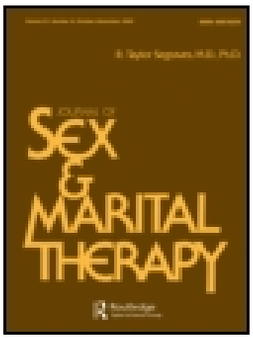

Routledge
Taylor & Francis Group

Journal of Sex & Marital Therapy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/usmt20

The Myth of "Reliable Research" in Pediatric Gender Medicine: A critical evaluation of the Dutch Studies—and research that has followed

E. Abbruzzese, Stephen B. Levine & Julia W. Mason

To cite this article: E. Abbruzzese, Stephen B. Levine & Julia W. Mason (2023): The Myth of "Reliable Research" in Pediatric Gender Medicine: A critical evaluation of the Dutch Studies—and research that has followed, Journal of Sex & Marital Therapy, DOI: 10.1080/0092623X.2022.2150346

To link to this article: https://doi.org/10.1080/0092623X.2022.2150346

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

Published online: 02 Jan 2023.

Submit your article to this journal

Article views: 24245

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=usmt20

JOURNAL OF SEX & MARITAL THERAPY
https://doi.org/10.1080/0092623X.2022.2150346

Routledge
Taylor & Francis Group

ARTICLE COMMENTARY

OPEN ACCESS

Check for updates

The Myth of "Reliable Research" in Pediatric Gender Medicine: A critical evaluation of the Dutch Studies—and research that has followed

E. Abbruzzese[a], Stephen B. Levine[b] and Julia W. Mason[c]

[a]Society for Evidence-Based Gender Medicine, Twin Falls, ID, United States; [b]Department of Psychiatry, Case Western Reserve University, Cleveland, OH, United States; [c]Calcagno Pediatrics, Gresham, OH, United States

ABSTRACT

Two Dutch studies formed the foundation and the best available evidence for the practice of youth medical gender transition. We demonstrate that this work is methodologically flawed and should have never been used in medical settings as justification to scale this "innovative clinical practice." Three methodological biases undermine the research: (1) subject selection assured that only the most successful cases were included in the results; (2) the finding that "resolution of gender dysphoria" was due to the reversal of the questionnaire employed; (3) concomitant psychotherapy made it impossible to separate the effects of this intervention from those of hormones and surgery. We discuss the significant risk of harm that the Dutch research exposed, as well as the lack of applicability of the Dutch protocol to the currently escalating incidence of adolescent-onset, non-binary, psychiatrically challenged youth, who are preponderantly natal females. "Spin" problems—the tendency to present weak or negative results as certain and positive—continue to plague reports that originate from clinics that are actively administering hormonal and surgical interventions to youth. It is time for gender medicine to pay attention to the published objective systematic reviews and to the outcome uncertainties and definable potential harms to these vulnerable youth.

Introduction

In our recent paper on informed consent for youth gender transition, we recognized a serious problem: the field has a penchant for exaggerating what is known about the benefits of the practice, while downplaying the serious health risks and uncertainties (Levine et al., 2022a). As a result, a false narrative has taken root. It is that "gender-affirming" medical and surgical interventions for youth are as benign as aspirin, as well-studied as penicillin and statins, and as essential to survival as insulin for childhood diabetes—and that the vigorous scientific debate currently underway is merely "science denialism" motivated by ignorance, religious zeal, and transphobia (Drescher et al., 2022; McNamara et al., 2022; Turban, 2022). This highly politicized and fallacious narrative, crafted and promoted by clinician-advocates, has failed to withstand scientific scrutiny internationally, with public health authorities in Sweden, Finland, and most recently England doing a U-turn on pediatric gender transitions in the last 24 months (COHERE

CONTACT Stephen B. Levine sbl2@case.edu Department of Psychiatry, Case Western Reserve University, 23425 Commerce Park, STE 104, Cleveland, OH 44106-7078, United States

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

(Council for Choices in Health Care), 2020; Socialstyrelsen [National Board of Health and Welfare], 2022; National Health Service (NHS), 2022a). In the U.S., however, medical organizations so far have chosen to use their eminence to shield the practice of pediatric "gender affirmation" from scrutiny. In response to mounting legal challenges, these organizations have been exerting their considerable influence to insist the science is settled (American Medical Association (AMA), 2022). We argued that this stance stifles scientific debate, threatens the integrity and validity of the informed consent process—and ultimately, hurts the very patients it aims to protect.

To demonstrate problems in existing research, we discussed two seminal studies that gave rise to the now-common practice of performing gender transitions on young people by giving them puberty blockers, cross-sex hormones, and "gender-affirming" surgery (de Vries et al., 2011; de Vries et al., 2014). We argued that these Dutch studies suffer from such profound limitations that they should never have been used as justification for propelling these interventions into general medical practice. We called for rigorous clinical research into the interventions known as "gender-affirming" care before these interventions are further scaled. Until such research is available, we urged clinicians to disclose the profound uncertainties regarding the outcomes of this treatment pathway to enable patients and families to make better-informed decisions about their care.

Our assertions drew a response from the first author of these Dutch studies (de Vries, 2022).[1] de Vries dismissed much of our criticism as a mere "misunderstanding" of their gender clinic's process. While de Vries acknowledged some of the limitations in the Dutch research, she asserted that these gaps have since been sufficiently remedied by subsequent research from others in the field, rendering the practice of pediatric gender transition as proven beneficial, and ready to be widely scaled in general medical practice.

Having carefully examined de Vries' counterarguments, we failed to find a single instance where our "misunderstanding" could explain away the significant problems that we pointed out. In this article, we justify our position that neither the Dutch research, nor the research that followed, is fit for shaping policy or treatment decisions regarding gender dysphoric youth at the population level. We present our response to de Vries in three sections. *First,* we provide a more complete justification for our assertions of the significant flaws in the foundational Dutch research. *Second,* we demonstrate that the claims that subsequent research remedied the deficiencies in the prior research are untrue. *Third*, we provide recommendations for research structure to yield reliable, trustworthy information. We conclude with a sense of urgency to avoid future harms by reminding readers of the intrinsic value of high-quality science.

Before we embark on outlining the critical methodological limitations of the Dutch research, we would like to make it clear that it is not our intention to discredit the Dutch clinicians' past work. The quality of the Dutch studies, while unacceptably low by today's standards, is commensurate with clinical and research practices in the 1990s. The key problem in pediatric gender medicine is not the lack of research rigor in the *past*—it is the field's *present-day* denial of the profound problems in the existing research, and an unwillingness to engage in high quality research requisite in evidence-based medicine.

Evidence-based medicine vs empirical-based medicine

When the Dutch clinicians launched the practice of pediatric gender transition, it was not uncommon for medical professionals to practice medicine based on "empirical evidence," relying on expert opinion and often backed by only minimal research (Drisko & Friedman, 2019). The term "evidence-based medicine" and its focus on quality comparative clinical research to determine optimal treatment only emerged in the 1990s (Guyatt, 1993). The Dutch researchers began to medically transition gender dysphoric adolescents in the late 1980s–early 1990s—just as medicine was starting to undergo this major paradigm shift.

Examining the Dutch research from today's vantage point, their gender-transitioning of youth is most consistent with the "innovative practice" framework. This framework allows clinicians

to implement untested but promising interventions for a condition which, if left untreated, might have dire outcomes; when existing treatment options seem ineffective; and when the number of affected patients is small (Brierley & Larcher, 2009; Earl, 2019). The number of adolescents suffering from gender dysphoria in the 1990s was exceedingly small. Evidence was starting to demonstrate that gender reassignment undertaken in adulthood failed to resolve trans people's mental health problems (Cohen-Kettenis & Van Goozen, 1997). The Dutch clinicians hoped that the "less positive results among adults" (p. 266) would be remedied with early adolescent gender transition. In this context, the methodological deficiencies in the foundational Dutch research ought not to be viewed as a *failure*. It was never their goal to generate *reliable reproducible research*. In fact, the many irregularities, which we elucidate below, reflect the Dutch *success* at rapidly evolving their approaches to reach a point of *technical excellence*: convincing physical transformations of adolescent bodies that satisfied the young patients (Biggs, 2022). These clinicians were "flying the plane while building the plane," and their published research merely reflects this messy clinical reality.

The "innovative practice" model of care is a double-edged sword. On the one hand, it rapidly advances the medical field. On the other hand, it is capable of hurting individuals and societies by promoting a nonbeneficial or harmful intervention. For these reasons, it is an ethical requirement that as soon as viability of a new intervention is demonstrated under the "innovative practice" framework, the research must move into high-quality clinical research settings capable of demonstrating that the benefits outweigh the risks. This step is imperative because it prevents "runaway diffusion"—the phenomenon whereby the medical community mistakes a small innovative experiment as a proven practice, and a potentially nonbeneficial or harmful practice "escapes the lab," rapidly spreading into general clinical settings (Earl, 2019).

"Runaway diffusion" is exactly what has happened in pediatric gender medicine. "Affirmative treatment" with hormones and surgery rapidly entered general clinical practice worldwide, without the necessary rigorous clinical research to confirm the hypothesized robust and lasting psychological benefits of the practice. Nor was it ever demonstrated that the benefits were substantial enough to outweigh the burden of lifelong dependence on medical interventions, infertility and sterility, and various physical health risks. The studies also failed to quantify the risk to "false positives"—that is, those gender dysphoric youth whose distress would have remitted with time without resorting to irreversible medical and surgical interventions.

The speed of the "runaway diffusion" accelerated exponentially when pediatric gender dysphoria/transgender identity went from a relatively rare phenomenon before 2015, to one that impacts as many as 1 in 10–20 young people in the Western world (American College Health Association [ACHA], 2022; Johns et al., 2019; Kidd et al. 2021). The current politicization of transgender healthcare has provided further fuel to the rapid proliferation of youth gender reassignment. A proposal by the U.S. government to mandate healthcare entities to provide "gender-affirming" interventions to minors, or risk claims of "discrimination" and loss of federal healthcare funding is yet another example of "runaway diffusion" (Health and Human Services [HHS], 2022; Keith, 2022).

The difficult task of reversing runaway diffusion begins with a systematic review of evidence, follows with updating treatment guidelines, and culminates with de-implementation of unproven or harmful practices, known as "practice reversals" (Herrera-Perez et al., 2019; Prasad, 2011; Prasad & Ioannidis, 2014). *Systematic reviews of evidence* play a uniquely important role in this process. Rather than arbitrarily selecting studies and simply restating their results and conclusions, systematic reviews of evidence analyze *all of the available evidence* meeting pre-specified criteria and *scrutinize the studies* for methodological bias and errors, issuing an overarching conclusion about what's known about the effects of an intervention based on the totality of the evidence (Higgins et al., 2022). A "practice reversal" of pediatric gender transitions has already begun. Several recent international systematic reviews of evidence have concluded that the practice of pediatric gender transition rests on *low to very low quality evidence*—meaning that the benefits reported by the existing studies are unlikely to be true due to profound problems in the study designs (National

Institute for Health and Care Excellence (NICE), 2020a, 2020b; Pasternack et al., 2019; SBU (Swedish Agency for Health Technology Assessment and Assessment of Social Services), 2022). Following these systematic reviews of evidence, three European countries—Sweden, Finland and England—have begun to articulate new and much more cautious treatment guidelines for gender dysphoric youth, which prioritize noninvasive psychosocial interventions while sharply restricting the provision of hormones and surgery (COHERE (Council for Choices in Health Care), 2020; Socialstyrelsen [National Board of Health and Welfare], 2022; NHS, 2022a).

Paradoxically, this international reckoning has had almost no influence on the U.S. gender medicine establishment. When Florida's Medical Board, following an overview of existing systematic reviews (Brignardello-Peterson & Wiercioch, 2022), took on the question of regulating pediatric gender medicine and invited the proponents of pediatric gender transitions to reconcile their stance with the recent European developments, these clinician advocates were either unaware of the European changes, or minimized their extent and significance (Janssen, 2022 00:46:43; McNamara, 2022 01:45:27). More generally, when faced with questions about the rapidly growing numbers of youth subjected to highly invasive and often irreversible interventions based on *low to very low quality evidence*, the field of U.S. pediatric gender medicine has chosen to throw its weight behind two indefensible and contradictory claims: (1) that "low quality evidence" is a misleading technical term which actually describes high quality reliable research; and (2) that true high quality research can only come from randomized placebo-controlled trials, which are unattainable and unethical (Drescher, 2022; McNamara et al., 2022). We refuted these misleading claims in our recent publication (Levine et al., 2022b).

As we begin our discussion of the profound limitations in the two foundational Dutch studies that have propelled the practice of pediatric gender transition into mainstream clinical practice worldwide, we are aware that we are mounting a serious challenge to the research that has been viewed by many as the "gold standard" in the field. Questioning this assumption, we welcome further debate. A quote from philosopher Karl Popper, perceptively invoked by Balon (2022), is particularly apt: "the growth of knowledge depends entirely on disagreement."

I. The "Dutch studies" are deeply flawed

There is no argument that the Dutch experience, and in particular two Dutch studies—de Vries et al. (2011), and de Vries et al. (2014)—forms the foundation of the practice of youth gender transition. It is evident when examining prevailing treatment guidelines. The Endocrine Society's statements regarding the potential benefits of puberty blockers and cross-sex hormones in gender dysphoric adolescents are supported only by references to these two studies (Hembree et al., 2017, p. 12, p. 16). Similarly, the World Professional Association for Transgender Health (WPATH) "Standards of Care" guidelines version 7 (SOC 7)—the version under which the practice of medicalization of gender dysphoric youth became widespread—only references the Dutch experience (Coleman et al., 2012). Despite several newer studies available, the proponents of gender affirmation still correctly emphasize that "the best longitudinal data we have on transgender youth comes primarily out of the Dutch clinic…the Dutch studies in the Dutch model of care. That's the prevailing model that most of the American clinics have based their care upon" (Janssen, 2022, 00:47:42). de Vries in her response to us, also agrees with this: "…indeed, as of today, the Dutch papers, and especially the de Vries et al., 2014 study, are still used as main evidence for provision of early medical intervention including puberty blockers in transgender youth (de Vries et al., 2014)" (de Vries, 2022, p. 2).

