Progression of Gender Dysphoria in Children and Adolescents: A Longitudinal Study

Stephanie Wagner, MD, MPH,[a,b] Leonidas Panagiotakopoulos, MD,[a] Rebecca Nash, MPH,[b] Andrew Bradlyn, PhD,[c] Darios Getahun, MD, PhD,[d,e] Timothy L. Lash, DSc, MPH,[b] Douglas Roblin, PhD,[f] Michael J. Silverberg, PhD, MPH,[g] Vin Tangpricha, MD, PhD,[a,h] Suma Vupputuri, PhD,[f] Michael Goodman, MD, MPH[b]

abstract

BACKGROUND AND OBJECTIVES: The progression of gender-expansive behavior to gender dysphoria and to gender-affirming hormonal treatment (GAHT) in children and adolescents is poorly understood.

METHODS: A cohort of 958 gender-diverse (GD) children and adolescents who did not have a gender dysphoria–related diagnosis (GDRD) or GAHT at index were identified. Rates of first GDRD and first GAHT prescription were compared across demographic groups.

RESULTS: Overall, 29% of participants received a GDRD and 25% were prescribed GAHT during the average follow-up of 3.5 years (maximum 9 years). Compared with youth assigned male sex at birth, those assigned female sex at birth were more likely to receive a diagnosis and initiate GAHT with hazard ratio (95% confidence interval) estimates of 1.3 (1.0–1.7), and 2.5 (1.8–3.3), respectively. A progression to diagnosis was more common among those aged ≥15 years at initial presentation compared with those aged 10 to 14 years and those aged 3 to 9 years (37% vs 28% vs 16%, respectively). By using the youngest group as a reference, the adjusted hazard ratios (95% confidence interval) for a GDRD were 2.0 (1.3–3.0) for age 10 to 14 years and 2.7 (1.8–3.9) for age ≥15 years. Racial and ethnic minorities were less likely to receive a diagnosis or be prescribed GAHT.

CONCLUSIONS: This study characterized the progression of GD behavior in children and adolescents. Less than one-third of GD youth receive an eventual GDRD, and approximately one-quarter receive GAHT. Female sex at birth, older age of initial GD presentation to medical care, and non-Hispanic white race and ethnicity increased the likelihood of receiving diagnosis and treatment.

WHAT'S KNOWN ON THIS SUBJECT: The progression of gender-expansive behavior to gender dysphoria and to gender-affirming hormonal treatment in children and adolescents is poorly understood. We do not yet know rates of conversion, diagnosis, or speed of progression to diagnosis.

WHAT THIS STUDY ADDS: In this study, we characterized the progression of gender-diverse behavior in children and adolescents.

Understanding the natural history of gender-expansive behavior is an increasingly important issue in caring for gender-diverse (GD) children and adolescents. Changing definitions, evolving theories of gender identity development, and availability of new data affect current understanding of the optimal care required to support GD youth.

The term "gender identity" refers to a wide range of individual self-identifications that may encompass various degrees of maleness or femaleness or a complete rejection of binary gender categories. GD individuals are those whose gender identity does not fully match their recorded sex at birth.[1] "Gender-variant behavior" is a related term that describes behaviors that contrast with what society may term as "typical" or "sex-typed."[1] The language pertaining to individuals who experience distress with their assigned sex has changed in recent decades, and it continues to evolve, particularly with the introduction of the term "gender dysphoria" in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* to emphasize distress secondary to GD identity rather than pathologizing GD identity itself,[1–4] and the diagnosis of gender dysphoria as part of the *International Classification of Diseases, 11th Revision*, in which distress is no longer required, but must give access to gender-affirming medical treatment if necessary.[5] To date, no comprehensive longitudinal studies have characterized the progression of GD behaviors to gender dysphoria or related diagnosis.

The optimal age of starting gender-affirming therapies is an area of ongoing discussion. A strong scientific debate concerns optimal time for intervention because of some evidence indicating that childhood dysphoria might not still be equally present in adolescence[6] and other articles indicating that chances of attenuation of the gender nonconformity are considered to be much higher in the prepubertal compared with the pubertal youth.[7] Guidelines issued by the Endocrine Society and the World Professional Association for Transgender Health (WPATH) recommend psychosocial support and possible social transition for prepubescent children.[8–10] As children enter the period of early development of secondary sex characteristics, defined as Tanner stage 2,[11,12] guidelines recommend the use of gonadotropin-releasing hormone (GnRH) agonists to suppress continued puberty. The goals of hormone suppression are to (1) minimize dysphoria, (2) allow the safe passing of time before more definite decisions are made, and (3) facilitate gender affirmation later in life.[9,10] Both guidelines suggest initiating gender-affirming hormonal treatments (GAHTs) at ~16 years of age but acknowledge that in many cases it may be appropriate to initiate earlier.[9,10,13] Access and affordability of these interventions vary across countries, health systems, and insurance plans.[8]

Practice guidelines for puberty suppression and GAHT are primarily based on consensus rather than high-quality empirical data.[9,10,13–18] The authors of these guidelines point out a lack of strong evidence in identifying the optimal age at which gender-affirming treatments should be initiated.[13,14,16]

Most available studies addressing the gender-affirming care offered to GD children and adolescents were based in individual clinics and included relatively small numbers (range: 25–187) of participants and nonuniformity in the design of subjects studied.[19–21] Moreover, little is known about the typical time course from initial presentation of GD behavior in children to initiation of gender-affirming care.[18,20–28] With the knowledge that the GD population is growing and increasing proportions of GD individuals are presenting at an earlier age, there is a need for large-scale longitudinal studies investigating patterns and determinants of GD-specific care in children and adolescents.[29–33] The purpose of this study was to examine the likelihood and predictors of receiving a GD-specific diagnosis and GAHT among youth who express gender-variant behavior. We used data from a large cohort of GD youth who received care within 3 integrated health care systems in the United States.

