Epidemiology of Childhood Hyperthyroidism in France: A Nationwide Population-Based Study

Marie Simon,[1] Annabel Rigou,[1] Joëlle Le Moal,[1] Abdelkrim Zeghnoun,[1] Alain Le Tertre,[1] Perrine De Crouy-Chanel,[1] Florentia Kaguelidou,[2,3] and Juliane Leger[4,5,6]

[1]Santé publique France, Environmental Health Direction, Saint-Maurice 94415, France; [2]Department of Pediatric Pharmacology and Pharmacogenetics, Hôpital Universitaire Robert Debré, Assistance Publique-Hôpitaux de Paris, F-75019, Paris, France; [3]Clinical Investigations Center, INSERM CIC 1426; Université Paris Diderot, Sorbonne Paris Cité, UMR-1123 ECEVE, F-75019, Paris, France; [4]Department of Pediatric Endocrinology and Diabetes, Reference Centre for Endocrine Growth and Development Diseases, Hôpital Universitaire Robert Debré, Assistance Publique-Hôpitaux de Paris F-75019, Paris, France; [5]Université Paris Diderot, Sorbonne Paris Cité, F-75019, Paris, France; and [6]UMR 1141, DHU Protect, Institut National de la Santé et de la Recherche Médicale (Inserm), F-75019 Paris, France

Context: Hyperthyroidism affects all age groups, but epidemiological data for children are scarce.

Objective: To perform a nationwide epidemiological survey of hyperthyroidism in children and adolescents.

Design: A cross-sectional descriptive study.

Setting: Identification of entries corresponding to reimbursements for antithyroid drugs in the French national insurance database.

Participants: All cases of childhood hyperthyroidism (6 months to 17 years of age) in 2015.

Main Outcome Measures: National incidence rate estimated with a nonlinear Poisson model and spatial distribution of cases.

Results: A total of 670 cases of childhood hyperthyroidism were identified. Twenty patients (3%) had associated autoimmune or genetic disease, with type 1 diabetes and Down syndrome the most frequent. The annual incidence for 2015 was 4.58/100,000 person-years (95% CI 3.00 to 6.99/100,000). Incidence increased with age, in both sexes. This increase accelerated after the age of 8 in girls and 10 in boys and was stronger in girls. About 10% of patients were affected before the age of 5 years (sex ratio 1.43). There was an interaction between age and sex, the effect of being female increasing with age: girls were 3.2 times more likely to be affected than boys in the 10 to 14 years age group and 5.7 times more likely to be affected in the 15 to 17 years age group. No conclusions about spatial pattern emerged.

Conclusion: These findings shed light on the incidence of hyperthyroidism and the impact of sex on this incidence during childhood and adolescence. The observed incidence was higher than expected from the results published for earlier studies in Northern European countries. (***J Clin Endocrinol Metab*** **103: 2980–2987, 2018)**

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received 2 February 2018. Accepted 23 May 2018.
First Published Online 28 May 2018

Abbreviations: ATD, antithyroid drug; DCIR, Données de consummation inter-régimes; EDC, endocrine-disrupting chemical; GD, Graves disease; IR, incidence rate; NIR, numéro d'inscription au repertoire; PMSI, Programme de médicalisation des systèmes d'information; SNIIRAM, Système national d'information inter-régimes de l'Assurance maladie.

Hyperthyroidism in children is predominantly caused by Graves disease (GD), an autoimmune disorder caused by stimulation of the TSH receptor by autoantibodies specific for this receptor. Increases in the incidence of autoimmune diseases, including type 1 diabetes mellitus and celiac disease, in children and adolescents have been reported in recent years, and a similar increase has been suggested for hyperthyroidism (1–3). Like other autoimmune disorders, GD is thought to result from a complex interaction among genetic background, environmental factors (*e.g.*, iodine status), and the immune system. Moreover, patients often present more than one autoimmune disorder (4).

Endocrine-disrupting chemicals (EDCs) have been identified as potential environmental risk factors. A specific epidemiological program making use of existing databases to monitor nationwide trends for health indicators in the context of EDC exposure (5) in France was launched in 2014 (6). Autoimmune disorders frequently affect the thyroid gland, and thyroid disorders are specific outcomes of interest, as many EDCs act as thyroid disruptors. A few studies have identified molecules that induce increases in thyroid hormone levels: chlorinated pesticides (7–10), bisphenol A (11, 12), polybrominated flame retardants (13, 14), and perfluorinated chemicals (8). In addition to interacting directly with the thyroid gland, thyroid disruptors are thought to interact with the autoimmune system (15–19), potentially triggering thyroid autoimmune diseases, such as GD.

GD, a disease displaying female predominance, is rare and severe in children and adolescents. It is thought to account for 5% (20) of diagnosed cases of hyperthyroidism throughout life, but epidemiological data are limited. Previous studies were subject to selection bias, due to the selection criteria applied to the study populations: outpatients only (21), willingness of the pediatrician to report data (22), and limited study area (23). Moreover, population-based data are required in areas with adequate iodine levels.

GD may adversely affect the health and development of the child, and patients therefore require long-term follow-up. The first-line treatment of newly diagnosed children is based on the use of antithyroid drugs (ATDs) prescribed by clinicians. No other indication for these drugs is known in children, and this treatment is, therefore, specific. Alternative treatments, such as radioiodine and thyroidectomy, are proposed as a second line of treatment in cases of failure to control the disease after ATD treatment (4).

We therefore conducted a nationwide study to estimate the incidence of hyperthyroidism and describe its epidemiology in children and adolescents and analyze spatial trends in France in 2015.

Methods

Data source

We performed a population-based study with the French National Health Insurance Information System, Système national d'information inter-régimes de l'Assurance maladie (SNIIRAM) (24). This system comprehensively covers the entire population living in France (>66 million inhabitants) (25, 26). It records, in a specific database known as Données de consummation inter-régimes (DCIR; interscheme consumption data), anonymous and exhaustive data about patient reimbursement for health care expenditure, including reimbursements for drugs prescribed by clinicians. Drugs are identified in the database according to their Anatomical Therapeutic Classification codes. These data are relayed in real time from the various health insurance schemes. Another database, Programme de médicalisation des systèmes d'information (PMSI), included in SNIIRAM, provides medical information for all patients admitted to public and private hospitals in France, including discharge diagnoses recorded as International Statistical Classification of Diseases and Related Health Problems, 10th revision codes. Demographic data [age, sex, place of residence (département code), and affiliated scheme] are available from the DCIR and PMSI databases. All of the data for a given patient present in the SNIIRAM (DCIR and PMSI) can be linked through anonymous social security identification numbers [numéro d'inscription au repertoire (NIRs)].

The DCIR holds data for 3 years plus the current year. Data for the years 2013, 2014, 2015, and the first 6 months of July 2016 were, therefore, available to us at the time of the study.

The National Institute of Statistics and Economic Studies supplied population data.

Access to the SNIIRAM databases was authorized by the Institute of Health Data (Institut des Données de Santé) and the French data protection authority (Commission Nationale de l'Informatique et des Libertés).

Study population

The selected study population consisted of all individuals under the age of 18 years (age defined as the number of full years of life completed) living in France. Incident cases of hyperthyroidism were defined as the first reimbursement for ATDs (propylthiouracil, benzylthiouracil, carbimazole, and thiamazole) in 2015 recorded in the DCIR, with no reimbursement for ATDs in the 2 preceding years (2013 and 2014), to exclude the resumption of ATD treatment after relapses. Such relapses are common in children with hyperthyroidism, being observed in ~70% to 80% of all cases after an initial 2-year course of ATD treatment, with ~75% of patients relapsing within 6 months of the end of treatment and 95% relapsing within 18 months (27, 28).

We excluded three categories of patients: children under the age of 6 months in 2015, mostly considered to correspond to cases of neonatal transient thyrotoxicosis due to maternal GD (29). Given the difficulties distinguishing between incident cases and relapses, we excluded patients who had undergone surgery (total or subtotal thyroidectomy) or radioiodine treatment in 2015 (n = 1) by linking incident cases of hyperthyroidism to hospital stays including a medical act encoded KCFA010, KCFA009, KCFA005, KCFA007, KCNL003, and KCNL004 (French Common Classification of Medical Procedures codes). We ensured that only the population dwelling in France was

targeted by excluding records for which the area of residence (département) was not reported or for which the area of residence cited was located outside France (n = 1).

We identified the autoimmune diseases or related conditions associated with hyperthyroidism by linking incident cases of hyperthyroidism with cases of type 1 diabetes mellitus (International Statistical Classification of Diseases and Related Health Problems, 10th revision, E100 to E109), Turner syndrome (Q960 to Q964, Q968, and Q969), Down syndrome (Q900 to Q902 and Q909), DiGeorge syndrome (D821), celiac disease (K900), Addison disease (E271), or idiopathic thrombocytopenic purpura (D693) by cross-referencing with hospital data (PMSI).

Statistical analysis

We described the characteristics of the population and distribution of cases by sex and age group. The sex ratio (girls vs boys) was calculated for the total population and by age group. We explored month-to-month variations graphically to detect a possible seasonal influence on the disease that might be related to environmental exposure. We focused initially on the month of diagnosis and then on birth month.

We estimated incidence rates (IRs) and 95% CIs by Poisson regression analysis using population size as the offset. We used generalized linear models to take into account nonlinear relationships between IRs and the explanatory variables, age and sex. We introduced the age variable into the model as a spline function with three degrees of freedom to take the possibility of a nonlinear relationship into account, and we explored the interaction between age and sex. We assessed the validity of the model by analyzing the distribution of the residuals, which were found to be essentially normally distributed. No trend was seen, and the dispersion of residual values around the mean remained constant.

We estimated IRs for the overall population aged from 6 months to 17 years and by age and sex using the number of predicted cases in 2015 and midyear population estimates for children in France for 2015 (14.8 million children).

For comparison of our results with published values, we converted the age-specific IR values provided by the model into IR values for each age group. The rates for age treated as a continuous variable were calculated from Poisson regression model predictions. For the estimation of IRs by age group, we calculated the weighted mean of IRs for each age in the age group. The weighting used was the inverse of the variance of IR by age. The age groups of interest were 6 months to 4 years, 5 to 9 years, 10 to 14 years, and 15 to 17 years. We also estimated overall IRs for children under 15 years of age (12.3 million children).

We analyzed the spatial distribution of cases in France at the département level (a French administrative region, equivalent to a county), with four models of structured and unstructured spatial heterogeneity based on Poisson regression (30).

Statistical analyses were performed with SAS Guide 7.1 (SAS Institute, Cary, NC) and R for Poisson regression. The model for IR estimation was developed with the generalized linear model procedure implemented in the MGCV package (31) and the INLA model for spatial analyses using R software (32).

Results

Study population and patient characteristics

In total, we identified 670 children newly treated for hyperthyroidism in 2015: 157 boys (23.4%) and 513 girls (76.6%). A female preponderance was documented throughout childhood and adolescence, with a female-to-male sex ratio of 3.27. The ages of the affected children ranged from 6 months to 17 years, and mean age was higher in girls than in boys (Table 1).

The principal ATDs used to treat hyperthyroidism in children and adolescent in 2015 were carbimazole (63% cases), followed by thiamazole (31%), propylthiouracil (5%), and benzylthiouracil (1%). During the observation period, most of the patients (n = 626; 93%) used only one drug. However, 44 patients switched once (n = 40; 6%) or twice (n = 4; 1%) between treatments.

Twenty patients (3%) had associated autoimmune conditions: n = 1 in the 6 months to 4 years, n = 4 in the 5 to 9 years, n = 10 in the 10 to 14 years, and n = 5 in the 15 to 17 years age groups. Table 2 describes the types of autoimmune disease observed by age group and sex. The number of incident cases remained stable from month to month, regardless of whether the analysis was based on month of diagnosis or birth month (data not shown).

Incidence estimation

Figure 1 provides a comprehensive representation of the model. A statistically noteworthy increase in incidence with age was observed after the age of 5 years. Indeed, the inclusion of age or the interaction between age and sex in the Poisson regression model significantly decreased the residual deviance. This effect was visible on the plot, as the regression lines for girls and boys had different slopes, particularly for teenagers.

The increase in IR was more marked from the age of 8 years in girls and 10 years in boys, with a female preponderance observed at all ages. The CIs for IRs were larger for the youngest children, with considerable overlap between the sexes, making it difficult to draw any firm conclusions.

Table 3 presents the sex-specific IRs calculated for several age groups, based on Poisson model predictions. The overall IR for 2015 was 4.58/100,000 person-years (95% CI 3.00 to 6.99/100,000). It was 3.4 times higher

Table 1. Characteristics of the Subjects at the Time of First Reimbursement for ATD Treatment of Hyperthyroidism in Children and Adolescents in France for 2015

	Total (N = 670)	Boys (n = 157; 23.43%)	Girls (n = 513; 76.57%)
Mean age, y	12.5 (4.68)	11.2 (5.36)	13.5 (4.34)
Age group			
6 mo to 4 y	68 (10.20)	28 (17.83)	40 (7.80)
5–9 y	75 (11.20)	24 (15.29)	51 (9.94)
10–14 y	215 (32.10)	56 (35.67)	159 (30.99)
15–17 y	312 (46.60)	49 (31.21)	263 (51.27)

Data are mean (SD).

Table 2. Associated Autoimmune Diseases in Children and Adolescents First Treated for Hyperthyroidism in 2015 in France by Age Group and Sex

	Associated Disease		
	T1DM (N = 13)[a]	Celiac Disease (N = 2)[a]	Down Syndrome (N = 7)[a]
Age group			
6 mo to 4 y	—	—	1
5–9 y	4	1	1
10–14 y	7	—	3
15–17 y	2	1	2
Sex			
Boys	3	0	2
Girls	10	2	5

No cases of Addison disease, idiopathic thrombocytopenic purpura, Turner syndrome, or DiGeorge syndrome were reported.

Abbreviation: T1DM, type 1 diabetes mellitus.

[a]One patient had T1DM, celiac disease, and Down syndrome.

in girls than in boys [IRs: 7.08 (95% CI 4.85 to 10.30) vs 2.07 (95% CI 1.03 to 4.16)]. The overall IR for 2015 was markedly lower in children under the age of 15 years at 2.91/100,000 person-years (95% CI 2.05 to 4.13/100,000) and was 2.4 times higher in girls than in boys for this age group [IRs: 4.11 (2.91 to 5.83) vs 1.71 (1.00 to 2.93)]. In comparisons by age group, IRs were markedly higher in teenage girls than in teenage boys: 3.2 times higher for the 10 to 14 years age group and 5.7 times higher for the 15 to 17 years age group.

Spatial distribution

Crude IRs by département varied from 0/100,000 person-years to 12.5/100,000 person-years. The spatial models used did not fit our data well, and the results obtained were inconclusive. Some neighboring départements had discrepant incident case numbers, making it impossible to estimate IRs with the necessary level of precision for all départements.

Discussion

In this nationwide French study, our main objective was to improve understanding of the epidemiology of hyperthyroidism in children. We estimated the incidence of this disease from an indicator based on drug reimbursements. In 2015, the IR of hyperthyroidism in children was 4.58/100,000 person-years (2.91/100,000 person-years in children under 15), highlighting the rarity of this disease in children and adolescents. We described IR by age considered as a continuous variable using a nonlinear model to ensure that the most complete information was retained. The IR in girls was markedly higher than that in boys, and this difference increased strongly with age, particularly during the teenage years. The results of this study also extend our knowledge of the incidence of this disease in very young children, because ~10% of the patients in this cohort began receiving treatment before the age of 5 years.

Epidemiological data concerning childhood hyperthyroidism are scarce. Studies have been carried out in Northern Europe and China. Nationwide results were reported, except for the Chinese studies, and all of these studies reported incidence by age and sex for children under the age of 15 years. A Danish study based on the Danish National Patient Registry estimated the IR of childhood hyperthyroidism at 1.83/100,000 person-years for 2008 to 2012 (95% CI 1.47 to 2.25/100,000), the

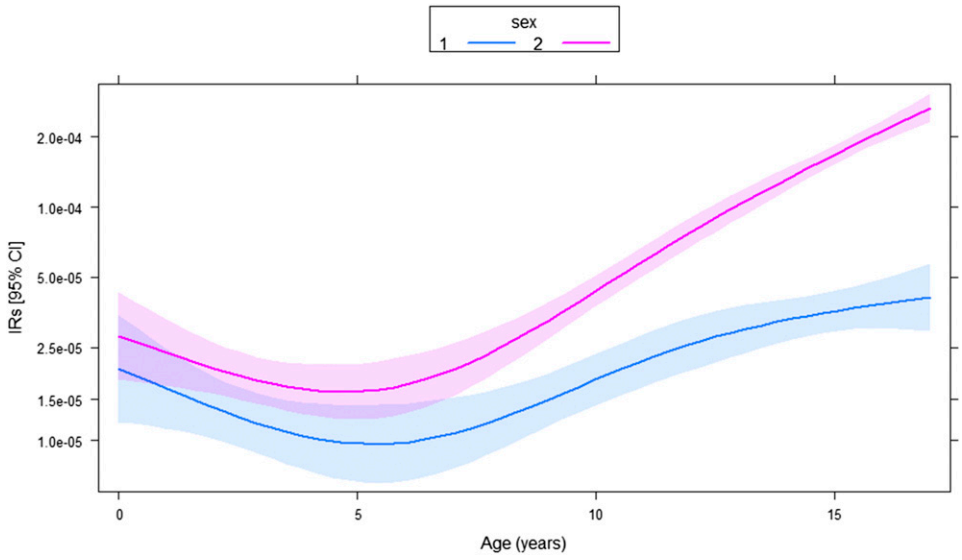


Figure 1. Predicted IRs by age and sex.

Table 3. Sex-Specific IRs and 95% CIs in Children and Adolescents First Treated for Hyperthyroidism in 2015 by Age Group

	Total		Boys		Girls	
2015	IR[a]	95% CI	IR[a]	95% CI	IR[a]	95% CI
6 mo to 17 y	4.58	3.00–6.99	2.07	1.03–4.16	7.08	4.85–10.30
6 mo to 14 y	2.91	2.05–4.13	1.71	1.00–2.93	4.11	2.91–5.83
6 mo to 4 y	1.53	1.14–2.05	1.30	1.19–1.42	1.80	1.54–2.10
5–9 y	1.85	1.48–2.33	1.26	1.17–1.34	2.25	1.99–2.53
10–14 y	5.28	4.52–6.15	2.48	2.32–2.65	7.89	6.90–9.02
15–17 y	12.03	10.06–14.38	3.78	3.49–4.10	21.53	18.96–24.45

Poisson regression analysis was used to estimate IR and CIs.

[a]Number of cases per 100,000 person-years.

study period closest to that for our study. The incidence reported in the Danish study was lower than that reported in this study by a factor of 1.6, but was calculated 5 years ago, and the CI overlaps with that for our findings. The Swedish studies that most closely resembled ours in those cases were selected based on diagnosis and treatment. It involved a retrospective analysis of medical records and computer-based registers. The authors of the Swedish study estimated the IR at 2.7/100,000 person-years, a value slightly lower than that obtained in this study (2.91/100,000), but with overlapping CIs, suggesting that it is consistent with our results. However, the Swedish study was performed 15 years before our study and included children under the age of 16 years. The two studies are not, therefore, strictly comparable. In a study performed in the Hong Kong area from 1989 to 1998 (33), with a registry of patients with childhood GD, the overall IR was found to be 5.0/100,000 person-years (95% CI 2.6 to 8.8) for children under the age of 15 years. From 1994 to 1998, the IR was 6.5/100,000 person years (95% CI 3.7 to 10.5), a figure much higher than that reported in this study, with no overlap of CIs, and for an earlier period. The authors speculated that their result might reflect high local levels of iodine. The last study, conducted in the United Kingdom and Ireland in 2004 to 2005 (22), considered all new cases of thyrotoxicosis in childhood, collected prospectively from pediatricians only. The overall IR was 0.9 per 100,000 person-years (95% CI 0.8 to 1.1), one-third lower than reported in this study, with no overlap between the CIs of the two studies. However, case reporting was almost certainly not exhaustive, because it depended on the voluntary participation of pediatricians.

Overall, our results are slightly higher, but within the range of previous results, except for those for Hong Kong. However, all of these studies were performed several years or even decades ago, and, to our knowledge, no updated data are available. As previously observed in the studies performed in Denmark, Sweden, and Hong Kong (3, 23, 33), the results suggest that the incidence of childhood hyperthyroidism in France may have increased, but further data are required to confirm this finding.

Other studies have also reported an increase in the frequency of hyperthyroidism with age, peaking during adolescence and affecting girls to a much greater extent than boys. However, the IRs by age group obtained in this study were higher than those reported in other European studies, for all age groups.

In our study, the sex ratio in children under the age of 15 years (girls vs boys, 2.37) was lower than those reported in Denmark (4.3) and China (9.7). A few studies have focused on sex ratio, and, to our knowledge, changes in sex ratio with age have never been investigated for childhood hyperthyroidism. The reason for this sex-related difference is unknown, but autoimmune diseases are generally more frequent in girls, due at least partly to the effects of estrogen (34, 35). Further studies are warranted to explore the crucial role of estrogen, particularly at the time at which serum estrogen levels increase during puberty in girls and, to a lesser extent, in males (34, 36).

Our findings also provide precise information about the prevalence of associated autoimmune conditions in children with hyperthyroidism. They are consistent with those of previous studies in children reporting an association of GD with type 1 diabetes, celiac disease, and Down syndrome (2, 28, 37–39). However, the epidemiological studies reporting these associations were based on a very limited study population of children with hyperthyroidism and on studies focusing mostly in adult patients, in whom the most common coexisting autoimmune diseases were similar to those reported in this study (40).

The strengths of our study include its population-based design, with a large study population including all of the incident cases of hyperthyroidism in children for which treatment was initiated in 2015 in France. As this serious disease is mostly well detected and always

managed, mostly with drugs in the first line, our indicator provides an exhaustive reflection of the cases of childhood hyperthyroidism treated in France. This study applies a nonlinear model to childhood hyperthyroidism to ensure the retention of as complete a set of information as possible for the explanatory variables. We were also able to study the associated autoimmune conditions at the time of hyperthyroidism diagnosis in the various groups defined based on age and sex.

However, this study also presents several inherent limitations. We cannot exclude the possibility that some of the patients in remission experienced a relapse of hyperthyroidism >24 months after the end of ATD treatment, as reported in adult studies, in which a few relapses have been reported to occur as much as 5 years after the end of ATD treatment (41, 42), although such late relapses are considered unlikely and have never been described in children. It is also possible that some of our patients with hyperthyroidism will go on to develop other autoimmune disorders later in life.

Given the rarity of this disease in very young children, the absolute numbers of very young patients were very low, and this may have affected the estimation of incidence. Moreover, the etiological diagnosis of hyperthyroidism in each individual patient was not validated with patient charts and by differential diagnosis. Exceptional causes of nonautoimmune hyperthyroidism, such as those related to activating mutations of the TSH receptor gene, with severe forms in children diagnosed early in life (43), may have led to incidence being overestimated in very young children.

Another limitation of the indicator used in this study is the potential existence of duplicate cases generated by differences in the identification number (NIR) under which the child is registered. In most cases, the NIR assigned to each child is unique and remains associated with them throughout their lifetime. Otherwise, children may be recorded under several different NIRs, such as those of their parents (both parents for example) to facilitate access to care, leading to possible duplicates and the overestimation of cases. Duplicates undoubtedly accounted for only a very small proportion of the population and are unlikely to have influenced the results. In our study, 8% of the cases were recorded under two different NIRs and 92% under a unique, single NIR. Of the 8% of cases recorded under two different NIRs, 4.5% followed treatment under both NIRs. We therefore assessed the potential overestimation of hyperthyroidism cases at 0.3% (0.08 × 0.04). We therefore believe that our IR estimates are robust.

This study was also limited by the availability of data for only a short period (3 years plus the year underway), restricting data analyses to a single year. We were, therefore, unable to study temporal trends, a prerequisite for any demonstration of an increase in incidence in France.

We now plan to use archived data from SNIIRAM to study cases over a 10-year period. This approach will make it possible to analyze temporal and spatial trends and develop hypotheses concerning possible causes (*e.g.*, the role of EDC exposure).

Conclusion

In conclusion, this population-based study provides epidemiological data for hyperthyroidism in children and adolescents and the estimates of its incidence in France, with robust results. It provides updated results and original data for incidence by age and sex, highlighting the increase in incidence and sex ratio in teenagers. French IRs are slightly higher than expected from the results published for earlier studies in Northern European countries, and further studies are required to determine whether incidence is actually increasing in France. If such an increase is confirmed, it will be interesting to see whether it also concerns the youngest children, as shown, over the last decade for other autoimmune diseases, such as type 1 diabetes (2, 44). The association of hyperthyroidism with other autoimmune conditions and hypotheses concerning links to environmental factors should be investigated over a longer study period.

Acknowledgments

The authors thank colleagues at Santé publique France for help with this study: Marjorie Boussac, Edwige Bertrand, Mélina Le Barbier, and Sébastien Denys.

Correspondence and Reprint Requests: Joëlle Le Moal, MD, MPH, Santé publique France, Environmental Health Direction, 12 rue du Val d'Osne, Saint-Maurice 94415, France. E-mail: joelle.lemoal@santepubliquefrance.fr.

Disclosure Summary: The authors have nothing to disclose.

References

1. Dydensborg S, Toftedal P, Biaggi M, Lillevang ST, Hansen DG, Husby S. Increasing prevalence of coeliac disease in Denmark: a linkage study combining national registries. *Acta Paediatr*. 2012; **101**(2):179–184.
2. Mayer-Davis EJ, Lawrence JM, Dabelea D, Divers J, Isom S, Dolan L, Imperatore G, Linder B, Marcovina S, Pettitt DJ, Pihoker C, Saydah S, Wagenknecht L; SEARCH for Diabetes in Youth Study. Incidence trends of type 1 and type 2 diabetes among youths, 2002-2012. *N Engl J Med*. 2017;**376**(15):1419–1429.
3. Havgaard Kjær R, Smedegård Andersen M, Hansen D. Increasing incidence of juvenile thyrotoxicosis in Denmark: a nationwide study, 1998-2012. *Horm Res Paediatr*. 2015;**84**(2):102–107.