The two main Dutch studies in question, de Vries et al., 2011, and de Vries et al., 2014 (from here on, "the Dutch studies") convincingly demonstrated that hormonal and surgical interventions can successfully change the phenotypical appearance of secondary sex characteristics of adolescents and young adults. What the studies *failed* to show, however, is that these physical changes resulted in meaningful psychological improvements significant enough to justify the adverse effects of the treatment—including the *certainty* of sterility.

Besides the lack of a control group and a small final sample of 55 cases, with key outcomes available for as few as 32 individuals, there are *three major areas of concern* that render these studies unfit for clinical or policy decision-making.

A. **High risk of bias**: The Dutch studies suffer from multiple sources of bias which undermine confidence into the reported "benefits." The subject selection assured that only the most successful cases at each treatment stage were included in reported results. The linchpin finding of "resolution of gender dysphoria" is entirely invalid, since the home-grown gender dysphoria scale and its scoring mechanism were reversed after treatment, essentially guaranteeing a significant post-surgical drop in "gender dysphoria" scores. The finding of modest psychological benefits was compromised by the conflation of medical interventions with psychotherapy, making it impossible to determine whether gender reassignment, therapy, or the psychological maturation that occurs with the passage of time led to these few modest "improvements."
B. **Incompleteness of evidence regarding physical health risks**: The Dutch studies did not evaluate *physical health* outcomes of "gender-affirmative" treatments, even though adverse effects of hormonal interventions on bone and brain had been hypothesized from the start (and were confirmed by subsequent research). Even without setting out to assess the risks, the Dutch research inadvertently revealed that the rate of short-term morbidity and mortality associated with "gender-affirming" interventions may be as high as 6%-7%.
C. **Poor generalizability/applicability to current cases**: Today, most youth suffer from post-pubertal onset of gender dysphoria and significant mental illness—two clinical presentations the Dutch *explicitly disqualified* from their studies. As such, none of the Dutch findings are applicable to most of the youth seeking treatment today.

de Vries (2022) disputed only our assertion that the studies suffer from *high risk of bias* and therefore their findings of benefits are unreliable. She did not comment on our arguments that the research *failed to assess physical health risks* and *were not generalizable* to the majority of currently presenting cases. It is unclear if this silence indicates agreement or disagreement. Below, we address each of our points in greater detail, concluding with an additional observation about the overall lack of equipoise—genuine uncertainty about which treatment options are superior (London, 2017), which limits the utility of the Dutch research beyond describing a small-scale "innovative practice."

A. High risk of bias in the Dutch research

de Vries rejected our assertion that the Dutch findings suffer from a high risk of bias and insisted that we mistook the study protocol's careful process of establishing study eligibility for "bias." To clarify, we use the term "risk of bias" in a strict methodological sense. It refers to a systematic error, or deviation from the "truth" in study results (Boutron et al., 2022; Socialstyrelsen [National Board of Health and Welfare], 2022). Observational research conducted in the context of ongoing clinical care is often subject to risk of bias (Nguyen et al., 2021), which is one of the main reasons why rigorous clinical research using robust research designs must follow. In the case of the Dutch studies, we identified three major sources of bias, or systematic error, involving: (1) case selection; (2) measurement of outcomes; and (3) confounding.

1. Bias in case selection: Only the "best-case scenario" cases made it into the Dutch studies' "completers"

Because of an unusual case selection and reporting methodology, the Dutch studies inadvertently reported on only their best-case outcomes at each of the three phases of treatment (puberty blockers, cross-sex hormones, and surgery)—while failing to report the outcomes of the less positively affected, or even harmed, cases. de Vries disagreed with this assertion, continuing to insist that "participation was based on consecutive referral" (de Vries, 2022, p. 4).

Below, we present evidence that the claim of consecutive referral-based *prospective case selection* is not technically accurate. The actual case selection for the original sample of 70 puberty-blocked cases (de Vries et al., 2011) was *retrospective* and inadvertently biased toward including cases with favorable outcomes. The outcome reporting methodology in the second and final Dutch study (de Vries et al., 2014), which evaluated the final outcomes post-surgery, further biased the results toward reporting on the most favorable cases.

de Vries et al., 2011 ("puberty blocker" study). The 70 cases comprising the entire sample for the "puberty blocker" study (de Vries et al., 2011) were *retrospectively, non-randomly selected* from a larger group of consecutively referred 111 cases. According to both the original study, and de Vries' response to us, to participate in the "puberty blocker" study, a study subject already had to be starting the *next phase* of treatment with cross-sex hormones:

> Of the 196 consecutively referred adolescents…111 (those below age 16) had started puberty suppression… In the 2011 study we evaluated the first 70 of those 111 who were about to start with the next step of their treatment, affirming hormones, around the age of 16 years. (de Vries, 2022, p. 4)[2]

Using the start date of the *next phase* of treatment (cross-sex hormones) as the defining inclusion criterion for the study of the *prior phase* of the treatment (puberty blockers) introduced serious bias.

First, had any of the original 111 study subjects been harmed by puberty blockers or chosen to stop the treatment, they would never have advanced to the next phase, and thus, they had no chance of being included in the puberty blocker study, skewing the sample. *Second*, since the Dutch considered the puberty suppression phase both a treatment and a *diagnostic phase* (Cohen-Kettenis & van Goozen, 1998), the more complex cases may have remained in the puberty blocked phase longer. As de Vries' predecessors explained, subjects for whom the psychotherapist or parents had doubts, or where "the personal situation of the youngster" was more complicated, were delayed from starting cross-sex hormone treatment, which was the first stage the Dutch researchers considered to have an "irreversible" effect (Gooren & Delemarre-van de Waal, 1996, p. 11). This would further skew "the first 70 of those 111 who were about to start with the next step of their treatment, affirming hormones" (de Vries, 2022, p. 4)—the entire puberty blocker study sample—toward the most clinically straightforward and stable cases.

Third, such an unusual case selection methodology may have skewed the sample toward an older age than was stipulated by the protocol. Since to be eligible for the "puberty blocker" study, a subject had to have been deemed ready to start the next phase of cross-sex hormones, which *required a minimum age of 16* (accroding to the Dutch protocol version published in 2012, de Vries, 2012), all else being equal, older subjects had a greater chance of being included than younger ones. This may explain why the sample of 70 selected subjects was on average, age 15 when started on puberty blockers rather than age 12 as outlined by the protocol, which introduced another source of systematic error, by biasing the sample toward subjects with greater physical and cognitive maturity.

Given that the 2011 Dutch study's main goal was to evaluate the novel use of *puberty blockers* for gender dysphoria in a prospective cohort study (de Vries et al., 2011), the study should have enrolled, and reported the outcomes of, *all of the intent to treat* cases based on the date of eligibility to start *puberty suppression*—not cross-sex hormones.

It is notable that the only attempt to replicate the 2011 Dutch study results with more than a handful of cases took place in the UK but failed (Carmichael et al., 2021), with the conclusion of "no changes in psychological function" (p. 1). We suspect the key reason for this failure was the fact that the UK researchers truly *prospectively* selected "sequentially eligible" cases for treatment (Carmichael et al., 2021, p. 4) and as a result, ended with a diverse range of outcomes, including worsening of problems among female subjects during puberty blockade (Biggs, 2020). In contrast, the Dutch *retrospective* case selection methodology (misunderstood as prospective) inadvertently resulted in skewing the sample toward the best-case-scenario puberty-blocked cases. In our view, such case selection methodology invalidates the 2011 study conclusions of

psychological benefits of puberty suppression—or, as research methodologists would say, puts this finding at a "critical risk of bias."

de Vries et al, 2014 (post-surgery study). Skewing the sample toward the best-case scenario cases is even more apparent in the 2014 study, which reported on post-surgical outcomes and assessed the entire "gender-affirmative" treatment pathway (de Vries et al., 2014). The 70 participants who began the 2014 study, already biased toward more positive outcomes, shrank to 55. Fifteen subjects were dropped from the study and relabeled "nonparticipants." This subset, however, was not random, but instead heavily skewed toward subjects who experienced serious problems, including 3 who developed severe diabetes and obesity and 1 death following surgical complications. There is also considerable uncertainty about the outcomes of the 5 of 70 subjects (refusal, failure to return questionnaire, and dropping out of care) who, after several years of close contact with the research team, were unwilling to engage further:

> Nonparticipation (n = 15, 11 transwomen and 4 transmen) was attributable to not being 1 year postsurgical yet (n = 6), refusal (n = 2), failure to return questionnaires (n = 2), being medically not eligible *(e.g., uncontrolled diabetes, morbid obesity) for surgery (n = 3)*, dropping out of care (n =1), and *1 transfemale died after her vaginoplasty owing to a postsurgical necrotizing fasciitis* [emphasis added]. (de Vries et al., 2014, p. 697)

In her response, de Vries repeated the assertion that because a statistical comparison of the 15 "nonparticipants" to the 55 "participants" revealed no significant difference in their *pretreatment* baseline characteristics, "the results of the 2014 study can be generalized with substantial trust to the complete group of 70" (de Vries, 2022, pp. 4–5). We strongly disagree. The "participant" and "nonparticipant" cohorts are demonstrably different: while 100% of the 55 "participants" had successful gender reassignment according to the study reporting, at least 27% of the "nonparticipant" group (4/15: 1 death and 3 cases of diabetes) did not. Not only is a statistical analysis of such small subgroups massively underpowered to detect differences, ***no*** statistical analysis of *pretreatment* data suggesting "similarity" can negate the reality of the markedly different *post-treatment* outcomes in two groups. Nor is it clear why the research team made the unusual decision to stop the study early, before the remaining 6 participants had a chance to complete the 1-year post-surgical follow-up.

The "missing" Dutch study on the effect of cross-sex hormones. The second and final Dutch study (de Vries et al., 2014) combined the cross-sex hormone and post-surgical treatment results into a single set of outcomes. This conflation may have made some sense at the time, as all the hormonally-treated patients were *required* to undergo surgery (removal of breasts, ovaries, uterus, penis, testes, and construction of a neovagina) by the protocol. When surgery is not required, only 25–35% of transgender-identified adults appear to seek "gender-affirming" surgical procedures (Nolan et al., 2019). According to recently published data, this number is even smaller for youth: for every teen treated surgically, there are 15 treated *only* with cross-sex hormones (Respaut & Terhune, 2022). The inability of the Dutch research to elucidate the outcomes of cross-sex hormone treatments (separate from surgery) has been noted by NICE, which appropriately excluded the 2014 Dutch study from its systematic review of evidence (NICE, 2020b).

It is unknown whether the 4.3% of the sample (n = 3) that experienced obesity and diabetes sometime before the surgery was a result of the hormonal treatment; this rate appears to be double the expected rate for pediatric populations in the Netherlands at the time (Rotteveel et al., 2007; Schönbeck et al., 2011). Nor is it known if the cross-sex hormones contributed to the one subject who discontinued treatment due to other medical or psychological problems. Other research suggest that testosterone may actually *increase* dysphoria in female gender-dysphoric individuals (Olson-Kennedy, Warus, et al., 2018).

2. Bias in measurement of outcomes: The finding of "resolution of gender dysphoria" is invalid

The linchpin result of the Dutch studies is the reported *resolution of gender dysphoria*, as measured by the Utrecht Gender Dysphoria Scale (UGDS) (Steensma, Kreukels, et al., 2013). de

Vries agreed with us on this point: "the main finding remains the resolution of gender dysphoria" (de Vries, 2022, p. 3). According to the final Dutch study, the UGDS *gender dysphoria* scores plummeted, from a near-maximum score of 54 (maximum of 60) at baseline, to the near-minimum score of 16 (minimum of 12) after the final surgery (de Vries et al., 2014).

Rather than a true "resolution" of *gender dysphoria*, however, this spectacular drop was an artifact of switching the scale from "female" to "male" versions (and vice versa) before and after treatment, prompting a problematic *reversal* in the scoring. We argued that this fact alone invalidates the study's main conclusion of the resolution of gender dysphoria (Levine et al., 2022a). While de Vries conceded the use of the UGDS scale post-treatment was "not ideal" because "the UGDS was not…designed to be used after treatment," she asserted that it "does not imply that UGDS 'falsely' measured the improvement in GD [gender dysphoria]" (de Vries, 2022, p. 4). We think it is vitally important for the scientific community to recognize that the UGDS scale use was not merely "not ideal"—but that it *entirely invalidated* the Dutch study's main finding.

The following hypothetical scenario clearly demonstrates the problem. A severely gender dysphoric, cross-sex identified female patient is asked to answer two of the UGDS questions: "Every time someone treats me like a girl I feel hurt" and "Every time someone treats me like a boy I feel hurt" (Items 2 on the "female" and the "male" versions of the UGDS scale, respectively). It is likely that the patient would *strongly agree* with the first statement, and *strongly disagree* with the second. The first answer would lead to the score of "5" on the UGDS gender dysphoria scale, indicating the highest possible level of gender dysphoria. The second answer—which is effectively the same answer—would result in the score of "1" indicating the lowest possible gender dysphoria. This is because unlike the first question, which belongs to the "female" battery of questions, the second question belongs to the "male" battery of questions and effectively assumes the subject to be male—hence, the lack of distress of being associated with "maleness" receives the minimum "gender dysphoria" score.

If we now consider that only the "female" scale was used for gender dysphoric females at baseline but was then switched to the "male" scale after the final surgery (and vice-versa for male subjects), it becomes clear that the remarkable drop in "gender dysphoria" the UGDS scale registered after surgery entirely results from switching the scale. The *same* gender dysphoric individual, effectively answering the *same* question (albeit linguistically inverted), in the *same* way results in either the maximum or the minimum "gender dysphoria" score—depending on which sexed version of the scale was used. We reproduced both the "male" and the "female" versions of the UGDS scale in Table 1 so that others can easily observe how switching the scale "sex" version consistently leads to a "drop" of the gender dysphoria score, regardless of any treatment effect.

When defending the choice to reverse the UGDS scale (de Vries, 2022), de Vries pointed out—and we agree—that it would make no sense to ask postoperative natal males to rate a statement such as "I dislike having erections" (Table 1, UGDS-M, item 11), since they no longer have penises. We empathize with the Dutch researchers' plight, as they found themselves without a valid tool to measure the construct of "gender dysphoria" after treatment. It is equally nonsensical, however, to ask natal males to rate statements such as, "I hate menstruating because it makes me feel like a girl" (Table 1, UGDS-F, item 10)—and it makes even less sense to report "resolution of gender dysphoria" because they don't "hate menstruating."