METHODS

This study uses data from the Study of Transition, Outcomes, and Gender (STRONG) cohort. The STRONG cohort includes participants from Kaiser Permanente (KP) integrated health care systems in Georgia, Northern California, and Southern California. The 3 KP sites collectively provide comprehensive health care to ~9 million individuals.[34,35] The current transgender health care protocols at KP follow the Endocrine Society and WPATH guidelines, which include mental health support, GnRH agonists, feminizing and masculinizing hormones, and surgery.[36] The Emory University Rollins School of Public Health served as the coordinating center. The study protocol received approvals from the institutional review boards of all 4 institutions (3 KP sites and the coordinating center) with exemption of informed consent. The details of STRONG cohort ascertainment and data collection were described in previous publications.[17,32,37] Participants were identified in the electronic health record from January 1, 2006, to December 31, 2014, by searching for keywords in free-text clinical notes reflecting GD behaviors.

Each participant's index date was defined on the basis of the first evidence of GD behaviors mentioned in the notes, as evidenced in the presence of keywords such as "transgender," and "gender identity." The full list of keywords can be found in our previous publication.[32] All notes were reviewed to confirm eligibility. The analytic data set was limited to participants who were aged <18 years at index date, had at least 1 follow-up appointment, and had evidence of GD behavior, as reflected in the keywords, but did not have a diagnosis related to gender dysphoria and had not received any GAHT. Participants whose sex recorded at birth could not be determined ($n = 14$) were excluded from the analyses because sex recorded at birth was considered a key variable in the analyses. Two types of events of interest were ascertained during follow-up: an assignment of the first gender dysphoria–related diagnosis (GDRD) and receipt of GAHT. The GDRDs were based on the *International Classification of Diseases, Ninth Revision* and included codes for conditions such as transsexualism (302.5) and gender identity disorder in children (302.6). The *Diagnostic and Statistical Manual of Mental Disorders* uses the *International Classification of Diseases* coding scheme with direct match between the 2 systems. For this reason, any *International Classification of Diseases* codes used in this analysis would be equivalent to the same diagnosis in the contemporaneous *Diagnostic and Statistical Manual of Mental Disorders*. Cohort ascertainment and follow-up were undertaken before the health plans switched to *International Classification of Diseases, 10th Revision* codes. GAHT receipt was determined from pharmacy records and date of therapy initiation was based on the first prescription for a puberty suppression medication or cross-sex hormonal therapy.

The follow-up for each participant extended from the index date until the event of interest (diagnostic code or first ordered GAHT prescription, depending on the analysis), disenrollment from KP, or end of follow-up (December 31, 2014). The data were examined by using time-to-event analyses to take into account censoring and variable duration of follow-up.

Kaplan Meier curves were constructed to compare timing and occurrence of GDRD and GAHT initiation across subgroups of participants. The independent variables in these analyses included age category at index date (categorized as 3–9, 10–14 and ≥15–17 years), recorded sex at birth, and race and ethnicity (non-Hispanic white individuals versus individuals of other races and ethnicities). These age groups represent 3 distinct populations: children who have not started puberty, based on national averages (age 3–9 years); children after the onset of puberty who may be considered as candidates for puberty suppression (age 10–14 years); and teenagers who might be offered feminizing or masculinizing hormones on the basis of current treatment guidelines (age ≥15 years).[36]

Multivariable Cox proportional hazards models were used to evaluate the associations of all 3 independent variables (site, sex recorded at birth, age, and race and ethnicity) considered individually and simultaneously with each event of interest. The results of the Cox models were expressed as crude and adjusted hazard ratios (HRs) and corresponding 95% confidence intervals (CIs). All models were evaluated for validity of proportional hazard assumptions by inspecting log-log curves. If proportional hazard assumptions were violated, stratified Cox models were used. The data analyses were performed by using SAS software version 9.4 (SAS Institute, Inc, Cary, NC).

RESULTS

After applying eligibility criteria (Fig 1), 958 children were included in the final analysis data set (Table 1). Of those, 431 individuals were assigned male sex at birth (AMAB) and 527 individuals were assigned female sex at birth (AFAB). Children aged <10 years at index date represented 21% of the total cohort, 30% of the AMAB group and 14% of the AFAB group. In both the AMAB and AFAB groups, non-Hispanic white individuals made up >45% of the children. A majority of individuals (63%) received care at KP Northern California. Of the total analytic cohort, 29% of participants received a GDRD and 25% were prescribed GAHT during an average follow-up of 3 years (maximum 9 years). Among 677 cohort members without a diagnosis, 74 (11%) received GAHT treatment, whereas, among 281 children and adolescent with a diagnosis, 162 (58%) initiated GAHT.

Compared with the AFAB group, a lower proportion of AMAB children and adolescents received a GDRD (24% vs 33%) or initiated GAHT (14% vs 33%) during follow-up (Table 1). When time to diagnosis
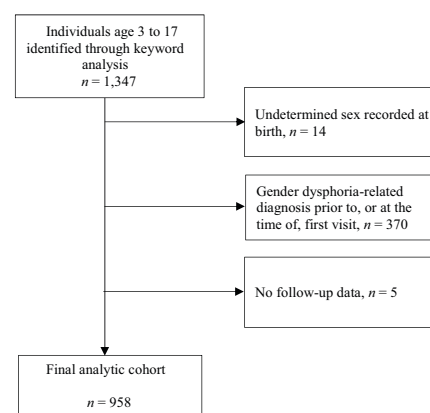
FIGURE 1
Inclusion and exclusion of individuals.

