaction improved at 24 months compared to baseline, these findings should not be described without serious qualifications. For example, males showed no improvement in any outcome except for “appearance congruence,” which was measured, as pointed out in the Review, on a scale that has never been validated in minors. Originally, the Review reported that appearance congruence improvement was “the only statistically significant finding” but omitted the qualifying phrase “in males.” The Review acknowledged the statistically significant improvements in females. A *statistically* significant improvement, however, should not be confused with improvement that is *clinically* significant or meaningful. In Chen et al., the mean Beck Depression Inventory score improved over 24 months from 16.01 to 13.85 (63-point scale); the mean Revised Children’s Manifest Anxiety Scale improved from 59.84 to 57.32 (T-score, where 50 is the population average and 10 is one standard deviation); and the mean life satisfaction score improved from 40.03 to 44.68 (T-score) on a subscale of the NIH Toolbox Emotion Battery. These are small improvements of questionable significance to clinicians and patients. For the Beck Depression Inventory, for example, researchers have suggested a 17.5% decrease from baseline score may represent a “minimal clinically important difference.” The mean decrease of 2.16 points on this (63-point) scale in Chen et al. (2023) does not, according to this criterion, meet the minimal threshold of clinical importance. Given the known and plausible harms of these interventions, even if such minor benefits were established via well - conducted studies (e.g., randomized controlled trials), the risk/benefit profile of hormonal interventions would remain unfavorable. There are many other problems with Chen et al., including the fact that follow-up data for mental health outcome measures were unavailable for 31–34% of participants (introducing selection bias), the shifting hypotheses between the preregistered protocol and the publication, and the failure to report many preregistered outcomes such as suicidality, self-harm, and gender dysphoria. Most crucially, the uncontrolled observational design precludes any conclusion about whether cross-sex hormone (CSH) treatment *caused* any improvement. Nonetheless, that did not prevent Chen et al. from erroneously describing the results on the first page in explicitly causal language: CSH “improved appearance congruence and psychosocial functioning”.

<sup>97</sup> “Dowshen et al. claim that, because Chen et al. (2023) purportedly had very positive findings and because “other cohort studies [report] improvements in psychosocial functioning after treatment ...” the Review selectively misuses research”.

<sup>98</sup> “Here, Dowshen et al. cite five additional papers in support of their claim about improved psychological functioning. Pre-post studies in this field have shown inconsistent results with respect to psychological improvements (see Chapter 4). Their chosen references include Achille et al. (2020), in which (after regression analysis) depression improved only in males (the opposite finding from Chen et al.); Chelliah et al., (2024), in which the reported improvement was a small decrease in the mean Quick Inventory of Depressive Symptomatology (QIDS) score from 10.7 to 8.2 (on a 42-point scale); and a study in which there was a small improvement in self-reported depression but not in the clinician-reported depression outcome measure (Kuper et al., 2020).

The last of the five citations (Olson-Kennedy, Wang et al., 2025) is a paper reporting data derived from the same patient cohort as Chen et al. (2023). It is misleading to imply that this is one of the “other” cohort studies, when in fact it is the same cohort as that reported in Chen et al. (It is unclear why Trans Youth Care researchers are spreading the outcome data from this research project over multiple publications. Future systematic reviewers should be vigilant for “salami slicing.”

The problems in these studies illustrate why it is inappropriate to rely on low-quality observational studies when rigorous SRs are available”.

<sup>99</sup> “Dowshen et al. criticize the Review for omitting a reference to Nunes-Moreno et al. (2025).”

<sup>100</sup> “Although Nunes-Moreno et al. leverages a large multicenter dataset, it has similar limitations as previously reported observational studies. Consideration of this study would not have changed the conclusion of prior SRs on PBs and CSH, nor the conclusion of Appendix 4’s umbrella review. In evidence-based medicine, strength of evidence is determined by quality, not quantity, of studies”.