In her response, de Vries pointed to the validation research of the UGDS dysphoria scale (de Vries, 2022; Steensma, Kreukels, et al., 2013). To the best of our knowledge, this work has never appeared in a peer-reviewed publication. In our opinion, this UGDS validation research missed a key opportunity to identify the threat to validity of using the UGDS scale in post-gender reassignment context, which should have become apparent to the Dutch research team by 2013 when the validation paper was published. The greater community of international gender clinicians relying on the Dutch pioneering experience was not alerted to the need to find another instrument that can provide a valid pre-post "gender dysphoria" measure. Instead, the validation

Table 1. Utrecht Gender Dysphoria Scale, Adolescent Version (de Vries, Cohen-Kettenis, & Delemarre-van de Waal, 2006). Response categories are *agree completely, agree somewhat, neutral, disagree somewhat, disagree completely*.

UGDS-F (female) Response categories are: agree completely, agree somewhat, neutral, disagree somewhat, disagree completely. Items 1, 2, 4–6 and 10–12 are scored from 5 to 1; items 3 and 7–9 are scored from 1 to 5.	UGDS-M (male) Response categories are: agree completely, agree somewhat, neutral, disagree somewhat, disagree completely. Items are all scored from 5 to 1.
1. I prefer to behave like a boy.	1. My life would be meaningless if I would have to live as a boy.
2. Every time someone treats me like a girl I feel hurt.	2. Every time someone treats me like a boy I feel hurt.
3. I love to live as a girl.	3. I feel unhappy if someone calls me a boy.
4. I continuously want to be treated like a boy.	4. I feel unhappy because I have a male body.
5. A boy's life is more attractive for me than a girl's life.	5. The idea that I will always be a boy gives me a sinking feeling.
6. I feel unhappy because I have to behave like a girl.	6. I hate myself because I'm a boy.
7. Living as a girl is something positive for me.	7. I feel uncomfortable behaving like a boy, always and everywhere.
8. I enjoy seeing my naked body in the mirror.	8. Only as a girl my life would be worth living.
9. I like to behave sexually as a girl.	9. I dislike urinating in a standing position.
10. I hate menstruating because it makes me feel like a girl.	10. I am dissatisfied with my beard growth because it makes me look like a boy.
11. I hate having breasts.	11. I dislike having erections.
12. I wish I had been born as a boy.	12. It would be better not to live than to live as a boy.

research buttressed the problematic practice of using UGDS to measure the level of gender dysphoria after gender reassignment by stating: "From follow-up studies it was already known that gender dysphoria, as measured by the UGDS, disappeared post gender reassignment. These qualities make the instrument useful for clinical and research purposes" (Steensma, Kreukels, et al., 2013, p. 56). This statement is misleading, as the finding of the "disappearance" of gender dysphoria post-gender reassignment in the past "follow-up" research came from studies that also switched the sexed scale versions post-treatment, as Dr. de Vries pointed out in her response to us (de Vries, 2022).

Thus, in a spectacular display of circular reasoning, the scale validation research claimed that the follow-up research endorsed the use of the inverted UGDS scale version post gender reassignment, while the follow-up research defended this unusual practice by pointing to the validation research. de Vries doubled down on this circular reasoning in her response to our critique (de Vries, 2022):

> Levine et al. (2022) questions whether the improvement in gender dysphoria does then not stem from this switching, and not from the treatment? However, this seems turning the matter around. What the measure shows, the disappearance or resolution of gender dysphoria, is what the gender affirming treatment is aimed to resolve. (pp. 3–4)

At least three research groups noted the critical threat to the validity of the finding of "resolution of gender dysphoria" due to the switching of the scale (Biggs, 2022; McGuire et al., 2020; van de Grift et al., 2017). McGuire et al. (2020) explicitly stated, "Because the original UGDS is composed of two scales, it is impossible to determine if this is a real difference in gender dysphoria between groups or if this is an artifact of measurement error (p. 195).

The likely meaning of the "plummeting" gender dysphoria scores. What, if anything, did the "plummeting" gender dysphoria scores post scale-flipping signal, if not the "disappearance of gender dysphoria" claimed by the Dutch researchers? We posit that the UGDS scale can only measure the construct which it was originally designed and validated to measure—the level of incongruence between natal sex and gender identity leading to the provision of the DSM diagnosis (Cohen-Kettenis & van Goozen, 1997; Iliadis et al., 2020; Steensma, Kreukels, et al., 2013). This is true whether the scale is used before or after treatment, and whether the "treatment" in question is "gender-affirmation" with hormones and surgeries, psychotherapy, or mere "watchful waiting," with the scale administered at various time points.

The fact that after gender reassignment, the UGDS scores were low on the opposite-sex scale indicates that the subjects would have scored high on the natal sex scale, which corresponds to a *persistence in transgender identity*. This is the only plausible interpretation of the "plummeting" UGDS scores that survives in the context of the scale questions and the linguistic and numerical gymnastics the scale underwent in the post-gender-reassignment context. The finding of persistence of transgender identity is not unexpected, especially since the Dutch researchers selected subjects with lifelong extreme cross-sex identification and follow-up was only 1.5 years post-surgery. What it does *not* mean is that the feeling of "incongruence" resolved. This point is underscored by the long-term follow-up data on male-to-female Dutch transitioners, presented at the WPATH 2022 Symposium by Dr. van der Meulen (Steensma et al., 2022). Nearly a quarter of the participants have felt that their bodies were still too masculine, and over half have experienced shame for the "operated vagina" and fearful their partner will find out their post-surgical status—despite registering low "gender dysphoria" UGDS scores (Steensma et al., 2022).

3. Bias from confounding: Psychotherapy was comingled with medical interventions

Although the Dutch research is frequently commended for having demonstrated "psychological improvements," an examination of the outcomes reveals that standard measures of psychological functioning such as anxiety, depression, anger, and global function showed very little clinically significant change after treatment (Levine et al., 2022a). de Vries acknowledged that a number of psychological measures showed no meaningful change, but insisted that the "more robust" measures, such as Child Behavior Check List (CBCL) and Youth Self Report (YSR), *did* show clinically relevant changes (de Vries, 2022, p. 3). She also noted that post-intervention, the sample of gender dysphoric youth in the Dutch research functioned at a similarly high level as their non-dysphoric peers, which was also an indicator of success. We have three observations about this response.

First, the impressive drop in the percentage of cases in the "clinical" range for CBCL and YSR (de Vries et al., 2014) was only apparent after *dichotomizing* these scales into the "clinical" (problematic) versus "non-clinical" ranges. In comparison, the sample's *average* post-intervention score changes on these scales were much more modest. For example, while the 2014 Dutch study points out that the "percent in the clinical range dropped from 30% to 7% on the YSR/ASR," which looks like an impressive reduction, the *average* t-scores had a modest drop of from 54.72 before treatment, to 48.53 after surgery (de Vries et al., 2014, p. 702). Further, both before and after t-scores were less than 60—typically interpreted as having no clinically significant symptoms (Achenbach & Rescorla, 2001). This suggests the reported improvements in CBCL and YSR came from relatively small score changes, which are of limited clinical significance, even if in the process the clinical threshold is crossed for some cases.

Second, while de Vries points to the *post-treatment* similarity in function of the gender-dysphoric group to the general population as evidence of treatment success, it is not known how different the groups were from the general population *pretreatment*. According to earlier research by Cohen-Kettenis and van Goozen (1997), which presumably utilized similar selection criteria, "when both pre- and posttest group means were compared with Dutch normative data, *all scores turned out to be within the average range* [emphasis added]" (p. 269). Smith et al. (2001) confirm this and explicitly state that both pretreatment and post-treatment, the group of gender dysphoric youth selected for the interventions were "normal functioning" as compared to their age peers in the Netherlands (Smith et al., 2001, p. 477). If the sample used in the two Dutch studies, which was recruited several years later but used the same careful case selection criteria, bears resemblance to the sample described by this earlier Dutch research, then the reported post-treatment similarities in psychological function between the "treated" group and the general population of peers should not be attributed to gender reassignment.

Third, and perhaps most relevant to this discussion, is the question of whether *any* of the reported changes in post-treatment psychological function scores, clinically significant or not, can be reasonably attributed to gender reassignment—or if these changes were influenced by confounding factors not accounted for in the research design. As noted by the authors of the

CBCL and YSR scales that de Vries says she favors, "improvement in scores from before to after services does not prove that the services were responsible for improvement. Other explanations are possible, such as (a) children's problems tend to decrease as they get older; (b) the people providing the data may report improvements because they believe that the services helped, and (c) the test-retest attenuation effect (a general tendency for people to report fewer problems at a second assessment)" (Achenbach & Rescorla, 2001, p. 183).

In addition to the general sources of confounding in uncontrolled studies relying on "before and after" measures, a vital source of confounding in the Dutch studies has been hiding in plain sight: All the subjects received psychotherapy at the same time they were undergoing gender reassignment. This comingling of interventions makes it impossible to determine which of the interventions "worked."

Psychotherapy was a key element in the Dutch protocol. Contrary to the now-common but erroneous assertion by the U.S. gender medicine establishment that psychotherapy for gender dysphoria is akin to "conversion" and should be avoided or even banned (Cantor, 2020), the Dutch studies reveal that psychotherapy was a key element of the protocol. According to the Dutch protocol, "[i]n cases involving confusion about gender feelings, psychotherapy and peer support can be helpful in *resolving the confusion and coming to self-acceptance* [emphasis added]" (de Vries, Cohen-Kettenis & Delemarre-van de Waal, 2006, p. 87). Not only was psychotherapy thought to be beneficial, but apparently it was a core part of the intervention: "…the adolescents were all regularly seen by one of the clinic's psychologists or psychiatrists. Psychological or social problems could thus be timely addressed" (de Vries et al., 2011, p. 2281). The researchers acknowledge that psychotherapy "…may have contributed to the psychological well-being of these gender dysphoric adolescents" (de Vries et al., 2011, p. 2281).

A discussion of the utility of psychotherapy to ameliorate gender dysphoria and related psychological distress is outside the scope of this article, other than to point out that the results of at least two studies suggest that psychological interventions are associated with improvements in two of the outcome domains—*gender dysphoria* (van de Grift et al., 2017) and *global function* (Costa et al., 2015)—absent any medical interventions.

B. Incompleteness of evidence regarding risks

Failure to consider the physical health risks of "gender-affirming" endocrine and surgical interventions is another methodological weakness of the Dutch studies. This omission is surprising since the Dutch team hypothesized that hormonal interventions might adversely impact bone and brain development several years before their seminal studies commenced (Delemarre-van de Waal & Cohen-Kettenis, 2006, p. 134). As discussed earlier, the Dutch studies did, however, report on the cases that were reclassified from "participants" to "non-participants," and listed the reasons for the nonparticipation, which revealed a possible 6–7% rate of associated adverse events.

Several studies since have confirmed likely adverse health effects of hormonal interventions, although their long-term impact on future health is not yet known. Research suggests that youth treated with puberty blockers develop problems with bone density accrual (Biggs, 2021; Nokoff et al., 2022) and that bone density may be impaired even after treatment with cross-sex hormones is initiated (Klink et al., 2015). Other research suggests heightened insulin resistance (Nokoff et al., 2021), elevated blood pressure, elevated triglycerides, and impaired liver function (Olson-Kennedy, Okonta, et al., 2018). Cross-sex hormone administration places adolescents in the medical category of early life indicators of future cardiovascular disease (Jacobs et al., 2022).

These adverse changes, already evident after a relatively short period of hormonal interventions, do not bode well for long-term health, since "gender-affirming" hormones are prescribed with the presumption of ongoing, lifelong treatment essential for maintaining a masculinized or feminized appearance. It is likely that other medical risks will emerge in the future. Patients and their families cannot make informed decisions about a treatment when the physical health

risks are assumed to be minimal and not reported, and only the potential psychological benefits are considered.

C. Poor generalizability/applicability to currently presenting cases

Given the dramatic change in the epidemiology of youth gender dysphoria which occurred after the studies were published (Levine et al., 2022a), the question of the applicability of the Dutch research to the current clinical dilemmas is one of the most important questions to interrogate in the field of pediatric gender medicine today.

Generalizability/applicability questions whether "available research evidence can be directly used to answer the health and healthcare question at hand" (Schünemann et al., 2022). We asserted and continue to assert that the Dutch studies are not applicable/generalizable to most gender dysphoric youth presenting today. This is evidenced by two facts: (1) the most common profile of youth seeking gender transition today is an adolescent with postpubertal emergence of a transgender identity and significant uncontrolled mental health comorbidities; (2) the Dutch researchers explicitly disqualified such patients from their studies because of their concern that the risks of early gender transition might outweigh the benefits.

1. Most of today's adolescents have postpubertal onset of trans identity and comorbid mental illness

Until about a decade ago, most patients seen by gender clinics were very young boys who wished to be girls and most of these children subsequently lost their cross-sex identification before reaching adulthood (Hembree et al., 2017; Ristori & Steensma, 2016; Singh et al., 2021). Today, the majority are female adolescents (de Graaf et al., 2018; Kaltiala-Heino et al., 2018; Zhang et al., 2021) with previously gender-normative childhoods whose trans identity emerged around or after puberty (Hutchinson et al., 2020; Zucker, 2019). Many suffer from significant preexisting mental illness such as depression and anxiety or neurocognitive challenges such as autism spectrum disorder (ASD) or attention deficit hyperactivity disorder (ADHD) (Becerra-Culqui et al., 2018; de Graaf et al., 2021; Kaltiala-Heino et al., 2015; Kozlowska et al., 2021; Strang et al., 2018; Thrower et al., 2020).

The presentation of adolescent-onset gender dysphoria is not entirely new—what's new is its scale. As with many trends, the change occurred "gradually, then suddenly." While there was evidence of it in the mid-2000s, around 2014–2015 the presentation of pediatric gender dysphoria in the Western world sharply shifted, from childhood-onset that skewed toward males, to adolescent-onset with a preponderance of females with mental health problems (Aitken et al., 2015; de Graaf et al., 2018). The Dutch researchers began their experiments with pediatric gender transition well before this demographic shift began to dominate clinical presentations of youth gender dysphoria.