4. Léger J, Kaguelidou F, Alberti C, Carel JC. Graves' disease in children. *Best Pract Res Clin Endocrinol Metab*. 2014;**28**(2): 233–243.
5. French Government. Stratégie nationale sur les perturbateurs endocriniens. Available at: www.ecologique-solidaire.gouv.fr/sites/default/files/Strategie Nationale Perturbateurs Endocriniens 29 avril 2014.pdf. Accessed 22 June 2018.
6. Le Moal J, Sharpe RM, Jørgensen N, Levine H, Jurewicz J, Mendiola J, Swan SH, Virtanen H, Christin-Maître S, Cordier S, Toppari J, Hanke W; HURGENT Network. Toward a multi-country monitoring system of reproductive health in the context of endocrine disrupting chemical exposure. *Eur J Public Health*. 2016;**26**(1):76–83.
7. Croes K, Den Hond E, Bruckers L, Govarts E, Schoeters G, Covaci A, Loots I, Morrens B, Nelen V, Sioen I, Van Larebeke N, Baeyens W. Endocrine actions of pesticides measured in the Flemish environment and health studies (FLEHS I and II). *Environ Sci Pollut Res Int*. 2015;**22**(19):14589–14599.
8. de Cock M, de Boer MR, Lamoree M, Legler J, van de Bor M. Prenatal exposure to endocrine disrupting chemicals in relation to thyroid hormone levels in infants - a Dutch prospective cohort study. *Environ Health* 2014;**13**:106.
9. Freire C, Koifman RJ, Sarcinelli P, Rosa AC, Clapauch R, Koifman S. Long term exposure to organochlorine pesticides and thyroid function in children from Cidade dos Meninos, Rio de Janeiro, Brazil. *Environ Res*. 2012;**117**:68–74.
10. Meeker JD. Exposure to environmental endocrine disruptors and child development. *Arch Pediatr Adolesc Med*. 2012;**166**(10): 952–958.
11. Kim JH, Rozek LS, Soliman AS, Sartor MA, Hablas A, Seifeldin IA, Colacino JA, Weinhouse C, Nahar MS, Dolinoy DC. Bisphenol A-associated epigenomic changes in prepubescent girls: a cross-sectional study in Gharbiah, Egypt. *Environ Health*. 2013;**12**(1):33.
12. Zoeller RT, Bansal R, Parris C. Bisphenol-A, an environmental contaminant that acts as a thyroid hormone receptor antagonist in vitro, increases serum thyroxine, and alters RC3/neurogranin expression in the developing rat brain. *Endocrinology*. 2005; **146**(2):607–612.
13. Andra SS, Makris KC. Thyroid disrupting chemicals in plastic additives and thyroid health. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev*. 2012;**30**(2):107–151.
14. Turyk ME, Persky VW, Imm P, Knobeloch L, Chatterton R Jr, Anderson HA. Hormone disruption by PBDEs in adult male sport fish consumers. *Environ Health Perspect*. 2008;**116**(12):1635–1641.
15. Brent GA. Environmental exposures and autoimmune thyroid disease. *Thyroid*. 2010;**20**(7):755–761.
16. Brucker-Davis F, Hiéronimus S, Fénichel P. [Thyroid and the environment]. *Presse Med*. 2016;**45**(1):78–87.
17. Burek CL, Talor MV. Environmental triggers of autoimmune thyroiditis. *J Autoimmun*. 2009;**33**(3-4):183–189.
18. Dietert RR. Developmental immunotoxicity, perinatal programming, and noncommunicable diseases: focus on human studies. *Adv Med* 2014;**2014**:867805.
19. Winans B, Humble MC, Lawrence BP. Environmental toxicants and the developing immune system: a missing link in the global battle against infectious disease? *Reprod Toxicol*. 2011;**31**(3): 327–336.
20. Abraham-Nordling M, Bystrom K, Torring O, Lantz M, Berg G, Calissendorff J, Nystrom HF, Jansson S, Jorneskog G, Karlsson FA, Nystrom E, Ohrling H, Orn T, Hallengren B, Wallin G. Incidence of hyperthyroidism in Sweden. *Eur J Endocrinol* 2011;**165**(6): 899–905.
21. Lavard L, Ranlov I, Perrild H, Andersen O, Jacobsen BB. Incidence of juvenile thyrotoxicosis in Denmark, 1982-1988. A nationwide study. *Eur J Endocrinol* 1994;**130**(6):565–568.
22. Williamson S, Greene SA. Incidence of thyrotoxicosis in childhood: a national population based study in the UK and Ireland. *Clin Endocrinol (Oxf)*. 2010;**72**(3):358–363.
23. Forssberg M, Arvidsson CG, Engvall J, Lindblad C, Snellman K, Aman J. Increasing incidence of childhood thyrotoxicosis in a population-based area of central Sweden. *Acta Paediatr*. 2004; **93**(1):25–29.
24. Moulis G, Lapeyre-Mestre M, Palmaro A, Pugnet G, Montastruc JL, Sailler L. French health insurance databases: what interest for medical research? *Rev Med Interne*. 2015;**36**(6): 411–417.
25. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, Maura G, Caillol H, Tajahmady A, Coste J, Gissot C, Weill A, Fagot-Campagna A. Value of a national administrative database to guide public decisions: From the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. *Rev Epidemiol Sante Publique*. 2017;**65**(Suppl 4): S149–S167.
26. Goldberg M. [Promoting the use of the French national health database (SNIIRAM)]. *Rev Epidemiol Sante Publique*. 2017; **65**(Suppl 4):S141–S143.
27. Léger J, Gelwane G, Kaguelidou F, Benmerad M, Alberti C; French Childhood Graves' Disease Study Group. Positive impact of long-term antithyroid drug treatment on the outcome of children with Graves' disease: national long-term cohort study. *J Clin Endocrinol Metab*. 2012;**97**(1):110–119.
28. Kaguelidou F, Alberti C, Castanet M, Guitteny MA, Czernichow P, Léger J; French Childhood Graves' Disease Study Group. Predictors of autoimmune hyperthyroidism relapse in children after discontinuation of antithyroid drug treatment. *J Clin Endocrinol Metab*. 2008;**93**(10):3817–3826.
29. Léger J. Management of fetal and neonatal Graves' disease. *Horm Res Paediatr*. 2017;**87**(1):1–6.
30. Besag J, York J, Mollie A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Statist Math*. 1991;**43**(1): 1–20.
31. Wood SN. *Generalized Additive Models: An Introduction with R*. Boca Raton, FL: Chapman&Hall/CRC ed.; 2006.
32. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: Foundation for Statistical Computing; 2014.
33. Wong GWK, Cheng PS. Increasing incidence of childhood Graves' disease in Hong Kong: a follow-up study. *Clin Endocrinol (Oxf)*. 2001;**54**(4):547–550.
34. Dragin N, Bismuth J, Cizeron-Clairac G, Biferi MG, Berthault C, Serraf A, Nottin R, Klatzmann D, Cumano A, Barkats M, Le Panse R, Berrih-Aknin S. Estrogen-mediated downregulation of AIRE influences sexual dimorphism in autoimmune diseases. *J Clin Invest*. 2016;**126**(4):1525–1537.
35. Whitacre CC. Sex differences in autoimmune disease. *Nat Immunol*. 2001;**2**(9):777–780.
36. Chailurkit LO, Aekplakorn W, Ongphiphadhanakul B. The relationship between circulating estradiol and thyroid autoimmunity in males. *Eur J Endocrinol* 2013;**170**(1):63–67.
37. De Luca F, Corrias A, Salerno M, Wasniewska M, Gastaldi R, Cassio A, Mussa A, Aversa T, Radetti G, Arrigo T. Peculiarities of Graves' disease in children and adolescents with Down's syndrome. *Eur J Endocrinol* 2010;**162**(3):591–595.
38. Lombardo F, Messina MF, Salzano G, Rabbone I, Lo Presti D, Calcaterra V, Aversa T, De Luca F, Wasniewska M. Prevalence, presentation and clinical evolution of Graves' disease in children and adolescents with type 1 diabetes mellitus. *Horm Res Paediatr*. 2011;**76**(4):221–225.
39. Miskiewicz P, Gos-Zajac A, Kurylowicz A, Plazinska TM, Franaszczyk M, Bartoszewicz Z, Kondracka A, Pirko-Kotela K, Rupinski M, Jarosz D, Regula J, Ploski R, Bednarczuk T. Hla Dq2 haplotype, early onset of Graves disease, and positive family history of autoimmune disorders are risk factors for developing celiac disease in patients with Graves disease. *Endocr Pract*. 2015;**21**(9): 993–1000.

40. Boelaert K, Newby PR, Simmonds MJ, Holder RL, Carr-Smith JD, Heward JM, Manji N, Allahabadia A, Armitage M, Chatterjee KV, Lazarus JH, Pearce SH, Vaidya B, Gough SC, Franklyn JA. Prevalence and relative risk of other autoimmune diseases in subjects with autoimmune thyroid disease. *Am J Med*. 2010; **123**(2):183.e1-9.
41. Benker G, Reinwein D, Kahaly G, Tegler L, Alexander WD, Fassbinder J, Hirche H; The European Multicentre Trial Group of the Treatment of Hyperthyroidism with Antithyroid Drugs. Is there a methimazole dose effect on remission rate in Graves' disease? Results from a long-term prospective study. *Clin Endocrinol (Oxf)*. 1998;**49**(4):451–457.
42. Vitti P, Rago T, Chiovato L, Pallini S, Santini F, Fiore E, Rocchi R, Martino E, Pinchera A. Clinical features of patients with Graves' disease undergoing remission after antithyroid drug treatment. *Thyroid*. 1997;**7**(3):369–375.
43. Léger J, Carel JC. Hyperthyroidism in childhood: causes, when and how to treat. *J Clin Res Pediatr Endocrinol*. 2013;**5**(Suppl 1): 50–56.
44. Bruno G, Maule M, Biggeri A, Ledda A, Mannu C, Merletti F, Songini M; Sardinian Group for Diabetes Epidemiology. More than 20 years of registration of type 1 diabetes in Sardinian children: temporal variations of incidence with age, period of diagnosis, and year of birth. *Diabetes*. 2013;**62**(10):3542–3546.

ORIGINAL ARTICLE

A Copeptin-Based Approach in the Diagnosis of Diabetes Insipidus

W. Fenske, J. Refardt, I. Chifu, I. Schnyder, B. Winzeler, J. Drummond, A. Ribeiro-Oliveira, Jr., T. Drescher, S. Bilz, D.R. Vogt, U. Malzahn, M. Kroiss, E. Christ, C. Henzen, S. Fischli, A. Tönjes, B. Mueller, J. Schopohl, J. Flitsch, G. Brabant, M. Fassnacht, and M. Christ-Crain

ABSTRACT

BACKGROUND

The indirect water-deprivation test is the current reference standard for the diagnosis of diabetes insipidus. However, it is technically cumbersome to administer, and the results are often inaccurate. The current study compared the indirect water-deprivation test with direct detection of plasma copeptin, a precursor-derived surrogate of arginine vasopressin.

METHODS

From 2013 to 2017, we recruited 156 patients with hypotonic polyuria at 11 medical centers to undergo both water-deprivation and hypertonic saline infusion tests. In the latter test, plasma copeptin was measured when the plasma sodium level had increased to at least 150 mmol per liter after infusion of hypertonic saline. The primary outcome was the overall diagnostic accuracy of each test as compared with the final reference diagnosis, which was determined on the basis of medical history, test results, and treatment response, with copeptin levels masked.

RESULTS

A total of 144 patients underwent both tests. The final diagnosis was primary polydipsia in 82 patients (57%), central diabetes insipidus in 59 (41%), and nephrogenic diabetes insipidus in 3 (2%). Overall, among the 141 patients included in the analysis, the indirect water-deprivation test determined the correct diagnosis in 108 patients (diagnostic accuracy, 76.6%; 95% confidence interval [CI], 68.9 to 83.2), and the hypertonic saline infusion test (with a copeptin cutoff level of >4.9 pmol per liter) determined the correct diagnosis in 136 patients (96.5%; 95% CI, 92.1 to 98.6; P<0.001). The indirect water-deprivation test correctly distinguished primary polydipsia from partial central diabetes insipidus in 77 of 105 patients (73.3%; 95% CI, 63.9 to 81.2), and the hypertonic saline infusion test distinguished between the two conditions in 99 of 104 patients (95.2%; 95% CI, 89.4 to 98.1; adjusted P<0.001). One serious adverse event (desmopressin-induced hyponatremia that resulted in hospitalization) occurred during the water-deprivation test.

CONCLUSIONS

The direct measurement of hypertonic saline–stimulated plasma copeptin had greater diagnostic accuracy than the water-deprivation test in patients with hypotonic polyuria. (Funded by the Swiss National Foundation and others; ClinicalTrials.gov number, NCT01940614.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Christ-Crain at the Department of Endocrinology, Diabetes and Metabolism, University Hospital Basel, Petersgraben 4, 4031 Basel, Switzerland, or at mirjam.christ-crain@unibas.ch, or to Dr. Fenske at the Department of Endocrinology and Nephrology, University of Leipzig, Liebigstrasse 21, 04103 Leipzig, Germany, or at wiebkekristin.fenske@medizin.uni-leipzig.de.

*Drs. Fenske and Refardt contributed equally to this article.

N Engl J Med 2018;379:428-39.
DOI: 10.1056/NEJMoa1803760

The determination of a specific diagnosis in patients with polyuria and low plasma osmolality (i.e., hypotonic polyuria) is a frequent problem in clinical practice. In the absence of osmotic diuresis, polyuria can result from one of three fundamentally different conditions[1]: insufficient production and secretion of the antidiuretic hormone arginine vasopressin (central diabetes insipidus), diminished renal sensitivity to the antidiuretic activity of arginine vasopressin (nephrogenic diabetes insipidus), or primary excessive fluid intake (primary polydipsia).

It is important to differentiate these entities because treatments differ substantially, and incorrect strategies may lead to severe complications.[2,3] The indirect water-deprivation test measures the maximal urine concentration during prolonged withholding of oral liquids and the renal response to administered desmopressin.[4] It is conceptually simple, but difficulties in interpretation are common, mainly because any water diuresis may compromise the renal medullary concentration gradient[1,5,6] and promote a downregulation of kidney aquaporin-2 water channels, which could potentially affect the diagnostic value of these urinary measures.[5] Previous attempts to improve the diagnosis of polyuric disorders with direct measurement of circulating arginine vasopressin[7-10] failed to gain traction in clinical practice, largely because of the technical difficulties of measuring arginine vasopressin.[11-14]

Copeptin, the C-terminal segment of the arginine vasopressin prohormone, is an arginine vasopressin surrogate with high ex vivo stability that is easy to measure.[12,15,16] In previous studies, we reported outcome data that suggested that measurement of osmotically stimulated copeptin might be useful in differentiating the various causes of hypotonic polyuria.[5,17-19] The current study assessed the diagnostic performance of a test measuring copeptin that was osmotically stimulated by water deprivation or by hypertonic saline infusion as compared with the indirect water-deprivation test.

Methods

Study Design and Patients

This international, multicenter, prospective study was conducted at 11 tertiary medical centers in Switzerland, Germany, and Brazil from July 2013 to June 2017; the 3-month follow-up visits were completed by September 2017. We recruited 156 patients 16 years of age or older with hypotonic polyuria (a urine output of >50 ml per kilogram of body weight during a 24-hour period, with a urine osmolality <800 mOsm per kilogram) or with a confirmed diagnosis of central diabetes insipidus. Three patients were excluded from the analyses because they were found to have nephrogenic diabetes insipidus, and 12 patients were excluded for other reasons (Fig. S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org).

The local ethics committees at all centers approved the study protocol (available at NEJM.org). Written informed consent was obtained from all patients or from a legal guardian, when applicable. Laboratory measurement of copeptin was funded by Thermo Fisher Scientific, which had no other role in the study; there was no other commercial support for the study. All the authors vouch for the accuracy and completeness of the data and for the fidelity of the study to the protocol.

Procedures at Baseline

The water-deprivation and hypertonic saline infusion tests were performed on separate days. After a detailed medical history was obtained, a standardized clinical and biochemical evaluation was performed. Magnetic resonance imaging (MRI) of the head was performed at the discretion of the attending physician, although it was recommended in all patients if imaging had not been performed within 3 months before study enrollment. Diuretic or antidiuretic medications were discontinued for at least 24 hours before each test, and smoking and alcohol were prohibited for at least 12 hours before each test.

Test Protocols

Indirect Water-Deprivation Test

As is standard for the water-deprivation test,[4] a 17-hour fluid restriction started at midnight, or at 6 a.m. in patients with known or suspected complete diabetes insipidus. Every 2 hours, vital signs and body weight were monitored and urine was collected for measurement of volume and osmolality. Blood samples were obtained at 8 a.m. and immediately before the administration of desmopressin (1 hour before the end of the test).

For safety reasons, the water-deprivation test was stopped early in patients who met one of the following criteria: a decrease in body weight of more than 3%, symptoms of orthostatic hypotension with an increase in heart rate or a decrease in mean arterial blood pressure of more than 15%, or an increase in plasma sodium level of 150 mmol or more per liter. At 4 p.m., or when the test was stopped, each patient received 2 μg of desmopressin intravenously, and a final urine specimen for osmolality measurement was obtained at least 60 minutes thereafter.

Hypertonic Saline Infusion Test

Patients underwent the hypertonic saline infusion test between 8 a.m. and 11 a.m., as described previously.[16] An initial 250-ml bolus infusion of 3% saline was administered, and the infusion was continued at a rate of 0.15 ml per kilogram per minute. Blood samples for the measurement of plasma osmolality and sodium, urea, and glucose levels were obtained every 30 minutes, and sodium levels were monitored by venous blood gas analysis until the target level of at least 150 mmol per liter was reached. Thereafter, a final blood sample for plasma copeptin measurement was obtained, and patients were given water orally (30 ml of water per kilogram) within 30 minutes, followed by a 500-ml infusion of 5% glucose within 40 to 60 minutes after the patients received water. For safety reasons, the plasma sodium level was measured again 1 hour after the start of the glucose infusion to ensure that the level was within the normal range before the patient was discharged.

ADVERSE EVENTS AND SYMPTOM BURDEN

Adverse events during both tests were strictly documented, and clinical symptoms were rated by patients according to a visual-analogue scale that ranged from 0 to 10, with 0 indicating no symptoms and 10 indicating the most severe symptoms imaginable. Additional details are provided in the Supplementary Appendix.

TEST INTERPRETATION AND PRELIMINARY DIAGNOSIS

After the patients had completed both tests, they were discharged from the hospital with a preliminary diagnosis and treatment recommendation that were based on best current clinical practice. A follow-up visit was scheduled for 3 months later to assess response to treatment and clinical outcome and to reevaluate the accuracy of the preliminary diagnosis.

DIAGNOSTIC CRITERIA

Indirect Water-Deprivation Test

In accordance with the original description of the indirect water-deprivation test[4] and the subsequent modification,[5,8,20] complete central diabetes insipidus was diagnosed in patients who had a maximum urine osmolality of less than 300 mOsm per kilogram and an increase in urine osmolality of more than 50% after administration of desmopressin. Partial central diabetes insipidus was diagnosed in patients who had a maximum urine osmolality of 300 to 800 mOsm per kilogram and an increase in urine osmolality of 9 to 50% after administration of desmopressin. Primary polydipsia was diagnosed in patients who had a maximum urine osmolality of 300 to 800 mOsm per kilogram and an increase in urine osmolality of less than 9% after administration of desmopressin.

Plasma Copeptin Stimulated by Water Deprivation

Previous data suggested that the diagnostic accuracy of the indirect water-deprivation test could be improved by the additional measurement of baseline (at 8 a.m.) as well as stimulated (before administration of desmopressin) plasma copeptin levels.[5] According to those results, prespecified cutoff levels were used: a ratio of stimulated copeptin (the change in copeptin level over 8 hours during water deprivation, in picomoles per liter) to plasma sodium (measured at the end of the test in millimoles per liter) of 0.02 pmol or more per liter indicated primary polydipsia, and a basal plasma copeptin level of less than 2.6 pmol per liter indicated complete central diabetes insipidus. A ratio of less than 0.02 pmol per liter indicated partial central diabetes insipidus.

Plasma Copeptin Stimulated by Hypertonic Saline Infusion

The diagnostic criteria for hypertonic saline–stimulated copeptin in distinguishing primary polydipsia from central diabetes insipidus were suggested previously by our group[19] and were used in this study. A plasma copeptin cutoff level of 4.9 pmol or less per liter indicated complete or

partial central diabetes insipidus, and a level greater than 4.9 pmol per liter indicated primary polydipsia.

FINAL REFERENCE DIAGNOSIS

In the absence of a diagnostic standard, the final reference diagnosis was determined after the study was completed by two independent board-certified experts in endocrinology, who were unaware of the copeptin levels, after careful consideration of each patient's medical history and clinical symptoms, the results of the water-deprivation test, the available laboratory and imaging data, and the therapeutic response at the 3-month follow-up. In the event of discordant diagnoses (which occurred in 4 of 144 patients), a third expert was consulted, and results were discussed until a consensus was reached.

LABORATORY MEASUREMENTS

Blood samples were obtained and processed for measurement of plasma copeptin and for routine laboratory measurements (urine and plasma osmolality, hematocrit, and plasma sodium, potassium, creatinine, urea, calcium, albumin, glucose, and hemoglobin levels). Plasma copeptin was measured centrally in one batch with the use of a commercial automated immunofluorescence assay (B.R.A.H.M.S KRYPTOR Copeptin proAVP, Thermo Scientific Biomarkers). Details are provided in the Supplementary Appendix.

STATISTICAL ANALYSIS

The primary end point was the overall diagnostic accuracy — the percentage of correctly diagnosed patients — in the differentiation of central diabetes insipidus from primary polydipsia. Only patients with a final diagnosis were included in the analysis; however, the three patients with nephrogenic diabetes insipidus were only descriptively assessed. Details concerning the full-analysis population and the per-protocol population, as well as additional statistical details, are provided in the Supplementary Appendix.

The primary objectives were first to determine whether the measurement of copeptin during hypertonic saline infusion and during water deprivation was superior to the indirect water-deprivation test, and then to determine whether copeptin measurement during hypertonic saline infusion was noninferior to copeptin measurement during water deprivation; the second objective would be tested only if superiority could be shown for the first objective. The primary hypothesis thus consisted of two components, with a two-step statistical testing procedure. Sample size was estimated for the noninferiority test: assuming a diagnostic accuracy of 90% for water-deprivation–stimulated copeptin[5] and a noninferiority margin of 10%, a total of 115 patients who could be evaluated would provide 90% power to establish the noninferiority of hypertonic saline–stimulated copeptin measurement to water-deprivation–stimulated copeptin measurement. To assess whether the diagnostic accuracy varied depending on the severity of central diabetes insipidus, a prespecified subgroup analysis was performed to assess the diagnostic accuracy of hypertonic saline–stimulated copeptin and water-deprivation–stimulated copeptin as compared with the indirect water-deprivation test in specifically distinguishing primary polydipsia from partial central diabetes insipidus.

RESULTS

BASELINE CHARACTERISTICS

Of the 141 patients (66% female) included in the analyses, 82 (58%) received a final diagnosis of primary polydipsia after all 3-month follow-ups were completed, and 59 (42%) received a diagnosis of central diabetes insipidus (Fig. S1 in the Supplementary Appendix). Among the 59 patients who received a diagnosis of central diabetes insipidus, complete central diabetes insipidus was diagnosed in 36 patients (61%), and partial central diabetes insipidus in 23 (39%).

There were significant differences between the groups in some baseline characteristics (Table 1). Results of MRI of the head were available for 97 patients. The hyperintense signal in the posterior region on T_1-weighted images, which is considered to be a physiological signal that indicates the pituitary arginine vasopressin content,[21,22] was absent in 70% of the patients with central diabetes insipidus, but it was also absent in 39% of the patients with primary polydipsia (Table 1).