TABLE 1 Selected Participant Characteristics by Recorded Sex at Birth

Participant Characteristics	Children Who Were AMAB, n (%)[a]	Children Who Were AFAB, n (%)[a]	Total, n (%)[a]
Age at index date, y			
3–9	131 (30)	74 (14)	205 (21)
10–14	128 (30)	172 (33)	300 (31)
≥15	172 (40)	281 (53)	453 (47)
Race and ethnicity			
Non-Hispanic white	196 (45)	277 (53)	473 (49)
Non-Hispanic Black	42 (10)	44 (8)	86 (9)
Non-Hispanic Asian American	30 (7)	47 (9)	77 (8)
Hispanic	128 (30)	130 (25)	258 (27)
Other	5 (1)	8 (2)	13 (1)
Unknown	30 (7)	21 (4)	51 (5)
Study Site			
Georgia	12 (3)	15 (3)	27 (3)
Northern California	264 (61)	340 (65)	604 (63)
Southern California	155 (36)	172 (33)	327 (34)
GDRD during follow-up			
Yes	105 (24)	176 (33)	281 (29)
No	326 (76)	351 (67)	677 (71)
GAHT initiation during follow-up			
Yes	60 (14)	176 (33)	236 (25)
No	371 (86)	351 (67)	722 (75)
Overall	431 (45)[b]	527 (55)[b]	958 (100)[b]

[a] Column percentages.
[b] Row percentages.

was compared across age groups among AMAB and AFAB children (Fig 2A and C, respectively), the youngest age group (3–9 years of age) was most likely to remain diagnosis-free to the end of follow-up for both (85% in AMAB and 72% in AFAB). Among AMAB participants, the proportion of those who received a diagnosis before the end of follow-up was lower in the middle age group (10–14 years) than in the oldest age group (≥15 years) (20% vs 35%, respectively) (Fig 2A) but this was not the case among AFAB participants (Fig 2C). Whereas few of the youngest children received GAHT by the end of the follow-up (5% in the AFAB group and 4% in the AMAB group), treatment was generally more delayed in older AMAB than older AFAB children (Fig 2B and D).

In the multivariable model (Table 2), the difference in diagnosis rates between AFAB and AMAB participants was attenuated (adjusted HR = 1.3; 95% CI: 1.0–1.7), but the difference in GAHT

receipt remained evident (adjusted HR = 2.5; 95% CI: 1.8–3.3). By contrast, the differences in diagnosis rates remained pronounced across the 3 age groups; using the youngest age group (3–9 years) as reference, the adjusted HRs (95% CI) were 2.0 (1.3–3.0) for age 10 to 14 years and 2.7 (1.8–3.9) for age ≥15 years. The proportional hazard assumption for the age variables was violated in the analyses that used GAHT initiation as the end point of interest, and for this reason the corresponding adjusted HR estimates across the age groups were not generated by the model (Table 2).

Relative to non-Hispanic white children, children of minority racial and ethnic groups were less likely to receive a GDRD (26% vs 33%) or be prescribed GAHT (21% vs 29%) during follow-up, and the time to diagnosis and GAHT were also different in the 2 groups (Fig 3). Controlling for other variables (Table 2), the difference in diagnosis rates was less evident (adjusted HR = 0.8; 95% CI: 0.6–1.0) whereas the

difference in GAHT receipt was greater (adjusted HR = 0.6; 95% CI: 0.5–0.8).

When the data were examined by site, rates of diagnosis appeared to be lower among children residing in Georgia compared with their California counterparts (Fig 4). The HR estimates, however, were imprecise because of the small size of the KP Georgia cohort. The proportional hazard assumption for site variable was violated for GAHT initiation, and for this reason we controlled for site in the stratified Cox model, but the adjusted HR estimates for this variable are not presented (Table 2).

DISCUSSION

This electronic health record-based cohort study nested in 3 large integrated health care systems revealed that gender-expansive behaviors in most children do not lead to a GDRD and initiation of GAHT. We observed that the rates of diagnosis and GAHT initiation differed across demographic categories of participants and tended to be higher in (1) older adolescents, (2) AFAB cohort members, and (3) non-Hispanic white people. Because of the limited sample size, differences across study sites could not be definitively tested and no conclusion could be reached. In our cohort, data have overrepresentation from subjects who were AFAB, and, although this is also the case in more recent publications,[38,39] the skew toward subjects who were AFAB was not as pronounced as in those studies. This is likely due to the fact that our cohort dates back to 2006, and thus the difference is not as pronounced as in studies including subjects from more recent years.