Finland's national pediatric gender program was among the first to sound the alarm regarding the changing epidemiology of gender dysphoria presentation in youth. In 2015, they began observing that the youth presenting for treatment were primarily females who "do not fit the commonly accepted image of a gender dysphoric minor" (Kaltiala-Heino et al., 2015). The Finnish researchers saw a new pattern of "severe psychopathology preceding onset of gender dysphoria," with 75% already in treatment for other psychiatric issues when their gender dysphoria emerged. By 2019, the Finnish gender program was in full-alarm mode: "Research on adolescent onset gender dysphoria is scarce, and optimal treatment options have not been established… The reasons for the sudden increase in treatment-seeking due to adolescent onset gender dysphoria/transgender identification are not known" (Kaltiala-Heino & Lindberg, 2019, p. 62). This changing epidemiology was noted by other Nordic countries as well (Kaltiala, Bergman, et al., 2020).

The novel presentation of youth gender dysphoria was also reported by the largest pediatric gender clinic in the world at the time, the UK's GIDS/Tavistock (de Graaf et al., 2018). The now-famous graph of the GIDS data shows a trickle of gender dysphoric youth in years past

turning into a tidal wave by 2015, with a significant overrepresentation of teen girls. Between 2009 and 2016, the number of gender dysphoric females increased more than 70 times (de Graaf et al., 2018). The UK researchers concluded:

> The steep increase in birth-assigned females seeking help from gender services across the age range highlights an emerging phenomenon. It is important to follow birth-assigned females' trajectories, to better understand the changing clinical presentations in gender-diverse children and adolescents and to monitor the influence of social and cultural factors that impact on their psychological well-being. (de Graaf et al., 2018, p. 4)

The number of gender dysphoric youth referrals in the UK doubled again between 2020–2021 and 2021–2022 (NHS, 2022b).

While U.S. population-level data are hard to come by due to the country's decentralized and highly fragmented health care system, recent research shows that the number of gender dysphoric teens has also sharply risen in recent years, with a nearly 70% increase just between 2020 and 2021 (Respaut & Terhune, 2022). Combined with U.S. medical chart data samples, which show that the composition of the population changed "from predominantly transfeminine to…predominantly transmasculine in children and adolescents" (Zhang et al., 2021, p. 390) and that over 70% of gender dysphoric youth had been diagnosed with ASD, ADHD and other mental health problems *before* their diagnosis of gender dysphoria (Becerra-Culqui et al., 2018), it is apparent that the U.S. has not been immune to this remarkable epidemiologic trend that has engulfed youth in the Western world.

This now-ubiquitous presentation of gender dysphoria in troubled adolescents with previously gender-normative childhoods lacks a DSM-5-TR descriptor (American Psychiatric Association [APA], 2022), leaving clinicians to refer to it by many names, including *adolescent-onset gender dysphoria*; *postpuberty adolescent-onset transgender history*; and *rapid-onset gender dysphoria (ROGD).* The latter term was introduced by a U.S. researcher (Littman, 2018). Despite the controversy that Littman's hypotheses generated in the gender medicine establishment (Marchiano, 2018*),* her research withstood a second round of rigorous peer review (Littman, 2020). Subsequent detransitioner research lent further support to the ROGD hypothesis, with patients themselves reporting "that their gender dysphoria began during or after puberty and that mental health issues, trauma, peers, social media, online communities, and difficulty accepting themselves as lesbian, gay, or bisexual were related to their gender dysphoria and desire to transition" (Littman, 2021, p. 15). Even WPATH, which in 2018 strongly objected to Littman's research (WPATH, 2018), conceded in its 2022 "Standards of Care 8" that while no one has attempted to replicate Littman's research, it is apparent that "[f]or a select subgroup of young people, susceptibility to social influence impacting gender may be an important differential to consider" (Coleman et al., 2022, p. S45).

The novel phenomenon of high numbers of young people declaring a transgender identity for the first time in adolescence, often in the context of preexisting mental illness and/or trauma and social difficulties, has been described by several other mental health clinicians (Hutchinson et al., 2020; Schwartz, 2021; Zucker 2019). The only exception to the trend of mentally struggling adolescents presenting with gender dysphoria is the Amsterdam gender clinic itself, which has also seen an influx of teens and the preponderance of girls, but apparently without the mental health problems (Arnoldussen et al., 2020). Nonetheless, writing for the American journal *Pediatrics*, de Vries recognized the emergence of this new clinical phenomenon, noting that "gender identity development is diverse, and a new developmental pathway is proposed involving youth with postpuberty adolescent-onset transgender histories" (de Vries, 2020, p. 1) and noting that "some case histories illustrate the complexities that may be associated with later-presenting transgender adolescents and describe that some eventually detransition (de Vries, 2020, p. 2).

2. The Dutch studies disqualified cases most commonly presenting today: Adolescents with recent-onset gender dysphoria, nonbinary identities, or mental illness

From the outset in the late 1990s when the Dutch researchers first began to report on the results of youth gender transitions, they made it clear that their focus was exclusively on youth with

complete cross-sex identification "from toddlerhood onwards" (Cohen-Kettenis & van Goozen, 1998, p. 1). Furthermore, there was a strict requirement of psychological stability:

> First, they must have shown a *lifelong extreme and complete crossgender identity/role* [emphasis added]. Around puberty these feelings and behaviors must have become more rather than less pronounced. Second, they must be *psychologically stable* [emphasis added] (with the exception of depressed feelings, which often are a consequence of their living in the unwanted gender role) and function socially without problems (e.g., have a supportive family, do well at school). (Cohen-Kettenis & van Goozen, 1997, p. 265)

Of note, youth with non-binary identities, common today (Green et al., 2022), were *ineligible* for medical interventions according to the Dutch protocol, and instead needed psychotherapy: "adolescents… whose wish for sex reassignment seems to originate from factors other than a genuine and complete cross-gender identity are *served best by psychological interventions* [emphasis added] (de Vries et al., 2006, pp. 87–88).

Thus, the Dutch protocol explicitly *excluded* the characteristics of adolescents presenting to clinics in recent years—those whose trans-identities emerged around puberty; non-binary presentations without the wish for a complete cross-sex reassignment; or cases of gender dysphoria accompanied by significant uncontrolled mental illness. The high level of psychological functioning of the Dutch cohort *at baseline* serves as evidence that these selection criteria were indeed followed at the time (de Vries et al., 2011). The fact that "gender-affirming" interventions are now provided to the very segment that was explicitly excluded from the eligibility in the foundational studies is alarming.

D. Failure to consider alternatives (lack of research equipoise)

The Dutch researchers began their research into treatments of gender-dysphoric adolescents with the *foregone conclusion* that children who had life-long gender dysphoria and who continue to be cross-sex identified as adolescents would inevitably grow up to be transgender-identified adults. This assumption, based on "expert observations" from a handful of cases (O'Malley & Ayad, 2022; Cohen-Kettenis & van Goozen, 1997), has never been tested in rigorous comparative research. Further, the research team assumed that the only feasible treatment for these adolescents is early gender transition, and that psychotherapy alone is ineffective—also without testing this assumption through research. This violates the key requirement of equipoise in research—the principle that clinical investigators must approach research with genuine uncertainty regarding diagnostic, prevention, and treatment options—and allocate individuals to interventions in a manner that allows for generation of new knowledge (Freedman, 1987; London, 2017).

In fact, as de Vries' response to us emphasizes, the Dutch researchers continue to hold such firm belief into the beneficial nature of gender reassignment for youth, that they are far more concerned with the risk of "nontreatment" with hormones and surgery than they are with the possibility that the youth undergoing transition may not have needed such drastic interventions (de Vries, 2022, p. 3). However, some of the earlier research on the "non-treated" gender-variant and gender dysphoric adolescents challenges the assumptions of the permanence of trans identity in teens.

1. Non-treatment of "referred" adolescents with significant mental illness

Because of the careful case selection, the Dutch protocol rejected some youth from eligibility for gender reassignment due to serious "psychological or environmental problems" (Smith et al., 2001, p. 473). According to the study that followed the trajectories of these youth, the majority no longer wished to undergo gender transition once they reached *adulthood.*

Smith et al. (2001) reported that individuals rejected from gender reassignment in adolescence found noninvasive ways to deal with their gender dysphoria, and gender dysphoria significantly diminished. Upon follow-up 1–7 years later, only 22% of the rejected subjects (6/27) underwent gender reassignment as adults, while 78% refrained from it. Among those who remained medically untreated and participated in follow-up research, a remarkable 79% (11/14) "*did not feel*

any regrets about having refrained from SR [sex reassignment] or being rejected….” Only 7% (1 of 14) expressed strong regret (Smith et al., 2001, p. 477).

Data from the study by Smith et al. (2001) raise the possibility that the majority of those rejected from hormonal interventions not only were unharmed by waiting but benefited from “nontreatment” with gender reassignment in adolescence. Unlike the medically and surgically treated subjects, the “rejects” completed uninterrupted physical and psychological development, avoided sterility, maintained their sexual function, eliminated their risk of iatrogenic harm from surgery, and avoided the need for decades of dependence on cross-sex hormones. These cases also demonstrate that the assumption that “adolescents do not desist” was not true even at the time the Dutch team first introduced gender transitions of youth. It is even less true now, with research showing 10-30% rates of medical detransition among those who were trans-identified in adolescence and young adulthood (Boyd et al., 2022; Hall et al., 2021; Roberts et al., 2022). The long-term follow-up data on the Dutch adolescent transitioner cohort recently presented at the WPATH 2022 Symposim (Steensma et al., 2022) also suggest that the rate of cross-sex identification was not as stable as originally expected, with a sizable percentage reporting one or more instances of identity changes after treatment completion, especially among the individuals on the autistic spectrum (Steensma et al., 2022).

2. Non-treatment of “gender variant” youth in a community sample

Another study, also from the Netherlands, that took place before the practice of pediatric gender transition became widespread (Steensma, van der Ende, et al., 2013), also sheds light on what happens when childhood and adolescent gender-variance remains medically untreated. This large prospective longitudinal study based on a community sample (n = 879) found that about 6% of children (n = 51) ages 7–8 in a community sample were identified as “gender variant.” At follow-up 24 years later, when the subjects were on average in their early 30s, *not a single individual* from the previously “gender-variant” subgroup of 51 children sought to undergo gender reassignment, despite the availability of these services.

There are three noteworthy observations in this study. *First*, the rate of “gender variance” of 6% reported in the community sample is remarkably similar to the current rate of transgender identification in U.S. youth of 2–9% (Johns et al., 2019; Kidd et al. 2021). *Second*, the gender-variant children were roughly 8–15 times more likely to grow up to be gay, lesbian, or bisexual adults compared to gender-normative youth. Gender variance is a common precursor to future homosexuality (Korte et al., 2008) and in fact in the Dutch studies, 97% of youth were gay, lesbian, or bisexual relative to their natal sex (de Vries et al., 2011). *Third*, only *one* of the 879 individuals in the sample underwent a male-to-female gender reassignment as an adult—and the individual had *not* been deemed “gender-variant” as a child (Steensma, van der Ende, et al., 2013, p. 2729). This challenges the current focus on medical interventions at increasingly younger ages.

The fact that none of the “gender variant” children in the sample sought gender reassignment as adults, when the study was published in 2013, merits scrutiny. These children would have been coming “of age” just a few years before the Dutch researchers conceived of the notion of *juvenile transsexual* and began to offer gender reassignment to adolescents. Thus, these children just missed the clinical shift in the Dutch practice—and perhaps not coincidentally, apparently all avoided the lifelong medical burden of living as a gender-reassigned individual.

The title of de Vries’ commentary, *Ensuring Care for Transgender Adolescents Who Need It* (de Vries, 2022) prompts us to pose two questions. First, has the availability of the Dutch protocol itself created the “need?” Second, absent clear criteria to separate a young person’s “wish” from a “need,” will research rigor be required to demonstrate that the benefits outweigh the risks?

II. Newer research claiming benefits of youth gender transition is even more flawed

de Vries acknowledged that the Dutch research suffers from some limitations but insisted that newer research has sufficiently addressed these problems. She criticized us for not including a

review of newer studies that "consistently demonstrate improved or stable psychological functioning, body image, or treatment satisfaction varying from three months to up to two years from the initiation of treatment" (de Vries, 2022, p. 5). We are familiar with the seven studies de Vries mentions—as well as a number of other recent studies. What these studies "consistently demonstrate" is the art of *spin*—a well-documented problem in biomedical research where researchers "distort the interpretation of results and mislead readers so that results are viewed in a more favorable light" (Chiu et al., 2017). Due to length concerns, we discuss only three examples— Carmichael et al. (2021), Costa et al. (2015), and Tordoff et al. (2022). Most of the current research on the purported benefits of "gender-affirming care" suffers from similar limitations.

The UK study of puberty blockers by Carmichael et al. (2021), which attempted to replicate the Dutch puberty blocker study's findings of psychological improvements (de Vries et al., 2011), *failed to demonstrate psychological improvements*, conceding that its results are "in contrast to the Dutch study" (Carmichael et al., 2021, p. 19). The study found problems in bone mass density accrual among puberty-blocked youth. These problematic findings take on a decisively positive spin in the study conclusions, which refocus the reader on the positive "overall patient experience of changes on GnRHa treatment"; dismiss bone density problems as merely "consistent with suppression of growth"; and camouflage the failure to replicate the psychological benefits of puberty suppression by simply stating, "we identified no changes in psychological function" (Carmichael et al., 2021, p. 2). de Vries aided in the positive interpretation of the results by recasting the lack of improvement in psychological function following puberty suppression, as a *positive* finding of "stable psychological function" (de Vries 2022, p. 5)—yet it has never been demonstrated that psychological function of gender dysphoric adolescents with high baseline mental health function, as was required by the study criteria, would be expected to deteriorate absent intervention.

Spin also characterizes Costa et al. (2015), which compared psychosocial functioning of gender dysphoric youth who were puberty-suppressed to those who were delayed for medical treatment and received only psychotherapy. By the end of the 18-month study period, both groups ended up in the same psychosocial functional range using the Children's Global Assessment Scale (CGAS): 61–70 (out of 100 points), corresponding to "[s]ome difficulty in a single area, but generally functioning pretty well" (Shaffer, 1983). This study can hardly be cited as evidence of the superiority of the medical approach and in fact points to the viability of providing noninvasive therapy as an alternative to puberty suppression. Yet, the authors focus their abstract on the fact that the puberty-blocked group had higher function after puberty suppression than before, ignoring the fact that both the puberty-suppressed and the psychologically-treated only groups improved and there was no statistically-significant difference betwen the two by the end of the study period (Biggs, 2019). Questions regarding the extent to which improvements in self-reported psychological measures could be due to the placebo effect of puberty blockers have been recently raised (Clayton, 2022).