PRIMARY OUTCOME

The overall diagnostic accuracy of the hypertonic saline infusion test was significantly higher than that of the indirect water-deprivation test

Table 1. Baseline Characteristics of the Patients with Hypotonic Polyuria, According to Final Reference Diagnosis.*

Characteristic	Central Diabetes Insipidus			Primary Polydipsia (N=82)
	Complete Diabetes Insipidus (N=36)	Partial Diabetes Insipidus (N=23)	All Central Diabetes Insipidus (N=59)	
Median age (IQR) — yr	48 (39–53)	43 (30–48)	45 (33–53)	32 (24–44)
Female sex — no. (%)	23 (64)	15 (65)	38 (64)	55 (67)
Median body-mass index (IQR)†	29 (25–31)	26 (23–32)	28 (24–31)	24 (21–26)
Clinical symptoms				
Polydipsia — median liters consumed/day (IQR)	6.0 (5.0–8.6)	5.5 (4.0–6.5)	6.0 (4.8–8.0)	5.0 (4.5–6.9)
Polyuria — median liters of urine/day (IQR)	6.0 (4.9–8.0)	4.0 (3.5–6.0)	5.0 (4.0–7.8)	5.0 (4.0–6.0)
Nocturia — no. (%)	35 (97)	21 (91)	56 (95)	56 (68)
Median no. of events/night (IQR)	3 (2–4)	3 (2–5)	3 (2–4)	1 (0–3)
Drinking at night — no. (%)	33 (92)	21 (91)	54 (92)	51 (62)
Median liters consumed/night (IQR)	1.5 (1.0–2.0)	1.0 (0.5–2.0)	1.3 (0.9–2.0)	0.2 (0–0.6)
Preference for cold drinks — no. (%)	26 (72)	18 (78)	44 (75)	49 (60)
Preference for water — no./total no. (%)	32/34 (94)	19/22 (86)	51/56 (91)	73/80 (91)
Sudden onset of symptoms — no. (%)	22 (61)	15 (65)	37 (63)	18 (22)
Persistent symptoms — no. (%)	34 (94)	22 (96)	56 (95)	64 (78)
Polyuria preceding manifestation of polydipsia — no. (%)	3 (8)	2 (9)	5 (9)	3 (4)
Onset of symptoms coinciding with stressful emotional or physical event — no./total no. (%)	2/36 (6)	7/23 (30)	9/59 (15)	15/81 (19)
Medical history — no. (%)				
Brain tumor	17 (47)	12 (52)	29 (49)	5 (6)
Transsphenoidal surgery	20 (56)	10 (44)	30 (51)	2 (2)
Anterior pituitary insufficiency	25 (69)	12 (52)	37 (63)	2 (2)
Psychiatric disorder	10 (28)	0	10 (17)	22 (27)

MRI characteristics — no./total no. (%)				
MRI performed during the study period	33/36 (92)	22/23 (96)	55/59 (93)	42/82 (51)
Hyperintense signal in posterior pituitary absent	21/26 (81)	12/21 (57)	33/47 (70)	14/36 (39)
Pituitary stalk enlarged	6/30 (20)	3/22 (14)	9/52 (17)	1/39 (3)
Hyperintense signal in posterior pituitary absent and pituitary stalk enlarged	3/26 (12)	2/21 (10)	5/47 (11)	1/36 (3)
Laboratory data				
Median plasma sodium — mmol/liter (IQR)	142 (140–143)	143 (142–144)	142 (141–144)	141 (139–142)
Median plasma osmolality — mOsm/kg (IQR)	291 (286–296)	294 (291–301)	292 (288–298)	287 (283–291)
Median urine osmolality — mOsm/kg (IQR)	176 (95–231)	421 (252–460)	228 (116–380)	408 (237–576)

* There were significant between-group differences in age, body-mass index, nocturia, drinking at night, sudden onset of symptoms, persistency of symptoms, history of brain tumor, history of transsphenoidal surgery, anterior pituitary insufficiency, absent hyperintense signal in posterior pituitary, enlargement of pituitary stalk, plasma sodium, plasma osmolality, and urine osmolality. IQR denotes interquartile range, and MRI magnetic resonance imaging.
† The body-mass index is the weight in kilograms divided by the square of the height in meters.

(96.5% [95% confidence interval {CI}, 92.1 to 98.6] vs. 76.6% [95% CI, 68.9 to 83.2]; P<0.001) (Table 2). The diagnostic accuracy of the hypertonic saline infusion test was also clearly superior to that of the indirect water-deprivation test when only patients with partial central diabetes insipidus were compared with patients with primary polydipsia (95.2% [95% CI, 89.4 to 98.1] vs. 73.3% [95% CI, 63.9 to 81.2]; adjusted P<0.001) (Table 2). Additional details on the test results are provided in Table S1 in the Supplementary Appendix.

The copeptin level measured after hypertonic saline infusion more accurately distinguished primary polydipsia from central diabetes insipidus than the water-deprivation test with or without copeptin measurement (Fig. 1). Additional details on the course of copeptin levels during hypertonic saline infusion are provided in Fig. S2 in the Supplementary Appendix. The prespecified hypertonic saline–stimulated copeptin cutoff level of more than 4.9 pmol per liter[19] had a 93.2% sensitivity (95% CI, 83.5 to 98.1) and 100% specificity (95% CI, 95.5 to 100.0) to discriminate between primary polydipsia and central diabetes insipidus (Table 2), with a receiver-operating-characteristic area under the curve for this discrimination of 0.97 (95% CI, 0.93 to 1.00). The most accurate copeptin cutoff level was 6.5 pmol per liter (derived post hoc), which had a diagnostic accuracy of 97.9% (95% CI, 93.9 to 99.6), sensitivity of 94.9% (95% CI, 85.9 to 98.9), and specificity of 100% (95% CI, 95.5 to 100.0) (Fig. 2 and Table 2).

The overall diagnostic accuracy of water-deprivation–stimulated copeptin (with use of the prespecified ratio of stimulated copeptin to plasma sodium described above) in distinguishing primary polydipsia from central diabetes insipidus[5] was significantly lower than that of the indirect water-deprivation test (44.0% [95% CI, 35.7 to 52.5] vs. 76.6% [95% CI, 68.9 to 83.2]) (Table 2). When the prespecified morning copeptin cutoff level of less than 2.6 pmol per liter after overnight water deprivation was used to identify patients with complete central diabetes insipidus,[5] the diagnostic accuracy was 78.4% (95% CI, 70.6 to 84.9). Plasma copeptin values after overnight water deprivation and the associated receiver-operating-characteristic area under the curve are provided in Figure S3 in the Supplementary Appendix.

Table 2. Diagnostic Performance of the Indirect Water-Deprivation Test, the Water-Deprivation–Stimulated Copeptin Test, and the Hypertonic Saline–Stimulated Copeptin Test.*

Comparison and Test	Diagnostic Accuracy		Sensitivity		Specificity		Positive Predictive Value		Negative Predictive Value	
	% (95% CI)	no./total no.	% (95% CI)	no./total no.	% (95% CI)	no./total no.	% (95% CI)	no./total no.	% (95% CI)	no./total no.
Primary polydipsia vs. central diabetes insipidus										
Indirect water-deprivation test	76.6 (68.9–83.2)	108/141	86.4 (75.0–94.0)	51/59	69.5 (58.4–79.2)	57/82	67.1 (55.4–77.5)	57/76	87.7 (77.2–94.5)	57/65
Water-deprivation–stimulated copeptin†	44.0 (35.7–52.5)	62/141	98.3 (90.9–100.0)	58/59	5.0 (1.4–12.3)	4/80	43.3 (34.8–52.1)	58/134	80.0 (28.4–99.5)	4/5
Hypertonic saline–stimulated copeptin with >4.9 pmol/liter cutoff‡	96.5 (92.1–98.6)	136/141	93.2 (83.5–98.1)	55/59	100.0 (95.5–100.0)	81/81	100.0 (93.5–100.0)	55/55	95.3 (88.4–98.7)	81/85
Hypertonic saline–stimulated copeptin with >6.5 pmol/liter cutoff§	97.9 (93.9–99.6)	137/140	94.9 (85.9–98.9)	56/59	100.0 (95.5–100.0)	81/81	100.0 (93.6–100.0)	56/56	96.4 (89.9–99.3)	81/84
Primary polydipsia vs. partial central diabetes insipidus										
Indirect water-deprivation test	73.3 (63.9–81.2)	77/105	87.0 (66.4–97.2)	20/23	69.5 (58.4–79.2)	57/82	44.4 (29.6–60.0)	20/45	95.0 (86.1–99.0)	57/60
Water-deprivation–stimulated copeptin†	25.2 (17.2–34.8)	26/103	95.7 (78.1–99.9)	22/23	5.0 (1.4–12.3)	4/80	22.4 (14.6–32.0)	22/98	80.0 (28.4–99.5)	4/5
Hypertonic saline–stimulated copeptin with >4.9 pmol/liter cutoff‡	95.2 (89.4–98.1)	99/104	82.6 (61.2–95.0)	19/23	100.0 (95.5–100.0)	81/81	100.0 (82.4–100.0)	19/19	95.3 (88.4–98.7)	81/85
Hypertonic saline–stimulated copeptin with >6.5 pmol/liter cutoff§	97.1 (91.8–99.4)	101/104	87.0 (66.4–97.2)	20/23	100.0 (95.5–100.0)	81/81	100.0 (83.2–100.0)	20/20	96.4 (89.9–99.3)	81/84

* A total of 141 patients underwent both the water-deprivation test and the hypertonic saline–stimulated copeptin test. For the assessment of the primary outcome of diagnostic accuracy, missing copeptin measurements were imputed as false results (two for the water-deprivation test and one for the hypertonic saline infusion test). Sensitivity, specificity, positive predictive value, and negative predictive value were calculated according to the number of patients with complete data (141 for the indirect water-deprivation test, 139 for water-deprivation–stimulated copeptin, and 140 for hypertonic saline–stimulated copeptin).

† A ratio of stimulated copeptin (the change in copeptin level over 8 hours during water deprivation in picomoles per liter) to plasma sodium (measured at the end of the test in millimoles per liter) of 0.02 pmol or more per liter indicated primary polydipsia, and a basal plasma copeptin level less than 2.6 pmol per liter indicated central diabetes insipidus. A ratio of less than 0.02 pmol per liter indicated partial central diabetes insipidus.

‡ A prespecified copeptin level of more than 4.9 pmol per liter indicated primary polydipsia, and a level of 4.9 pmol or less per liter indicated complete or partial central diabetes insipidus.

§ In a secondary analysis, a copeptin level of more than 6.5 pmol per liter indicated primary polydipsia, and a level of 6.5 pmol or less per liter indicated complete or partial central diabetes insipidus.

The New England Journal of Medicine

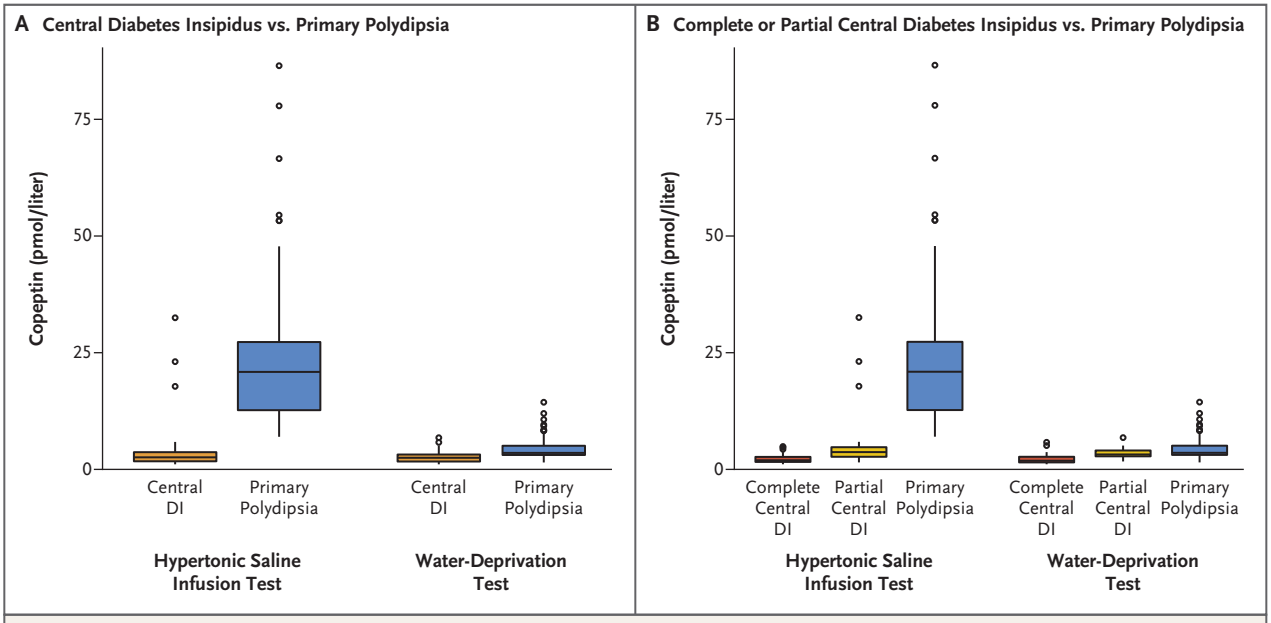


Figure 1. Stimulated Copeptin Levels in Response to the Hypertonic Saline Infusion and Water-Deprivation Tests in Patients with Hypotonic Polyuria.

Shown are stimulated copeptin levels in response to the hypertonic saline infusion test and water-deprivation test in patients with hypotonic polyuria that was caused by central diabetes insipidus as compared with primary polydipsia (Panel A) and in patients with hypotonic polyuria that was caused by complete central diabetes insipidus or partial central diabetes insipidus as compared with primary polydipsia (Panel B). The horizontal line in each box represents the median, the lower and upper boundaries of the boxes the interquartile range, the ends of the whisker lines the minimum and maximum values within 1.5 times the interquartile range, and the dots outliers. DI denotes diabetes insipidus.

SECONDARY OUTCOMES AND BURDEN OF TESTS

Patients rated the overall burden of the water-deprivation test higher than that of the hypertonic saline infusion test (median score on the visual-analogue scale, 6 [interquartile range, 4 to 7] vs. 5 [interquartile range, 3 to 6]) and the overall tolerability (i.e., convenience of the test and patients' comfort level during the test) lower (38% of patients preferred the water-deprivation test, whereas 62% preferred the hypertonic saline infusion test) (Table 3). The plasma sodium level increased to more than 155 mmol per liter in 12 patients during hypertonic saline infusion (in 6 patients with a final diagnosis of primary polydipsia, 5 with complete central diabetes insipidus, and 1 with partial central diabetes insipidus), as compared with 2 patients during water deprivation (both patients had complete central diabetes insipidus). All 12 patients were female and had baseline plasma sodium levels of 140 to 144 mmol per liter. Additional information on the course of plasma sodium level during hypertonic saline infusion is provided in Figure S4 in the Supplementary Appendix.

Nine adverse events occurred during hypertonic saline infusion, and seven during water deprivation. One serious adverse event was reported: desmopressin-induced hyponatremia after the water-deprivation test, which resulted in hospitalization of the patient (Table 3).

DISCUSSION

This prospective, multicenter study showed that measurement of hypertonic saline–stimulated copeptin was superior to the indirect water-deprivation test in distinguishing polyuria due to primary polydipsia from polyuria due to central diabetes insipidus. However, the postulated superiority of water-deprivation–stimulated copeptin to the indirect water-deprivation test could not be confirmed in this study. The diagnostic accuracy of the indirect water-deprivation test of approximately 70% in our study, which is consis-

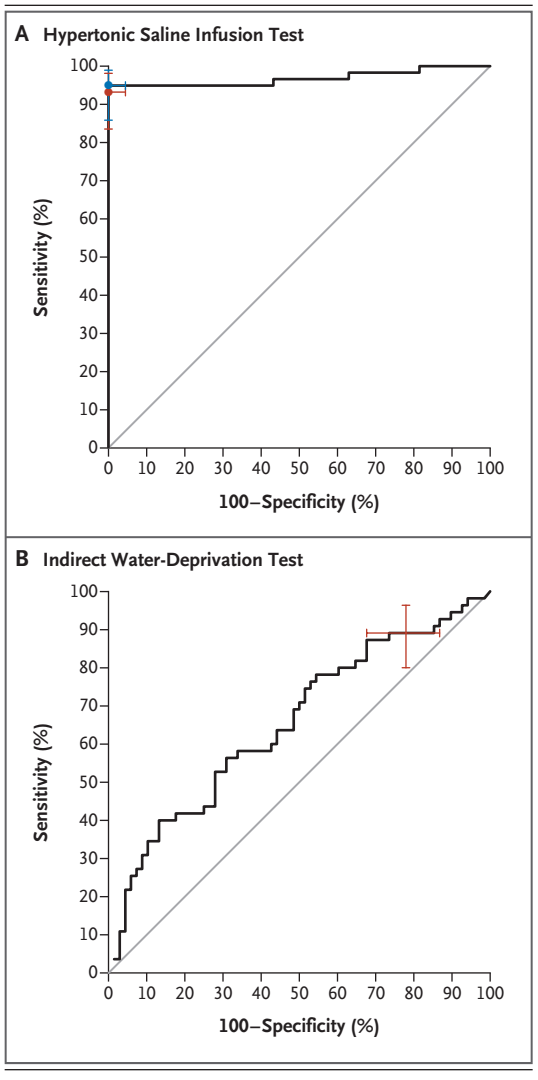


Figure 2. Receiver-Operating-Characteristic (ROC) Curves for the Hypertonic Saline Infusion Test and the Indirect Water-Deprivation Test.

Shown are the ROC curves for the discriminative accuracy of the hypertonic saline infusion test and the indirect water-deprivation test. Panel A shows the discriminative accuracy of hypertonic saline–stimulated copeptin levels in differentiating primary polydipsia from central diabetes insipidus (area under the curve [AUC], 0.97; 95% CI, 0.93 to 1.00). The copeptin cutoff level of 4.9 pmol per liter (prespecified) is indicated in red, and the cutoff of 6.5 pmol per liter (derived post hoc) is indicated in blue. I bars in Panel A indicate 95% confidence intervals. Panel B shows the discriminative accuracy of the indirect water-deprivation test (which measures the change in urine osmolality before and after administration of desmopressin) in differentiating primary polydipsia from central diabetes insipidus (AUC, 0.65; 95% CI, 0.56 to 0.75). I bars in Panel B indicate the 95% confidence intervals for the sensitivity and specificity of the indirect water-deprivation test at a 9% cutoff for the increase in urine osmolality after administration of desmopressin. The gray diagonal lines represent the results that would be expected by chance alone.

tent with previous findings in smaller studies,[5,8] resulted in approximately 30% of patients with primary polydipsia incorrectly receiving a diagnosis of central diabetes insipidus. All the patients in the current study whose diagnosis was misclassified according to the indirect water-deprivation test received the correct diagnosis according to the results of the hypertonic saline infusion test with the prespecified copeptin cutoff of 4.9 pmol per liter.

As reported previously,[5,6,8,23-25] indirect measures of renal arginine vasopressin activity do not accurately discriminate primary polydipsia from central diabetes insipidus, which has challenged the use of the indirect water-deprivation test as the diagnostic standard.[5,8,19] An essential limitation of urinary measures is the variably reduced maximal urinary concentration capacity inherent in all forms of chronic polyuria.[8] Moreover, enhanced renal sensitivity to even low levels of circulating arginine vasopressin in patients with central diabetes insipidus may complicate the interpretation of indirect tests.[26]

Our data confirm that additional measurements such as the basal plasma sodium level[27,28] or the urine-to-plasma osmolality ratio after fluid restriction[29] are of limited diagnostic value. In addition, in our study, the measurement of water-deprivation–stimulated copeptin levels did not improve diagnostic discrimination (73% of patients did not achieve hyperosmolality after 16 hours of fluid deprivation). Finally, although some clinical criteria (e.g., the presence of certain diseases,[30] additional clinical presentations,[2,31] and findings on MRI of the head[21,22,32]) are sometimes recommended to help in making specific diagnoses in patients with polyuria,[2] evidence to support their diagnostic value is insufficient and was not supported by our results.

Consequently, a test method that provides high diagnostic specificity, particularly for the critical distinction of primary polydipsia from central diabetes insipidus, is needed. After the early report on hypertonic saline administration by Hickey and Hare in 1944, in which indirect measures of renal function were used to detect

Table 3. Adverse Effects and Events and Test Burden.*

Variable	Water-Deprivation Test		Hypertonic Saline Infusion Test	
	no. (%)	*median VAS score (IQR)*	*no. (%)*	*median VAS score (IQR)*
Adverse effects				
Thirst	138 (98)	7 (5–9)	141 (100)	8 (6–9)
Vertigo	52 (39)	4 (2–5)	94 (68)	5 (3–7)
Headache	83 (63)	4 (2–5)	94 (67)	4 (2–6)
Nausea	47 (36)	2 (2–4)	69 (50)	4 (2–7)
Malaise	78 (59)	4 (3–6)	96 (69)	5 (3–7)
Adverse events				
Symptomatic overstimulation of plasma sodium	1 (<1)		2 (<1)	
Shivering	1 (<1)		3 (<1)	
Headache requiring pain medication	1 (<1)		0	
Diarrhea	0		1 (<1)	
Emesis after oral water intake	NA		1 (<1)	
Prolonged time until plasma sodium normalization after hypertonic saline infusion	NA		2 (<1)	
Hyponatremia due to excess water retention after administration of desmopressin	4 (<1)†		NA	
Test characteristics‡				
Symptom burden		6 (4–7)		5 (3–6)
Preference	49 (38)		79 (62)	

* A total of 141 patients underwent both tests. Scores on the visual-analogue scale (VAS) range from 0 to 10, with 0 indicating no symptoms and 10 indicating the most severe symptoms imaginable. NA denotes not applicable.
† One event led to hospitalization.
‡ The mean duration of the water-deprivation test was more than 20 hours, and the mean duration of the hypertonic saline infusion test was 3.1 hours.

the release of arginine vasopressin,[33] Zerbe and Robertson further developed the method by introducing direct measurement of plasma arginine vasopressin to differentiate primary polydipsia from diabetes insipidus.[8] Given the technical constraints of arginine vasopressin quantitation[5,11,34,35] and the highly variable functional sensitivity and specificity of the few assays approved for clinical use,[8,34] we designed our study to determine whether the measurement of copeptin is more reliable and easier to process and whether it can be standardized as a test that would detect the osmotic arginine vasopressin reserve.

The copeptin assay is designed to overcome the technical and functional caveats inherent in the arginine vasopressin assay,[8,12] and it appears to have the diagnostic potential not only to identify nephrogenic diabetes insipidus,[5,19,36] but also to distinguish central diabetes insipidus from primary polydipsia.[19] Building on our previous work, in which hypertonic saline infusion was initiated after fluid deprivation,[19] the current prospective validation study used a simplified protocol[16]: the test started with a saline bolus, which was followed by an infusion (at a rate according to each patient's body weight), thereby providing a more potent and prompt osmotic stimulus. This modified protocol attained a better outcome that validated the prespecified copeptin cutoff of 4.9 pmol per liter[19] and, excluding post hoc analysis, yielded the highest diagnostic accuracy for the entire population (96.5%), as well as for the critical distinction between mild forms of

arginine vasopressin deficiency and primary polydipsia (95.2%).

We note caveats with respect to the use of hypertonic saline infusion in the clinical evaluation of patients with polyuria. More adverse effects were reported with the hypertonic saline infusion test than with the water-deprivation test. The hypertonic saline infusion test required close monitoring of sodium levels to ascertain a diagnostically meaningful increase in plasma sodium within the hyperosmotic range[34,37] while preventing a marked increase, to which female patients appeared more vulnerable than male patients in this study.

Our study has limitations and strengths. One limitation is that there is no diagnostic standard for hypotonic polyuria. Here, we constructed criteria for reference diagnoses that were based on the full set of clinical data, the results of the indirect water-deprivation test, and the response of each patient to individual therapy at a 3-month follow-up visit, in accordance with clinical practice.[5,19] The simultaneous evaluation of the diagnostic accuracy of the indirect water-deprivation test and the use of those results in final decision making may have resulted in an incorporation bias. However, if this bias happened at all, it may have resulted in an overestimation of the diagnostic performance of the water-deprivation test. The strengths of the study involve the international multicenter design, the prospective validation of prespecified cutoff levels for hypertonic saline–stimulated copeptin release, and a relatively large sample size of patients with diabetes insipidus and primary polydipsia.

In conclusion, this prospective evaluation of patients with hypotonic polyuria validated hypertonic saline–stimulated copeptin measurement as a diagnostic method that appeared to be superior to the indirect water-deprivation test in distinguishing central diabetes insipidus from primary polydipsia.

Supported by a grant (SNF-162608, to Dr. Christ-Crain) from the Swiss National Foundation and the University Hospital Basel, Switzerland, a grant (FKZ: 01EO1501, to Dr. Fenske) from the Federal Ministry of Education and Research, Germany, and a grant (AOBJ: 624808, to Dr. Fenske) from Deutsche Forschungsgemeinschaft, Germany. Laboratory measurement of copeptin was funded by Thermo Fisher Scientific.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank all the participants for their contribution to our study; Peter Kopp for the critical review of an earlier version of the manuscript; Clara Sailer for her help in analyzing and presenting the data; Stefanie Hahner, Anna Schrader, and Christian Haas for their help in completing the study; and the support staff, study, and laboratory personnel at University Hospital Basel, Switzerland; University Hospital Leipzig, Germany; University Hospital Würzburg, Germany; Faculdade de Medicina da Universidade Federal de Minas Gerais, Brazil; Kantonsspital St. Gallen, Switzerland; Inselspital Bern, Switzerland; Luzerner Kantonsspital, Switzerland; Kantonsspital Aarau, Switzerland; University Hospital Munich, Germany; University Hospital Hamburg-Eppendorf, Germany; and University Hospital Lübeck, Germany.

This article is dedicated to Bruno Allolio from the University of Würzburg, Germany, who died in 2015.

APPENDIX

The authors' full names and academic degrees are as follows: Wiebke Fenske, M.D., Ph.D., Julie Refardt, M.D., Irina Chifu, M.D., Ingeborg Schnyder, M.D., Bettina Winzeler, M.D., Juliana Drummond, M.D., Antônio Ribeiro-Oliveira, Jr., M.D., Ph.D., Tilman Drescher, M.D., Stefan Bilz, M.D., Deborah R. Vogt, Ph.D., Uwe Malzahn, Ph.D., Matthias Kroiss, M.D., Ph.D., Emanuel Christ, M.D., Ph.D., Christoph Henzen, M.D., Stefan Fischli, M.D., Anke Tönjes, M.D., Beat Mueller, M.D., Jochen Schopohl, M.D., Jörg Flitsch, M.D., Georg Brabant, M.D., Martin Fassnacht, M.D., and Mirjam Christ-Crain, M.D., Ph.D.

The authors' affiliations are as follows: the University of Leipzig, Department of Endocrinology and Nephrology (W.F., A.T.), and Leipzig University Medical Center, Integrated Research and Treatment Center Adiposity Diseases (W.F.), Leipzig, the Department of Internal Medicine I, Division of Endocrinology and Diabetes, University Hospital, University of Würzburg (I.C., M.K., M.F.), and the Clinical Trial Center (U.M.) and Central Laboratory (M.F.), University Hospital Würzburg, Würzburg, Medizinische Klinik und Poliklinik IV, Ludwig-Maximilians-Universität, Munich (J.S.), the Department of Neurosurgery, University Hospital Hamburg-Eppendorf, Hamburg (J.F.), and Experimental and Clinical Endocrinology, University Hospital Lübeck, Lübeck (G.B.) — all in Germany; the Clinic of Endocrinology, Diabetology and Metabolism (J.R., I.S., B.W., E.C., M.C.-C.) and the Clinical Trial Unit (D.R.V.), Department of Clinical Research, University of Basel and University Hospital Basel, Basel, the Department of Endocrinology, Kantonsspital St. Gallen, St. Gallen (T.D., S.B.), the Department of Endocrinology, Inselspital Bern, Bern (E.C.), the Department of Endocrinology, Luzerner Kantonsspital, Lucerne (C.H., S.F.), and the Division of Endocrinology, Diabetology and Metabolism, Medical University Clinic, Kantonsspital Aarau, Aarau (B.M.) — all in Switzerland; and Faculdade de Medicina da UFMG, Universidade Federal de Minas Gerais (J.D., A.R.-O.), and Hermes Pardini Institute (J.D.), Belo Horizonte, Brazil.