<sup>101</sup> “Dowshen et al. claim the Review “provides no evidence for its assertion that puberty-pausing medications and hormone therapy are harmful to TGD youth,” and that the Review “even states that evidence of harms is ‘sparse’”

<sup>102</sup> “The full quotation is: Evidence for harms associated with pediatric medical transition in systematic reviews is also sparse, but this finding should be interpreted with caution.”. Chapter 6 of the Review explains why consideration of harms associated with PMT needs to go beyond evidence from SRs. As Guyatt and colleagues remark, “Many, if not most, systematic reviews fail to address some key outcomes, particularly harms, associated with an intervention.” For example,

---

the fact that studies have not reported infertility data (and therefore SRs have been unable to capture it) does not mean infertility can be ignored in a comprehensive evidence appraisal. Far from “[providing] no evidence” of harms, Chapter 7 of the Review presents detailed indirect evidence derived from basic science, endocrinology, and developmental physiology, demonstrating plausible and biologically expected harms”.

<sup>103</sup> “Dowshen et al. assert that the Review’s comments on the lack of long-term outcome data are “misleading.” They mention two Dutch studies providing “over 20 years of follow-up data,” and an American study providing “up to 10 years” of follow-up data”

<sup>104</sup> “The first study, van der Loos, Klink et al. (2023), evaluated treatment trajectories. It is misleading to describe this study as a supplying “20 years of follow-up data.” Here, 20 years refers to the intake period (1997 to 2018). With respect to follow-up after hormone initiation, the study’s median follow-up was 4.6 years.

The second Dutch study, van der Loos, Vlot et al. (2023), evaluated bone density of 25 males and 50 females treated with PBs followed by CSH. As discussed in Chapter 7 of the Review, this study found that Z-scores returned to pre treatment baseline at follow-up.<sup>42</sup> A major limitation of this study is the 40% non- participation rate.

The American study, Olson et al. (2024), assessed satisfaction and regret after initiation of hormones. It is misleading to describe this as a 10-year follow-up, as the mean follow-up was 4.86 years after starting PBs and 3.4 years after starting CSH (median follow-up was not reported and may have been considerably lower). Notably, the study may not be representative, as participants in this cohort were completely socially transitioned before puberty (some were as young as age two at the time of social transition; average age at social transition was 6.4943). Further, the absence of physiological or psychiatric outcome data in this study is a critical limitation.

As the Review explains (Section 13.4), satisfaction and regret, though important data points, are not valid proxies for evaluating the justification for PMT. A Letter to the Editor responding to Olson et al. (2024) points out that “Patient satisfaction is generally considered a complementary measure of health care quality and is typically assessed after the safety and effectiveness of the intervention are established”.

<sup>105</sup> “Dowshen et al. suggest that PBs are safe for “TGD youth” because “they have been safely and effectively used for decades to treat cisgender youth with medical conditions such as precocious puberty”

<sup>106</sup> “Chapter 7 of the Review contrasts the use of PBs in treating central precocious puberty (CPP) with their use for GD. There are three key differences. First, purpose: CPP is a physical pathology and PBs are used to stop abnormally timed puberty. GD is not a physical pathology and PBs in this case are used to stop normally timed puberty.

Second, diagnosis: CPP is diagnosed using objective tests such as blood work, and the natural history of the condition is well-understood, whereas the diagnosis of GD relies on subjective criteria and has poor predictive validity.

Third, prognosis: In CPP cases, PBs are stopped and puberty resumes. For pediatric GD cases, over 90% of youth treated with PBs proceed to CSH, and for these patients puberty (properly defined<sup>45</sup>) does not resume.

The senior author of Dowshen et al., McNamara, recently acknowledged that PBs should not be assessed as a standalone intervention, but rather as a component of a single treatment modality comprised of both PB and CSH.<sup>46</sup> By her own admission, then, it makes little sense to assert that PBs are safe for use in GD on the grounds that they are safe for use in CPP. The combined-use pathway (PBs followed by CSH) presents a fundamentally different risk profile. For example, infertility is not an expected treatment outcome when PBs are used for CPP, but it is an expected outcome when PBs are used prior to or alongside CSH for pediatric GD.