The spin of Tordoff et al. (2022) is dramatic. This study claimed that puberty blockers and "gender-affirming" hormones produced a 60% reduction in depression after only one year. However, this conclusion is in stark contrast to the raw data: at baseline, 59% of the yet-to-be treated patients had *moderate to severe depression*; by the end of the study at 12 months, 56% were still moderately to severely depressed, despite receiving hormone treatment (Supplementary material of eTable 3 Tordoff et al., 2022). This unchanged rate of depression became an "observed 60% lower odds of depression" via a methodology that *inferred* the "improvement" in the *treated cases* from the reported "worsening" in the *untreated cases*. Indeed, the untreated cases in the study had depression rates of 86% by the end of the study period (n = 7), compared to 56% of the treated cases (n = 57), seemingly supporting the conclusion that treatment with hormones alleviates depression.

However, by basing their conclusion about the relative success of the "treated" on the finding of lack of success among the "untreated" cases, the researchers failed to consider that

they lost an astounding 80% of their "untreated" cohort by the end of the study (28 of 35); in contrast, over 80% of the "treated" cohort (57 of 69) remained enrolled. The high dropout rate in "untreated" subjects makes intuitive sense: the study took place in a gender clinic setting, the primary purpose of which is provision of gender transition services. Youth whose distress was ameliorated without the use of hormones would have little reason to stay enrolled in the clinic and participate in the ongoing research. However, what this also suggests is that the highest functioning "untreated" youth dropped out of the study. Thus, the entire conclusion that because "untreated" cases faired so poorly on measures of depression, anxiety, or suicidality, it must be that hormones given to the "treated" cases "worked," is invalid. There are other problems in the study, including the fact that the use of psychiatric medications was not accounted for in the analysis. The university was aware of the problems with this research but chose to remain silent because the study's optimistic conclusions were so well received by national news media outlets (Rantz, 2022).

These examples demonstrate why we do not share de Vries' optimism that the newer studies conducted since the publication of the two seminal Dutch studies provide any additional confidence in, or support for, the practice of youth gender transitions. Most of the current research into the practice of pediatric transition continues in the context of gender clinic settings, which are actively providing gender transition to willing youth. Such low-quality observational research not only lacks the ability to control for the multiple sources of bias due to limitations in research design, but also is often led by clinicians with vested intellectual, professional, and financial conflicts of interest (Prasad, 2013).

III. Suggestions for future research

We were pleased to learn that de Vries has been awarded a substantial research grant to continue to study the effects of the Dutch protocol (Amsterdam UMC, 2022a). We welcome her decision to study the effects of the Dutch protocol on the novel cohort of youth whose trans identity only emerged in adolescence, as we agree that it is important to know "whether medical treatment is …useful for this group or whether there are too many risks… such as regret afterwards" (Amsterdam UMC, 2022b).

However, we think the time has come to reexamine the entire 25 years of Dutch experience using rigorous methodologies, to answer the critical questions about the full range of risks and benefits of the Dutch protocol. We offer five suggestions relating to both past and future research:

1. Conduct comprehensive retrospective research

There have been over 6600 referrals to the Amsterdam gender clinic alone between 2000 and 2019 (Steensma et al., 2022), with likely additional referrals to the other Dutch gender clinics over the same time period, as well as new referrals since 2019. A retrospective chart review of these referred patients, supplemented by the data from the Dutch health and civil records registries (Registers in The Netherlands 2022) could allow researchers to reexamine its quarter-century of experience of gender transition of youth and their outcomes in a way that is methodologically sound. The analysis should include outcomes of *all* patients diagnosed with gender dysphoria as children, adolescents, or young adults, rather than focusing only on those who chose to pursue medical interventions and explicitly agreed to participate in research. This retrospective review should seek to examine the outcomes of medical transition, psychotherapy, and no intervention. The effects of each step of the Dutch protocol should be disaggregated to gain a better understanding of the benefits and risks at each stage, and the results should be analyzed by natal sex and the age of gender dysphoria onset as validated by medical records.

2. Focus on comparative outcomes

The importance of *comparative* research to determine optimal treatments has been known since the 1990s (Guyatt, 1993). Comparing "before" and "after" psychological outcomes tends to overstate benefits due to number of factors, including "regression to the mean" (Knapp, 2016). Gender dysphoric youth often seek help at the peak of their distress. That many such "extreme" situations tend to naturally revert to a milder state even without an intervention is a well-recognized clinical and statistical phenomenon. While randomization is still the gold standard to reliably estimate treatment effects, when it is not possible (as is the case with retrospective research), researchers should consider utilizing quasi-experimental research designs (Harris et al., 2006). Recent post-hoc analysis of the effects of "gender-affirming" surgery, which utilized propensity-score matching to construct comparator groups, is an example of such analysis (Bränström & Pachankis, 2020c).

3. Track a full range of health outcomes utilizing objective measures whenever possible

The current exclusive focus on psychological and sexual functioning and self-reports is insufficient. Research should include a more objective evaluation of the effects of gender reassignment interventions on bone, brain, cardiovascular health, malignancies, and overall morbidity and all-cause mortality. As mentioned earlier, retrospective chart reviews of the referred patient cohorts, supplemented with relevant data from the Dutch health and civil records registries, should provide sufficient information to estimate the longer-term impact of hormonal and surgical interventions on morbidity and mortality, while also documenting the incidence of osteoporosis, cardiovascular disease, and cancer, as well as rates of mental illness and suicidality/suicide.

4. Pre-specify primary and secondary outcome measures and consistently track them

The primary outcomes of pediatric gender reassignment have been a moving target. In 1997, the Dutch researchers stated that the decision to start gender transition had as its goal to improve the "psychological problems of untreated adolescents" (Delemarre-van de Waal & Cohen-Kettenis, 2006, p. 132), since transitions undertaken in adulthood were already adequately relieving the feeling of gender incongruence itself. In her commentary, however, de Vries stated that psychological function may not the "best indicator for the benefits of such treatment" and that "measures that assess what makes life most worth living…" are most appropriate (de Vries, 2022, p. 3). Yet in a recent interview, she stated that the best indicator of treatment benefits is "satisfaction with care" (O'Malley & Ayad, 2022, 54:36). Primary outcome measures that serve as the rationale for the intervention must be clearly stated, justified, and consistently tracked.

If relief of "gender dysphoria" is still considered a primary outcome by the Dutch research team, a new measure of gender dysphoria that can be validated in both the pre- and the post-treatment settings is urgently needed, as the UGDS scale's use post-treatment is invalid. The updated UGDS-GS scale (McGuire et al., 2020) currently favored by de Vries (de Vries, 2022), appears to be a derivative of the earlier UGDS scale, and therefore may suffer from similar limitations when used in post-gender-reassignment settings.

5. Focus on long-term outcomes

Until recently, the long-term outcomes on the cohort of 70/55 cases have been an unanswered question. It was partially answered in a recent WPATH Symposium presentation by the Dutch team, comprised of presentations by Drs. de Rooy, Asseler, van der Meulen, van der Miesen, and Steensma (Steensma et al., 2022). As we look forward to seeing these preliminary findings elucidated in the upcoming peer-reviewed publications, we note several concerns.

First, it appears that the follow-up research combined the earlier-treated cohorts with the later-treated ones. We hope to see the outcomes of the 70/55 cases reported separately from other cases, so that the original cohort's outcomes can be quantified. *Second*, only half of the treated cases engaged in follow-up research (Bazelon, 2022; Steensma et al., 2022). This can bias the results, as individuals who experience more difficulties with their gender transition are less likely to engage with the physicians who treated them (Vandenbussche, 2022). Much follow-up research that reports positive outcomes relies on self-reported data compromised by high dropout rates (D'Angelo, 2018). In contrast, research that utilizes medical records and objective outcome measures shows much less optimistic outcomes (Dhejne et al., 2011; Bränström & Pachankis, 2020a, 2020b, 2020c). To mitigate the non-response bias, the Dutch research team should leverage chart data for all the referred patients, and report objective health outcomes for the *entire cohort* that was treated.

Third, we are concerned by the apparent dismissal of reproductive regret, which affected more than a quarter of the patients (according to the data presented by Asseler), as merely a problem of the past when sterilizing surgery was a requirement (Steensma et al., 2022). The current treatment protocol of blocking puberty at Tanner stage 2 followed by cross-sexhormones, endorsed by the Endocrine Society (Hembree et al., 2017) and WPATH (Coleman et al., 2022), will most likely lead to chemical sterility, just as the prior surgical protocol led to permanent surgically-induced sterility. There are currently no effective, established methods to preserve fertility of individuals whose gametes have not matured (Rosenthal, 2021).

Fourth, the reported relationship difficulties reported by Asseler, with over 60% of individuals in their early to mid-30's still single, also deserve serious consideration. The apparent sexual difficulties reported by male-to-female transitioners by van der Meulen (around 70% have problems with libido, have pain during sex, or have problems with achieving orgasm), combined with reproductive challenges, may be contributing to this outcome. *Fifth*, the team's preliminary optimistic conclusions that early puberty blockade did not worsen sexual function appears to be based on a problematic combining Tanner stages 2 and 3. The development of sexual organs and fertility is significantly more advanced in Tanner stage 3, compared to stage 2. Whether or not the high rate of sexual problems found in the transitioned population may be related to blocking puberty at Tanner stage 2 needs to be investigated.

These newly reported data underscore an urgent need to determine whether the benefits of medical interventions outweigh the now much better understood risks.

Concluding thoughts

The question, "Just because we can, should we?" is not unique to pediatric gender medicine. What makes this arena exceptional is the radical, irreversible nature of "gender-affirming" medical and surgical interventions desired by the exponentially growing numbers of youth in the Western world. The recent changes announced by WPATH SOC 8—specifically the removal of minimum age limits for medical and surgical treatments, and the elimination of the "distress" requirement by switching from DSM-5-TR to ICD-11 diagnostic criteria (Coleman et al., 2022; Robles García & Ayuso-Mateos, 2019; World Health Organization, 2019)—takes the field further in a truly extraordinary direction whereby *any desired body modification* desired by a child or a young person becomes automatically "medically necessary."

Another unique aspect of the gender medicine field is that a number of clinicians tasked with caring for gender-distressed have taken on the role of political campaigners—and in doing so, have traded wisdom and nuance for blunt activism (Kuper et al., 2022; McNamara et al., 2022). Their insistence that today's gender-dysphoric teens are tomorrow's transgender adults, and that their future happiness and mere survival hinges on early access to gender reassignment, is demonstrably false. While still reported as "rare" by the gender medicine establishment (Coleman et al., 2022; McNamara et al., 2022), the rate of medical detransition is already 10%-30% just a few years following transition (Boyd et al., 2022; Hall et al., 2021; Roberts et al., 2022). These

numbers are likely to rise in the future as regret historically has taken over a decade to materialize (Dhejne et al., 2014). Not all of those who detransitioned will consider themselves harmed, but many will—and a number already have (Vandenbussche, 2022; Littman, 2021).

When clinician-activists misuse the eminence of their institutions and medical societies to deny or obfuscate important facts about pediatric gender transition—that puberty blockers are prescribed to peri-pubertal children as young as 8–9; that mastectomies are commonly provided to teens; that the wave of detransition is rising and already far exceeds what's been historically recorded; and that no other pediatric intervention of similarly drastic nature has ever been delivered at scale based such low quality of evidence (McNamara et al., 2022)—they may succeed in scoring a political or legal "victory" in the short-term, but they also contribute to the longer-term erosion of public trust in the medical profession. They also inadvertently contribute to medical harm.

The scale of the potential harm can be fully appreciated if one considers that an astounding 1 in 10–20 middle school, high school, and college students in the West currently claim a transgender identity (ACHA, 2022; Johns et al., 2019; Kidd et al. 2021). Adolescent mental health in general is at an all-time low (Centers for Disease Control and Prevention [CDC], 2022). Lesbian, gay and bisexual youth and those on the autism spectrum (Bradley, 2022) are at particularly high risk of refracting their gender-non-conformity through the prism of transgender identity. Youth referrals for gender reassignment have risen already several thousand percent in the last decade, and nearly doubled between 2020/2021 and 2021/2022 (NHS, 2022b; Respaut & Terhune, 2022). If these young patients' sense of urgency is confused with certainty about their future happiness, while a flawed evidence base is mistaken for proven safety and effectiveness of youth gender reassignment, harm at scale will ensue.

As physicians are increasingly instructed to widely adopt "gender identity screening" of adolescents to "facilitate and increase…the delivery of gender-affirming" interventions (Lau et al., 2021, p. 1) and are misled about the (very low) quality of research, an analogy of the opioid epidemic powerfully emerges. The gender medicine field must reflect on the parallels between the pain as the "fifth vital sign," the misuse of research (Porter & Jick, 1980; Zhang, 2017), the pressure to meet patient demands, and the role of powerful special interests during the height of the opioid epidemic—and the trends in pediatric gender medicine today.

The field of gender medicine has a short time to self-correct before a growing number of authorities step in and impose guardrails to safeguard youth. Public health authorities in Finland, Sweden, and most recently England have already done just that, sharply deviating from the WPATH's poorly evidenced recommendations in "SOC 7" (Dahlen et al., 2021), with no apparent intention to follow the updated "SOC 8" either (COHERE (Council for Choices in Health Care), 2020; Socialstyrelsen [National Board of Health and Welfare], 2022; NHS, 2022a). NHS England's decision to close GIDS/Tavistock—the world's biggest pediatric gender clinic—and to place the care of gender-distressed youth in established clinical settings that "maintain a broad clinical perspective," provide "strong links to mental health services," and do not "exceptionalise gender identity issues," (Cass, 2022; NHS, 2022b) is a vote of no-confidence in the WPATH-endorsed "gender-affirming" approach that dominates the "gender clinic" model of care.