REFERENCES

1. Christ-Crain M, Fenske W. Copeptin in the diagnosis of vasopressin-dependent disorders of fluid homeostasis. Nat Rev Endocrinol 2016;12:168-76.

2. Fenske W, Allolio B. Current state and future perspectives in the diagnosis of diabetes insipidus: a clinical review. J Clin Endocrinol Metab 2012;97:3426-37.

3. Qureshi S, Galiveeti S, Bichet DG, Roth J. Diabetes insipidus: celebrating a century of vasopressin therapy. Endocrinology 2014;155:4605-21.

4. Miller M, Dalakos T, Moses AM, Fellerman H, Streeten DH. Recognition of partial defects in antidiuretic hormone secretion. Ann Intern Med 1970;73:721-9.
5. Fenske W, Quinkler M, Lorenz D, et al. Copeptin in the differential diagnosis of the polydipsia-polyuria syndrome — revisiting the direct and indirect water deprivation tests. J Clin Endocrinol Metab 2011;96:1506-15.
6. Epstein FH, Kleeman CR, Hendrikx A. The influence of bodily hydration on the renal concentrating process. J Clin Invest 1957;36:629-34.
7. Moses AM, Streeten DH. Differentiation of polyuric states by measurement of responses to changes in plasma osmolality induced by hypertonic saline infusions. Am J Med 1967;42:368-77.
8. Zerbe RL, Robertson GL. A comparison of plasma vasopressin measurements with a standard indirect test in the differential diagnosis of polyuria. N Engl J Med 1981;305:1539-46.
9. Milles JJ, Spruce B, Baylis PH. A comparison of diagnostic methods to differentiate diabetes insipidus from primary polyuria: a review of 21 patients. Acta Endocrinol (Copenh) 1983;104:410-6.
10. Baylis PH, Gaskill MB, Robertson GL. Vasopressin secretion in primary polydipsia and cranial diabetes insipidus. Q J Med 1981;50:345-58.
11. Robertson GL, Mahr EA, Athar S, Sinha T. Development and clinical application of a new method for the radioimmunoassay of arginine vasopressin in human plasma. J Clin Invest 1973;52:2340-52.
12. Morgenthaler NG, Struck J, Alonso C, Bergmann A. Assay for the measurement of copeptin, a stable peptide derived from the precursor of vasopressin. Clin Chem 2006;52:112-9.
13. Preibisz JJ, Sealey JE, Laragh JH, Cody RJ, Weksler BB. Plasma and platelet vasopressin in essential hypertension and congestive heart failure. Hypertension 1983;5:I-129–I-138.
14. Wun T. Vasopressin and platelets: a concise review. Platelets 1997;8:15-22.
15. Balanescu S, Kopp P, Gaskill MB, Morgenthaler NG, Schindler C, Rutishauser J. Correlation of plasma copeptin and vasopressin concentrations in hypo-, iso-, and hyperosmolar states. J Clin Endocrinol Metab 2011;96:1046-52.
16. Fenske WK, Schnyder I, Koch G, et al. Release and decay kinetics of copeptin vs AVP in response to osmotic alterations in healthy volunteers. J Clin Endocrinol Metab 2018;103:505-13.
17. Winzeler B, Zweifel C, Nigro N, et al. Postoperative copeptin concentration predicts diabetes insipidus after pituitary surgery. J Clin Endocrinol Metab 2015;100:2275-82.
18. Christ-Crain M, Morgenthaler NG, Fenske W. Copeptin as a biomarker and a diagnostic tool in the evaluation of patients with polyuria-polydipsia and hyponatremia. Best Pract Res Clin Endocrinol Metab 2016;30:235-47.
19. Timper K, Fenske W, Kühn F, et al. Diagnostic accuracy of copeptin in the differential diagnosis of the polyuria-polydipsia syndrome: a prospective multicenter study. J Clin Endocrinol Metab 2015;100:2268-74.
20. Verbalis JG. Disorders of body water homeostasis. Best Pract Res Clin Endocrinol Metab 2003;17:471-503.
21. Fujisawa I. Magnetic resonance imaging of the hypothalamic-neurohypophyseal system. J Neuroendocrinol 2004;16:297-302.
22. Kurokawa H, Fujisawa I, Nakano Y, et al. Posterior lobe of the pituitary gland: correlation between signal intensity on T1-weighted MR images and vasopressin concentration. Radiology 1998;207:79-83.
23. De Wardener HE, Herxheimer A. The effect of a high water intake on the kidney's ability to concentrate the urine in man. J Physiol 1957;139:42-52.
24. Alexander CS, Filbin DM, Fruchtman SA. Failure of vasopressin to produce normal urine concentration in patients with diabetes insipidus. J Lab Clin Med 1959;54:566-71.
25. Harrington AR, Valtin H. Impaired urinary concentration after vasopressin and its gradual correction in hypothalamic diabetes insipidus. J Clin Invest 1968;47:502-10.
26. Block LH, Furrer J, Locher RA, Siegenthaler W, Vetter W. Changes in tissue sensitivity to vasopressin in hereditary hypothalamic diabetes insipidus. Klin Wochenschr 1981;59:831-6.
27. Rose BD, Post TW. Clinical physiology of acid-base and electrolyte disorders. 5th ed. New York: McGraw-Hill, 2001:748-67.
28. Robertson GL. Diabetes insipidus. Endocrinol Metab Clin North Am 1995;24:549-72.
29. Dashe AM, Cramm RE, Crist CA, Habener JF, Solomon DH. A water deprivation test for the differential diagnosis of polyuria. JAMA 1963;185:699-703.
30. Sailer CO, Winzeler B, Nigro N, et al. Characteristics and outcomes of patients with profound hyponatraemia due to primary polydipsia. Clin Endocrinol (Oxf) 2017;87:492-9.
31. Loh JA, Verbalis JG. Disorders of water and salt metabolism associated with pituitary disease. Endocrinol Metab Clin North Am 2008;37:213-34.
32. Robertson GL. Diabetes insipidus: differential diagnosis and management. Best Pract Res Clin Endocrinol Metab 2016;30:205-18.
33. Hickey RC, Hare K. The renal excretion of chloride and water in diabetes insipidus. J Clin Invest 1944;23:768-75.
34. Robertson GL. The regulation of vasopressin function in health and disease. Recent Prog Horm Res 1976;33:333-85.
35. Kluge M, Riedl S, Erhart-Hofmann B, Hartmann J, Waldhauser F. Improved extraction procedure and RIA for determination of arginine8-vasopressin in plasma: role of premeasurement sample treatment and reference values in children. Clin Chem 1999;45:98-103.
36. Vergier J, Fromonot J, Alvares De Azevedo Macedo A, et al. Rapid differential diagnosis of diabetes insipidus in a 7-month-old infant: the copeptin approach. Arch Pediatr 2018;25:45-7.
37. Robertson GL, Shelton RL, Athar S. The osmoregulation of vasopressin. Kidney Int 1976;10:25-37.

Diagnosis of GH Deficiency as a Late Effect of Radiotherapy in Survivors of Childhood Cancers

Jad G. Sfeir,[1,2] Nana Esi N. Kittah,[2] Shrikant U. Tamhane,[1,2] Sina Jasim,[1,3] Wassim Chemaitilly,[4] Laurie E. Cohen,[5] and M. Hassan Murad[1]

[1]Evidence-Based Practice Center, Mayo Clinic, Rochester, Minnesota 55905; [2]Division of Endocrinology, Diabetes, Metabolism and Nutrition, Mayo Clinic, Rochester, Minnesota 55905; [3]Division of Endocrinology, Metabolism and Lipid Research, Washington University in St. Louis, School of Medicine, St. Louis, Missouri 63130; [4]Department of Pediatric Medicine, Division of Endocrinology, St. Jude Children's Research Hospital, Memphis, Tennessee 38105; and [5]Division of Endocrinology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115

Background: Limited guidance exists for selecting a laboratory method for diagnosing GH deficiency (GHD) when it occurs as a late effect of radiotherapy in childhood cancer survivors (CCSs).

Methods: We searched Medline, Embase, Cochrane Central Register of Controlled Trials and Database of Systematic Reviews, and Scopus for studies evaluating GHD that used IGF-1 or IGF–binding protein 3 (IGFBP-3) measurements compared with GH dynamic testing.

Results: We included 15 studies [IGF-1 (8 studies) and IGFBP-3 (7 studies)] enrolling 477 patients. Comparator tests varied widely. Overall, both IGF-1 and IGFBP-3 had suboptimal diagnostic accuracy but were strongly correlated. The use of both tests simultaneously in the same cohort did not improve the diagnostic accuracy. Despite high variability in the testing protocols, dynamic tests remained the most accurate for appropriately identifying patients with GHD. The insulin tolerance test (ITT) appears to be the most accepted reference test when used alone or in combination with arginine; however, standardized testing strategies among practice groups are absent. GHRH and arginine stimulation performed almost similarly to the ITT; however, in one study GHRH with arginine stimulation had 66% sensitivity and 88% specificity compared with the ITT. Insufficient data were available to assess the accuracy of serial GH testing (nocturnal or over 24 hours).

Conclusion: The diagnostic accuracy of various dynamic tests for GHD in CCSs appears to follow the same patterns as those in non-CCSs. Interpreting GHRH stimulation is a challenge given the primarily hypothalamic dysfunction in CCSs. IGF-1 and IGFBP-3 perform poorly in this population. (***J Clin Endocrinol Metab*** **103: 2785–2793, 2018)**

Clinical manifestations of GH deficiency (GHD) vary by age and pubertal stage and can be nonspecific, and a GHD diagnosis must be confirmed via appropriate laboratory evaluation. However, hormonal testing remains imprecise, with highly variable accuracy. Low serum levels of IGF-1 or IGF–binding protein 3 (IGFBP-3) appear to suggest the diagnosis in adults and can be used to diagnose GHD in children with structural abnormalities, known hypopituitarism, or underlying genetic causes (1). However, normal IGF-1 levels do not exclude the diagnosis and thus mandate further investigation.

Because of the pulsatile secretion of GH, multiple sampling is not practical in the clinical setting; dynamic hormonal testing remains the primary modality used to diagnose GHD. The insulin tolerance test (ITT) has been long considered the gold standard test to diagnose GHD in adults; however, its use has been limited because of the associated risks, mainly hypoglycemia and seizures. GH

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received 31 May 2018. Accepted 4 June 2018.
First Published Online 29 June 2018

Abbreviations: ALL, acute lymphoblastic leukemia; CCS, childhood cancer survivor; GHD, GH deficiency; HP, hypothalamic/pituitary; hpGRF1, human pancreatic GH–releasing factor; IGFBP-3, IGF–binding protein 3; ITT, insulin tolerance test; TBI, total body irradiation.

stimulation testing with a multitude of other agents has been evaluated in this setting. However, these tests have significant intrinsic false-positive rates. Recent observational data revealed that the levodopa/arginine combination and glucagon stimulation testing are the most widely used in the United States (27% and 25%, respectively) (2). The combination of GHRH and arginine stimulation provided 95% sensitivity and 91% specificity compared with the ITT in the general population. The four other stimulation tests using arginine alone, clonidine, levodopa, and arginine plus levodopa in combination did not perform as well (3).

Because of the advances in care for childhood malignancies, an estimated 0.2% of US adults in their second or third decade of life are childhood cancer survivors (CCSs) (4). Improved survival rates have increased the recognition of late-onset complications related to cancer treatments, especially radiotherapy, commonly referred to as late effects. The lifetime prevalence of endocrine late effects in CCSs has been estimated at 50% (5). Hypothalamic/pituitary (HP) irradiation is the primary risk factor for GHD, with a reported prevalence of 46.5% in CCSs treated with this modality (5).

Although GHD is the most common, and often only, endocrine late effect experienced by CCSs exposed to HP radiotherapy, data pertaining to laboratory testing modalities for GHD in this population are scarce and provide only limited guidance. Most of the literature on GHD diagnosis is derived from people who were not CCSs despite suspected differences due to the location and nature of radiation-induced damage (6). Therefore, an Endocrine Society taskforce was charged to develop guidance on the management of HP and growth disorders in CCSs. To support this guideline, we systematically reviewed the available evidence for GHD diagnostic testing in CCSs after HP radiation. We specifically aimed to evaluate the utility of IGF-1 and IGFBP-3 in this population and the relative significances of the various dynamic GH testing methods.

Methods

We performed a systematic review of evaluating screening for GHD by using IGF-1 or IGFBP-3 measurements compared with GH dynamic testing and diagnosing GHD by using various GH dynamic tests. The approach of this systematic review followed standards set by the Cochrane Collaboration and are reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (7).

Eligibility criteria

We searched for cohort studies and case series evaluating the diagnostic accuracy of provocative, dynamic, and static testing when compared with reference standard tests. We included studies reporting >10 patients undergoing any kind of GH testing.

Study identification

A comprehensive search of several databases was conducted from each database's inception to 1 March 2016 in any language. The databases included Ovid Medline In-Process & Other Non-Indexed Citations, Ovid Medline, Ovid Embase, Ovid Cochrane Central Register of Controlled Trials, Ovid Cochrane Database of Systematic Reviews, and Scopus. The search strategy was designed and conducted by a medical reference librarian with input from the investigators. Controlled vocabulary supplemented with keywords was used to search for studies on GHD diagnosis in CCSs.

Subsequently, references for the studies included in the original search were screened by the reviewers for further inclusions. Finally, additional references were obtained directly from the Endocrine Society taskforce members.

Selection of studies

Reviewers working independently and in duplicate reviewed all the retrieved abstracts and selected full-text manuscripts for eligibility. Disagreements regarding the full-text screening were resolved by the consensus of two reviewers.

Data collection and management

Working independently and in duplicate, reviewers used a standardized Web-based form to collect information from each eligible study. We recorded the age and sex of the included population, the GHD testing methods, and the cutoffs used for detection and diagnosis. The outcome of interest was the performances of the various GH testing methods against each study's gold standard.

Risk of bias and quality of evidence

Reviewers working independently and in duplicate used the Quality Assessment of Diagnostic Accuracy Studies 2 tool to assess the risk of bias for each study included (8). The quality of evidence was evaluated via the Grades of Recommendation, Assessment, Development and Evaluation approach (9).

Results

Search results

A total of 211 citations were retrieved. After abstract review, 177 full-text studies were reviewed for a final inclusion of 15 studies. Figure 1 summarizes the flow of the study selection and exclusion criteria.

Patient characteristics

Table 1 shows detailed characteristics of the included patients. Overall, 477 patients, including children, adolescents, and young adults with a male preponderance (n = 266, 56%), were evaluated in 15 studies. Notably, 5 studies did not report the age at diagnosis, and 3 did not report the age at the time of GHD evaluation; thus, a median age for the entire population could not be accurately estimated. In addition, 2 studies did not report the sex proportions of their cohorts.

The most common childhood cancers were acute lymphoblastic leukemia (ALL) (n = 177) and medulloblastoma (n = 90). All patients received HP, cranial, or craniospinal radiation. Some studies enrolled healthy controls and performed the same index testing in both groups. The latter group of studies did not appropriately evaluate

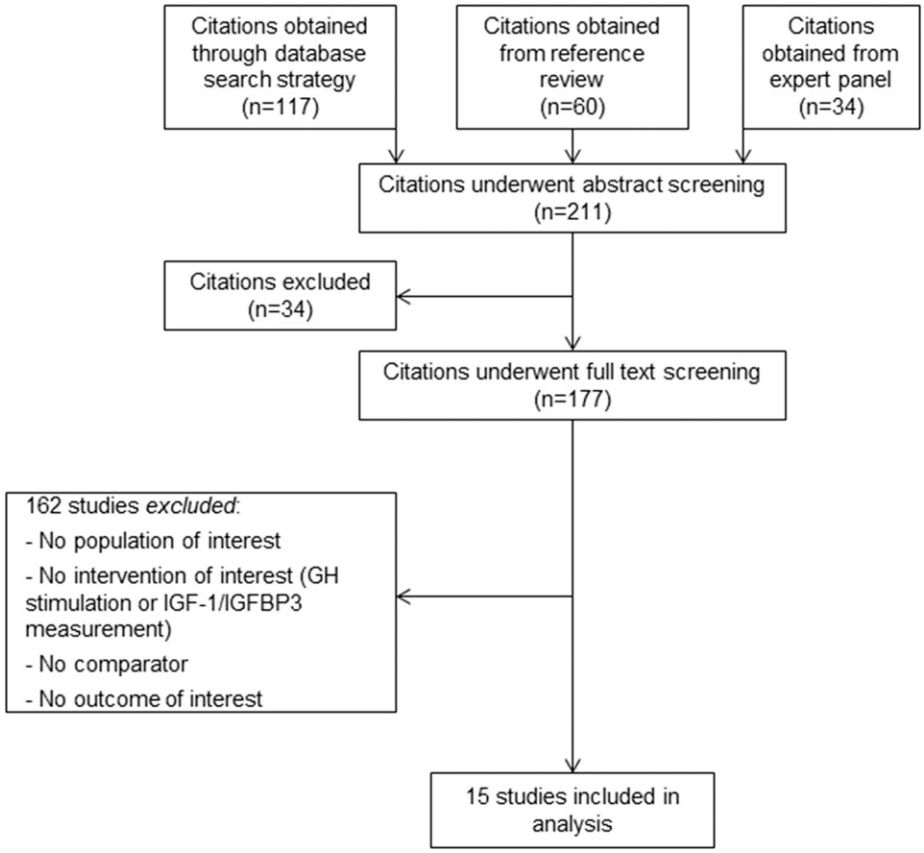


Figure 1. Study flowchart.

the performances of the index tests because they lacked a reference diagnostic test.

Study characteristics

Arginine stimulation and the ITT were used in 9 studies each, mostly as comparator tests. Two studies combined GHRH with arginine. GHRH stimulation was used in isolation in 4 other studies, one of which used a formulation of human pancreatic GH–releasing factor (hpGRF1). Finally, one study evaluated GH secretion after exercise. Nonprovocative tests included IGF-1 in 10 studies, IGFBP-3 in 9 studies, and continuous GH profiles in 3 studies (24-hour profiling in 2 and nocturnal sampling in 1). The individual study characteristics are listed in Table 2.

There was high variability in the confirmatory test performed in each study and in the testing strategies. The major analytical challenge was the lack of a gold standard to diagnose GHD. A number of studies used the ITT as a reference, and other studies compared the performance of multiple tests against each other.

Provocative tests

GH provocative test protocols

Insulin. After an overnight fast, insulin was injected intravenously at a dosage of 0.07– to 0.2 U/kg body weight, and hypoglycemia was subsequently confirmed after serial glucose measurements, with nadir glucose levels reported between 30 and 50 mg/dL. Blood was drawn at 18- to 30-minute intervals after the injection for a duration of 90 to 120 minutes for GH measurements (6, 10–16). Lisset *et al.* (17) reported the achievement of "satisfactory hypoglycemia" but no time intervals or durations of the blood draws.

GHRH. Human recombinant GHRH was administered intravenously at a dosage of 1 μg/kg after an overnight fast, and GH measurements were taken every 15 to 30 minutes for a duration of 60 to 120 minutes (6, 15, 16, 18, 19). Only one study reported a maximal dosage of 50 μg (15).

Arginine. Arginine used in isolation or in combination with GHRH was administered at a dosage of either 20 g/m^2 or 0.5 g/kg via intravenous infusion over 30 minutes, and subsequent GH measurements were taken every 15 to 30 minutes. When used in combination with GHRH, the infusion was started immediately after GHRH administration, and the duration varied from 60 to 120 minutes (6, 10, 12, 16, 17). Only one study reported a maximal dosage of 30 g.

Levodopa. GH measurements after the oral administration of levodopa at 290 mg/m^2 body surface area were performed every 30 minutes for durations of 90 to

Table 1. Patient Characteristics

Study	Country	Age at Time of Radiation or Diagnosis Median (Range), y	Age at Time of Study Median (Range), y	Male (%)	Time From Radiation Median (Range), y	Type of Tumors (n)	Total Number of Subjects	Site of Radiation	Radiation Dose Range (Gy)
Shalet *et al.* 1976 (10)	United Kingdom	N/A	10 (7–14)	6 (66.67)	4 y (2–10)	Medulloblastoma (4) Astrocytoma (2) Glioma pons (1) Glioma optic nerve (1) Ependymoma (1)	9	Cranial or craniospinal	10 to 42.5
Ahmed and Shalet 1984 (11)	United Kingdom	8 y (2–14)	15.5 (11–23)	N/A	7 (3–18)	Glioma (4) Medulloblastoma (3) Ependymoma (2) Pineal tumor (1)	10	Cranial or craniospinal	28.5 (27–47.5)
Kirk *et al.* 1987 (12)	Australia	N/A	11 (4–20)	25 (54.3)	5.8 (3.5–9.2)	ALL (46)	46	Cranial	24
Moëll *et al.* 1989 (18)	Sweden	2.6 (1.1–7.1)	10.5 (8.4–13.4)	0	4.6 (2.9–7.3)	ALL (13)	13	Cranial	20 to 24
Sklar *et al.* 1993 (22)	USA	N/A	9.4 (6.7–16)	9 (60)	2.7 (2–7)	Medulloblastoma (5) ALL (4) Rhabdomyosarcoma (3) Acute myeloid leukemia (1) Ependymoma (1) Glioblastoma (1)	15	HP	18 to >60
Talvensaari *et al.* 1994 (19)	Finland	4.0 (0.3–14.9)	18.0 (10.8–26.4)	11 (44)	13.2 (8.9–18.6)	Leukemia or lymphoma	25	Cranial	15 to 46
Brennan *et al.* 1998 (13)	United Kingdom	6·9 (1.7–16)	23 (18.8–33)	16 (50)	17.8 (6.8–28.6)	ALL (32)	32	Cranial	18 to 25
Tillmann *et al.* 1998 (14)	United Kingdom	N/A	7–24.3	14 (50)	0.4–14.2	Medulloblastoma (7) Rhabdomyosarcoma (3) Glioma (3) Astrocytoma (2) Pineal teratoma (1) ALL (12)	28	Cranial or craniospinal	14 to 55
Cicognani *et al.* 1999 (20)	Italy	4.6 (2.7–10.2)	5.3 (3.3–10.9)	7 (58)	N/A	Medulloblastoma (4) Rhabdomyosarcoma (4) Ewing sarcoma (1) Ependymoma (1) Sarcoma (1) Carcinoma (1)	12	Cranial	30 to 72
Schmiegelow *et al.* 1999 (23)	Denmark	6.3 (1.7–16.5)	N/A	15 (79)	3.4 (1.2–6.3)	Medulloblastoma (10) Ependymoma (3) Astrocytoma (2) Germ cell tumor (2) Glioma (1) Choroid plexus carcinoma (1)	19	Cranial or craniospinal	30.6 to 54.0
Weinzimer *et al.* 1999 (21)	USA	9.83 ± 3.1	N/A	51 (71)	N/A	Primitive neuroectodermal tumor (31) Craniopharyngioma (16) Hypothalamic/chiasmatic glioma (12) Nasopharyngeal Rhabdomyosarcoma (5) Germinoma (3) Cerebellar glioma (1) Histiocytosis (1) Pinealoma (1) Prolactinoma (1) Retinoblastoma (1)	72	Cranial	N/A
Schmiegelow *et al.* 2000 (15)	Denmark	8.9 (0.8–14.9)	N/A	42 (68)	11 (2.0–28.0)	Astrocytoma (26) Medulloblastoma (19) Ependymoma (5) Germ cell tumor (3) Glioma (2) Pinealoma (1) Hemangiopericytoma (1) Primitive neuroectodermal tumor (1) Nonhistologically verified (4)	62	Cranial or craniospinal	22.4 to 56.2
Lissett *et al.* 2001 (17)	United Kingdom	N/A	10 (3–16)	48 (59)	3.1 (0.1–13.5)	Medulloblastoma (28) ALL (27) Astrocytoma (5) Germinoma (5) Ependymoma (4) Glioma (3) Rhabdomyosarcoma (3) Acute myeloid leukemia (2) Histiocytosis X (1) Lymphosarcoma (1) Other (2)	81	HP	30 (12–50)
Björk *et al.* 2005 (16)	Sweden	4.7 (1.2–18)	26 (20–33)	22 (51)	21 (8.9–28)	ALL (43)	43	Cranial	18 to 30
Ham *et al.* 2005 (6)	USA	7.5 ± 3.5	21.8 (17–24)	N/A	13.8 (7.8–19.6)	Medulloblastoma (10)	10	Craniospinal	23 to 36

Abbreviation: N/A, not applicable.