Our results need to be viewed in the context of similar findings reported previously in European
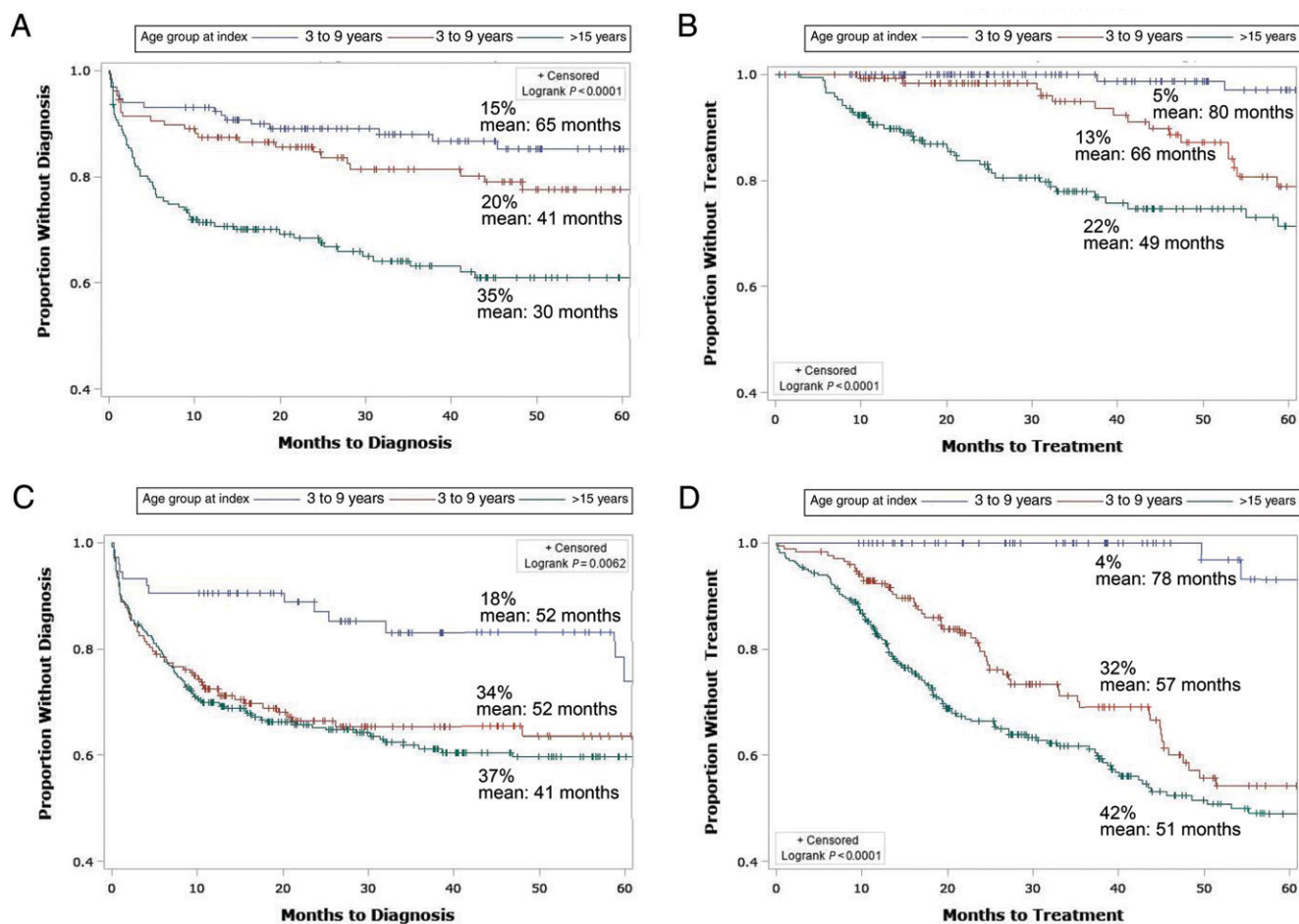
FIGURE 2
Kaplan-Meier time-to-event analysis by age at index date stratified by sex recorded at birth. A, AMAB outcome: GDRD. B, AMAB outcome: GAHT. C, AFAB outcome: GDRD. D, AFAB outcome: GAHT.

research.[21,28] Two similarly designed, but nonoverlapping, studies performed a follow-up assessment of children treated for gender dysphoria at a specialized clinic in the Netherlands.[21,28] The first study included 77 children who had been referred to a gender-specific clinic between 1989 and 2005 for gender dysphoria at age <12 years at initial presentation.[21] After an average follow-up of 10 years, 27% of the initial cohort continued experiencing gender dysphoria; however, this result may have been affected by the relatively high (30%) proportion of participants who did not respond to the survey. The authors also reported that individuals with gender dysphoria had more extreme gender dysphoria observed during childhood and were more likely to meet criteria for a gender dysphoria diagnosis during childhood.[21]

The second Dutch study originated from the same clinic but sampled a different group of adolescents (n = 127) between 2000 and 2008.[28] As in the earlier study, participants received a GDRD at <12 years old and were followed-up at 15 years old or older. Approximately 37% of adolescents in the overall cohort still experienced gender dysphoria at follow-up, although a high percentage of nonrespondents (22%) were counted in the study as no longer experiencing gender dysphoria. It is noteworthy that this study did not record pubertal Tanner staging at enrollment; thus, although participants were all <12 years old, some could have been pubertal and some others prepubertal. Factors associated with unalleviated gender dysphoria included more pronounced dysphoria symptoms and older age at presentation. In addition, continuous gender dysphoria was more common among individuals who were AFAB.[28]

Both Dutch studies found that most participants did not experience gender dysphoria beyond puberty. This result is consistent with our observation that less than one-third

TABLE 2 Results of Cox Proportional Models Evaluating Associations Between Participant Characteristics and Each Event of Interest

Independent Variables	Crude HR	95% CI	Adjusted HR	95% CI
Outcome: GDRD				
Assigned sex at birth				
Male (AMAB)	1.0	Reference	1.0	Reference
Female (AFAB)	1.5	1.2–1.9	1.3	1.0–1.7
Race and/or ethnicity				
Non-Hispanic white children	1.0	Reference	1.0	Reference
Other groups	0.8	0.6–1.0	0.8	0.6–1.0
Age at index date, y				
3–9	1.0	Reference	1.0	Reference
10–14	2.0	1.3–3.0	2.0	1.3–3.0
≥15	2.8	1.9–4.1	2.7	1.8–3.9
Study site				
Georgia	1.0	Reference	1.0	Reference
Northern California	1.0	0.8–1.2	1.0	0.8–1.3
Southern California	0.5	0.2–1.3	0.5	0.2–1.3
Outcome: GAHT initiation[a]				
Recorded sex at birth				
Male (AMAB)	1.0	Reference	1.0	Reference
Female (AFAB)	3.0	2.2–4.0	2.5	1.8–3.3
Race and ethnicity				
Non-Hispanic white children	1.0	Reference	1.0	Reference
Other groups[b]	0.6	0.5–0.8	0.6	0.5–0.8

[a] Model was stratified on age and study site because of violation of proportional hazards assumption for these 2 variables; for this reason, the results are controlled for age and study site, but no HR estimates are provided.
[b] Includes persons with unknown race and ethnicity.

of children presenting with GD behaviors received a GDRD and only approximately one-quarter initiated hormone therapy during follow-up. The Dutch researchers also reported a greater likelihood of unalleviated gender dysphoria in children who presented at an older age and among AFAB participants, both results in agreement with our findings.