The American medical establishment appears to be taking a different approach. Rather than acknowledging the problems with the gender-affirmation model of care, there is an apparent effort underway to retrospectively redefine what "gender-affirmation" is. Originally defined as comprised of the provision of hormones and surgery to youth (Table 2, Rafferty, 2018), more recently gender affirmation has been positioned as merely "holistic care." The American Academy of Pediatrics recently made a surprising and welcome statement that hormones and surgery are not the preferred treatment for gender dysphoric youth, and that in fact "for the vast majority of children, it recommends the opposite" (Szilagyi, 2022). Whether this statement will be followed by earnest efforts to restrict the provision of highly invasive interventions to exceptional situations and to endorse non-invasive psychosocial interventions as first line of treatment—instead of inappropriately conflating psychotherapy for gender dysphoria with "conversion"—remains to be seen.

The former era of eminence-based, expert-opinion-led medicine, under which the innovative clinical practice of pediatric gender transition proliferated, has been replaced by a new standard, *evidence-based medicine*, which demands rigor in the research that underpins population-level treatment recommendations (Sackett et al., 1996; Zimerman, 2013). Our analysis of the Dutch protocol has been written with three goals in mind. *First*, we wanted to definitively refute the claims that the foundational Dutch research represents "solid prospective research" that provides reliable evidence of net benefits of youth gender transition. In fact, it is much better described as case series—one of the lowest levels of evidence available (Dekkers et al., 2012, Mathes & Pieper, 2017). *Second*, we aimed to demonstrate that the type of non-comparative, short-term research that the gender medicine establishment continues to pursue is incapable of generating reliable information. And *third and most importantly*, we wanted to remind the medical community that medicine is a double-edged sword capable of both much good and much harm. The burden of proof—demonstrating that a treatment does more good than harm—is *on those promoting the intervention*, not on those concerned about the harms. Until gender medicine commits to conducting high quality research capable of reliably demonstrating the preponderance of benefits over harms of these invasive interventions, we must be skeptical of the enthusiasm generated by headlines claiming that yet another "gender study" proved benefits of transitioning youth. This time-honored concern about risk/benefit ratio is a sobering reminder that the history of medicine is replete with examples of "cures" which turned out to far more harmful than the "disease."

Notes

1. de Vries also served as a peer-reviewer of our original paper, Levine et al. (2022a).
2. While not central to our argument, de Vries' claim that the selection of the 111 participants from the original 196 was based only on the researchers' interest in those age 16 and under is contradicted by the data. According to Table 1 in de Vries et al. (2011), there was at least one natal female participant who was 18.6 years old when the puberty blockers were initiated. Although selection criteria of the 111 from 196 may have introduced additional bias, we are most concerned with bias in the subsequent selection of 70 from the 111.

Acknowledgements

The Society for Evidence-Based Gender Medicine (SEGM) paid for this publication to have Open Access.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

References

Achenbach, T. M., & Rescorla, L. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, & Families.

Aitken, M., Steensma, T. D., Blanchard, R., VanderLaan, D. P., Wood, H., Fuentes, A., Spegg, C., Wasserman, L., Ames, M., Fitzsimmons, C. L., Leef, J. H., Lishak, V., Reim, E., Takagi, A., Vinik, J., Wreford, J., Cohen-Kettenis, P. T., de Vries, A. L. C., Kreukels, B. P. C., & Zucker, K. J. (2015). Evidence for an altered sex ratio in clinic-referred adolescents with gender dysphoria. *The Journal of Sexual Medicine*, *12*(3), 756–763. doi:10.1111/jsm.12817

American College Health Association (ACHA). (2022). National College Health Assessment III. Undergraduate Student Reference Group. Data Report. Spring 2022.

American Medical Association (AMA). (2022). *AMA reinforces opposition to restrictions on transgender medical care*. Retrieved November 15, 2022, from https://www.ama-assn.org/press-center/press-releases/ama-reinforces-opposition-restrictions-transgender-medical-care

American Psychiatric Association (APA). (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). doi:10.1176/appi.books.9780890425787

Amsterdam UMC. (2022a, July 1). *Three Vidi grants for Amsterdam Public Health researchers*. Retrieved September 3, 2022 from https://www.amsterdamumc.org/en/research/institutes/amsterdam-public-health/news/three-vidi-grants-for-amsterdam-public-health-researchers.htm

Amsterdam UMC. (2022b, July 7). *Three Vidi grants for Amsterdam Public Health researchers*. Retrieved September 3, 2022 from https://www.amsterdamumc.org/nl/vandaag/800.000-euro-voor-onderzoek-naar-transgenderzorg-jongeren.htm

Arnoldussen, M., Steensma, T. D., Popma, A., van der Miesen, A. I. R., Twisk, J. W. R., & de Vries, A. L. C. (2020). Re-evaluation of the Dutch approach: Are recently referred transgender youth different compared to earlier referrals? *European Child & Adolescent Psychiatry*, *29*(6), 803–811. doi:10.1007/s00787-019-01394-6

Balon R. (2022). Commentary on Levine et al.: Festina Lente (Rush Slowly). *Journal of Sex & Marital Therapy*, *48*(8), 775–778. doi:10.1080/0092623X.2022.2055686

Bazelon, E. (2022, June 15). The battle over gender therapy. The New York Times. https://www.nytimes.com/2022/06/15/magazine/gender-therapy.html

Becerra-Culqui, T. A., Liu, Y., Nash, R., Cromwell, L., Flanders, W. D., Getahun, D., Giammattei, S. V., Hunkeler, E. M., Lash, T. L., Millman, A., Quinn, V. P., Robinson, B., Roblin, D., Sandberg, D. E., Silverberg, M. J., Tangpricha, V., & Goodman, M. (2018). Mental health of transgender and gender nonconforming youth compared with their peers. *Pediatrics*, *141*(5), e20173845. doi:10.1542/peds.2017-3845\

Biggs, M. (2019). A Letter to the Editor Regarding the Original Article by Costa et al.: Psychological support, puberty suppression, and psychosocial functioning in adolescents with gender dysphoria. *The Journal of Sexual Medicine*, *16*(12), 2043. doi:10.1016/j.jsxm.2019.09.002

Biggs, M. (2020). Gender dysphoria and psychological functioning in adolescents treated with GnRHa: Comparing Dutch and English prospective studies. *Archives of Sexual Behavior*, *49*(7), 2231–2236. doi:10.1007/s10508-020-01764-1

Biggs, M. (2021). Revisiting the effect of GnRH analogue treatment on bone mineral density in young adolescents with gender dysphoria. *Journal of Pediatric Endocrinology and Metabolism*, *34*(7), 937–939. doi:10.1515/jpem-2021-0180

Biggs, M. (2022). The Dutch protocol for Juvenile transsexuals: Origins and evidence. *Journal of Sex & Marital Therapy*. Advance online publication. doi:10.1080/0092623X.2022.2121238

Boyd, I. L., Hackett, T., & Bewley, S. (2022). Care of transgender patients: A general practice quality improvement approach. *Healthcare*, *10*(1):121. doi:10.3390/healthcare10010121

Boutron I, Page MJ, Higgins JPT, Altman DG, Lundh A, & Hróbjartsson A. (2022). Chapter 7: Considering bias and conflicts of interest among the included studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, & Welch VA (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions version 6.3*. London: Cochrane. (updated February 2022). Available from www.training.cochrane.org/handbook.

Bradley, S. J. (2022). Understanding vulnerability in girls and young women with high-functioning autism spectrum disorder. *Women*, *2*(1), 64–67. doi:10.3390/women2010007

Bränström, R., & Pachankis, J. E. (2020a). Reduction in mental health treatment utilization among transgender individuals after gender-affirming surgeries: A total population study. *American Journal of Psychiatry*, *177*(8), 727–734. doi:10.1176/appi.ajp.2019.19010080

Bränström, R., & Pachankis, J. E. (2020b). Correction to Bränström and Pachankis. (2020). *American Journal of Psychiatry*, *177*(8), 734–734. doi:https://doi.org/10.1176/appi.ajp.2020.1778correction

Bränström, R., & Pachankis, J. E. (2020c). Toward rigorous methodologies for strengthening causal inference in the association between gender-affirming care and transgender individuals' mental health: Response to letters. *American Journal of Psychiatry*, *177*(8), 769–772. doi:10.1176/appi.ajp.2020.20050599

Brierley, J., & Larcher, V. (2009). Compassionate and innovative treatments in children: A proposal for an ethical framework. *Archives of Disease in Childhood*, *94*(9), 651–654. doi:10.1136/adc.2008.155317

Brignardello-Peterson, R., & Wiercioch, W. (2022). *Effects of gender affirming therapies in people with gender dysphoria: Evaluation of the best available evidence*. https://ahca.myflorida.com/letkidsbekids/docs/AHCA_GAPMS_June_2022_Attachment_C.pdf

Cantor, J. M. (2020). Transgender and gender diverse children and adolescents: Fact-checking of AAP policy. *Journal of Sex & Marital Therapy*, *46*(4), 307–313. doi:10.1080/0092623X.2019.1698481

Carmichael, P., Butler, G., Masic, U., Cole, T. J., De Stavola, B. L., Davidson, S., Skageberg, E. M., Khadr, S., & Viner, R. M. (2021). Short-term outcomes of pubertal suppression in a selected cohort of 12 to 15 year old young people with persistent gender dysphoria in the UK. *PLOS ONE*, *16*(2), e0243894. doi:10.1371/journal.pone.0243894

Cass, H. (2022). *Entry 8—Beyond the headlines*. Retrieved August 19, 2022, from https://cass.independent-review.uk/entry-8-beyond-the-headlines/

Centers for Disease Control and Prevention (CDC). (2022). *New CDC data illuminate youth mental health threats during the COVID-19 pandemic*. https://www.cdc.gov/media/releases/2022/p0331-youth-mental-health-covid-19.html

Chiu, K., Grundy, Q., & Bero, L. (2017). 'Spin' in published biomedical literature: A methodological systematic review. *PLOS Biology*, *15*(9), e2002173. doi:10.1371/journal.pbio.2002173

Clayton, A. (2022). Gender-affirming treatment of gender dysphoria in youth: A perfect storm environment for the placebo effect—the Implications for research and clinical practice. *Archives of Sexual Behavior. Online ahead of print*. doi:10.1007/s10508-022-02472-8

Cohen-Kettenis, P. T., & van Goozen, S. H. M. (1997). Sex reassignment of adolescent transsexuals: A follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(2), 263–271. doi:10.1097/00004583-199702000-00017

Cohen-Kettenis, P. T., & van Goozen, S. H. M. (1998). Pubertal delay as an aid in diagnosis and treatment of a transsexual adolescent. *European Child & Adolescent Psychiatry*, *7*(4), 246–248. doi:10.1007/s007870050073

COHERE (Council for Choices in Health Care). (2020). *Palveluvalikoimaneuvoston Suositus: Alaikäisten Sukupuoli-identiteetin Variaatioihin Liittyvän Dysforian Lääketieteelliset Hoitomenetelmät. [Recommendation of the Council for Choices in Health Care in Finland: Medical Treatment Methods for Dysphoria Related to Gender Variance in Minors.]*. https://segm.org/Finland_deviates_from_WPATH_prioritizing_psychotherapy_no_surgery_for_minors

Coleman, E., Bockting, W., Botzer, M., Cohen-Kettenis, P., DeCuypere, G., Feldman, J., Fraser, L., Green, J., Knudson, G., Meyer, W. J., Monstrey, S., Adler, R. K., Brown, G. R., Devor, A. H., Ehrbar, R., Ettner, R., Eyler, E., Garofalo, R., Karasic, D. H., … Zucker, K. (2012). Standards of care for the health of transsexual, transgender, and gender-nonconforming people, Version 7. *International Journal of Transgenderism*, *13*(4), 165–232. doi:10.1080/15532739.2011.700873

Coleman, E., Radix, A. E., Bouman, W. P., Brown, G. R., de Vries, A. L. C., Deutsch, M. B., Ettner, R., Fraser, L., Goodman, M., Green, J., Hancock, A. B., Johnson, T. W., Karasic, D. H., Knudson, G. A., Leibowitz, S. F., Meyer-Bahlburg, H. F. L., Monstrey, S. J., Motmans, J., Nahata, L., … Arcelus, J. (2022). Standards of care for the health of transgender and gender diverse people, Version 8. *International Journal of Transgender Health*, *23*(sup1), S1–S259. doi:10.1080/26895269.2022.2100644

Costa, R., Dunsford, M., Skagerberg, E., Holt, V., Carmichael, P., & Colizzi, M. (2015). Psychological support, puberty suppression, and psychosocial functioning in adolescents with gender dysphoria. *The Journal of Sexual Medicine*, *12*(11), 2206–2214. doi:10.1111/jsm.13034

Dahlen, S., Connolly, D., Arif, I., Junejo, M. H., Bewley, S., & Meads, C. (2021). International clinical practice guidelines for gender minority/trans people: Systematic review and quality assessment. *BMJ Open*, *11*(4), e048943. doi:10.1136/bmjopen-2021-048943

D'Angelo, R. (2018). Psychiatry's ethical involvement in gender-affirming care. *Australasian Psychiatry*, *26*(5), 460–463. doi:10.1177/1039856218775216

Dekkers, O. M., Egger, M., Altman, D. G., & Vandenbroucke, J. P. (2012). Distinguishing case series from cohort studies. *Annals of Internal Medicine*, *156*(1_Part_1), 37. doi:10.7326/0003-4819-156-1-201201030-00006