Table 2. Study Characteristics

Study	Type of Study	Index Test	Dose Administered (Index)	GH Cutoff Value (Index)	Comparator Test	Dose Administered (Comparator)	GH Cutoff Value (Comparator)
Shalet *et al.* 1976 (10)	Cross-sectional	Arginine stimulation	Bovril 20 g/ 1.5 m^2 oral	<6.67 μg/L	ITT	0.1–0.2 U/kg IV	<6.67 μg/L
Ahmed and Shalet 1984 (11)	Cohort	hpGRF1			ITT	0.2 U/kg IV	>6.67 μg/L
					Arginine stimulation	20 g/m^2 IV	>6.67 μg/L
Kirk *et al.* 1987 (12)	Cohort	Arginine stimulation		<10 μg/L	24-h GH		
		ITT		<10 μg/L			
		Exercise		<10 μg/L			
Moëll *et al.* 1989 (18)	Cross-sectional	GHRH stimulation IGF-1	1 μg/kg	<4.67 μg/L	24-h GH		Area under the curve (μg/L/24 h)
Sklar *et al.* 1993 (22)	Cohort	IGFBP-3		Age-specific normal	Clonidine and levodopa stimulation	Clonidine 150 μg/m^2 oral	<10 μg/L
		IGF-1					
Talvensaari *et al.* 1994 (19)	Case control	GHRH stimulation IGFBP-3	1 μg/kg	<10 μg/L	Nocturnal GH		<2.35 μg/L/h
Brennan *et al.* 1998 (13)	Case control	IGF-1 IGFBP-3			ITT and arginine stimulation	Insulin 0.2 U/kg IV Arginine 20 g/m^2 IV	<3 μg/L to both
Tillmann *et al.* 1998 (14)	Cross-sectional	IGF-1		<−2 SD	Arginine stimulation Glucagon stimulation		<7.5 μg/L
		IGFBP-3		<−1.5 SD	ITT Clonidine stimulation		
Cicognani *et al.* 1999 (20)	Cohort	IGF-1		<−2 SD	Arginine stimulation	0.5 g/kg IV	<8 μg/L
		IGFBP-3		<−2 SD	Levodopa stimulation	290 mg/m^2 oral	<8 μg/L
Schmiegelow *et al.* 1999 (23)	Cohort	IGFBP-3					
Weinzimer *et al.* 1999 (21)	Cohort	IGF-1		<−2 SD	One or more GH stimulation tests[a]	Arginine 0.5 g/kg Clonidine 0.1 mg	<7.0 mg/L
		IGFBP-3		<−2 SD		Levodopa 500 mg/1.74 m^2 Propranolol/ glucagon Propranolol 40 mg (20 mg if <20 kg and 10 mg if <10 kg) and glucagon 0.5 mg	
Schmiegelow *et al.* 2000 (15)	Cross sectional	GHRH stimulation	1 μg/kg		ITT and arginine stimulation		
Lissett *et al.* 2001 (17)	Cross sectional	Arginine stimulation	20 μg/m^2	<8 μg/L	ITT	0.2 U/kg IV	<8 μg/L
Björk *et al.* 2005 (16)	Cross-sectional	GHRH + Arginine stimulation	GHRH 1 μg/kg arginine 0.5 g/kg	<9 μg/L	ITT	0.1–0.2 U/kg IV	<3 μg/L
Ham *et al.* 2005 (6)	Cohort	GHRH + Arginine stimulation IGF-1 IGFBP-3	GHRH 1 μg/kg arginine 0.5 g/kg	<9 μg/L	ITT	0.07–0.1 U/kg IV	<3 μg/L

[a] Children were considered to have GHD if they met both of the following criteria: growth velocity *z* score <−2 or clinical deceleration of growth velocity, plus failing one or more GH stimulation tests.

120 minutes (20, 21). One study did not report dosing, intervals, or follow-up duration (22).

Clonidine. Clonidine was administered at a dosage of either 150 μg/m^2 or 0.1 mg orally once, and GH was measured every 30 minutes for a duration of 120 minutes (21, 22). Tillmann *et al.* (14) did not report their stimulation testing strategies.

GH assays

Although most studies used the double antibody RIA technique for the measurement of GH (10–13, 15, 17–19, 21, 22), three studies used a dissociation-enhanced lanthanide fluorescence immunoassay instead (16, 20, 21). However, all assays used international standard preparations as references (International Reference Preparation 66/217 before 1990 and World Health Organization 80/505 after 1990).

Cutoff values

The various cutoff values used by the studies evaluated ranged from 4.67 μg/L to 10 μg/L and are summarized in Table 2. The GH values have been converted to a single unit of μg/L for better comparison.

Performance

Because of the lack of a single comparator used in all or most of the studies, meta-analyses of sensitivity and specificity could not be performed for the index tests. Despite high variability in the testing protocols, dynamic tests remained the most accurate in appropriately identifying patients with GHD, and the ITT was the most commonly used reference test. The ability of GHRH stimulation, with or without arginine, to diagnose GHD across different studies was equivocal; in one study, GHRH in combination with arginine stimulation was 66% sensitive and 88% specific in comparison with the ITT (16).

IGF-1 and IGFBP-3

Overall, IGF-1 and IGFBP-3 had poor diagnostic accuracy (6, 14). When calculations were possible, the sensitivity and specificity of IGF-1 varied from 47% to 66% and 77% to 100%, respectively (14, 22). IGFBP-3 had a low sensitivity of 20%. However, most studies showed a strong correlation

between IGF-1 and IGFBP-3 (13, 19–23). The use of simultaneous testing in the same cohort did not increase the diagnostic accuracy of either test alone (20).

GH secretion profiles

Insufficient data were available to assess the accuracy of serial blood sampling (collected both nocturnally and over a 24-hour period) for the determination of GH.

Risk of bias (methodological quality) and quality of evidence

The methodological quality of the included studies was moderate overall. However, the cohorts could not be combined, and the index tests could not be compared given the high interstudy variability in the testing strategies, cutoff levels, and, most importantly, comparator tests. The quality of evidence (certainty in diagnostic estimates) was moderate because of the small sample size and inability to aggregate data.

Discussion

Although a rare disorder (estimated prevalence of 1 in 4000 adults), GHD is the most common endocrinopathy in survivors of childhood central nervous system tumors and in those exposed to HP radiation. The latter includes survivors of childhood ALL exposed to prophylactic or therapeutic central nervous system radiotherapy, patients receiving conditioning total body irradiation (TBI) before stem cell transplantation, and patients with non-HP intracranial solid tumors, such as nasopharyngeal carcinoma and retinoblastoma. The availability of recombinant human GH and its beneficial effects on adult height prospects highlight the importance of diagnosing GHD in these patients. Minimizing false-positive tests for GHD is also important given the safety concerns surrounding the use of GH in CCSs due to its known promitogenic and proliferative properties *in vitro* (5, 24).

The risk of GHD increases in a time-dependent fashion in relation to radiotherapy, and its prevalence thus increases with longer follow-up durations. GHD is also radiation dose-dependent, with the highest risk after HP radiation with ≥30 Gy; however, risks also occur after 18 to 30 Gy, after TBI with ≥10 Gy in one fraction, and after TBI with ≥12 Gy in multiple fractions. Additional risk factors include HP tumor involvement, young age at diagnosis, and use of tyrosine kinase inhibitors or anti-CTLA4 monoclonal antibodies (5, 25). In fact, the increasing use of immune checkpoint inhibitors to treat malignancies such as malignant melanoma has shed light on the increased incidence of immune-mediated endocrine disorders, such as hypophysitis, thyroiditis, and autoimmune diabetes mellitus.

GHD diagnosis in both adult and pediatric general populations is challenging, mostly because of the lack of standardization of diagnostic tests, their respective cutoff values, and assay variability. Similarly, a consensus on the optimal testing method to diagnose GHD in CCSs is lacking. GHD typically presents with a decreased growth velocity compared with sex- and age-adjusted reference values (less than −2 SD over 1 year or −1.5 SD over 2 years). However, confirmatory testing is essential but requires a good understanding of the technical considerations and challenges of these tests.

GH provocative tests

GH stimulation tests aim to demonstrate inappropriately low peak levels of GH after physical or chemical stimulation. Physical stimulation has long been eliminated from routine use because of its lack of precision. Pharmacological stimulation uses GHRH, arginine, clonidine, glucagon, insulin, or levodopa. Kirk *et al.* (12) used exercise stimulation in parallel with other provocative tests, but this study is listed in our evaluation for completeness. Similarly, a study by Ahmed *et al.* (11) is included despite its lacking the modern use of hpGRF1 stimulation.

The ITT has traditionally been used as the gold standard, and most index tests evaluated were compared with some version of the ITT. GHRH stimulation, with or without arginine, had substantial clinical use before November 2008, when it became no longer commercially available in the United States. Glucagon stimulation has since increased in popularity for use in patients with contraindications to the ITT but was used in only one CCS study.

Provocative test protocols

Overall, homogeneity was observed regarding the protocols used to stimulate GH responses in dynamic tests, which was most notable in terms of dosing regimens and administration route. Differences were related mostly to the time of follow-up to detect peak GH responses, varying from 60 to 180 minutes. There were also some differences in the blood sampling intervals, which varied from 15 to 30 minutes. Notably, two studies measured the baseline GH twice, both at 15 to 30 minutes before the initiation of dynamic testing and at the start of the test, whereas all other protocols used only the measurement at time zero.

Sex steroid priming

The current guidelines for GHD testing in children and adolescents recommend priming prepubertal children of pubertal age with sex steroids before provocative GH testing to better differentiate GHD from constitutional delay (26); this recommendation is based largely on observational data in boys. There has been no evaluation of sex steroid priming in CCSs, a population that is enriched for delayed puberty due to hypogonadotropic hypogonadism and primary gonadal insufficiency, and none of the studies evaluated used such a protocol before GH stimulation.

GH assays

A multitude of test methods are currently used for GH measurement. Interassay variability was reported at 25% in

an audit of 96 UK clinical laboratories. In Germany, a comparative study determined that a GHD diagnosis depends on the assay used in ~30% of children undergoing GH stimulation. Most of this variability results from the presence of multiple GH variants in serum, including polymeric and fragmented forms, which are differentially detected in various assays (24). As with all hormones, protein-bound and free forms of GH can also affect testing results. Mass spectrometry has been suggested to overcome a proportion of this variability, but the most recent consensus statements and guidelines highlight the importance of using a single reference material across laboratories to harmonize GH measurements (27).

The studies we evaluated spanned three decades. During this time, shifts from polyclonal to monoclonal antibody use in the RIA and from purified human GH reference solutions to recombinant solutions occurred, leading to higher specificity in the currently used assays. These differences affect the comparative interpretations of the tests and the cutoffs used for diagnosis, although they do not influence the evaluation of individual tests in individual studies.

Cutoff levels

Controversies remain regarding the optimal GH cutoff levels for GHD diagnosis via a specific provocative test and the point at which GH replacement is indicated. Historically, a cutoff value of 5 μg/L correlated with GHD phenotypes, particularly pertaining to height velocity and adult height (24). Allowing for assay differences, including the introduction of mass spectrometry, a recent study from Germany ascertained a total GH of 7 μg/L as the diagnostic cutoff. A 2016 survey of 48 pediatric departments and 57 biochemistry departments in the UK found that the cutoff values used to diagnose GHD in children ranged from 5 to 10 μg/L, with the overwhelming majority using a cutoff of <7 μg/L (28). A similar pattern was seen in our review of published studies on CCSs.

Response to GH stimulation is also affected by body mass index, as determined by blunted responses in obese and overweight patients. In adults, this difference is reflected by the reliance on IGF-1 for diagnosis. Multiple provocative studies in children use differential cutoffs based on body weight or body mass index, but none of these studies include CCSs. Thus, using similar adjustments in this population could be accomplished only by extrapolation and would not account for the overall metabolic effects of HP dysfunction as opposed to isolated GHD or GHD due to other etiologies. Furthermore, GHD due to hypothalamic dysfunction can be missed by provocative testing that includes GHRH stimulation, because these patients can have false-negative responses, particularly early in the disease process.

Test performance

In addition to the aforementioned assay variability and test interpretations, the lack of a reference comparative test has led to variability in the reported reproducibilities and accuracies of the various provocative tests. The ITT has largely been considered the gold standard, and other provocative tests are recommended only when contraindications to the ITT occur (*i.e.,* seizure disorders and cardiovascular disease). The false-positive rates of the various dynamic tests range from 8.9% to 49% (29). False-negatives are also common, with glucagon stimulation testing being reported to miss ≤58% of patients with GHD diagnosed based on both the ITT and arginine stimulation tests (30).

Based on data gathered from our evaluation of various studies, the diagnostic accuracy patterns of all the index tests appear to mimic those seen in the general population. Arginine used in isolation has had poor reproducibility, with up to two-thirds of patients having different diagnostic outcomes upon retesting with the same stimulation test (31). In the general population, both the same test in the same child and two different tests in the same child have poor reproducibility. Retrospectively, in a large French cohort study, the ITT was the only test whose confidence interval, albeit very wide, crossed the satisfactory threshold value when performed in duplicate (32). None of the provocative tests evaluated in our report were tested for reproducibility over time.

GH secretion profiles

Pituitary GH secretion is pulsatile, with most pulses occurring overnight and thus limiting the utility of random GH sampling. Obtaining 12- or 24-hour GH profiles has previously been used but has very limited clinical use today given the frequent sampling necessary and the need for hospital admission. Furthermore, both the sensitivity and specificity of this protocol have been questioned, because overlap in secretory profiles occurs between healthy individuals and those with GHD (33, 34).

IGF-1 and IGFBP-3

Under the control of GH, the liver produces circulating IGF-1 bound to numerous binding proteins, including IGFBP-3. Both IGF-1 and IGFBP-3 exist in a steady state at constant concentrations and thus offer the advantage of random serum sampling for measurement. However, normative values need to be adjusted not only for chronological age but also for pubertal age and Tanner stage. Similar to GH, assay variability and standardization problems also exist for World Health Organization solutions (27). In the general population, IGF-1 is known to have a high specificity (≤95%); however, a normal level does not exclude GHD but rather warrants additional testing based on clinical suspicion. IGFBP-3 performance has been similar and does not add value to the IGF-1 measurement (35). In CCSs, however, the use of either of these markers does not appear promising, because their diagnostic accuracies are quite poor even when used in combination.

Diagnosis

The availability of recombinant GH treatment has made establishing a diagnosis crucial to appropriately identifying patients who would benefit from such therapy. Given the lack of strong evidence supporting treatment in patients with isolated GHD or idiopathic short stature, guidelines have recommended the use of two provocative tests to establish the diagnosis in children. However, CCSs have established cranial pathologies or have been subjected to HP radiation and thus need a lower testing threshold based on these guidelines. Subsequently, a failure to stimulate GH in any one provocative test is sufficient to establish the diagnosis due to high pretest diagnostic suspicion, especially in the presence of other HP deficits, with the caveat that a poor growth rate may be caused by skeletal growth impairment independent of GHD after treatment with agents such as TBI, imatinib mesylate, and *cis*-retinoic acid (36–40).

This caveat highlights the importance of testing reproducibility and accuracy in this population. The ITT, GHRH (with or without arginine), and glucagon have been recommended by the Endocrine Society, in this order, for the diagnosis of GHD in adults. In CCSs, however, the use of GHRH is not recommended for the reasons discussed earlier. Glucagon stimulation has not been investigated enough in this population, and the ITT thus remains the only reliable dynamic test even though its performance is based largely on the general population and historical experience.

Finally, reliance on nonprovocative tests in this population is not recommended. GH profiling has been replaced by less labor-intensive and more cost-efficient testing; IGF-1 and IGFBP-3 performed poorly in this population, in contrast to the general pediatric and adult populations tested for GHD, and thus cannot be used to establish diagnosis.

Limitations

The studies included in this report span four decades and represent an array of testing and clinical consensuses, which has led to variability in reporting results and outcomes. Thus, calculating the diagnostic accuracies of the various tests evaluated based on the available primary data was not possible. Additionally, there is no gold standard for diagnosing GHD to which to compare any of these tests.

Conclusion

Evaluating the GH axis in CCSs is an opportunity to treat these patients with available and effective replacement therapy. The controversy regarding the benefit-to-risk ratios of such therapies highlights the importance of appropriate patient selection and accuracy of testing for GHD. Based on this systematic review, dynamic testing remains the cornerstone for GHD diagnosis. Additional research on this population is needed to establish the best possible test. In the meantime, however, reliance on the ITT (as the gold standard) appears to be appropriate, with recognition that this test is not feasible at many institutions.

Acknowledgments

Financial Support: Funded by a grant from the Endocrine Society to M.H.M.

Correspondence and Reprint Requests: M. Hassan Murad, MD, MPH, Evidence-Based Practice Center, Mayo Clinic, 200 First Street SW, Rochester, Minnesota 55905. E-mail: murad.mohammad@mayo.edu.

Disclosure Summary: W.C. is a consultant for Pfizer. L.E.C. is a site principal investigator for Versartis and Ascendis and received an honorarium from Scherer Clinical Communications (grant from Novo Nordisk). The remaining authors have nothing to disclose.

References

1. Molitch ME, Clemmons DR, Malozowski S, Merriam GR, Vance ML; Endocrine Society. Evaluation and treatment of adult growth hormone deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2011;**96**(6):1587–1609.
2. Cook DM, Yuen KC, Biller BM, Kemp SF, Vance ML; American Association of Clinical Endocrinologists. American Association of Clinical Endocrinologists medical guidelines for clinical practice for growth hormone use in growth hormone–deficient adults and transition patients: 2009 update. *Endocr Pract*. 2009;**15**(suppl 2):1–29.
3. Biller BM, Samuels MH, Zagar A, Cook DM, Arafah BM, Bonert V, Stavrou S, Kleinberg DL, Chipman JJ, Hartman ML. Sensitivity and specificity of six tests for the diagnosis of adult GH deficiency. *J Clin Endocrinol Metab*. 2002;**87**(5):2067–2079.
4. Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. *CA Cancer J Clin*. 2014;**64**(2):83–103.
5. Chemaitilly W, Cohen LE. Diagnosis of endocrine disease: endocrine late-effects of childhood cancer and its treatments. *Eur J Endocrinol*. 2017;**176**(4):R183–R203.
6. Ham JN, Ginsberg JP, Hendell CD, Moshang T Jr. Growth hormone releasing hormone plus arginine stimulation testing in young adults treated in childhood with cranio-spinal radiation therapy. *Clin Endocrinol (Oxf)*. 2005;**62**(5):628–632.
7. McGrath TA, Alabousi M, Skidmore B, Korevaar DA, Bossuyt PMM, Moher D, Thombs B, McInnes MDF. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. *Syst Rev*. 2017;**6**(1):194.
8. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;**155**(8):529–536.
9. Murad MH. Clinical practice guidelines: a primer on development and dissemination. *Mayo Clin Proc*. 2017;**92**(3):423–433.
10. Shalet SM, Beardwell CG, Morris-Jones P, Bamford FN, Ribeiro GG, Pearson D. Growth hormone deficiency in children with brain tumors. *Cancer*. 1976;**37**(2 suppl):1144–1148.
11. Ahmed SR, Shalet SM. Hypothalamic growth hormone releasing factor deficiency following cranial irradiation. *Clin Endocrinol (Oxf)*. 1984;**21**(5):483–488.

12. Kirk JA, Raghupathy P, Stevens MM, Cowell CT, Menser MA, Bergin M, Tink A, Vines RH, Silink M. Growth failure and growth-hormone deficiency after treatment for acute lymphoblastic leukaemia. *Lancet*. 1987;**1**(8526):190–193.
13. Brennan BM, Rahim A, Mackie EM, Eden OB, Shalet SM. Growth hormone status in adults treated for acute lymphoblastic leukaemia in childhood. *Clin Endocrinol (Oxf)*. 1998;**48**(6):777–783.
14. Tillmann V, Shalet SM, Price DA, Wales JK, Pennells L, Soden J, Gill MS, Whatmore AJ, Clayton PE. Serum insulin-like growth factor-I, IGF binding protein-3 and IGFBP-3 protease activity after cranial irradiation. *Horm Res*. 1998;**50**(2):71–77.
15. Schmiegelow M, Lassen S, Poulsen HS, Feldt-Rasmussen U, Schmiegelow K, Hertz H, Müller J. Growth hormone response to a growth hormone-releasing hormone stimulation test in a population-based study following cranial irradiation of childhood brain tumors. *Horm Res*. 2000;**54**(2):53–59.
16. Björk J, Link K, Erfurth EM. The utility of the growth hormone (GH) releasing hormone-arginine test for diagnosing GH deficiency in adults with childhood acute lymphoblastic leukemia treated with cranial irradiation. *J Clin Endocrinol Metab*. 2005;**90**(11): 6048–6054.
17. Lissett CA, Saleem S, Rahim A, Brennan BM, Shalet SM. The impact of irradiation on growth hormone responsiveness to provocative agents is stimulus dependent: results in 161 individuals with radiation damage to the somatotropic axis. *J Clin Endocrinol Metab*. 2001;**86**(2):663–668.
18. Moëll C, Garwicz S, Westgren U, Wiebe T, Albertsson-Wikland K. Suppressed spontaneous secretion of growth hormone in girls after treatment for acute lymphoblastic leukaemia. *Arch Dis Child*. 1989;**64**(2):252–258.
19. Talvensaari KK, Lanning M, Pääkkö E, Tapanainen P, Knip M. Pituitary size assessed with magnetic resonance imaging as a measure of growth hormone secretion in long term survivors of childhood cancer. *J Clin Endocrinol Metab*. 1994;**79**(4): 1122–1127.
20. Cicognani A, Cacciari E, Pession A, Pasini A, De Iasio R, Gennari M, Alvisi P, Pirazzoli P. Insulin-like growth factor-I (IGF-I) and IGF-binding protein-3 (IGFBP-3) concentrations compared to stimulated growth hormone (GH) in the evaluation of children treated for malignancy. *J Pediatr Endocrinol Metab*. 1999;**12**(5):629–638.
21. Weinzimer SA, Homan SA, Ferry RJ, Moshang T. Serum IGF-I and IGFBP-3 concentrations do not accurately predict growth hormone deficiency in children with brain tumours. *Clin Endocrinol (Oxf)*. 1999;**51**(3):339–345.
22. Sklar C, Sarafoglou K, Whittam E. Efficacy of insulin-like growth factor binding protein 3 in predicting the growth hormone response to provocative testing in children treated with cranial irradiation. *Acta Endocrinol (Copenh)*. 1993;**129**(6):511–515.
23. Schmiegelow M, Lassen S, Weber L, Poulsen HS, Hertz H, Müller J. Dosimetry and growth hormone deficiency following cranial irradiation of childhood brain tumors. *Med Pediatr Oncol*. 1999; **33**(6):564–571.
24. Chinoy A, Murray PG. Diagnosis of growth hormone deficiency in the paediatric and transitional age. *Best Pract Res Clin Endocrinol Metab*. 2016;**30**(6):737–747.
25. Chemaitilly W, Li Z, Huang S, Ness KK, Clark KL, Green DM, Barnes N, Armstrong GT, Krasin MJ, Srivastava DK, Pui CH, Merchant TE, Kun LE, Gajjar A, Hudson MM, Robison LL, Sklar CA. Anterior hypopituitarism in adult survivors of childhood cancers treated with cranial radiotherapy: a report from the St Jude Lifetime Cohort study. *J Clin Oncol*. 2015;**33**(5):492–500.
26. Grimberg A, DiVall SA, Polychronakos C, Allen DB, Cohen LE, Quintos JB, Rossi WC, Feudtner C, Murad MH; Drug and Therapeutics Committee and Ethics Committee of the Pediatric Endocrine Society. Guidelines for growth hormone and insulin-like growth factor-I treatment in children and adolescents: growth hormone deficiency, idiopathic short stature, and primary insulin-like growth factor-I deficiency. *Horm Res Paediatr*. 2016;**86**(6): 361–397.
27. Clemmons DR. Consensus statement on the standardization and evaluation of growth hormone and insulin-like growth factor assays. *Clin Chem*. 2011;**57**(4):555–559.
28. Chesover AD, Dattani MT. Evaluation of growth hormone stimulation testing in children. *Clin Endocrinol (Oxf)*. 2016;**84**(5): 708–714.
29. Ghigo E, Bellone J, Aimaretti G, Bellone S, Loche S, Cappa M, Bartolotta E, Dammacco F, Camanni F. Reliability of provocative tests to assess growth hormone secretory status. Study in 472 normally growing children. *J Clin Endocrinol Metab*. 1996;**81**(9): 3323–3327.
30. Secco A, di Iorgi N, Napoli F, Calandra E, Ghezzi M, Frassinetti C, Parodi S, Casini MR, Lorini R, Loche S, Maghnie M. The glucagon test in the diagnosis of growth hormone deficiency in children with short stature younger than 6 years. *J Clin Endocrinol Metab*. 2009; **94**(11):4251–4257.
31. Hilczer M, Smyczynska J, Stawerska R, Lewinski A. Stability of IGF-I concentration despite divergent results of repeated GH stimulating tests indicates poor reproducibility of test results. *Endocr Regul*. 2006;**40**(2):37–45.
32. Carel JC, Tresca JP, Letrait M, Chaussain JL, Lebouc Y, Job JC, Coste J. Growth hormone testing for the diagnosis of growth hormone deficiency in childhood: a population register-based study. *J Clin Endocrinol Metab*. 1997;**82**(7):2117–2121.
33. Lanes R. Diagnostic limitations of spontaneous growth hormone measurements in normally growing prepubertal children. *Am J Dis Child*. 1989;**143**(11):1284–1286.
34. Zadik Z, Chalew SA, Gilula Z, Kowarski AA. Reproducibility of growth hormone testing procedures: a comparison between 24-hour integrated concentration and pharmacological stimulation. *J Clin Endocrinol Metab*. 1990;**71**(5):1127–1130.
35. Gordon MB, Levy RA, Gut R, Germak J. Trends in growth hormone stimulation testing and growth hormone dosing in adult growth hormone deficiency patients: results from the Answer Program. *Endocr Pract*. 2016;**22**(4):396–405.
36. Growth Hormone Research Society. Consensus guidelines for the diagnosis and treatment of growth hormone (GH) deficiency in childhood and adolescence: summary statement of the GH Research Society. *J Clin Endocrinol Metab*. 2000;**85**(11):3990–3993.
37. Brownstein CM, Mertens AC, Mitby PA, Stovall M, Qin J, Heller G, Robison LL, Sklar CA. Factors that affect final height and change in height standard deviation scores in survivors of childhood cancer treated with growth hormone: a report from the childhood cancer survivor study. *J Clin Endocrinol Metab*. 2004; **89**(9):4422–4427.
38. Rastogi MV, Stork L, Druker B, Blasdel C, Nguyen T, Boston BA. Imatinib mesylate causes growth deceleration in pediatric patients with chronic myelogenous leukemia. *Pediatr Blood Cancer*. 2012; **59**(5):840–845.
39. Tauer JT, Hofbauer LC, Jung R, Gerdes S, Glauche I, Erben RG, Suttorp M. Impact of long-term exposure to the tyrosine kinase inhibitor imatinib on the skeleton of growing rats. *PLoS One*. 2015;**10**(6):e0131192.
40. Noyes JJ, Levine MA, Belasco JB, Mostoufi-Moab S. Premature epiphyseal closure of the lower extremities contributing to short stature after cis-retinoic acid therapy in medulloblastoma: a case report. *Horm Res Paediatr*. 2016;**85**(1):69–73.