We also note that Dowshen et al.’s description of CPP patients as “cisgender youth” is inaccurate. Any child diagnosed with CPP will be a candidate for PBs, irrespective of how he or she identifies. CPP and GD are two distinct clinical scenarios and the implication that concerns about PB use in one scenario but not the other are due to identity-based discrimination is seriously misleading.”.

<sup>107</sup> “Dowshen et al. cite “a recent comprehensive review commissioned by the Utah state legislature and completed by experts at the University of Utah,” which “concluded “that puberty-pausing medications and hormone therapy can also be used safely in TGD youth.”

<sup>108</sup> “The “Utah Review”<sup>47</sup> cited by Dowshen et al. was published after the HHS Review and therefore was not included in its analysis. We discuss the Utah Review in our reply to the APA”.

<sup>109</sup> “Dowshen et al. say that a “central premise” of the Review “is the unsupported claim that gender identity among adolescents is inherently unstable”

---

<sup>110</sup> “First, the Review is concerned with gender dysphoria, not “gender identity,” and it did not adopt any claim about the stability of the latter as a “central premise.” Second, as discussed in the Review, the assumed permanence of adolescent (in contrast to childhood) GD has served as the basis for the Dutch Protocol, but this assumption was never based on credible evidence. In the Review, we discussed other, more recent evidence suggesting that for a significant number of children and adolescents, gender dysphoria appears to be a transient phenomenon. The burden of proof for PBs, CSH, and surgeries as treatments for adolescent GD—a mental health condition—rests on those advocating for these interventions”

<sup>111</sup> “Dowshen et al. object to the Review’s broadly positive treatment of (“exploratory”) psychotherapy, claiming that it is “an ill-defined practice that aims to change a young person’s identity, which is akin to conversion therapy.” “Decades of evidence,” they say, “demonstrate that conversion practices are both ineffective and dangerous for the psychological health of transgender [youth].” Dowshen et al. allege that the Review includes a “recommendation” for “nonevidence-based conversion practices”.

<sup>112</sup> “Dowshen et al. fail to engage with Section 14.5.2.1, which anticipates and refutes this charge... Prominent PMT advocate Dr. Jack Turban, director of the Gender Psychiatry Program at UCSF, has said that “conversion efforts and exploratory psychotherapy are distinct, mutually exclusive practices.”<sup>50</sup> And even WPATH—the leading organization supporting PMT—said, in a statement condemning the HHS Review, that “[they] unequivocally oppose” “[equating] conversion therapy with psychotherapy” for “youth who are exploring their gender identity”.

<sup>113</sup> “Dowshen et al. criticize the Review for its discussion of deteriorating standards and the collapse of medical safeguarding in the U.S. They object to the inclusion of “unverified” whistleblower accounts, characterizing the whistleblowers as “individuals not directly involved in clinical decision-making for patients.””

<sup>114</sup> “Contrary to Dowshen et al.’s characterization, most of the whistleblowers (Chapter 11) are practicing clinicians who have treated this patient population. Their accounts are highly relevant for understanding the clinical realities of pediatric gender medicine in the United States”.

<sup>115</sup> “15. Dowshen et al. further suggest that the whistleblower accounts should be discounted because WPATH and ES guidelines require “an interdisciplinary team that performs a comprehensive biopsychosocial assessment” prior to initiation of PBs/CSH or referral for surgery”.

<sup>116</sup> “The problem is that accounts of whistleblowers describe clinicians offering inappropriate treatments within the context of “multidisciplinary” (or “interdisciplinary”) teams and “assessments” (see Chapter 11 and Section 14.3”.