Delemarre-van de Waal, H. A., & Cohen-Kettenis, P. T. (2006). Clinical management of gender identity disorder in adolescents: A protocol on psychological and paediatric endocrinology aspects. *European Journal of Endocrinology*, *155*(suppl_1), S131–S137. doi:10.1530/eje.1.02231

de Graaf, N. M., Giovanardi, G., Zitz, C., & Carmichael, P. (2018). Sex ratio in children and adolescents referred to the gender identity development service in the UK (2009–2016). *Archives of Sexual Behavior*, *47*(5), 1301–1304. doi:10.1007/s10508-018-1204-9

de Graaf, N. M., Huisman, B., Cohen-Kettenis, P. T., Twist, J., Hage, K., Carmichael, P., Kreukels, B. P. C., & Steensma, T. D. (2021). Psychological functioning in non-binary identifying adolescents and adults. *Journal of Sex & Marital Therapy*, *47*(8), 773–784. doi:10.1080/0092623X.2021.1950087

de Vries, A. L. C., & Cohen-Kettenis, P. T. (2012). Clinical management of gender dysphoria in children and adolescents: The Dutch approach. *Journal of Homosexuality*, *59*(3), 301–320. doi:10.1080/00918369.2012.653300

de Vries, A. L. C. (2020). Challenges in timing puberty suppression for gender-nonconforming adolescents. *Pediatrics*, *146*(4), e2020010611. doi:10.1542/peds.2020-010611

de Vries, A. L. C. (2022). Ensuring care for transgender adolescents who need it: Response to 'Reconsidering Informed Consent for Trans-Identified Children, Adolescents and Young Adults'. *Journal of Sex & Marital Therapy*. Advance online publication. doi:10.1080/0092623X.2022.2084479

de Vries, A. L. C., Cohen-Kettenis, P. T., & Delemarre-van de Waal, H. (2006). Clinical Management of gender dysphoria in adolescents. *International Journal of Transgenderism*, *9*(3–4), 83–94. doi:10.1300/J485v09n03_04

de Vries, A. L. C., McGuire, J. K., Steensma, T. D., Wagenaar, E. C. F., Doreleijers, T. A. H., & Cohen-Kettenis, P. T. (2014). Young adult psychological outcome after puberty suppression and gender reassignment. *Pediatrics*, *134*(4), 696–704. doi:10.1542/peds.2013-2958

de Vries, A. L. C., Steensma, T. D., Doreleijers, T. A. H., & Cohen-Kettenis, P. T. (2011). Puberty suppression in adolescents with gender identity disorder: A prospective follow-up study. *The Journal of Sexual Medicine*, *8*(8), 2276–2283. doi:10.1111/j.1743-6109.2010.01943.x

Dhejne, C., Lichtenstein, P., Boman, M., Johansson, A. L. V., Långström, N., & Landén, M. (2011). Long-term follow-up of transsexual persons undergoing sex reassignment surgery: Cohort study in Sweden. *PLoS ONE*, *6*(2), e16885. doi:10.1371/journal.pone.0016885

Dhejne, C., Öberg, K., Arver, S., & Landén, M. (2014). An analysis of all applications for sex reassignment surgery in Sweden, 1960–2010: Prevalence, incidence, and regrets. *Archives of Sexual Behavior*, *43*(8), 1535–1545. doi:10.1007/s10508-014-0300-8

Drescher, J. (2022). Informed consent or scare tactics? A response to Levine et al.'s "Reconsidering Informed Consent for Trans-Identified Children, Adolescents, and Young Adults." *Journal of Sex & Marital Therapy*. Advance online publication. doi:10.1080/0092623X.2022.2080780

Drisko, J. W., & Friedman, A. (2019). Let's clearly distinguish evidence-based practice and empirically supported treatments. *Smith College Studies in Social Work*, *89*(3–4), 264–281. doi:10.1080/00377317.2019.1706316

Earl, J. (2019). Innovative practice, clinical research, and the ethical advancement of medicine. *The American Journal of Bioethics*, *19*(6), 7–18. doi:10.1080/15265161.2019.1602175

Freedman, B. (1987). Equipoise and the ethics of clinical research. *New England Journal of Medicine*, *317*(3), 141–145. doi:10.1056/NEJM198707163170304

Gooren, L., & Delemarre-van de Waal, H. (1996). The feasibility of endocrine interventions in Juvenile transsexuals. *Journal of Psychology & Human Sexuality*, *8*(4), 69–74. doi:10.1300/J056v08n04_05

Green, A. E., DeChants, J. P., Price, M. N., & Davis, C. K. (2022). Association of gender-affirming hormone therapy with depression, thoughts of suicide, and attempted suicide among transgender and nonbinary youth. *Journal of Adolescent Health*, *70*(4), 643–649. doi:10.1016/j.jadohealth.2021.10.036

Guyatt, G. H. (1993). Users' guides to the medical literature: II. How to use an article about therapy or prevention A. Are the results of the study valid? *JAMA*, *270*(21), 2598. doi:10.1001/jama.1993.03510210084032

Hall, R., Mitchell, L., & Sachdeva, J. (2021). Access to care and frequency of detransition among a cohort discharged by a UK national adult gender identity clinic: Retrospective case-note review. *BJPsych Open*, *7*(6), e184. doi:10.1192/bjo.2021.1022

Harris, A. D., McGregor, J. C., Perencevich, E. N., Furuno, J. P., Zhu, J., Peterson, D. E., & Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in medical informatics. *Journal of the American Medical Informatics Association*, *13*(1), 16–23. doi:10.1197/jamia.M1749

Health and Human Services (HHS). (2022). Nondiscrimination in health programs and activities—Proposed rule. 87 FR 47824, pp. 47824–47920. Docket Number HHS-OS-2022-0012. Docket RIN 0945-AA17 https://www.federalregister.gov/documents/2022/08/04/2022-16217/nondiscrimination-in-health-programs-and-activities

Hembree, W. C., Cohen-Kettenis, P. T., Gooren, L., Hannema, S. E., Meyer, W. J., Murad, M. H., Rosenthal, S. M., Safer, J. D., Tangpricha, V., & T'Sjoen, G. G. (2017). Endocrine treatment of gender-dysphoric/gender-incongruent persons: An endocrine society* clinical practice guideline. *The Journal of Clinical Endocrinology & Metabolism*, *102*(11), 3869–3903. doi:10.1210/jc.2017-01658

Herrera-Perez, D., Haslam, A., Crain, T., Gill, J., Livingston, C., Kaestner, V., Hayes, M., Morgan, D., Cifu, A. S., & Prasad, V. (2019). A comprehensive review of randomized clinical trials in three medical journals reveals 396 medical reversals. *ELife*, *8*, e45183. doi:10.7554/eLife.45183

Higgins JPT., Thomas J., Chandler J., Cumpston M., Li T, Page MJ., Welch VA (Eds). (2022). Cochrane Handbook for Systematic Reviews of Interventions *version 6.3*. London: Cochrane. (updated February 2022). Available from www.training.cochrane.org/handbook

Hutchinson, A., Midgen, M., & Spiliadis, A. (2020). In support of research into rapid-onset gender dysphoria. *Archives of Sexual Behavior*, *49*(1), 79–80. doi:10.1007/s10508-019-01517-9

Iliadis, S. I., Axfors, C., Friberg, A., Arinell, H., Beckman, U., Fazekas, A., Frisen, L., Sandström, L., Thelin, N., Wahlberg, J., Södersten, M., & Papadopoulos, F. C. (2020). Psychometric properties and concurrent validity of the Transgender Congruence Scale (TCS) in the Swedish setting. *Scientific Reports*, *10*(1), 18701. doi:10.1038/s41598-020-73663-3

Janssen, A. (2022). 10/28/22 Florida boards of medicine and osteopathic medicine joint rules/legislative committee rule workshop. https://thefloridachannel.org/videos/10-28-22-florida-boards-of-medicine-and-osteopathic-medicine-joint-rules-legislative-committee-rule-workshop/ Accessed November 12, 2022

Jacobs, D. R., Jr, Woo, J. G., Sinaiko, A. R., Daniels, S. R., Ikonen, J., Juonala, M., Kartiosuo, N., Lehtimäki, T., Magnussen, C. G., Viikari, J., Zhang, N., Bazzano, L. A., Burns, T. L., Prineas, R. J., Steinberger, J., Urbina, E. M., Venn, A. J., Raitakari, O. T., & Dwyer, T. (2022). Childhood cardiovascular risk factors and adult cardiovascular events. *The New England Journal of Medicine*, *386*(20), 1877–1888. doi:10.1056/NEJMoa2109191

Johns, M. M., Lowry, R., Andrzejewski, J., Barrios, L. C., Demissie, Z., McManus, T., Rasberry, C. N., Robin, L., & Underwood, J. M. (2019). Transgender identity and experiences of violence victimization, substance use, suicide risk, and sexual risk behaviors among high school students—19 States and Large Urban School Districts, 2017. *Morbidity and Mortality Weekly Report*, *68*(3), 67–71. doi:10.15585/mmwr.mm6803a3

Kaltiala, R., Bergman, H., Carmichael, P., de Graaf, N. M., Egebjerg Rischel, K., Frisén, L., Schorkopf, M., Suomalainen, L., & Waehre, A. (2020). Time trends in referrals to child and adolescent gender identity services: A study in four Nordic countries and in the UK. *Nordic Journal of Psychiatry*, *74*(1), 40–44. doi:10.1080/08039488.2019.1667429

Kaltiala-Heino, R., & Lindberg, N. (2019). Gender identities in adolescent population: Methodological issues and prevalence across age groups. *European Psychiatry*, *55*, 61–66. doi:10.1016/j.eurpsy.2018.09.003

Kaltiala-Heino, R., Bergman, H., Työläjärvi, M., & Frisen, L. (2018). Gender dysphoria in adolescence: Current perspectives. *Adolescent Health, Medicine and Therapeutics*, *9*, 31–41. doi:10.2147/AHMT.S135432

Kaltiala-Heino, R., Sumia, M., Työläjärvi, M., & Lindberg, N. (2015). Two years of gender identity service for minors: Overrepresentation of natal girls with severe problems in adolescent development. *Child and Adolescent Psychiatry and Mental Health*, *9*(1), 9. doi:10.1186/s13034-015-0042-y

Keith, K. (2022, July 27). HHS proposes revised ACA anti-discrimination rule. *Health Affairs Forefront*. Advance online publication. doi:10.1377/forefront.20220727.369815

Kidd, K. M., Sequeira, G. M., Douglas, C., Paglisotti, T., Inwards-Breland, D. J., Miller, E., & Coulter, R. W. S. (2021). Prevalence of gender-diverse youth in an urban school district. *Pediatrics*, *147*(6), e2020049823. doi:10.1542/peds.2020-049823

Klink, D., Caris, M., Heijboer, A., van Trotsenburg, M., & Rotteveel, J. (2015). Bone mass in young adulthood following gonadotropin-releasing hormone analog treatment and cross-sex hormone treatment in adolescents with gender dysphoria. *The Journal of Clinical Endocrinology & Metabolism*, *100*(2), E270–E275. doi:10.1210/jc.2014-2439

Knapp, T. R. (2016). Why is the one-group pretest–posttest design still used? *Clinical Nursing Research*, *25*(5), 467–472. doi:10.1177/1054773816666280

Korte, A., Goecker, D., Krude, H., Lehmkuhl, U., Grüters-Kieslich, A., & Beier, K. M. (2008). Gender identity disorders in childhood and adolescence. *Deutsches Ärzteblatt International*, *105*(48), 834–841. doi:10.3238/arztebl.2008.0834

Kozlowska, K., Chudleigh, C., McClure, G., Maguire, A. M., & Ambler, G. R. (2021). Attachment Patterns in Children and Adolescents With Gender Dysphoria. *Frontiers in Psychology*, *11*, 582688. doi:10.3389/fpsyg.2020.582688

Kuper, L. E., Cooper, M. B., & Mooney, M. A. (2022). Supporting and advocating for transgender and gender diverse youth and their families within the sociopolitical context of widespread discriminatory legislation and policies. *Clinical Practice in Pediatric Psychology*, *10*(3), 336–345. doi:10.1037/cpp0000456

Lau, J. S., Kline-Simon, A., Sterling, S., Hojilla, J. C., & Hartman, L. (2021). Screening for gender identity in adolescent well visits: Is it feasible and acceptable? *Journal of Adolescent Health*, *68*(6), 1089–1095. doi:10.1016/j.jadohealth.2020.07.031

Levine, S. B., Abbruzzese, E., & Mason, J. W. (2022a). Reconsidering informed consent for trans-identified children, adolescents, and young adults. *Journal of Sex & Marital Therapy*, *48*(7), 706–727. doi:10.1080/0092623X.2022.2046221

Levine, S. B., Abbruzzese, E., & Mason, J. W. (2022b). What are we doing to these children? Response to Drescher, Clayton, and Balon Commentaries on Levine et al., 2022. *Journal of Sex & Marital Therapy*. Advance online publication. doi:10.1080/0092623X.2022.2136117

Littman, L. (2018). Parent reports of adolescents and young adults perceived to show signs of a rapid onset of gender dysphoria. *PLOS ONE*, *13*(8), e0202330. doi:10.1371/journal.pone.0202330

Littman, L. (2020). The use of methodologies in Littman (2018) is consistent with the use of methodologies in other studies contributing to the field of gender dysphoria research: Response to Restar (2019). *Archives of Sexual Behavior*, *49*(1), 67–77. doi:10.1007/s10508-020-01631-z

Littman, L. (2021). Individuals treated for gender dysphoria with medical and/or surgical transition who subsequently detransitioned: A survey of 100 detransitioners. *Archives of Sexual Behavior*, *50*(8), 3353–3369. doi:10.1007/s10508-021-02163-w

London, A. J. (2017). Equipoise in research: Integrating ethics and science in human research. *JAMA*, *317*(5), 525. doi:10.1001/jama.2017.0016

Marchiano, L. (2018, March 1). Transgenderism and the social construction of diagnosis. Quillette. https://quillette.com/2018/03/01/transgenderism-social-construction-diagnosis/

Mathes, T., & Pieper, D. (2017). Clarifying the distinction between case series and cohort studies in systematic reviews of comparative studies: Potential impact on body of evidence and workload. *BMC Medical Research Methodology*, *17*(1), 107. doi:10.1186/s12874-017-0391-8

McNamara, M. (2022). 10/28/22 Florida boards of medicine and osteopathic medicine joint rules/legislative committee rule workshop. https://thefloridachannel.org/videos/10-28-22-florida-boards-of-medicine-and-osteopathic-medicine-joint-rules-legislative-committee-rule-workshop/ Accessed November 12, 2022

McNamara, M., Lepore, C., & Alstott, A. (2022). Protecting transgender health and challenging science denialism in policy. *New England Journal of Medicine*, *387*(21), 1919–1921. doi:10.1056/NEJMp2213085

McGuire, J. K., Berg, D., Catalpa, J. M., Morrow, Q. J., Fish, J. N., Nic Rider, G., Steensma, T., Cohen-Kettenis, P. T., & Spencer, K. (2020). Utrecht gender dysphoria scale—Gender spectrum (UGDS-GS): Construct validity among transgender, nonbinary, and LGBQ samples. *International Journal of Transgender Health*, *21*(2), 194–208. doi:10.1080/26895269.2020.1723460