DICER1 Mutations Are Frequent in Adolescent-Onset Papillary Thyroid Carcinoma

Jonathan D. Wasserman,[1,2]* Nelly Sabbaghian,[3] Somayyeh Fahiminiya,[4] Rose Chami,[5] Ozgur Mete,[6] Meryl Acker,[1] Mona K. Wu,[3,7] Adam Shlien,[5,8] Leanne de Kock,[3,7]* and William D. Foulkes[3,4,7]*

[1]Division of Endocrinology, The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada; [2]Department of Pediatrics, University of Toronto, Toronto, Ontario M5S 1A1, Canada; [3]Lady Davis Institute, Segal Cancer Centre, Jewish General Hospital, Montreal, Quebec H3T 1E2, Canada; [4]Department of Medical Genetics, Research Institute of the McGill University Health Centre, Montreal, Quebec H4A 3J1, Canada; [5]Division of Pathology, Department of Pediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada; [6]Department of Pathology, University Health Network, University of Toronto, Toronto, Ontario M5G 2C4, Canada; [7]Department of Human Genetics, McGill University, Montreal, Quebec H3A 0C7, Canada; and [8]Program in Genetics and Genome Biology, SickKids Research Institute, Toronto, Ontario M5G 1X8, Canada

Context: Papillary thyroid carcinoma (PTC) is a common malignancy in adolescence and is molecularly and clinically distinct from adult PTC. Mutations in the *DICER1* gene are associated with thyroid abnormalities, including multinodular goiter and differentiated thyroid carcinoma.

Objective: In this study, we sought to characterize the prevalence of *DICER1* variants in pediatric PTC, specifically in tumors without conventional PTC oncogenic alterations.

Patients: Patients (N = 40) who underwent partial or total thyroidectomy and who were <18 years of age at the time of surgery were selected.

Design: The 40 consecutive thyroidectomy specimens (30 malignant, 10 benign) underwent genotyping for 17 PTC-associated variants, as well as full sequencing of the exons and exon-intron boundaries of *DICER1*.

Results: Conventional alterations were found in 12 of 30 (40%) PTCs (five $BRAF^{V600E}$, three *RET/PTC1*, four *RET/PTC3*). Pathogenic *DICER1* variants were identified in 3 of 30 (10%) PTCs and in 2 of 10 (20%) benign nodules, all of which lacked conventional alterations and did not recur during follow-up. *DICER1* alterations thus constituted 3 of 18 (16.7%) PTCs without conventional alterations. The three *DICER1*-mutated carcinomas each had two somatic *DICER1* alterations, whereas two follicular-nodular lesions arose in those with germline *DICER1* mutations and harbored characteristic second somatic RNase IIIb "hotspot" mutations.

Conclusions: *DICER1* is a driver of pediatric thyroid nodules, and *DICER1*-mutated PTC may represent a distinct class of low-risk malignancies. Given the prevalence of variants in children, we advocate for inclusion of *DICER1* sequencing and gene dosage determination in molecular analysis of pediatric thyroid specimens. **(*J Clin Endocrinol Metab* 103: 2009–2015, 2018)**

Thyroid carcinoma is the most common malignancy in adolescent and young adult women, and incidence is increasing across all age groups (1, 2). Nodular thyroid disease is also common in adolescence, and increasing access to cervical ultrasound has exposed high rates of previously subclinical thyroid nodules.

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received 12 December 2017. Accepted 16 February 2018.
First Published Online 21 February 2018

*These authors contributed equally to this study.

Abbreviations: miRNA, microRNA; PCR, polymerase chain reaction; PTC, papillary thyroid carcinoma.

Thyroid carcinoma in children and adolescents presents with several distinct clinical features when compared with adults. First, there is a substantially higher rate of regional and distally metastatic disease. In addition, recurrence rates are higher in this age group (3). Paradoxically, despite higher initial burden of disease, outcomes in children and adolescents are excellent (4). Certain histologic variants, including the diffuse sclerosing and solid variants of papillary thyroid carcinoma (PTC) are identified more commonly in children and adolescents (5, 6). Biologic and molecular mechanisms explaining differing clinical behavior of pediatric and adult PTC have thus far been elusive. A discrete set of conventional somatic variants in *BRAF*, *HRAS*, *NRAS*, and *KRAS* or gene fusions involving *RET* or *PAX8* are identified in up to 80% of adult thyroid carcinomas (7, 8). However, several recent studies assessing these variants in children have suggested a lower frequency of variants in these genes than in adults, resulting in a higher proportion of pediatric tumors lacking common oncogenic variants, so called "dark-matter" cases (Supplemental Table 1). The oncogenic drivers in this "dark-matter" subset remain unidentified.

Aberrant regulation of microRNA (miRNA) processing has been associated with multiple human malignancies (9–11), including differentiated and anaplastic thyroid carcinomas (12–15). The *DICER1* gene product, DICER1, is an endoribonuclease responsible for processing RNA into small interfering RNA and miRNA, which subsequently posttranscriptionally downregulate messenger RNAs. Pathogenic, typically truncating germline variants in *DICER1* cause DICER1 syndrome [Online Mendelian Inheritance in Man (OMIM) 601200], characterized by a spectrum of benign and malignant tumors of mainly pediatric onset. These include pleuropulmonary blastoma, ovarian Sertoli-Leydig cell tumor, cystic nephroma, pituitary blastoma, and multinodular goiter. Truncating *DICER1* mutations are accompanied by characteristic somatic missense mutations affecting highly specific metal ion-binding residues within the RNase IIIb domain in syndrome-related tumors (16).

In this study, we sought to define the prevalence of *DICER1* variants in 40 (30 malignant and 10 benign) consecutive pediatric thyroidectomy specimens (Table 1; Supplemental Table 2).

Materials and Methods

Patients and samples

This study was approved by the Hospital for Sick Children Research Ethics Board (project number 1000039078) and the institutional review board of the Research Institute of the McGill University Health Centre (project number MP-37-2017-2949). Patients undergoing partial or total thyroidectomy who were <18 years of age at the time of surgery were identified, and informed consent for participation in this study was obtained. All patients were assessed for a personal or family history of syndromic cancers. Patients were excluded if the indication for thyroidectomy was Graves disease without focal lesion or if consent was declined. Forty consecutive pediatric thyroidectomy specimens were obtained (30 malignant, 10 benign) from patients without a personal or family history of syndromic cancers (Table 1). Patients 3 and 6 had prior exposure to therapeutic ionizing radiation. No other patients had a history of radiation exposure or iodine deficiency, and no patients received thyroid hormone treatment prior to thyroidectomy.

Patients were followed postoperatively for a median of 22.4 months (range: 0.3 to 94.4 months). Carriers of *DICER1*-variant tumors were followed for a median of 21 months (6.3 to 94.4 months), whereas those with conventional-variant tumors were followed for 34.4 months (3.6 to 93.3 months) (Supplemental Table 3). The duration of follow-up was not statistically different for these two groups ($P = 0.27$). Patients were considered to have no evidence of disease at follow-up if there was no structural disease identified on ultrasound or radioactive iodine scan, and thyroglobulin was less than 0.5 μg/L with levothyroxine-suppressed thyroid-stimulating hormone. Elevated thyroglobulin without structural correlate was considered biochemically persistent disease, and evidence of persistent malignancy based on imaging was classified as structurally persistent disease. There were no patients who achieved no evidence of disease who subsequently demonstrated disease recurrence.

Table 1. Baseline Characteristics of the Cases Studied

	PTC Classic Variant	PTC Follicular Variant	PTC Diffuse Sclerosing Variant	PTC Solid Variant	Benign Lesions
Number of cases (N = 40)	17	6	6	1	10
Sex					
Female	8	5	4	1	10
Male	9	1	2	0	0
Mean age, y (range)	12.9 (5–17)	15.1 (11–17)	10.6 (8–13)	15	14.4 (11–18)
Size of dominant nodule, mm (range)	28.5 (9–60)				29 (7–40)
Nodal metastases present, n (%)	10 (59%)	0 (0%)	6 (100%)	0 (0%)	NA
Persistent/recurrent disease, n (%)	7 (41%)	0 (0%)	2 (33%)	0 (0%)	NA

Abbreviations: NA, not applicable.

Samples were snap-frozen in liquid nitrogen within 30 minutes of excision. DNA was extracted from tissue using the AllPrep DNA/RNA/miRNA Universal Kit (QIAGEN, Germantown, MD). Constitutional DNA was extracted from peripheral blood mononuclear cells using the Gentra Puregene Blood Kit (QIAGEN) in the Clinical Laboratory Improvement Amendments-College of American Pathologists molecular diagnostic laboratories at The Hospital for Sick Children.

Pathology review

Hematoxylin and eosin-stained pathology specimens were reviewed by two pathologists (O.M. and R.C.) and reported according to a standardized approach. In the case of discrepancy between reviewers, cases were discussed until consensus was achieved. One specimen (case 8) was unavailable for review, and the original pathologists' interpretation was used for the purposes of this study.

Genotyping of conventional PTC-associated alterations

Genotyping for the 17 most common thyroid carcinoma associated genetic alterations in *BRAF*, *HRAS*, *KRAS*, *NRAS*, *RET*, and *PAX8* (17) was performed by allele-specific quantitative polymerase chain reaction (PCR)/quantitative reverse transcription-PCR using the Thyroid Cancer Mutation Detection Kit and the Thyroid Cancer Fusion Detection Kit (Entrogen Inc., Woodland Hills, CA). Equivocal results were obtained for seven samples that underwent gene fusion analysis. These results were validated using a Digital Droplet PCR assay on a QX200 Digital Droplet PCR System (Bio-Rad, Hercules, CA) utilizing the same oligonucleotides used for the Thyroid Cancer Fusion Detection Kit assay (Entrogen Inc.). Variants screened for are detailed in the Supplemental Methods.

DICER1 sequencing

The full coding region and exon-intron boundaries of *DICER1* (NM_177438.2) were sequenced using a custom-designed Fluidigm Access Array (Fluidigm, Markham, Ontario, Canada) (18). Variants were called using FreeBayes version 0.4.1 via the Galaxy toolshed (www.usegalaxy.org/), and annotation of functional consequences was performed using wANNOVAR (http://wannovar.wglab.org/). Integrative Genomics Viewer (http://software.broadinstitute.org/software/igv) was used to manually visualize variants, which were then validated by Sanger sequencing. Matched-normal DNA was used to determine whether mutations were germline or somatic in origin.

Copy number variation analyses

CytoScanHD Array (Affymetrix, Santa Clara, CA) analysis was performed on 36 tumor samples to screen for genome-wide copy number alterations. DNA was hybridized on a CytoScanHD Array (Affymetrix), and normalization and analysis were performed using the Chromosome Analysis Suite (version 3.2.0.1252; r10346). In addition, CytoScanHD (Affymetrix) was performed in the same manner on peripheral blood mononuclear cell-derived DNA from two patients (patients 2 and 3) whose tumors were found to harbor copy number losses involving the *DICER1* locus to determine whether the identified changes were germline or somatic in origin.

Table 2. Cases With Conventional Somatic PTC-Associated Variants

Case	Age at Dx (y)	Sex	Gene	Variant	Histologic Variant	Tumor Classification	Anti-Tg (kIU/L)	Size of Dominant Nodule (mm)	ATA Risk	Current Status [Follow-Up Duration (mo)]
9	15	Female	*BRAF*	p.V600E	Classic	Malignant	<20	21	Low	BPD (36)
10	17	Male	*BRAF*	p.V600E	Classic	Malignant	<20	47	Intermediate	NED (19)
11	10	Male	*BRAF*	p.V600E	Classic	Malignant	54	Diffuse	High	SPD (90)
12	16	Male	*BRAF*	p.V600E	Classic	Malignant	<20	11	Intermediate	NED (11)
13	14	Male	*BRAF*	p.V600E	Classic	Malignant	<20	32	High	NED (6)
14	15	Female	*RET/PTC1*	Rearrangement	Classic	Malignant	120	50	High	SPD (32)
15	17	Female	*RET/PTC1*	Rearrangement	Follicular	Malignant	NA	9	Low	NED (5)
16	10	Male	*RET/PTC1*	Rearrangement	Classic	Malignant	2233	35	High	SPD (62)
17	12	Female	*RET/PTC3*	Rearrangement	Classic	Malignant	3348	55	High	NED (24)
18	8	Female	*RET/PTC3*	Rearrangement	DSV	Malignant	<20	Diffuse	High	SPD (25)
19	10	Female	*RET/PTC3*	Rearrangement	DSV	Malignant	<20	15[a]	High	NED (33)
20	11	Male	*RET/PTC3*	Rearrangement	DSV	Malignant	<10	Diffuse	High	SPD (19)

Abbreviations: ATA, American Thyroid Association; BPD, biochemically persistent disease; Dx, diagnosis; DSV, diffuse sclerosing variant; NA, not available; NED, no evidence of disease; SPD, structurally persistent disease; Tg, thyroglobulin.

[a] A discrete nodule was present with diffuse infiltration of the surrounding parenchyma by malignant cells.

Table 3. Characteristics of *DICER1*-Mutated Cases

			***DICER1* Screening Result**				
Case	**Age at Dx (y)**	**Sex**	**Variant(s)**	**Predicted Consequence**	**Germline vs Somatic Origin**	**ExAC MAF**	**PMH**
1	16.5	Female	c.2875A>T, p.K959*	Deleterious	Somatic	N/A	0
			c.5125G>A, p.D1709N	Deleterious	Somatic	N/A	
2	14	Female	c.5428G>T, p.D1810Y	Deleterious	Somatic	N/A	0
			LOH of *DICER1* (del chr14:94,043,795-104,822,229)	Deleterious	Somatic	N/A	
			c.1124C>G, p.P375R (rs148758903)	Likely benign	Germline[a]	0.0003545	
3	11.7	Female	c.5439G>C, p.E1813D	Deleterious	Somatic	N/A	ALL,TBI, HSCT
			LOH of *DICER1* (del chr14:78,529,021-100,616,514)	Deleterious	Somatic	N/A	
4	10	Male	c.4260_4262delGGA, p.E1420del (rs544960260)[b]	Likely benign	Not known	0.001588	0
5	15	Female	c.2997T>G, p.L999L (rs12018992)	Silent	Germline	0.005387	0
6	17.4	Female	c.20A>G, p.Q7R (rs117358479)	Likely benign	Not known	0.00181	ALL, TBI
7	12	Female	c.2535_2539delinsAATCAACTTCAAGCATT, p.T847delinsNFKHS[d]	Deleterious	Germline	Not in ExAC	0
			c.5438A>G, p.E1813G	Deleterious	Somatic	N/A	
8	16	Female	c.84dupT, p.G29Wfs*11	Deleterious	Germline	Not in ExAc	0
			c.5125G>A, p.D1709N	Deleterious	Somatic	N/A	

(*Continued*)

Results

Conventional PTC-associated alterations were identified in 12 of 30 (40%) malignant lesions: five cases with $BRAF^{V600E}$, three with *RET/PTC1*, and four with *RET/PTC3* (Table 2), consistent with published results (Supplemental Table 1). Sequencing of *DICER1* identified pathogenic mutations in 3 of 30 (10%) PTCs, all of which lacked conventional alterations, thus accounting for 16.7% (3 of 18) of "dark-matter" cases. Two of 10 (20%) benign lesions were also found to be *DICER1* mutated (Table 3; Fig. 1; Supplemental Table 2; Supplemental Figs. 1–4). The three malignant lesions each harbored two somatically acquired *DICER1* alterations, whereas the patients with *DICER1*-mutated benign nodules each carried a germline truncating mutation and an additional somatic RNase IIIb "hotspot" mutation in the nodule, as has been observed previously (19–22). One additional tumor (case 5) carried a synonymous *DICER1* variant, and two others (cases 4 and 6) each harbored a polymorphism classified as likely benign according to ClinVar (www.ncbi.nlm.nih.gov/clinvar/; variation IDs: 133971 and 133962, respectively). Neither case showed evidence of a second hit within the *DICER1* locus.

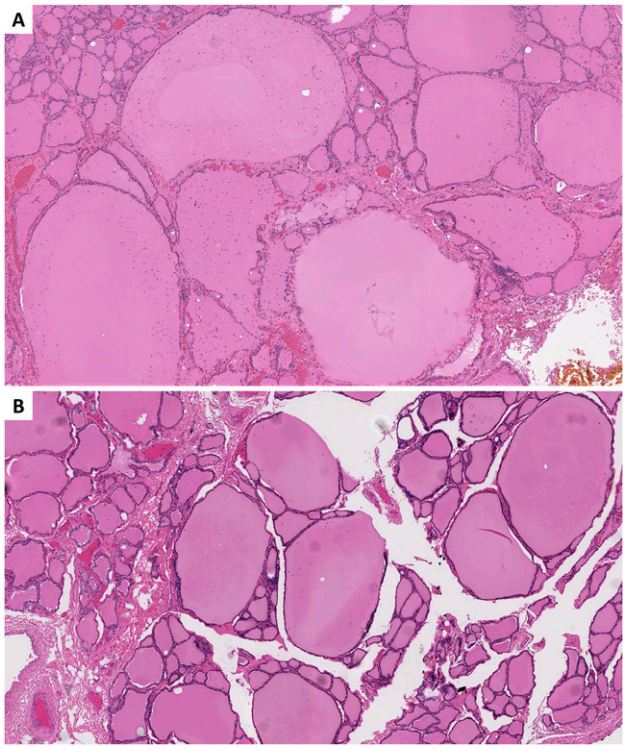


Figure 1. Representative *DICER1*-mutated thyroid neoplasms are illustrated. (A) Case 2: Encapsulated follicular-variant papillary carcinoma with minimal capsular invasion. The nontumorous thyroid parenchyma showed variable involutional changes characterized by dilated colloid containing follicles and slightly flattened epithelium. (B) Case 7: Multiple follicular nodular disease. The nonlesional thyroid parenchyma exhibited variable involutional changes. These findings are reminiscent of functional nodules despite a normal preoperative thyroid-stimulating hormone (0.64 mIU/L).

Discussion

The prevalence of pathogenic *DICER1* variants in this series of pediatric thyroid tumors was determined to be 12.5% (5 of 40) (Supplemental Table 3). Two large

Table 3. Continued

FH	Pre-Op TSH (mIU/L)/Free T4 (pmol/L)	Anti-Tg (kIU/L)	Tumor Pathology, Staging, and Risk					Current Status [Duration (mo)]
			Histologic Variant	Tumor Classification	Size of Dominant Nodule (mm)	AJCC Stage	ATA Risk	
0	0.81/13.8	<20	Encapsulated classical PTC	Malignant	60	$T_3N_0M_0$	Low	NED (62)
0	1.84/16.0	<20	Minimally invasive, encapsulated FV PTC	Malignant	37	$T_2N_0M_x$	Low	NED (41)
0	1.41/11.0	<10	Infiltrative classical PTC	Malignant	25	$T_2N_0M_x$	Low	NED (22)
0	Normal[c]	<10	Classical PTC with focal hobnail and tall cell change	Malignant	35	$T_2N_0M_0$	Low	NED (94)
MNG	0.96/13.0	<20	Minimally invasive solid-variant PTC	Malignant	13	$T_{1b}N_xM_x$	Low	NED (38)
0	2.05/10.8	Not done	(1) Minimally invasive FV PTC (2) Classical-variant papillary microcarcinoma	Malignant	12 3	$T_{1b}N_xM_x$	Low	NED (9)
MNG	0.64/11.9	Not done	Multifocal FND with papillary growth	Benign	28	N/A	—	NED (64)
0	1.28/21	Not done	FND	Benign	20	N/A	—	NED (61)

Abbreviations: ALL, acute lymphoblastic leukemia; AJCC, American Joint Committee on Cancer; ATA, American Thyroid Association; Dx, diagnosis; ExAC, Exome Aggregation Consortium; FH, family history; FND, follicular nodular disease; FV, follicular variant; HSCT, hematopoietic stem cell transplantation; LOH, loss of heterozygosity; MAF, minor allele frequency; MNG, multinodular goiter; N/A, not applicable; NED, no evidence of disease; PMH, personal medical history; TBI, total body irradiation; Tg, thyroglobulin; TSH, thyroid-stimulating hormone.

[a]The germline variant is encompassed by the deletion in the tumor and is therefore in trans with the p.D1810Y mutation.

[b]*In silico* predictions for variant rs544960260 (case 4): tolerated by Sorting Intolerant From Tolerant (SIFT) (score 0.11), benign by PolyPhen-2 (score 0), Combined Annotation Dependent Depletion (CADD)-scaled C-score: 3.877 (consistent with benign), and Rare Exome Variant Ensemble Learner (REVEL) score: 0.048 (consistent with benign).

[c]TSH was performed at a referring facility and was reported only as "normal." It was not repeated preoperatively.

[d]Functional analyses determined that this in-frame *DICER1* variant impedes processing of miRNAs (see Supplemental Fig. 4 for details).

genome-wide analyses have identified low rates of pathogenic *DICER1* variants in adult thyroid tumors. The Cancer Genome Atlas project identified pathogenic somatic *DICER1* variants in 2 of 401 (0.5%) PTCs (Supplemental Table 4) (7), whereas a subsequent study identified no *DICER1*-variant tumors among 125 PTC, and two tumors with *DICER1* variants in each of 30 minimally invasive follicular thyroid cancers and 25 follicular adenomas (8). *DICER1*-mutated PTCs are thus more frequent in the pediatric (3 of 30) than in the adult setting (2 of 526; $P = 0.0013$, two-sided Fisher exact test).

Although capsular or parenchymal invasion was observed among all *DICER1*-mutated thyroid carcinomas in our series, angioinvasion was only identified in one (case 3). None were associated with local or regional lymph node or distal metastases, and all are thus classified as low risk for persistent/recurrent disease according to the recent American Thyroid Association pediatric management guidelines (3). In contrast, four of five (80%) $BRAF^{V600E}$ tumors and six of seven (86%) *RET/PTC* tumors were associated with nodal metastases. Antithyroglobulin titers were undetectable in all individuals with *DICER1* tumors and were present in 7 of 27 (26%) of the remaining individuals with PTC, a proportion concordant with recent studies (23). Patients with *DICER1*-mutated tumors remained disease-free at a median of 21 months follow-up (range: 9.9 to 37.4 months) (Supplemental Table 3). None of the *DICER1* tumors had a chronic lymphocytic infiltrate, and four cases exhibited focal nonspecific thyroiditis. Lastly, in six of seven *DICER1*-mutated available specimens reviewed, the background thyroid parenchyma showed variable involutional changes similar to alterations associated with exposure to elevated thyroid hormone concentrations (Table 3; Fig. 1; Supplemental Fig. 1).

DICER1 alterations impact several miRNA species in thyroid lesions (24). Although multiple studies have demonstrated alterations in miRNA sequence and expression in thyroid malignancies (7, 25–27), their mechanistic implications and role in thyroid tumorigenesis remain to be fully

elucidated. It also remains to be seen whether alterations in other components of the miRNA processing machinery (*DROSHA*, *DGCR8*, *AGO2*, *etc.*) are altered in thyroid carcinoma.

Conclusion

DICER1 mutations in pediatric PTC are present at a frequency nearly 30 times that seen in adult PTC. *DICER1* malignancies comprised 16.7% of dark-matter tumors in this pediatric series and were not associated with thyroid autoimmunity. These data establish *DICER1* as a common oncogenic driver in American Thyroid Association pediatric low-risk PTC and broaden our understanding of the molecular pathogenesis of pediatric PTCs. Further delineation of the role of miRNA processing and aberrations in thyroid carcinogenesis may expand our insight into prognosis and potentially therapeutic approaches.

Acknowledgments

We thank Marc Fabian, PhD, for assistance with the *in vitro* cleavage assay.

Financial Support: This work was supported by the SickKids Foundation (to J.D.W.), the C^{17} Research Network, which was funded by the Childhood Cancer Canada Foundation and the Kids With Cancer Society (to J.D.W.), the Canadian Institute of Health Research Foundation Grant FDN-148390 (to W.D.F.), and the Vanier Canada Graduate Scholarship and DKG World Fellowship (to L.d.K.).

Author Contributions: W.D.F. and J.D.W. designed the study. W.D.F., J.D.W., and L.d.K. wrote the manuscript. N.S. and M.K.W. performed the *DICER1* genetic testing. S.F. performed bioinformatics analysis. O.M. and R.C. performed the pathology review. J.D.W., M.A., and A.S. performed oncogene genetic analysis, sample acquisition, and processing.

Correspondence and Reprint Requests: William D. Foulkes, MBBS, PhD, Department of Medical Genetics, Cancer Research Program, Research Institute of the McGill University Health Center, 1001 Décarie Boulevard, Montreal, Quebec H4A 3J1, Canada. E-mail: william.foulkes@mcgill.ca.

Disclosure Summary: The authors have nothing to disclose.