Perhaps the most important methodologic feature of our study compared with previous research is the use of system-wide cohort ascertainment that was not limited to a particular clinical center. Although many of the transgender members of KP receive specialized coordinated care through Multispecialty Transitions Clinics,

this was not an inclusion criterion in the study. Rather the present data set included any member with evidence of transgender status or gender diversity documented in the medical records. Thus, the resulting patient population represents an unselected group with diverse pathways to care. This may explain the relatively low percentage of GD children who received a GDRD, and the relatively high proportion of children who received care without a diagnosis. The deidentified data permitted inclusion of all eligible persons in the analyses because participation did not require subject opt-in. In addition, the keyword-based approach to identify eligible study participants offered a rare opportunity to evaluate the course of events in children at earlier stages of gender-variant behavior, which is rarely possible in specialized clinic-based studies.

The second methodologic feature that distinguishes our study from previously published studies conducted in the Netherlands is the more recent time period (1985–2008 vs 2006–2014). As the size and the composition of transgender populations are rapidly changing, especially in the youngest age
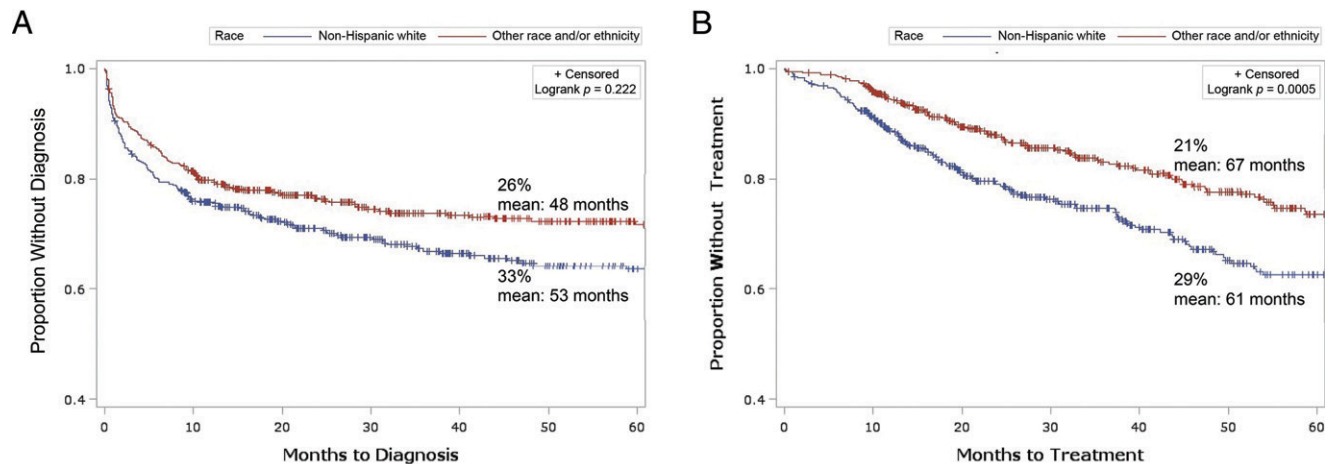

FIGURE 3
Kaplan-Meier time-to-event analysis for non-Hispanic white individuals versus individuals of other races and ethnicities. A, Outcome: GDRD. B, Outcome: GAHT.
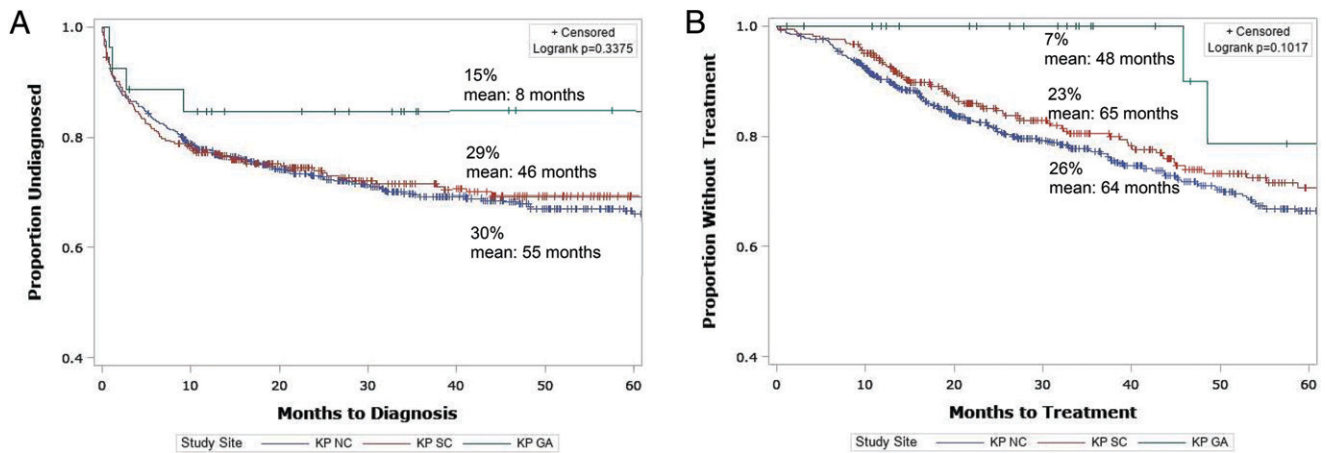
FIGURE 4
Kaplan-Meier time-to-event analysis by study site. A, outcome: GDRD. B, outcome: GAHT. KP GA, Kaiser Permanente Georgia; KP NC, Kaiser Permanente Northern California; KP SC, Kaiser Permanente Southern California.

groups,[40] it appears likely that the results may also be affected by sociopolitical and medical advances, increased access to medical care, less pronounced cultural stigma and evolving social norms with differential impact across generations.[41–44] On the other hand, greater awareness of the available gender-affirming care options, without adequate access to this care, may also exacerbate gender dysphoria.