National Health Service (NHS). (2022a, October 20). *Interim service specification for specialist gender dysphoria services for children and young people—Public consultation*. https://www.engage.england.nhs.uk/specialised-commissioning/gender-dysphoria-services/

National Health Service (NHS). (2022b, July 28). *Regional model for gender care announced for children and young people*. http://tavistockandportman.nhs.uk/about-us/news/stories/regional-model-for-gender-care-announced-for-children-and-young-people/

National Institute for Health and Care Excellence (NICE). (2020a). *Evidence review: Gonadotrophin releasing hormone analogues for children and adolescents with gender dysphoria*. https://cass.independent-review.uk/nice-evidence-reviews/

National Institute for Health and Care Excellence (NICE). (2020b). *Evidence review: Gender-affirming hormones for children and adolescents with gender dysphoria*. https://cass.independent-review.uk/nice-evidence-reviews/

Nguyen, V. T., Engleton, M., Davison, M., Ravaud, P., Porcher, R., & Boutron, I. (2021). Risk of bias in observational studies using routinely collected data of comparative effectiveness research: A meta-research study. *BMC Medicine*, *19*(1), 279. doi:10.1186/s12916-021-02151-w

Nokoff, N., Ma, N., Moreau, K., & Rothman, M. S. (2022). *Bone Mineral Density in Transgender Youth on Gender Affirming Therapies*. https://www.endocrine.org/news-and-advocacy/news-room/2022/longer-treatment-with-puberty-delaying-medication-leads-to-lower-bone-mineral-density

Nokoff, N. J., Scarbro, S. L., Moreau, K. L., Zeitler, P., Nadeau, K. J., Reirden, D., Juarez-Colunga, E., & Kelsey, M. M. (2021). Body composition and markers of cardiometabolic health in transgender youth on gonadotropin-releasing hormone agonists. *Transgender Health*, *6*(2), 111–119. doi:10.1089/trgh.2020.0029

Nolan, I. T., Kuhner, C. J., & Dy, G. W. (2019). Demographic and temporal trends in transgender identities and gender confirming surgery. *Translational Andrology and Urology*, *8*(3), 184–190. doi:10.21037/tau.2019.04.09

Olson-Kennedy, J., Okonta, V., Clark, L. F., & Belzer, M. (2018). Physiologic response to gender-affirming hormones among transgender youth. *Journal of Adolescent Health*, *62*(4), 397–401. doi:10.1016/j.jadohealth.2017.08.005

Olson-Kennedy, J., Warus, J., Okonta, V., Belzer, M., & Clark, L. F. (2018). Chest reconstruction and chest dysphoria in transmasculine minors and young adults: Comparisons of nonsurgical and postsurgical cohorts. *JAMA Pediatrics*, *172*(5), 431. doi:10.1001/jamapediatrics.2017.5440

O'Malley, S., & Ayad, S. (Hosts). (2022, March 7). Pioneers series: Where it all started. The Dutch Researchers Steensma & De Vries (No. 66) [Audio podcast episode]. Gender: A Wider Lens. https://gender-a-wider-lens.captivate.fm/episode/66-pioneers-series-where-it-all-started-the-dutch-researchers-steensma-de-vries

Pasternack, I., Söderström, I., Saijonkari, M., & Mäkelä, M. (2019). *Lääketieteelliset menetelmät sukupuolivariaatioihin liittyvän dysforian hoidossa. Systemaattinen katsaus. [Medical approached to treatment of dysphoria related to gender variations. A systematic review.]*. 106. https://app.box.com/s/y9u791np8v9gsunwgpr2kqn8swd9vdtx

Porter, J., & Jick, H. (1980). Addiction rare in patients treated with narcotics. *New England Journal of Medicine*, *302*(2), 123–123. doi:10.1056/NEJM198001103020221

Prasad, V. (2011). The frequency of medical reversal. *Archives of Internal Medicine*, *171*(18), 1675. doi:10.1001/archinternmed.2011.295

Prasad, V. (2013). Why randomized controlled trials are needed to accept new practices: 2 medical worldviews. *Mayo Clinic Proceedings*, *88*(10), 1046–1050. doi:10.1016/j.mayocp.2013.04.026

Prasad, V., & Ioannidis, J. P. (2014). Evidence-based de-implementation for contradicted, unproven, and aspiring healthcare practices. *Implementation Science*, *9*(1), 1748–5908. doi:10.1186/1748-5908-9-1

Rafferty, J. (2018). Ensuring comprehensive care and support for transgender and gender-diverse children and adolescents. *Pediatrics*, *142*(4), e20182162. doi:10.1542/peds.2018-2162

Rantz, J. (2022, August 23). Despite 'concerning' transgender study, UW kept quiet because of positive coverage. MyNorthwest.Com https://mynorthwest.com/3602854/rantz-despite-concerning-trans-study-uw-kept-quiet-because-of-positive-coverage/

Registers in The Netherlands. (2022). EIT Health Scandinavia, https://www.eithealth-scandinavia.eu/biobanksregisters/registers/netherlands/. Accessed 14 Dec.

Respaut, R., Terhune, C. (2022, October 6). Putting numbers on the rise in children seeking gender care. Retrieved October 14, 2022, from https://www.reuters.com/investigates/special-report/usa-transyouth-data/

Ristori, J., & Steensma, T. D. (2016). Gender dysphoria in childhood. *International Review of Psychiatry*, *28*(1), 13–20. doi:10.3109/09540261.2015.1115754

Roberts, C. M., Klein, D. A., Adirim, T. A., Schvey, N. A., & Hisle-Gorman, E. (2022). Continuation of gender-affirming hormones among transgender adolescents and adults. *The Journal of Clinical Endocrinology & Metabolism*, *107*(9), e3937–e3943. doi:10.1210/clinem/dgac251

Robles García, R., & Ayuso-Mateos, J. L. (2019). ICD-11 and the depathologisation of the transgender condition. *Revista de Psiquiatría y Salud Mental (English Edition)*, *12*(2), 65–67. doi:10.1016/j.rpsmen.2019.01.002

Rosenthal, S. M. (2021). Challenges in the care of transgender and gender-diverse youth: An endocrinologist's view. *Nature Reviews Endocrinology*, *17*(10), 581–591. doi:10.1038/s41574-021-00535-9

Rotteveel, J., Belksma, E. J., Renders, C. M., Hirasing, R. A., & Delemarre-Van de Waal, H. A. (2007). Type 2 diabetes in children in the Netherlands: The need for diagnostic protocols. *European Journal of Endocrinology*, *157*(2), 175–180. doi:10.1530/EJE-06-0754

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *BMJ*, *312*(7023), 71–72. doi:10.1136/bmj.312.7023.71

SBU (Swedish Agency for Health Technology Assessment and Assessment of Social Services). (2022). *Hormonbehandling vid könsdysfori—Barn och unga En systematisk översikt och utvärdering av medicinska aspekter [Hormone therapy at gender dysphoria—Children and young people A systematic review and evaluation of medical aspects]*. https://www.sbu.se/contentassets/ea4e698fa0c4449aaae964c5197cf940/hormonbehandling-vid-konsdysfori_barn-och-unga.pdf

Schönbeck, Y., Talma, H., van Dommelen, P., Bakker, B., Buitendijk, S. E., HiraSing, R. A., & van Buuren, S. (2011). Increase in prevalence of overweight in Dutch children and adolescents: A comparison of Nationwide Growth studies in 1980, 1997 and 2009. *PLoS ONE*, *6*(11), e27608. doi:10.1371/journal.pone.0027608

Schünemann HJ, Vist GE, Higgins JPT, Santesso N, Deeks JJ, Glasziou P, Akl EA, Guyatt GH. (2022). Chapter 15: Interpreting results and drawing conclusions. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ ,Welch VA (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions version 6.3*. London: Cochrane. (updated February 2022). Available from www.training.cochrane.org/handbook.

Schwartz, D. (2021). Clinical and ethical considerations in the treatment of gender dysphoric children and adolescents: When doing less is helping more. *Journal of Infant, Child, and Adolescent Psychotherapy*, *20*(4), 439–449. doi:10.1080/15289168.2021.1997344

Shaffer, D. (1983). A Children's Global Assessment Scale (CGAS). *Archives of General Psychiatry*, *40*(11), 1228. doi:10.1001/archpsyc.1983.01790100074010

Singh, D., Bradley, S. J., & Zucker, K. J. (2021). A follow-up study of boys with gender identity disorder. *Frontiers in Psychiatry*, *12*, 632784. doi:10.3389/fpsyt.2021.632784

Smith, Y. L. S., Van Goozen, S. H. M., & Cohen-Kettenis, P. T. (2001). Adolescents with gender identity disorder who were accepted or rejected for sex reassignment surgery: A prospective follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(4), 472–481. doi:10.1097/00004583-200104000-00017

Socialstyrelsen [National Board of Health and Welfare]. (2022). *Care of children and adolescents with gender dysphoria – Summary*. Retrieved July 22, 2022 from https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/kunskapsstod/2022-3-7799.pdf

Steensma, T. D., de Rooy, F. B. B., van der Meulen, I. S., Asseler, J. D., & van der Miesen, A. I. R. (2022, September 16–20). *Transgender Care Over the Years: First Long-Term Follow-Up Studies and Exploration of Sex Ratio in the Amsterdam Child and Adolescent Gender Clinic* [Conference presentation]. World Professional Association for Transgender Health Symposium, Montreal, QC, Canada.

Steensma, T. D., Kreukels, B. P., Jurgensen, M., Thyen, U., de Vries, A. L., Cohen-Kettenis, P. T. (2013). The utrecht gender dysphoria scale: A validation study. In: Steensma TD (Ed.), *From gender variance to gender dysphoria: Psychosexual development of gender atypical children and adolescents* (pp. 41–56). Amsterdam: Vrije Universiteit. https://research.vu.nl/ws/portalfiles/portal/42117766/table±of±contents.pdf

Steensma, T. D., van der Ende, J., Verhulst, F. C., & Cohen-Kettenis, P. T. (2013). Gender variance in childhood and sexual orientation in adulthood: A prospective study. *The Journal of Sexual Medicine*, *10*(11), 2723–2733. doi:10.1111/j.1743-6109.2012.02701.x

Strang, J. F., Meagher, H., Kenworthy, L., de Vries, A. L. C., Menvielle, E., Leibowitz, S., Janssen, A., Cohen-Kettenis, P., Shumer, D. E., Edwards-Leeper, L., Pleak, R. R., Spack, N., Karasic, D. H., Schreier, H., Balleur, A., Tishelman, A., Ehrensaft, D., Rodnan, L., Kuschner, E. S., … Anthony, L. G. (2018). Initial clinical guidelines for co-occurring autism spectrum disorder and gender dysphoria or incongruence in adolescents. *Journal of Clinical Child and Adolescent Psychology*, *47*(1), 105–115. doi:10.1080/15374416.2016.1228462

Szilagyi, M. (2022, August 21). Academy of pediatrics responds on trans treatment for kids. Wall Street Journal. https://www.wsj.com/articles/trans-gender-pediatric-aap-kids-children-care-surgery-affirm-treatment-11660942086

Thrower, E., Bretherton, I., Pang, K. C., Zajac, J. D., & Cheung, A. S. (2020). Prevalence of autism spectrum disorder and attention-deficit hyperactivity disorder amongst individuals with gender dysphoria: A systematic review. *Journal of Autism and Developmental Disorders*, *50*(3), 695–706. doi:10.1007/s10803-019-04298-1

Tordoff, D. M., Wanta, J. W., Collin, A., Stepney, C., Inwards-Breland, D. J., & Ahrens, K. (2022). Mental health outcomes in transgender and nonbinary youths receiving gender-affirming care. *JAMA Network Open*, *5*(2), e220978. doi:10.1001/jamanetworkopen.2022.0978

Turban, J. (2022, March 1). Opinion | Texas officials are spreading blatant falsehoods about medical care for transgender kids. Washington Post. Retrieved November 15, 2022, from https://www.washingtonpost.com/opinions/2022/03/01/texas-ken-paxton-greg-abbott-misinformation-transgender-medical-care/

van de Grift, T. C., Elaut, E., Cerwenka, S. C., Cohen-Kettenis, P. T., De Cuypere, G., Richter-Appelt, H., & Kreukels, B. P. C. (2017). Effects of medical interventions on gender dysphoria and body image: A follow-up study. *Psychosomatic Medicine*, *79*(7), 815–823. doi:10.1097/PSY.0000000000000465

Vandenbussche, E. (2022). Detransition-related needs and support: A cross-sectional online survey. *Journal of Homosexuality*, *69*(9), 1602–1620. doi:10.1080/00918369.2021.1919479

World Health Organization. (2019). International statistical classification of diseases and related health problems (11th ed.). https://icd.who.int/

World Professional Association for Transgender Health (WPATH). (2018). WPATH position on "Rapid-Onset Gender Dysphoria (ROGD)". Retrieved July 11, 2022, from https://www.wpath.org/media/cms/Documents/Public%20Policies/2018/9_Sept/WPATH%20Position%20on%20Rapid-Onset%20Gender%20Dysphoria_9-4-2018.pdf

Zhang, S. (2017, June 2). *The one-paragraph letter from 1980 that fueled the opioid crisis*. The Atlantic. https://www.theatlantic.com/health/archive/2017/06/nejm-letter-opioids/528840/

Zhang, Q., Rechler, W., Bradlyn, A., Flanders, W. D., Getahun, D., Lash, T. L., McCracken, C., Nash, R., Panagiotakopoulos, L., Roblin, D., Sandberg, D. E., Silverberg, M. J., Tangpricha, V., Vupputuri, S., & Goodman, M. (2021). Changes in size and demographic composition of transgender and gender non-binary population receiving care at integrated health systems. *Endocrine Practice*, *27*(5), 390–395. doi:10.1016/j.eprac.2020.11.016

Zimerman, A. (2013). Evidence-based medicine: A short history of a modern medical movement. *AMA Journal of Ethics*, *15*(1), 71–76. doi:10.1001/virtualmentor.2013.15.1.mhst1-1301

Zucker, K. J. (2019). Adolescents with gender dysphoria: Reflections on some contemporary clinical and research issues. *Archives of Sexual Behavior*, *48*(7), 1983–1992. doi:10.1007/s10508-019-01518-8