References

1. Pole JD, Zuk AM, Wasserman JD. Diagnostic and treatment patterns among children, adolescents, and young adults with thyroid cancer in Ontario: 1992-2010. *Thyroid*. 2017;**27**(8):1025–1033.
2. Canadian Cancer Society's Steering Committee on Cancer Statistics. *Canadian Cancer Statistics 2012*. Toronto, Canada: Canadian Cancer Society; 2012.
3. Francis GL, Waguespack SG, Bauer AJ, Angelos P, Benvenga S, Cerutti JM, Dinauer CA, Hamilton J, Hay ID, Luster M, Parisi MT, Rachmiel M, Thompson GB, Yamashita S; American Thyroid Association Guidelines Task Force. Management guidelines for children with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2015;**25**(7):716–759.
4. Golpanian S, Perez EA, Tashiro J, Lew JI, Sola JE, Hogan AR. Pediatric papillary thyroid carcinoma: outcomes and survival predictors in 2504 surgical patients. *Pediatr Surg Int*. 2016;**32**(3):201–208.
5. Koo JS, Hong S, Park CS. Diffuse sclerosing variant is a major subtype of papillary thyroid carcinoma in the young. *Thyroid*. 2009;**19**(11):1225–1231.
6. Nikiforov Y, Gnepp DR, Fagin JA. Thyroid lesions in children and adolescents after the Chernobyl disaster: implications for the study of radiation tumorigenesis. *J Clin Endocrinol Metab*. 1996;**81**(1):9–14.
7. Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell*. 2014;**159**(3):676–690.
8. Yoo SK, Lee S, Kim SJ, Jee HG, Kim BA, Cho H, Song YS, Cho SW, Won JK, Shin JY, Park J, Kim JI, Lee KE, Park YJ, Seo JS. Comprehensive analysis of the transcriptional and mutational landscape of follicular and papillary thyroid cancers. *PLoS Genet*. 2016;**12**(8):e1006239.
9. Markopoulos GS, Roupakia E, Tokamani M, Chavdoula E, Hatziapostolou M, Polytarchou C, Marcu KB, Papavassiliou AG, Sandaltzopoulos R, Kolettas E. A step-by-step microRNA guide to cancer development and metastasis. *Cell Oncol (Dordr)*. 2017;**40**(4):303–339.
10. Sarver AE, Li L, Kartha RV, Subramanian S. microRNAs in the malignant transformation process. *Adv Exp Med Biol*. 2015;**889**:1–21.
11. Jansson MD, Lund AH. MicroRNA and cancer. *Mol Oncol*. 2012;**6**(6):590–610.
12. Fuziwara CS, Kimura ET. MicroRNA deregulation in anaplastic thyroid cancer Biology. *Int J Endocrinol*. 2014;**2014**(743450):1–8.
13. Lee JC, Gundara JS, Glover A, Serpell J, Sidhu SB. MicroRNA expression profiles in the management of papillary thyroid cancer. *Oncologist*. 2014;**19**(11):1141–1147.
14. de Kock L, Sabbaghian N, Soglio DB, Guillerman RP, Park BK, Chami R, Deal CL, Priest JR, Foulkes WD. Exploring the association Between DICER1 mutations and differentiated thyroid carcinoma. *J Clin Endocrinol Metab*. 2014;**99**(6):E1072–E1077.
15. Rutter MM, Jha P, Schultz KA, Sheil A, Harris AK, Bauer AJ, Field AL, Geller J, Hill DA. DICER1 mutations and differentiated thyroid carcinoma: evidence of a direct association. *J Clin Endocrinol Metab*. 2016;**101**(1):1–5.
16. Foulkes WD, Priest JR, Duchaine TF. DICER1: mutations, microRNAs and mechanisms. *Nat Rev Cancer*. 2014;**14**(10):662–672.
17. Nikiforov YE, Ohori NP, Hodak SP, Carty SE, LeBeau SO, Ferris RL, Yip L, Seethala RR, Tublin ME, Stang MT, Coyne C, Johnson JT, Stewart AF, Nikiforova MN. Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: a prospective analysis of 1056 FNA samples. *J Clin Endocrinol Metab*. 2011;**96**(11):3390–3397.
18. de Kock L, Sabbaghian N, Plourde F, Srivastava A, Weber E, Bouron-Dal Soglio D, Hamel N, Choi JH, Park SH, Deal CL, Kelsey MM, Dishop MK, Esbenshade A, Kuttesch JF, Jacques TS, Perry A, Leichter H, Maeder P, Brundler MA, Warner J, Neal J, Zacharin M, Korbonits M, Cole T, Traunecker H, McLean TW, Rotondo F, Lepage P, Albrecht S, Horvath E, Kovacs K, Priest JR, Foulkes WD. Pituitary blastoma: a pathognomonic feature of germ-line DICER1 mutations. *Acta Neuropathol*. 2014;**128**(1):111–122.
19. Khan NE, Bauer AJ, Schultz KAP, Doros L, Decastro RM, Ling A, Lodish MB, Harney LA, Kase RG, Carr AG, Rossi CT, Field A, Harris AK, Williams GM, Dehner LP, Messinger YH, Hill DA, Stewart DR. Quantification of thyroid cancer and multinodular

goiter risk in the DICER1 syndrome: a family-based cohort study. *J Clin Endocrinol Metab*. 2017;**102**(5):1614–1622.
20. de Kock L, Bah I, Wu Y, Xie M, Priest JR, Foulkes WD. Germline and somatic DICER1 mutations in a well-differentiated fetal adenocarcinoma of the lung. *J Thorac Oncol*. 2015;**11**(3):e31–e33.
21. Wu MK, de Kock L, Conwell LS, Stewart CJ, King BR, Choong CS, Hussain K, Sabbaghian N, MacRae IJ, Fabian MR, Foulkes WD. Functional characterization of multiple DICER1 mutations in an adolescent. *Endocr Relat Cancer*. 2016;**23**(2):L1–L5.
22. de Kock L, Bah I, Revil T, Bérubé P, Wu MK, Sabbaghian N, Priest JR, Ragoussis J, Foulkes WD. Deep sequencing reveals spatially distributed distinct hot spot mutations in DICER1-related multinodular goiter. *J Clin Endocrinol Metab*. 2016;**101**(10): 3637–3645.
23. Görges R, Maniecki M, Jentzen W, Sheu SN, Mann K, Bockisch A, Janssen OE. Development and clinical impact of thyroglobulin antibodies in patients with differentiated thyroid carcinoma during the first 3 years after thyroidectomy. *Eur J Endocrinol*. 2005;**153**(1):49–55.
24. Rio Frio T, Bahubeshi A, Kanellopoulou C, Hamel N, Niedziela M, Sabbaghian N, Pouchet C, Gilbert L, O'Brien PK, Serfas K, Broderick P, Houlston RS, Lesueur F, Bonora E, Muljo S, Schimke RN, Bouron-Dal Soglio D, Arseneau J, Schultz KA, Priest JR, Nguyen VH, Harach HR, Livingston DM, Foulkes WD, Tischkowitz M. DICER1 mutations in familial multinodular goiter with and without ovarian Sertoli-Leydig cell tumors. *JAMA*. 2011; **305**(1):68–77.
25. Pishkari S, Paryan M, Hashemi M, Baldini E, Mohammadi-Yeganeh S. The role of microRNAs in different types of thyroid carcinoma: a comprehensive analysis to find new miRNA supplementary therapies. *J Endocrinol Invest*. 2018;**41**(3):269–283.
26. Wang T, Xu H, Qi M, Yan S, Tian X. miRNA dysregulation and the risk of metastasis and invasion in papillary thyroid cancer: a systematic review and meta-analysis. *Oncotarget*. 2017;**9**(4): 5473–5479.
27. Chruścik A, Lam AK. Clinical pathological impacts of microRNAs in papillary thyroid carcinoma: A crucial review. *Exp Mol Pathol*. 2015;**99**(3):393–398.

GH Therapy in Childhood Cancer Survivors: A Systematic Review and Meta-Analysis

Shrikant Tamhane,[1,2] Jad G. Sfeir,[1,2] Nana Esi N. Kittah,[2] Sina Jasim,[1,3] Wassim Chemaitilly,[4] Laurie E. Cohen,[5] and M. Hassan Murad[1]

[1]Evidence-Based Practice Center, Mayo Clinic, Rochester, Minnesota 55905; [2]Division of Endocrinology, Mayo Clinic, Rochester, Minnesota 55905; [3]Division of Endocrinology, Metabolism and Lipid Research, Washington University in St. Louis, School of Medicine, St. Louis, Missouri 63130; [4]Department of Pediatric Medicine, Division of Endocrinology, St. Jude Children's Research Hospital, Memphis, Tennessee 38105; and [5]Division of Endocrinology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115

Background: GH deficiency (GHD) is common among childhood cancer survivors (CCSs) with history of tumors, surgery, and/or radiotherapy involving the hypothalamus-pituitary region. We aimed to evaluate the effects of GH therapy (GHT) in CCSs on adult height, risk of diabetes mellitus, abnormal lipids, metabolic syndrome, quality of life, secondary tumors, and disease recurrence.

Methods: We searched multiple databases for randomized and observational studies. Pairs of reviewers independently selected studies and collected data. Random effects meta-analysis was used to pool outcomes across the studies.

Results: We included 29 observational studies at moderate to high risk of bias. Sixteen studies compared CCSs on GHT with those not on GHT (512 patients, GH dose: 0.3 to 0.9 IU/kg/week). GHT was significantly associated with height gain [standard deviation score, 0.61; 95% CI, 0.08 to 1.13] and was not significantly associated with the occurrence of secondary tumors [odds ratio (OR), 1.10; 95% CI, 0.72 to 1.67] or tumor recurrence (OR, 0.57; 95% CI, 0.31 to 1.02). Thirteen studies compared CCSs on GHT with normal age- or sex-matched controls or controls with idiopathic GHD or short stature. GHT was associated with either improved or unchanged risk of diabetes, lipid profiles, and metabolic syndrome. GHT was associated with improvements in quality of life.

Conclusion: CCSs treated with GHT gain height compared with the untreated controls. GHT may improve lipid profiles and quality of life and does not appear to increase the risk of diabetes or the development of secondary tumors, although close monitoring for such complications remains warranted due to uncertainty in the current evidence. (***J Clin Endocrinol Metab*** **103: 2794–2801, 2018)**

The survival rates of children treated for cancer have improved significantly over the years. GH deficiency (GHD) is common among these children, especially among childhood cancer survivors (CCSs) with tumors/surgery in the hypothalamic-pituitary (HP) region, CCSs exposed to HP radiation or CCSs exposed to cranial (CIR), craniospinal (CSI), or total body irradiation.

GHD is one of the most common endocrinopathies observed in CCSs with a history of central nervous system tumors. Studies have shown that cancer and its treatments negatively affect adult height and that CCSs may not fully recover their growth potential even after growth hormone therapy (GHT) (1). In addition to its impact on linear growth, GHD has been shown to increase cardiovascular risk and impair quality of life

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received 1 June 2018. Accepted 4 June 2018.
First Published Online 29 June 2018

Abbreviations: CCS, childhood cancer survivor; CIR, cranial radiation; GHD, GH deficiency; GHT, GH therapy; HP, hypothalamic-pituitary; QOL, quality of life; r-hGH, recombinant hGH.

(QOL); however, no studies assessing such outcomes have been conducted specifically in CCSs (2, 3). Given the toxicity of cancer treatments, the contribution of GHD to poor health outcomes in CCSs is debatable, and whether GHT reverses these findings is unknown. Pituitary-derived human GH was used between 1958 and 1984 to treat GHD. Recombinant hGH (r-hGH) was approved for clinical use in the United States in 1985 (4) and has been used to treat GHD in CCSs. There have been continued concerns that GHT may increase the risk of tumor recurrence, secondary tumors, and other adverse effects.

To support the Endocrine Society guidelines on the management of CCSs, we conducted this systematic review and meta-analysis to evaluate the effects of GHT in CCSs with tumors/surgery in the HP region and on CCSs subjected to CIR, CSI, or total body irradiation at a young age on adult height, risk of diabetes mellitus, lipid abnormalities, metabolic syndrome, QOL, secondary tumors, and disease recurrence.

Methods

The protocol for this systematic review was developed in collaboration with members of a taskforce from the Endocrine society. This report follows the standards set in the Preferred Reporting Items for Systematic Reviews and Meta-analysis statement (5).

Eligibility criteria

We searched for cohort studies, case series, randomized clinical trials, and meta-analysis evaluating outcomes related to adult height, risk of diabetes mellitus, abnormal lipids, metabolic syndrome, QOL, secondary tumors, and disease recurrence in CCSs receiving GHT compared with those not receiving GHT. There was no language restriction, and we excluded studies with missing data despite author contact. Because we anticipated the number of eligible studies to be small, we also sought additional studies that reported the outcomes of interest in CCSs receiving GHT if they had other control groups. We summarized such studies narratively and considered them as indirect supporting evidence.

Study identification

A comprehensive search that included the Ovid Medline In-Process & Other Non-Indexed Citations, Ovid EMBASE, Web of Science, and Scopus databases was conducted beginning from each database's inception to January 2016. The search strategy was designed by a medical reference librarian with input from the study investigators. Controlled vocabulary supplemented with keywords was used to search for studies evaluating the selected outcomes in CCSs. We consulted experts in the field, references from primary studies, and Google Scholar to identify studies missed by our search strategy. Studies that were referred by the expert panel through May 2017 were also assessed.

Selection of studies

Reviewers working independently and in duplicate reviewed all abstracts and selected full-text manuscripts for eligibility. Disagreements during the full-text screening were resolved by consensus.

Data collection and management

Working independently and in duplicate, reviewers used a standardized Web-based form to collect information from each eligible study. For each study, the baseline clinical features of the included population, such as age, type of tumor, radiation details, serum GH levels, GH dose, and duration of GHT, were recorded. The outcomes of interest were focused on adult height, risk of diabetes mellitus, abnormal lipids, metabolic syndrome, QOL, secondary tumors, disease recurrence, and mortality.

Risk of bias and quality of evidence

The risk of bias was assessed by reviewers working independently and in duplicate using a modified Newcastle-Ottawa scale (6) for observational studies. Disagreements were resolved by consensus. The quality of evidence (certainty in the estimate) was graded using the Grading of Recommendation, Assessment, Development, and Evaluation approach (7).

Summary measures and synthesis of results

We performed a meta-analysis of each of the outcomes of interest using a random effects model (8). The adjusted relative risks were used preferentially if available in the studies. We used the I^2 statistic to assess heterogeneity across individual studies, and $I^2 > 50\%$ indicated a large inconsistency across studies not explained by chance. Statistical analysis was performed using STATA, version 13.1 (StataCorp LP, College Station, TX).

Results

Search results

We included 29 observational studies (Fig. 1). Sixteen studies at moderate-to-high risk of bias (Supplemental Table 3) contributed data to the meta-analysis and had a control group of CCSs not treated with GHT. The characteristics of the studies and descriptions of the patients are included in Supplemental Tables 1 and 2, respectively. Thirteen additional studies on GHT in CCSs (Supplemental Table 3) were not included in the meta-analysis because they did not have the control group of interest (CCSs not treated with GHT) but rather had one of the following control groups: (1) idiopathic GHD, (2) non–GH-related short stature, (3) pituitary cause of GHD and GHT, or (4) normal age- or sex-matched general population.

Meta-analysis

The analysis included 512 patients who received an average GH dose of 0.3 to 0.9 IU/kg/week; using the conversion formula of 3.0 IU per 1 mg for r-hGH where applicable, the average dose received was 0.1 to 0.3 mg/kg/week (9). CCSs who were treated with GHT had significant height gains compared with those not treated with GHT [weighted mean difference, 0.61; 95% CI, 0.08 to 1.13; n = 6 studies; Fig. 2]. There was no statistically significant difference in the occurrence of secondary tumors with GHT

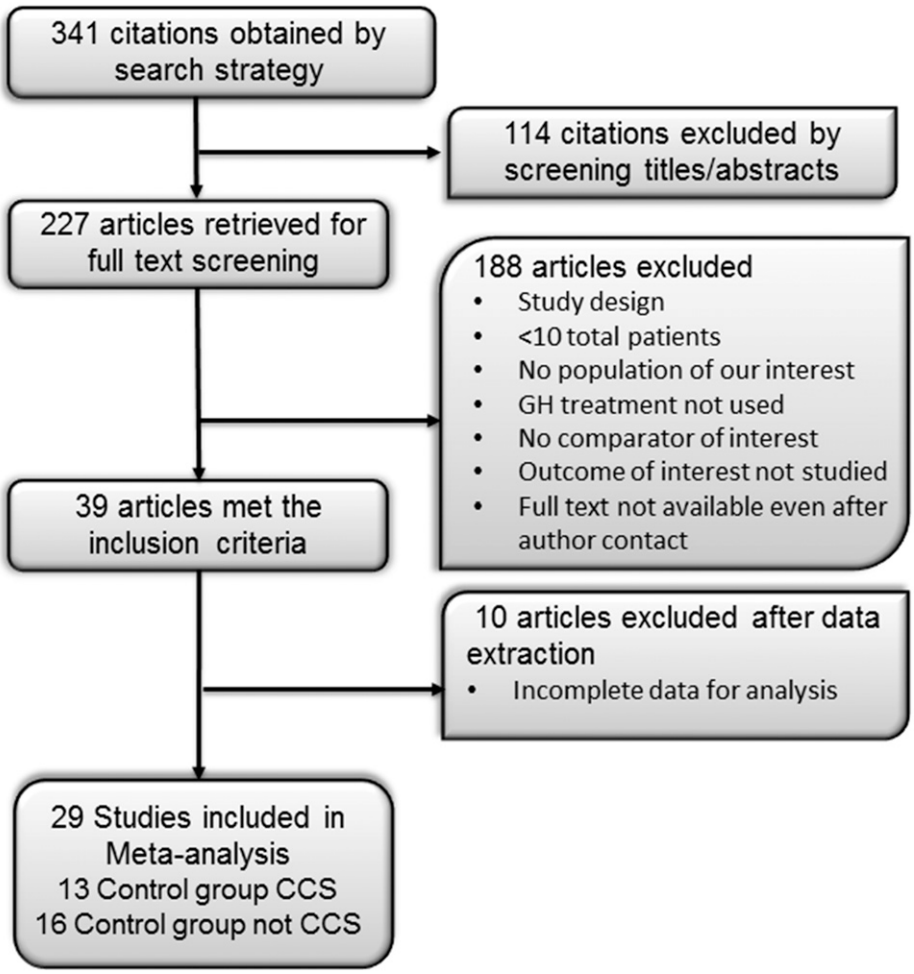


Figure 1. Study selection process.

(odds ratio, 1.10; 95% CI, 0.72 to 1.67; 5 studies; Fig. 3) and no increased risk of disease/tumor recurrence (odds ratio, 0.57; 95% CI, 0.31 to 1.02; 8 studies; Fig. 4).

Heterogeneity was substantial for height analysis ($I^2 > 90\%$) and, to a lesser extent, tumor recurrence analysis ($I^2 > 60\%$). The analysis of secondary tumors was homogeneous ($I^2 = 0\%$).

Studies comparing CCSs on GHT with other populations

Summary of the outcomes related to adult height, diabetes mellitus, lipid profile abnormalities, metabolic syndrome, QOL, secondary tumors, disease recurrence, and mortality are provided in Supplemental Tables 2.1 to 2.8.

The studies showed either an improvement or no difference in the risk of diabetes, lipid profile abnormalities, metabolic syndrome, and QOL in CCSs compared with the control groups.

Effect modifiers of height outcome

Adult height was positively correlated with the age at tumor therapy completion (10), height at the start of GHT (10, 11), height standard deviation score gain after the first year of GHT (10), age at diagnosis (1), age at irradiation (1, 12), target height (1, 13), dose of GHT (13), lower dose of radiation (14), and cotreatment with a GnRH agonist (12). Adult height was negatively correlated with CSI (10), spinal radiation (12) and dose (13), chemotherapy (12), and the presence of other endocrinopathies (13).

Methodological quality of studies and overall quality of evidence

The included studies had overall moderate risks of bias that were primarily related to the inability to control the analysis for confounders, making the pooled estimates unadjusted (Supplemental Table 3). The quality of evidence (*i.e.*, certainty in these estimates) was low because of the observational nature of the evidence, moderate risk of bias, and imprecision (small number of events).

Discussion

Main findings

GHT in CCSs is associated with a statistically significant gain in height and no apparent increases in the occurrence of secondary tumors or recurrence. There was

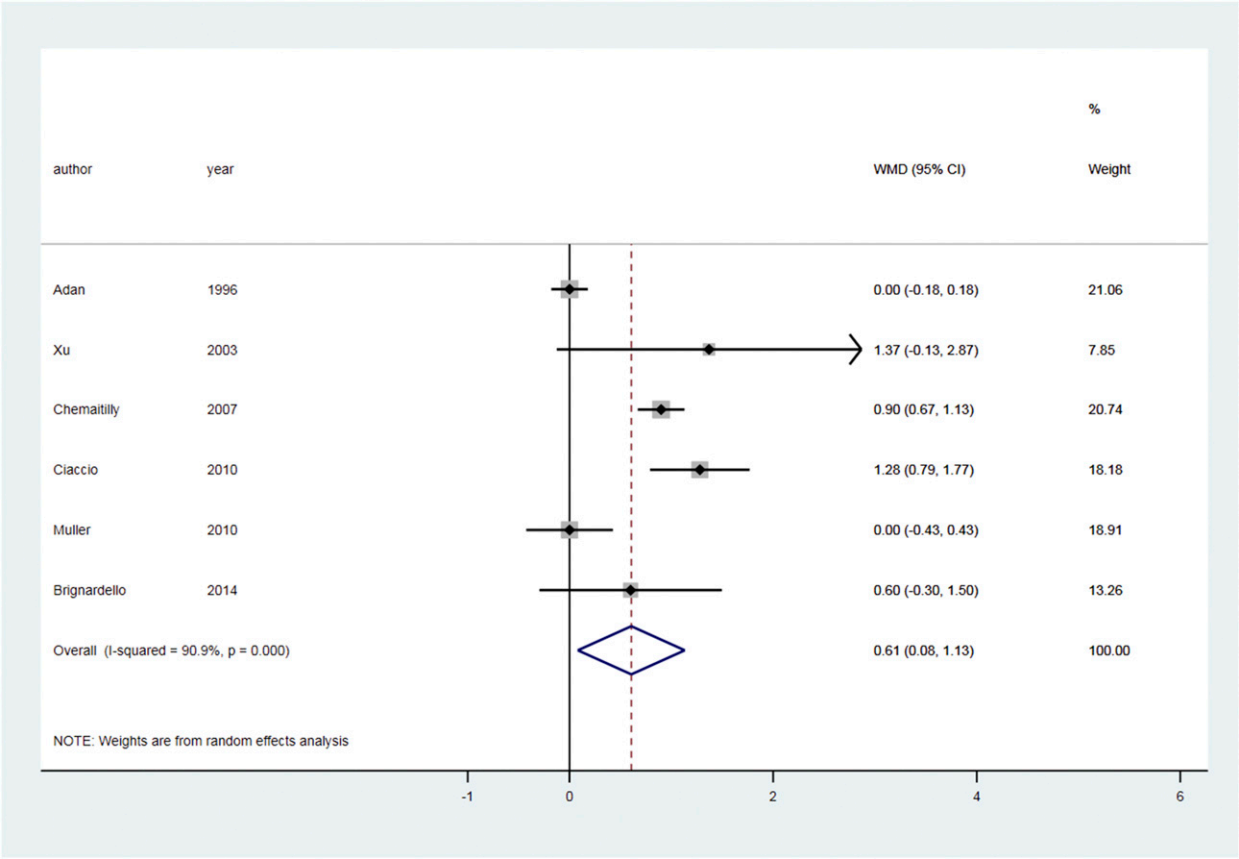


Figure 2. CCSs treated with GHT had a statistically significant gain in height compared with that in untreated survivors. If this estimate was expressed as a standardized mean difference (*i.e.*, in SD units), the results will be 0.95; 95% CI, 0.18, 1.72. WMD, weighted mean difference.

either an improvement or no difference in the risk of diabetes, lipid fractions, metabolic syndrome, and QOL.

The effect of GHT on height was heterogeneous as demonstrated by statistical measures of heterogeneity. One study (10) showed that GHT failed to restore adult height to the midparental height in nearly one-half of CCSs with radiation-induced GHD, especially in those irradiated at a young age or were short at the start of GHT. None of the patients in another study achieved his or her genetic potential height (15). The growth increment varied inversely to spontaneous GH secretion in one study (16). Another study noted that height velocity improved with GHT, with better results obtained in those who received CIR vs craniospinal radiation (17), likely because of a direct growth plate injury affecting the vertebral growth plates (18). Compared with patients with idiopathic GHD treated with GHT in the studies evaluated, CCSs treated with GHT (11, 16) showed either lower height gain or comparable growth velocity (19). In patients with medulloblastoma, adult heights but not sitting heights were improved with GHT (14, 20). Therefore, CCSs obviously gain variable improvements in height based on other factors, such as spinal injury (radiotherapy, scoliosis).

There is limited literature on the cardiovascular and metabolic effects of GHT in CCSs. Murray (21) showed that CCS patients with GHD had adverse lipid profiles, with significantly higher total cholesterol, low-density lipoprotein cholesterol, and triglyceride levels; however, after 12 months of GHT, there was no significant difference in the HbA1c and serum lipid values in these patients compared with those in the controls. However, small (but important) improvements were observed in body composition in the male subgroup and total cholesterol and triglyceride levels in the female subgroup. In another study on survivors of acute lymphoblastic leukemia (22), 5 years of GHT improved the plasma glucose levels, high-density lipoprotein cholesterol levels, ApoB/ApoA1 ratio, and prevalence of metabolic syndrome (compared with an untreated control group).

The QOL in CCSs on GHT improved dramatically by 3 months; this improvement was maintained at 12 months (21). Similar improvements across QOL domains were observed in CCSs on GHT compared with those in CCSs not treated with GHT at early (6 to 13 months) and long-term (24 to 77 months) follow-ups (23). The study on survivors of acute lymphoblastic leukemia (22), however, showed no substantial difference in QOL.