<sup>117</sup> “Dowshen et al. assert that guidelines which recommend PMT as the standard of care (e.g., WPATH and Endocrine Society guidelines) are “informed by the best available evidence, which demonstrates improved outcomes in mental health, psychological well-being, and suicidality.” Furthermore, Dowshen et al. claim the treatment approach recommended in WPATH’s guideline is “evidence-based” and “based on more than 70 systematic reviews.”

<sup>118</sup> “The best available evidence (i.e., from SRs) does not support Dowshen et al.’s assertion that psychological outcomes improve with PMT. Appendix 4’s umbrella review reveals that the effects on psychological outcomes are unknown. Chapters 9–11 of the Review identify serious problems with WPATH and Endocrine Society (ES) guidelines. With respect to the care of children and adolescents (Chapters 6 & 7), WPATH’s guideline is consensus-based, not evidence-based, as it is not based on evidence from SRs. Indeed, Standards of Care, Version 8 (SOC-8) states—falsely— that an SR of hormonal interventions in minors is “not possible.” Section 10.3.2 details how WPATH suppressed the publication of some systematic reviews it had commissioned to inform SOC-8, including reviews covering treatment of minors. This raises serious concerns about the scientific integrity of WPATH as a self-described healthcare organization”.

<sup>119</sup> “Dowshen et al. describe the WPATH/ES clinical practice guidelines (and other guidance that reference these) as “the existing standard of care” and imply that WPATH’s guidelines should be considered trustworthy in part because they are “widely endorsed,” “maintained since 1979,” “currently in their eighth edition,” “took almost a decade to develop,” and represent “the consensus recommendations of more than 100 experts in transgender health.”.

<sup>120</sup> “We would like to clarify that there is no accepted “existing standard of care” for treating pediatric patients with GD, and that guidelines/policies from around the world recommend very different treatment approaches.

With respect to Dowshen et al.’s list of attributes, none of them is recognized as relevant to a guideline’s trustworthiness. AGREE II, a widely used tool to assess trustworthiness of clinical practice guidelines, specifies six relevant domains: Scope and Purpose, Stakeholder Involvement, Rigor of Development, Clarity of Presentation, Applicability, and Editorial

---

Independence. The methodological rigor of a guideline’s development—specifically, whether it relied on SRs rather than expert consensus—is regarded as the most important of these domains.<sup>55</sup> Medicine should be evidence- based, not eminence-based”.

<sup>121</sup> “As with Dowshen et al. (2025), we have decided to treat Rider et al. (2025) as unsolicited peer review. Part A of this reply summarizes our responses to criticisms in Rider et al. that appear in other peer reviews and publications, namely Dowshen et al. and the peer review by the American Psychiatric Association (APA). Part B responds to the few points that do not appear in the other peer reviews. Part C discusses how Rider et al.’s commentary exemplifies problems common in this field: misrepresentation of research, inadequate citation practices, and poor understanding of evidence-based medicine (EBM) principles. The continued willingness of peer-reviewed journals to publish demonstrably false or misleading claims about evidence-based medicine and medical practices concerning child and adolescent health is deeply regrettable”.

<sup>122</sup> “Critiques previously addressed in replies to Dowshen et al. and/or the APA”.

<sup>123</sup> “Rider et al. object that “the authors [of the Review] were unnamed”.

<sup>124</sup> “As HHS initially explained, the identities of the contributors were temporarily withheld to “help maintain the integrity of this [peer review] process.” Withholding names in peer-review is standard practice in academic publishing. Please see point 2 in our reply to Dowshen et al., as well as Section 4 of our reply to the APA”.

<sup>125</sup> “Citing Dowshen et al. (2025), Rider et al. object that “more than a fifth” of the Review’s references “are from popular media articles, blogs, or social media”.