Another notable aspect of the current study is the diverse sample and ability to examine racial and/or ethnic disparities in the care of GD children. We observed that non-Hispanic white participants had an earlier progression to diagnosis and a more rapid initiation of GAHT compared with other racial and ethnic groups. Although relative sparsity of data for specific racial and/or ethnic minority groups precluded a more granular examination of differences by race and/or ethnicity, our findings are in broad agreement with other studies that have shown cultural biases adversely affecting accessing care for gender dysphoria.[45–47] There was also a suggestion that geographical location may play a role; however, the analyses lacked

power to evaluate modification by geographic region.

The observed disparities in receiving gender-affirming care may be the result of differences in attitudes toward gender diversity among children, their peers, and their families.[48–50] More broadly, a number of social factors, including political environment, culture, and religious taboos, may influence the degree to which GD children are accepted.[51–55] It is also possible that geographic region and/or ethnicity play a role in social acceptance and ability to access gender-affirming care.[56–58]

The difference in rates of diagnosis and especially treatment initiation between AMAB and AFAB children requires further investigation. Although some reports indicate that transgender boys and transgender girls experience differences in parental acceptance,[59] a more likely reason is that endogenous puberty occurs later in natal boys, and, because gender dysphoria typically worsens with pubertal progression, children who were AMAB might seek and receive gender-affirming care later than their AFAB counterparts.[60] Moreover, current Endocrine Society and WPATH guidelines recommend GAHT initiation once Tanner 2 stage

is reached; this occurs on average later in AMAB children compared with their AFAB peers.[11,12] One could also argue that the observed disparity in GAHT initiation between AMAB and AFAB participants is attributable to the differences in the cost of therapy. Whereas the use of GnRH agonists is routinely preferred in AMAB youth, in AFAB children and adolescents use of contraception offers a cheaper alternative. Recently, use of GnRH agonists has expanded as it is increasingly covered by insurance plans[61]; however, the cohort in the current study was followed to the end of 2014, which might have been just before this change in medical practice.

Although our findings indicate that most children with GD behaviors did not receive a GDRD and did not start GAHT during follow-up (3 years on average), these results cannot be interpreted as evidence of long-term outcomes. Thus, it is likely that the proportion of individuals who eventually start gender-affirming care will increase with extended follow-up; however, we expect that some GD study participants will remain dysphoria-free and not require intervention.

It is worth noting that the methodologic features of our study can be viewed as both its strengths and its weaknesses. Because the analyses were based exclusively on the information obtained from medical records, this design precluded collection of patient- and family-reported measures. As such, the degree of documentation by medical professionals can influence the index date and therefore the timing of follow-up initiation. Other limitations of our analyses include the lack of data on social environment or psychological support, pubertal status, and the inability to distinguish children who identify as transgender from those who present with nonbinary or other gender-nonconforming identities. In addition, GD children enrolled in integrated health care systems come primarily from families with health insurance and may not be representative of all GD youth in the United States. On the other hand, this cohort does include patients enrolled in Medicaid plans, allowing that at least some of the study participants come from populations with lower socioeconomic status. Although GAHT protocols do not differ substantially across study sites, the relatively long interval of data collection and follow-up (2006–2016) means that the therapeutic approaches changed over time, and thus children identified earlier in the study may have different pathways to care compared with their counterparts in more recent years.

CONCLUSIONS

Our analyses reveal that GD adolescents are more likely to receive a GDRD or hormone therapy compared with younger children. We also found that both diagnosis receipt and treatment initiation were more common among non-Hispanic white children and AFAB children relative to their respective counterparts. The observed differences by geographic locations require confirmation because of the limited sample size.

Taken together, these results indicate that, even in the presence of similar access to care, use and timing of services may differ across groups of GD children and adolescents. With respect to their clinical interpretation, our findings are more likely to inform primary health care providers who first encounter GD children rather than practitioners specializing in gender-affirming care. In future studies, researchers should explore the possible reasons for the observed differences by recruiting a cohort with a wider range of sociodemographic characteristics, especially with regards to race and ethnicity, and include data on parental perceptions of transgender care and pubertal staging. It would be interesting to know how many of the youth presenting to care for GAHT have already socially transitioned, as well as examine deeper the effect of therapy on psychological functioning. Perhaps the most important next step in this area of research is to compare health outcomes and quality of life among GD children and adolescents who began receiving care at different ages. These types of data are needed to inform clinical practice and facilitate development of evidence-based guidelines.

ABBREVIATIONS

AFAB: assigned female sex at birth
AMAB: assigned male sex at birth
CI: confidence interval
GAHT: gender-affirming hormonal treatment
GD: gender-diverse
GDRD: gender dysphoria–related diagnosis
GnRH: gonadotropin-releasing hormone
HR: hazard ratio
KP: Kaiser Permanente
STRONG: Study of Transition, Outcomes, and Gender
WPATH: World Professional Association for Transgender Health

REFERENCES

1. Pleak RR. Gender identity issues in youth: opportunities, terminologies, histories, and advancements. *Child Adolesc Psychiatr Clin N Am.* 2011;20(4):601–625

2. American Psychiatric Association DSM-5 Task Force. *Diagnostic and Statistical Manual of Mental Disorders.* 5th ed. Washington, DC: American Psychiatric Association; 2013

3. Cohen-Kettenis PT, Pfäfflin F. The DSM diagnostic criteria for gender identity disorder in adolescents and adults. *Arch Sex Behav.* 2010;39(2):499–513

4. Beek TF, Cohen-Kettenis PT, Kreukels BPC. Gender incongruence/gender dysphoria and its classification history. *Int Rev Psychiatry.* 2016;28(1):5–12

5. Goldberg D. A revised mental health classification for use in general medical settings: the ICD11-PHC. *Int Psychiatry.* 2011;8(1):1–3

6. Steensma TD, Cohen-Kettenis PT, Zucker KJ. Evidence for a change in the sex ratio of children referred for gender dysphoria: data from the center of expertise on gender dysphoria in Amsterdam (1988-2016). *J Sex Marital Ther.* 2018;44(7):713–715

7. de Vries ALC, McGuire JK, Steensma TD, Wagenaar ECF, Doreleijers TAH, Cohen-Kettenis PT. Young adult psychological outcome after puberty suppression and gender reassignment. *Pediatrics.* 2014;134(4):696–704

8. Shumer DE, Spack NP. Current management of gender identity disorder in childhood and adolescence: guidelines, barriers and areas of controversy. *Curr Opin Endocrinol Diabetes Obes.* 2013;20(1):69–73

9. Hembree WC, Cohen-Kettenis PT, Gooren L, et al. Endocrine treatment of gender-dysphoric/gender-incongruent persons: an endocrine society clinical practice guideline. *Endocr Pract.* 2017;23(12):1437

10. Coleman E, Bockting W, Botzer M, et al. Standards of care for the health of transsexual, transgender, and gender-nonconforming people, version 7. *Int J Transgend.* 2012;13(4):165–232

11. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. *Arch Dis Child.* 1969;44(235):291–303

12. Marshall WA, Tanner JM. Variations in the pattern of pubertal changes in boys. *Arch Dis Child.* 1970;45(239):13–23

13. Olson J, Garofalo R. The peripubertal gender-dysphoric child: puberty suppression and treatment paradigms. *Pediatr Ann.* 2014;43(6):e132–e137

14. Costa R, Carmichael P, Colizzi M. To treat or not to treat: puberty suppression in childhood-onset gender dysphoria. *Nat Rev Urol.* 2016;13(8):456–462

15. de Vries ALC, Cohen-Kettenis PT. Clinical management of gender dysphoria in children and adolescents: the Dutch approach. *J Homosex.* 2012;59(3):301–320

16. Khatchadourian K, Amed S, Metzger DL. Clinical management of youth with gender dysphoria in Vancouver. *J Pediatr.* 2014;164(4):906–911

17. Owen-Smith AA, Gerth J, Sineath RC, et al. Association between gender confirmation treatments and perceived gender congruence, body image satisfaction, and mental health in a cohort of transgender individuals. *J Sex Med.* 2018;15(4):591–600

18. Menino DD, Katz-Wise SL, Vetters R, Reisner SL. Associations between the length of time from transgender identity recognition to hormone initiation and smoking among transgender youth and young adults. *Transgend Health.* 2018;3(1):82–87

19. Smith YLS, van Goozen SHM, Kuiper AJ, Cohen-Kettenis PT. Transsexual subtypes: clinical and theoretical significance. *Psychiatry Res.* 2005;137(3):151–160

20. Steensma TD, Biemond R, de Boer F, Cohen-Kettenis PT. Desisting and persisting gender dysphoria after childhood: a qualitative follow-up study. *Clin Child Psychol Psychiatry.* 2011;16(4):499–516

21. Wallien MSC, Cohen-Kettenis PT. Psychosexual outcome of gender-dysphoric children. *J Am Acad Child Adolesc Psychiatry.* 2008;47(12):1413–1423

22. Zuger B. Early effeminate behavior in boys. Outcome and significance for homosexuality. *J Nerv Ment Dis.* 1984;172(2):90–97

23. Steensma TD, van der Ende J, Verhulst FC, Cohen-Kettenis PT. Gender variance in childhood and sexual orientation in adulthood: a prospective study. *J Sex Med.* 2013;10(11):2723–2733

24. Tuerk C. Considerations for affirming gender nonconforming boys and their families: new approaches, new challenges. *Child Adolesc Psychiatr Clin N Am.* 2011;20(4):767–777

25. Fast AA, Olson KR. Gender development in transgender preschool children. *Child Dev.* 2018;89(2):620–637

26. Golombok S, Rust J, Zervoulis K, Croudace T, Golding J, Hines M. Developmental trajectories of sex-typed behavior in boys and girls: a longitudinal general population study of children aged 2.5-8 years. *Child Dev.* 2008;79(5):1583–1593

27. Martin CL, Ruble DN. Patterns of gender development. *Annu Rev Psychol.* 2010;61:353–381

28. Steensma TD, McGuire JK, Kreukels BPC, Beekman AJ, Cohen-Kettenis PT. Factors associated with desistence and persistence of childhood gender dysphoria: a quantitative follow-up study. *J Am Acad Child Adolesc Psychiatry.* 2013;52(6):582–590

29. Becerra-Fernández A, Rodríguez-Molina JM, Asenjo-Araque N, et al. Prevalence, incidence, and sex ratio of transsexualism in the autonomous region of Madrid (Spain) according to healthcare demand. *Arch Sex Behav.* 2017;46(5):1307–1312

30. Wiepjes CM, Nota NM, de Blok CJM, et al. The Amsterdam cohort of gender dysphoria study (1972-2015): trends in prevalence, treatment, and regrets. *J Sex Med.* 2018;15(4):582–590

31. Blosnich JR, Brown GR, Shipherd Phd JC, Kauth M, Piegari RI, Bossarte RM. Prevalence of gender identity disorder and suicide risk among transgender veterans utilizing veterans health administration care. *Am J Public Health.* 2013;103(10):e27–e32

32. Quinn VP, Nash R, Hunkeler E, et al. Cohort profile: study of transition, outcomes and gender (STRONG) to assess health status of transgender people. *BMJ Open.* 2017;7(12):e018121

33. Aydin D, Buk LJ, Partoft S, Bonde C, Thomsen MV, Tos T. Transgender surgery in Denmark from 1994 to 2015: 20-year follow-up study. *J Sex Med.* 2016;13(4):720–725