<sup>126</sup> “To repeat part of our reply to Dowshen et al. (point 4), focusing on the ratio of peerreviewed to non-peer-reviewed sources is misguided. The important question is whether all relevant evidence is appropriately represented. The Review was tasked with evaluating both evidence and best practices. In addition to engaging with the relevant peer-reviewed scientific literature, the Review discusses non-peer-reviewed publications where appropriate. The Review’s appraisal of best practices examines clinical realities in the U.S., which are frequently documented in court filings and media outlets such as The New York Times, The Washington Post, The Boston Globe, Reuters, and The Economist”.

<sup>127</sup> “Rider et al. criticize the Review’s provenance (i.e. the January 2025 Executive Order directing the Secretary of Health and Human Services to commission the Review), describing the Review as an “[entity] with a political agenda targeting [patients, families, and providers]” and claiming that the Review “presents the White House’s political agenda as objective science”.

<sup>128</sup> “The Executive Order directed HHS to “publish a review of the existing literature on best practices for promoting the health of children who assert gender dysphoria.” If the findings of the Review were dictated by preexisting political agendas, it should be possible to identify errors within it. As explained in our response to Dowshen et al. and the APA, and as further demonstrated below, no such examples have been offered. Please see point 1 in our reply to Dowshen et al., as well as Section 4 of our reply to the APA”.

<sup>129</sup> “Rider et al. characterize pediatric medical transition (PMT) (three times) as “medically necessary,” claiming that “scientific evidence demonstrat[es] its safety and effectiveness in improving short- and long-term health outcomes for TGNB [transgender and nonbinary] adolescents”.

<sup>130</sup> “Rider et al. repeatedly exhibit a misunderstanding of basic EBM principles regarding quality (or certainty) of evidence. It is simply incorrect that scientific evidence “demonstrates” PMT’s “safety and effectiveness.” The Review’s umbrella review (Appendix 4) shows this quite clearly, and Rider et al. say nothing that casts doubt on its findings. Please see point 6 in our reply to Dowshen et al., regarding the inversion of the evidence hierarchy, and Section 1 of our reply to the APA, regarding the list of individual studies provided in their peer review. We discuss other examples of Rider et al.’s misunderstanding of EBM in points 8 and 9 below”.

<sup>131</sup> “Rider et al. criticize the Review’s engagement with the Cass Review. They claim the Review did not address the Cass Review’s alleged “omission of key findings from the broader literature,” or the fact that it has been “negatively critiqued and challenged repeatedly by professional organizations and individual experts in the field of pediatric gender care.” They also claim the HHS Review selectively quotes from the Cass Review’s conclusions”.

<sup>132</sup> “The Cass Review’s findings have been accepted by both major political parties in the U.K. and its recommendations are being implemented by the U.K.’s National Health Service. It is not surprising that gender clinicians and the professional associations that represent them would disparage a review that upended their favored treatment model in the U.K. We direct readers to comprehensive rebuttals to critiques of the Cass Review: see footnote 77 in Part C below and Section 3 of our reply to the APA, which also addresses Rider et al.’s claim that the Review selectively quotes from

---

the Cass Review. In short, the critiques are rife with demonstrable falsehoods and some appear motivated by legal goals rather than scientific ones. It is also worth noting that one of the references Rider et al. provide (Horton, 2024) in support of their claim that “thorough scientific and legal scholarship, as well as the critiques from field experts ... directly rebut the evaluation of evidence...” in the Cass Review was published a month before the final Cass Review was published”.

<sup>133</sup> “Rider et al. claim the Review “promotes a harmful practice known as ‘exploratory therapy’ ... which has been argued to be a form of conversion therapy encouraging a child or adolescent to accept the gender associated with their sex designated to them at birth”.