34. Gordon NP. *How Does the Adult Kaiser Permanente Membership in Northern California Compare with the Larger Community?* Oakland, CA: Kaiser Permanente Northern California Division of Research; 2006

35. Koebnick C, Langer-Gould AM, Gould MK, et al. Sociodemographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data. *Perm J.* 2012;16(3):37–41

36. Hembree WC, Cohen-Kettenis PT, Gooren L, et al. Endocrine treatment of gender-dysphoric/gender-incongruent persons: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab.* 2017;102(11):3869–3903

37. Becerra-Culqui TA, Liu Y, Nash R, et al. Mental health of transgender and gender nonconforming youth compared with their peers. *Pediatrics.* 2018;141(5):e20173845

38. Aitken M, Steensma TD, Blanchard R, et al. Evidence for an altered sex ratio in clinic-referred adolescents with gender dysphoria. *J Sex Med.* 2015;12(3):756–763

39. de Graaf NM, Carmichael P, Steensma TD, Zucker KJ. Evidence for a change in the sex ratio of children referred for gender dysphoria: data from the gender identity development service in London (2000-2017). *J Sex Med.* 2018;15(10):1381–1383

40. Goodman M, Adams N, Corneil T, Kreukels B, Motmans J, Coleman E. Size and distribution of transgender and gender nonconforming populations: a narrative review. *Endocrinol Metab Clin North Am.* 2019;48(2):303–321

41. Hatzenbuehler ML, Pachankis JE. Stigma and minority stress as social determinants of health among lesbian, gay, bisexual, and transgender youth: research evidence and clinical implications. *Pediatr Clin North Am.* 2016;63(6):985–997

42. Katz-Wise SL, Rosario M, Tsappis M. Lesbian, gay, bisexual, and transgender youth and family acceptance. *Pediatr Clin North Am.* 2016;63(6):1011–1025

43. Mavhandu-Mudzusi AH, Sandy PT. Religion-related stigma and discrimination experienced by lesbian, gay, bisexual and transgender students at a South African rural-based university. *Cult Health Sex.* 2015;17(8):1049–1056

44. Pham A, Morgan AR, Kerman H, et al. How are transgender and gender nonconforming youth affected by the news? a qualitative study. *J Adolesc Health.* 2020;66(4):478–483

45. Arnold E, Dhingra N. Health care inequities of sexual and gender minority patients. *Dermatol Clin.* 2020;38(2):185–190

46. Klein EW, Nakhai M. Caring for LGBTQ patients: methods for improving physician cultural competence. *Int J Psychiatry Med.* 2016;51(4):315–324

47. Suen LW, Lunn MR, Katuzny K, et al. What sexual and gender minority people want researchers to know about sexual orientation and gender identity questions: a qualitative study. *Arch Sex Behav.* 2020;49(7):2301–2318

48. Morris M, Cooper RL, Ramesh A, et al. Training to reduce LGBTQ-related bias among medical, nursing, and dental students and providers: a systematic review. *BMC Med Educ.* 2019;19(1):325

49. van Beusekom G, Baams L, Bos HMW, Overbeek G, Sandfort TGM. Gender nonconformity, homophobic peer victimization, and mental health: how same-sex attraction and biological sex matter. *J Sex Res.* 2016;53(1):98–108

50. van Beusekom G, Collier KL, Bos HMW, Sandfort TGM, Overbeek G. Gender nonconformity and peer victimization: sex and sexual attraction differences by age. *J Sex Res.* 2020;57(2):234–246

51. Barmania S, Aljunid SM. Transgender women in Malaysia, in the context of HIV and Islam: a qualitative study of stakeholders' perceptions. *BMC Int Health Hum Rights.* 2017;17(1):30

52. Barnett BS, Nesbit AE, Sorrentino RM. The transgender bathroom debate at the intersection of politics, law, ethics, and science. [published correction appears in *J Am Acad Psychiatry Law.* 2018;46(4):532]. *J Am Acad Psychiatry Law.* 2018;46(2): 232–241

53. Casey LS, Reisner SL, Findling MG, et al. Discrimination in the United States: Experiences of lesbian, gay, bisexual, transgender, and queer Americans. *Health Serv Res.* 2019;54(Suppl 2):1454–1466

54. Motmans J, Ponnet K, De Cuypere G. Sociodemographic characteristics of trans persons in Belgium: a secondary data analysis of medical, state, and social data. *Arch Sex Behav.* 2015;44(5):1289–1299

55. White Hughto JM, Murchison GR, Clark K, Pachankis JE, Reisner SL. Geographic and individual differences in healthcare access for U.S. transgender adults: a multilevel analysis. *LGBT Health.* 2016;3(6):424–433

56. Alencar Albuquerque G, de Lima Garcia C, da Silva Quirino G, et al. Access to health services by lesbian, gay, bisexual, and transgender persons: systematic literature review. *BMC Int Health Hum Rights.* 2016;16:2

57. Gonzales G, Henning-Smith C. Barriers to care among transgender and gender nonconforming adults. *Milbank Q.* 2017;95(4):726–748

58. The Lancet Diabetes Endocrinology. Transgender health: access to care under threat. *Lancet Diabetes Endocrinol.* 2018;6(6):427

59. Kane EW. "No way my boys are going to be like that!" parents' responses to children's gender nonconformity. *Gend Soc.* 2006;20(2):149–176

60. Cheng HL, Harris SR, Sritharan M, Behan MJ, Medlow SD, Steinbeck KS. The tempo of puberty and its relationship to adolescent health and well-being: a systematic review. *Acta Paediatr.* 2020;109(5):900–913

61. Lopez CM, Solomon D, Boulware SD, Christison-Lagay E. Trends in the

"off-label" use of GnRH agonists
among pediatric patients in the
United States. *Clin Pediatr (Phila)*.
2018;57(12):1432–1435