<sup>134</sup> “Here, Rider et al. rely on the work of lawyer Florence Ashley, who opposes requirements for mental health assessments prior to PMT initiation and advocates for the wide availability of PMT because it facilitates a minor’s “gender embodiment goals.” Ashley has repeatedly conflated psychotherapy for pediatric GD with conversion therapy. In addition, Rider et al. attempt to bolster their case with misleading citations. One reference, the United States Joint Statement, explicitly states that “Exploration of issues pertaining to gender identity and sexual orientation in a way that does not favor or presume a particular identity or experience, would not be considered conversion therapy.” Another reference is a United Nations report which concludes that conversion therapy “may constitute torture.” However, that report’s examples of conversion therapy include gay individuals being “blindfolded and pummeled with basketballs, bound with duct tape, rolled up into blankets and subjected to anti-gay slurs.” These and similar practices are indeed abhorrent, but they have no bearing whatsoever on talk therapy for minors with gender dysphoria (GD)”.

<sup>135</sup> “Rider et al. compare hormonal interventions for pediatric GD to the use of hormonal interventions for other pediatric medical conditions—e.g., central precocious puberty (CPP)—in “prepubescent and pubescent cisgender youth,” suggesting that “TGNB adolescents” are being unfairly singled out”.

<sup>136</sup> “This framing is extremely misleading. To summarize point 10 in our reply to Dowshen et al., Chapter 7 of the Review contrasts the use of puberty blockers (PBs) in treating CPP with their use for GD. When used for CPP, PBs arrest abnormally timed puberty (as opposed to normally timed puberty), and are not followed by administration of CSH, which may result in lifelong infertility and sexual dysfunction as well as other risks to health. Also, like Dowshen et al., Rider et al. inaccurately describe CPP patients as “cisgender youth.” Any child diagnosed with CPP will be a candidate for PBs, irrespective of how they identify. CPP and GD are two distinct clinical scenarios, and it is entirely wrong to suggest that concerns about PB use in one scenario but not the other are due to identity-based discrimination”.

<sup>137</sup> “Rider et al. criticize the Review for “omitting” context related to the ambient “political climate and proposed or existing legislative bans on GAMC [gender-affirming medical care] for TGNB adolescents and their caregivers.” Further, they claim this “distorts the application of evidence-based medicine”.

<sup>138</sup> “This criticism seems intended to deflect from the content of the Review by putting the focus on the ambient “political climate.” EBM is concerned with clinical decision-making based on the best available evidence for the safety and efficacy of treatments, not with the broader political climate or legislation. We agree that the political climate has made scientific debate very difficult, but we emphasize again that if errors appear in the Review, it should be possible to clearly identify them. Rider et al. identify no errors”.

<sup>139</sup> “Rider et al. claim the Review “overlook[s] bias in the systematic reviews [SRs] it cites and deemphasiz[es] other layers in the ‘hierarchy of evidence.’” According to Rider et al., “multifaceted data and studies across multiple levels in the ‘hierarchy of evidence’ comprise the robust body of evidence supporting [PMT]”.

<sup>140</sup> “With respect to alleged quality problems in SRs of PMT that have found very low certainty evidence, Rider et al. reference an analysis critiquing the Cass Review, Noone et al. (2025), that claimed two specific SRs were biased/flawed.<sup>63</sup> Using the Risk of Bias Assessment Tool for Systematic Reviews (ROBIS), the Review came to a different conclusion than Noone et al., finding that these two SRs on puberty blockers and cross-sex hormones were generally at low risk of bias. Further, even if those two SRs were excluded, the conclusion that the quality of the evidence for benefit of PMT is very low certainty would be unaffected. We refer Rider et al. to the peer review included in this Supplement by methodologists Dr. Trudy Bekkering and Professor Patrik Vankrunkelsven, which concluded that the Review’s umbrella review was conducted appropriately. If Rider et al. disagree with the Review’s analysis, it would have been helpful for them to explain why. Merely citing Noone et al. does not advance scientific understanding. Rider et al.’s claim that the Review “deemphasiz[es] other layers in the ‘hierarchy of evidence’” seems to be a suggestion that it should have ignored or minimized the findings of quality systematic reviews in favor of emphasizing conclusions reached by some individual studies (“other layers”). Doing so, however, would constitute an inversion of the hierarchy of evidence and a violation of a core principle of EBM. Low quality evidence of the kind favored by Rider et al. cannot be characterized as “robust”.

---

<sup>141</sup> “Rider et al. say that the Review “fails to acknowledge ... that most pediatric healthcare is guided by evidence of similar quality and strength as that supporting [PMT]”.

<sup>142</sup> “It is not true that “most pediatric healthcare is guided by evidence of similar quality and strength.” Dr. Hilary Cass, author of the Cass Review and a past president of the Royal College of Paediatrics and Child Health, observed that the evidence for the efficacy of PMT is very weak, even compared to other areas of pediatric medicine. But we need not appeal to the authority of Cass: her judgment is supported by Rider et al.’s own citation (Matheny Antommara et al., 2025). Matheny Antommara et al. analyzed 14 current pediatric clinical practice guidelines, finding that 58% were based on Level A or Level B evidence: “Level A evidence includes well-designed and -conducted randomized controlled trials; Level B randomized controlled trials with minor limitations or consistent evidence from multiple observational studies.” PMT is not supported by evidence at these levels: no randomized trials have been conducted, and the extant observational studies are generally low-quality. Therefore, Rider et al.’s claim that “most” pediatric healthcare is supported by evidence of a similarly low quality as that supporting PMT is false. There is a more important point. Quality of evidence, as assessed via a rigorous systematic review, can inform stakeholders regarding what is known about an intervention’s effectiveness. But it is not the only consideration in clinical decisionmaking for PMT or for any other intervention. Harms must also be considered, as well as the natural history of the condition and the risk/benefit profiles of alternative treatment options. Please also see Section 2 of our reply to the APA. Furthermore, Rider et al.’s discussion of the evidence verges on inconsistency. In one passage they refer to “substantial evidence of benefits” of PMT, implying there are studies furnishing high-quality evidence. But two paragraphs earlier they apparently concede that “most pediatric healthcare” and PMT both “[fall] short in the ‘hierarchy of evidence.’” Rider et al. cannot have it both ways”.

<sup>143</sup> “The Review discusses “rapid onset gender dysphoria” (ROGD); according to Rider et al., this is a “largely discredited diagnosis”.

<sup>144</sup> “The Review addresses the recent surge in adolescent females with GD and various attempts to explain the novel development of gender dysphoria in this clinical population. “ROGD” is simply a label for a new clinical phenomenon; contrary to Rider et al.’s assertion, ROGD was never presented as a “diagnosis.” For a discussion, we refer Rider et al. to Section 4.3.1.4 of the Review. Rider et al. give a citation to support the claim of “discreditation”; this is discussed in Part C below”.

<sup>145</sup> “Rider et al. claim that “the HHS review likens the field of GAC [“gender-affirming care”] to the Tuskegee syphilis study” and furthermore claim that this is not a legitimate comparison because parents provide consent for interventions that cause infertility, etc. in their assenting children, whereas the Tuskegee participants did not provide informed consent”.

<sup>146</sup> “This is a misreading. Rider et al. cite Section 13.2.4 of the Review, which merely observes that the Belmont Report was “published in 1978 in the wake of the U.S. Public Health Service’s Untreated Syphilis Study at Tuskegee.” No comparison between Tuskegee and “the field of GAC” is made or implied. The Review does reference Tuskegee one other time but in a different chapter. Advocates of PMT sometimes object to scrutiny of the practice on the grounds that the number of minors undergoing these interventions is relatively small. Tuskegee is cited in Section 11.2 as an example of a medical experiment widely recognized as profoundly unethical despite involving a relatively small number of people (and much smaller than the number of youth receiving PMT)”.

<sup>147</sup> “First, the commentary employs inflammatory rhetoric. Rider et al. allege that the Review has “little regard” for the “civil rights” of vulnerable youth and conflates psychotherapy for pediatric GD with conversion therapy, which it describes as akin to “torture.” It is inappropriate for a peer-reviewed journal to publish such extremely serious allegations, which impugn not only the Review but the moral character of its contributors, without evidence to support them... Second, as discussed above, Rider et al. build their argument against the Review upon a variety of informal fallacies and suspect reasoning. Examples include genetic fallacies (e.g., judging the Review’s content and conclusions based on the Executive Order that commissioned it), red herrings (e.g., purporting to critique the Review while instead focusing on political or legislative issues), and appeals to authority (e.g., claiming that PMT is beneficial because some medical organizations say so). Additionally, Rider et al. levy ad hominem attacks, seemingly implying that the Review’s contributors—presumably unknown to Rider et al. at time of writing—are “unqualified individuals with no expertise in the field of pediatric gender care... Rider et al.’s decision to substitute scholarly engagement with appeals to the authority of U.S.-based medical organizations is especially unfortunate given that the Review devotes entire sections to presenting evidence of how these organizations have misled their members, patients, and the public. Rider et al. never dispute any of this evidence. Another tactic used by Rider et al. is uncritical citation of a denunciatory statement as evidence that something—in this case, ROGD—is “discredited.” This is part of a widespread pattern of citation problems in the pediatric gender medicine literature. A statement that itself contains no bibliography or hyperlinks to relevant literature is cited

---

as factual; this is an example of “dead-end referencing.”... Rider et al. take for granted, for example, that children “as young as five years old” may be “transgender or nonbinary,” as if these categories could unproblematically be applied to an age group that has a rudimentary understanding of sex differences in terms of stereotypes. Presumably Rider et al. are following the lead of one of their co-authors, Dr. Diane Ehrensaft, who has taught that toddler actions, such as removing hair barrettes or unbuttoning onesies, can be “pre-verbal communication[s] about gender”... The “Woozle effect” is the frequent citation of an inadequate source to support a particular claim; this can create the illusion that the source is authoritative.<sup>80</sup> The Woozle effect plagues gender medicine research and advocacy, and makes an appearance in Rider et al. For instance, Rider et al. cite Chen et al. (2023) as part of the “scientific evidence demonstrating ... improvements in well-being and quality of life.” Here Rider et al. reproduce the words of Budge et al. (2024), which also cites Chen et al. as part of “existing research [which] demonstrates the effectiveness of [PMT],” through “improvement in well-being and quality of life.” Likewise, Dowshen et al. (2025) say that Chen et al. found improvements in “appearance congruence, positive effect [sic], life satisfaction, and depression and anxiety symptoms.”

None of these three papers mentions the elevated suicide rate in Chen et al.; the missing outcome measures; the improvements of questionable clinical importance in female patients; the fact that male patients did not improve in mental health, only in “appearance congruence” (on a scale that has not been validated in minors); or the alteration of the study’s central hypotheses between the written protocol and the published paper. Please see also point 5 in our reply to Dowshen et al., and Sections 4.3.4 and 6.2.3 of the Review. Rider et al. also twice cite Tordoff et al. (2022), another influential study exemplifying the Woozle effect. Tordoff et al. purport to show reductions in psychiatric morbidity following the provision of hormonal PMT interventions. According to Google Scholar, this study has been cited over 700 times—more than 200 times per year on average since it was published—despite the fact that an online supplementary table of the paper reveals no statistically significant improvement in patients receiving the interventions. The Review discusses Tordoff et al. in detail in Section 6.2.2. As researchers who have carefully observed pediatric gender medicine for years, we fully expect the Woozle effect to apply to Rider et al. (2025) and Dowshen et al. (2025), which will almost certainly be uncritically cited in future peer-reviewed articles as proof that the Review has been “debunked” despite the serious problems in these papers. We strongly urge peer reviewers and journal editors to attend more carefully to the lax scholarly norms in this field and to work to strengthen them”.