The risk of secondary tumors associated with GHT in CCSs was not statistically increased across all studies; however, considering the imprecision of the estimates, the CI did exclude an important increase in risk. Most of these

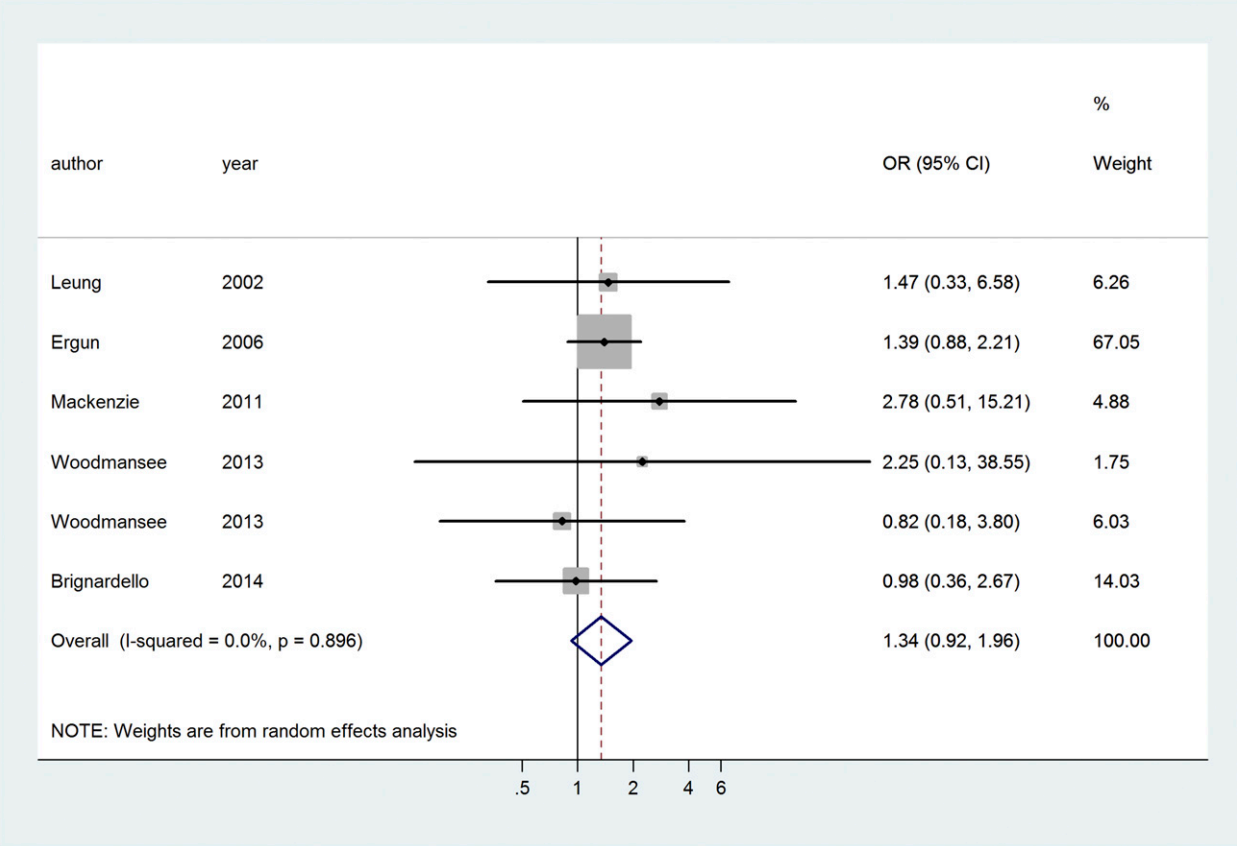


Figure 3. The occurrence of secondary tumors in CCSs treated with GHT was not significantly different from that in untreated survivors.

tumors were meningiomas (24, 25). The risk of secondary neoplasms in CCSs treated with GHT was lower after an extended follow-up (26) and became nonsignificant after adjusting for sex, age at primary diagnosis, CIR dose/time, and treatment type in a more recent report from the same cohort (27). Factors associated with meningioma development included female sex, young age at primary cancer diagnosis, and long periods since CIR (gliomas were associated with short periods since CIR). A report from the Pediatric Endocrine Society Drug and Therapeutic Committee suggested that GHT may increase the risk of subsequent neoplasms (28).

In the aforementioned Pediatric Endocrine Society report, there was also no substantial difference in the risk of recurrence in CCSs treated with GHT vs that in the controls. The outcomes of many individual studies (17, 24, 26, 29–31) have shown that the risk of recurrence in CCSs treated with GHT is not significantly different than that in CCSs not treated with GHT. Similar outcomes have been seen in medulloblastoma (32), craniopharyngioma (33), and acute lymphoblastic leukemia survivors (34). Two studies (26, 35) showed that the risk of recurrence after GHT was decreased compared with that in untreated patients. In these studies, however, the potential for bias in selecting patients with better prognosis for GHT and use of the last clinical contact as the date of exit from the study must be considered. The dose of GH and treatment modalities did not differ significantly between patients with and without recurrence (31), and there were no apparent increased risk of recurrence with the cumulative time that GHT was administered or the time elapsed since the treatment started (35). There have been no definitive studies on how long to wait after the completion of cancer therapy to start GHT; the Pediatric Endocrine Society guidelines suggest waiting for 12 months (9).

One study suggested a lower mortality for GH-treated patients compared with that for untreated patients, adjusting for potentially confounding prognostic variables (35). Mortality increased significantly with the time since the first GH treatment, and GH treatment had no effect on 3-year event-free survival rates. Increased mortality from cancer following GHT was reported in a recent report from the Safety and Appropriateness of Growth Hormone Treatments in Europe cohort; this increase was largely attributed to mortality from second cancers in CCSs (36). The study also reported that cancer mortality rates increased with increasing mean daily r-hGH doses, specifically in CCSs; these data, however, should be interpreted with caution because the cohort did not include CCSs who were not treated with GH and was not designed to specifically assess the risk of secondary neoplasia in CCSs.

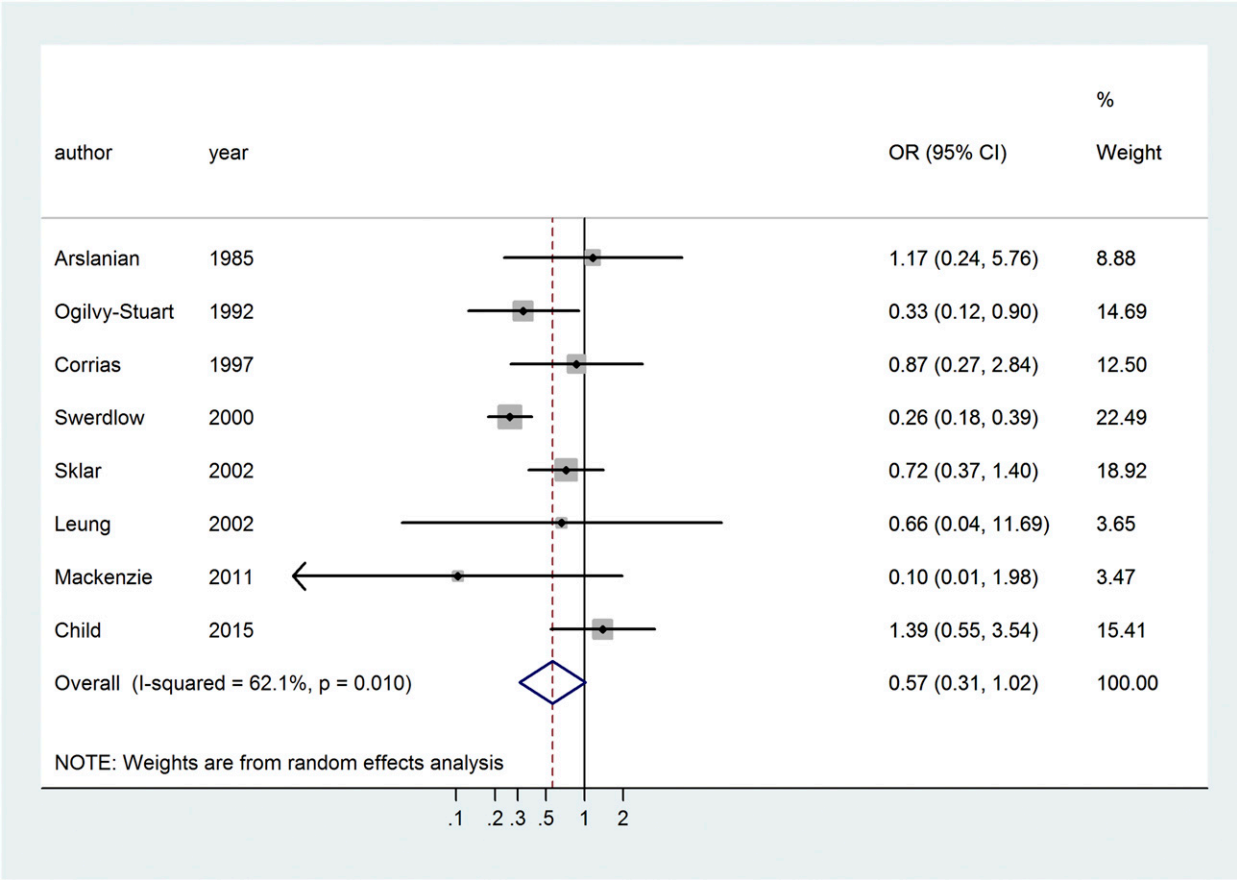


Figure 4. Disease/tumor recurrence in CCSs treated with GHT was not significantly different from that in untreated survivors.

Limitations and strengths

The strengths of this review relate to the comprehensive literature search, an *a priori*–established protocol and the duplicate process of study selection and appraisal. However, the available studies were observational with inherent limitations. Most studies assessing GHT have a relatively short follow-up duration (17, 30, 37–40), which can affect the outcome assessment dealing with the risk of recurrence and the incidence of secondary neoplasms. In the study by Swerdlow (36), however, there was no indication of increased cancer risk with longer follow-up durations. Furthermore, the relationship between the dose and duration response of GHT and cancer risks has been clearly evaluated in only a few studies (36, 41), which did not note increased cancer risks. The meta-analysis for outcomes related to height and recurrence risks has moderate to high heterogeneity.

Conclusion

GHT is effective in increasing height in CCSs with GHD. From the limited data available, no increased cardiovascular or metabolic risks were obviously associated with GHT, but short follow-up time must be considered a limitation. Although this meta-analysis did not show increased risks of recurrence or secondary neoplasms, additional studies on these risks are necessary.

Acknowledgments

Financial Support: The Endocrine Society (to M.H.M.).

Correspondence and Reprint Requests: M. Hassan Murad, MD, Evidence-Based Practice Center, Mayo Clinic, 200 First Street SW, Rochester, Minnesota 55905. E-mail: murad.mohammad@mayo.edu.

Disclosure Summary: W.C. has served as a consultant for Pfizer. L.E.C. is a site principal investigator for Versartis, Ascendis and received an honorarium from Scherer Clinical Communications (grant from Novo Nordisk).

References

1. Chemaitilly W, Boulad F, Heller G, Kernan NA, Small TN, O'Reilly RJ, Sklar CA. Final height in pediatric patients after hyperfractionated total body irradiation and stem cell transplantation. *Bone Marrow Transplant*. 2007;**40**(1):29–35.
2. Elbornsson M, Götherström G, Bosæus I, Bengtsson BA, Johannsson G, Svensson J. Fifteen years of GH replacement improves body composition and cardiovascular risk factors. *Eur J Endocrinol*. 2013;**168**(5):745–753.
3. Elbornsson M, Horvath A, Götherström G, Bengtsson BA, Johannsson G, Svensson J. Seven years of growth hormone (GH) replacement improves quality of life in hypopituitary patients with adult-onset GH deficiency. *Eur J Endocrinol*. 2017;**176**(2):99–109.

4. Cohen LE. Introduction: discovery of growth hormone and synthesis of recombinant human growth hormone. In: Cohen LE, ed. *Growth Hormone Deficiency*. Switzerland: Springer International Publishing; 2016:1–5.
5. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;**6**(7): e1000097.
6. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at: www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 10 September 2017.
7. Murad MH. Clinical practice guidelines: a primer on development and dissemination. *Mayo Clin Proc*. 2017;**92**(3):423–433.
8. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;**7**(3):177–188.
9. Grimberg A, DiVall SA, Polychronakos C, Allen DB, Cohen LE, Quintos JB, Rossi WC, Feudtner C, Murad MH; Drug and Therapeutics Committee and Ethics Committee of the Pediatric Endocrine Society. Guidelines for growth hormone and insulin-like growth factor-I treatment in children and adolescents: growth hormone deficiency, idiopathic short stature, and primary insulin-like growth factor-I deficiency. *Horm Res Paediatr*. 2016;**86**(6): 361–397.
10. Beckers D, Thomas M, Jamart J, Francois I, Maes M, Lebrethon MC, De Waele K, Tenoutasse S, De Schepper J. Adult final height after GH therapy for irradiation-induced GH deficiency in childhood survivors of brain tumors: the Belgian experience. *Eur J Endocrinol*. 2010;**162**(3):483–490.
11. Ranke MB, Price DA, Lindberg A, Wilton P, Darendeliler F, Reiter EO. Final height in children with medulloblastoma treated with growth hormone. *Horm Res*. 2005;**64**(1):28–34.
12. Gleeson HK, Stoeter R, Ogilvy-Stuart AL, Gattamaneni HR, Brennan BM, Shalet SM. Improvements in final height over 25 years in growth hormone (GH)-deficient childhood survivors of brain tumors receiving GH replacement. *J Clin Endocrinol Metab*. 2003;**88**(8):3682–3689.
13. Brownstein CM, Mertens AC, Mitby PA, Stovall M, Qin J, Heller G, Robison LL, Sklar CA. Factors that affect final height and change in height standard deviation scores in survivors of childhood cancer treated with growth hormone: a report from the childhood cancer survivor study. *J Clin Endocrinol Metab*. 2004; **89**(9):4422–4427.
14. Xu W, Janss A, Moshang T. Adult height and adult sitting height in childhood medulloblastoma survivors. *J Clin Endocrinol Metab*. 2003;**88**(10):4677–4681.
15. Herber SM, Dunsmore IR, Milner RD. Final stature in brain tumours other than craniopharyngioma: effect of growth hormone. *Horm Res*. 1985;**22**(1-2):63–67.
16. Lannering B, Albertsson-Wikland K. Improved growth response to GH treatment in irradiated children. *Acta Paediatr Scand*. 1989; **78**(4):562–567.
17. Corrias A, Picco P, Einaudi S, de Sanctis L, Besenzon L, Garrè ML, Brach del Prever A, de Sanctis C. Growth hormone treatment in irradiated children with brain tumors. *J Pediatr Endocrinol Metab*. 1997;**10**(1):41–49.
18. Clayton PE, Shalet SM. The evolution of spinal growth after irradiation. *Clin Oncol (R Coll Radiol)*. 1991;**3**(4):220–222.
19. Vassilopoulou-Sellin, Klein MJ, Moore BD III, Reid HL, Ater J, Zietz HA. Efficacy of growth hormone replacement therapy in children with organic growth hormone deficiency after cranial irradiation. *Horm Res*. 1995;**43**(5):188–193.
20. Ciaccio M, Gil S, Guercio G, Vaiani E, Alderete D, Palladino M, Warman DM, Rivarola MA, Belgorosky A. Effectiveness of rhGH treatment on adult height in GH-deficient childhood survivors of medulloblastoma. *Horm Res Paediatr*. 2010;**73**(4): 281–286.
21. Murray RD, Darzy KH, Gleeson HK, Shalet SM. GH-deficient survivors of childhood cancer: GH replacement during adult life. *J Clin Endocrinol Metab*. 2002;**87**(1):129–135.
22. Follin C, Thilén U, Osterberg K, Björk J, Erfurth EM. Cardiovascular risk, cardiac function, physical activity, and quality of life with and without long-term growth hormone therapy in adult survivors of childhood acute lymphoblastic leukemia. *J Clin Endocrinol Metab*. 2010;**95**(8):3726–3735.
23. Mukherjee A, Tolhurst-Cleaver S, Ryder WD, Smethurst L, Shalet SM. The characteristics of quality of life impairment in adult growth hormone (GH)-deficient survivors of cancer and their response to GH replacement therapy. *J Clin Endocrinol Metab*. 2005;**90**(3):1542–1549.
24. Mackenzie S, Craven T, Gattamaneni HR, Swindell R, Shalet SM, Brabant G. Long-term safety of growth hormone replacement after CNS irradiation. *J Clin Endocrinol Metab*. 2011;**96**(9): 2756–2761.
25. Ergun-Longmire B, Mertens AC, Mitby P, Qin J, Heller G, Shi W, Yasui Y, Robison LL, Sklar CA. Growth hormone treatment and risk of second neoplasms in the childhood cancer survivor. *J Clin Endocrinol Metab*. 2006;**91**(9):3494–3498.
26. Sklar CA, Mertens AC, Mitby P, Occhiogrosso G, Qin J, Heller G, Yasui Y, Robison LL. Risk of disease recurrence and second neoplasms in survivors of childhood cancer treated with growth hormone: a report from the Childhood Cancer Survivor Study. *J Clin Endocrinol Metab*. 2002;**87**(7):3136–3141.
27. Patterson BC, Chen Y, Sklar CA, Neglia J, Yasui Y, Mertens A, Armstrong GT, Meadows A, Stovall M, Robison LL, Meacham LR. Growth hormone exposure as a risk factor for the development of subsequent neoplasms of the central nervous system: a report from the childhood cancer survivor study. *J Clin Endocrinol Metab*. 2014;**99**(6):2030–2037.
28. Raman S, Grimberg A, Waguespack SG, Miller BS, Sklar CA, Meacham LR, Patterson BC. Risk of neoplasia in pediatric patients receiving growth hormone therapy–a report from the Pediatric Endocrine Society Drug and Therapeutics Committee. *J Clin Endocrinol Metab*. 2015;**100**(6):2192–2203.
29. Arslanian SA, Becker DJ, Lee PA, Drash AL, Foley TP Jr. Growth hormone therapy and tumor recurrence. Findings in children with brain neoplasms and hypopituitarism. *Am J Dis Child*. 1985; **139**(4):347–350.
30. Ogilvy-Stuart AL, Ryder WD, Gattamaneni HR, Clayton PE, Shalet SM. Growth hormone and tumour recurrence. *BMJ*. 1992; **304**(6842):1601–1605.
31. Darendeliler F, Karagiannis G, Wilton P, Ranke MB, Albertsson-Wikland K, Anthony Price D; On Behalf Of The Kigs International Board. Recurrence of brain tumours in patients treated with growth hormone: analysis of KIGS (Pfizer International Growth Database). *Acta Paediatr*. 2006;**95**(10): 1284–1290.
32. Packer RJ, Boyett JM, Janss AJ, Stavrou T, Kun L, Wisoff J, Russo C, Geyer R, Phillips P, Kieran M, Greenberg M, Goldman S, Hyder D, Heideman R, Jones-Wallace D, August GP, Smith SH, Moshang T. Growth hormone replacement therapy in children with medulloblastoma: use and effect on tumor control. *J Clin Oncol*. 2001;**19**(2):480–487.
33. Child CJ, Conroy D, Zimmermann AG, Woodmansee WW, Erfurth EM, Robison LL. Incidence of primary cancers and intracranial tumour recurrences in GH-treated and untreated adult hypopituitary patients: analyses from the Hypopituitary Control and Complications Study. *Eur J Endocrinol*. 2015;**172**(6): 779–790.
34. Leung W, Rose SR, Zhou Y, Hancock ML, Burstein S, Schriock EA, Lustig R, Danish RK, Evans WE, Hudson MM, Pui CH. Outcomes of growth hormone replacement therapy in survivors of childhood acute lymphoblastic leukemia. *J Clin Oncol*. 2002; **20**(13):2959–2964.

35. Swerdlow AJ, Reddingius RE, Higgins CD, Spoudeas HA, Phipps K, Qiao Z, Ryder WD, Brada M, Hayward RD, Brook CG, Hindmarsh PC, Shalet SM. Growth hormone treatment of children with brain tumors and risk of tumor recurrence. *J Clin Endocrinol Metab*. 2000;**85**(12):4444–4449.
36. Swerdlow AJ, Cooke R, Beckers D, Borgström B, Butler G, Carel JC, Cianfarani S, Clayton P, Coste J, Deodati A, Ecosse E, Gausche R, Giacomozzi C, Hokken-Koelega ACS, Khan AJ, Kiess W, Kuehni CE, Mullis PE, Pfaffle R, Sävendahl L, Sommer G, Thomas M, Tidblad A, Tollerfield S, Van Eycken L, Zandwijken GRJ. Cancer risks in patients treated with growth hormone in childhood: the SAGhE European Cohort Study. *J Clin Endocrinol Metab*. 2017;**102**(5):1661–1672.
37. Child CJ, Zimmermann AG, Woodmansee WW, Green DM, Li JJ, Jung H, Erfurth EM, Robison LL; HypoCCS International Advisory Board. Assessment of primary cancers in GH-treated adult hypopituitary patients: an analysis from the Hypopituitary Control and Complications Study. *Eur J Endocrinol*. 2011;**165**(2):217–223.
38. Bell J, Parker KL, Swinford RD, Hoffman AR, Maneatis T, Lippe B. Long-term safety of recombinant human growth hormone in children. *J Clin Endocrinol Metab*. 2010;**95**(1):167–177.
39. Wilton P, Mattsson AF, Darendeliler F. Growth hormone treatment in children is not associated with an increase in the incidence of cancer: experience from KIGS (Pfizer International Growth Database). *J Pediatr*. 2010;**157**(2):265–270.
40. Müller HL, Gebhardt U, Schröder S, Pohl F, Kortmann RD, Faldum A, Zwiener I, Warmuth-Metz M, Pietsch T, Calaminus G, Kolb R, Wiegand C, Sörensen N; study committee of KRANIOPHARYNGEOM 2000/2007. Analyses of treatment variables for patients with childhood craniopharyngioma–results of the multicenter prospective trial KRANIOPHARYNGEOM 2000 after three years of follow-up. *Horm Res Paediatr*. 2010;**73**(3):175–180.
41. Swerdlow AJ, Higgins CD, Adlard P, Preece MA. Risk of cancer in patients treated with human pituitary growth hormone in the UK, 1959-85: a cohort study. *Lancet*. 2002;**360**(9329):273–277.

Hypothalamic–Pituitary and Growth Disorders in Survivors of Childhood Cancer: An Endocrine Society* Clinical Practice Guideline

Charles A. Sklar,[1] Zoltan Antal,[1,2] Wassim Chemaitilly,[3] Laurie E. Cohen,[4] Cecilia Follin,[5] Lillian R. Meacham,[6] and M. Hassan Murad[7]

[1]Memorial Sloan-Kettering Cancer Center, New York, New York 10065; [2]Weill Cornell Medicine and New York Presbyterian Hospital, New York, New York 10065; [3]St. Jude Children's Research Hospital, Memphis, Tennessee 38105; [4]Boston Children's Hospital, Boston, Massachusetts 02115; [5]Skåne University Hospital, 221 85 Lund, Sweden; [6]Emory University School of Medicine, Atlanta, Georgia 30307; and [7]Mayo Clinic Evidence-Based Practice Center, Rochester, Minnesota 55905

*Cosponsoring Associations: European Society of Endocrinology and Pediatric Endocrine Society.

*Endorsing Association: The Pituitary Society.

Objective: To formulate clinical practice guidelines for the endocrine treatment of hypothalamic–pituitary and growth disorders in survivors of childhood cancer.

Participants: An Endocrine Society–appointed guideline writing committee of six medical experts and a methodologist.

Conclusions: Due to remarkable improvements in childhood cancer treatment and supportive care during the past several decades, 5-year survival rates for childhood cancer currently are >80%. However, by virtue of their disease and its treatments, childhood cancer survivors are at increased risk for a wide range of serious health conditions, including disorders of the endocrine system. Recent data indicate that 40% to 50% of survivors will develop an endocrine disorder during their lifetime. Risk factors for endocrine complications include both host (*e.g.*, age, sex) and treatment factors (*e.g.*, radiation). Radiation exposure to key endocrine organs (*e.g.*, hypothalamus, pituitary, thyroid, and gonads) places cancer survivors at the highest risk of developing an endocrine abnormality over time; these endocrinopathies can develop decades following cancer treatment, underscoring the importance of lifelong surveillance. The following guideline addresses the diagnosis and treatment of hypothalamic–pituitary and growth disorders commonly encountered in childhood cancer survivors. **(*J Clin Endocrinol Metab* 103: 2761–2784, 2018)**

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received 29 May 2018. Accepted 29 May 2018.
First Published Online 29 June 2018

Abbreviations: ACTH, adrenocorticotropic hormone; ACTHD, adrenocorticotropic hormone deficiency; AH, adult height; CNS, central nervous system; CPP, central precocious puberty; CRT, cranial radiation therapy; CSI, craniospinal irradiation; fT4, free T4; GHD, GH deficiency; GHT, GH treatment; GnRHa, gonadotropin-releasing hormone agonist; HP, hypothalamic–pituitary; ITT, insulin tolerance test; LH/FSHD, LH/FSH deficiency; RT, radiation treatment; SCFE, slipped capital femoral epiphysis; TBI, total body irradiation; T2DM, type 2 diabetes mellitus; TKI, tyrosine kinase inhibitor; TSH, thyroid-stimulating hormone; TSHD, thyroid-stimulating hormone deficiency.

List of Recommendations

Short stature/impaired linear growth in childhood cancer survivors

Diagnosis and monitoring of short stature/impaired linear growth in childhood cancer survivors

1.1 We recommend prospective follow-up of linear growth for childhood cancer survivors at high risk for short adult height, namely those exposed to cranial radiation therapy, craniospinal irradiation, or total body irradiation at a young age and those with a history of inadequate weight gain or prolonged steroid requirement. (1|⊕⊕⊕O)

1.2 We recommend measuring standing height and sitting height in childhood cancer survivors treated with radiation that included the spine (*i.e.*, total body irradiation, craniospinal irradiation, as well as radiation to the chest, abdomen, or pelvis). (1|⊕⊕OO)

Technical remark: Sitting height is measured directly using a sitting height stadiometer, and the lower segment can be determined by subtracting sitting height from standing height. Alternatively, the lower segment can be determined by measuring from the pubic symphysis to the floor, and the upper segment can be determined by subtracting leg length from height. The upper to lower segment ratio can then be calculated but differs depending on the method used and ethnicity. In situations where clinicians are unable to measure sitting height, measuring arm span and comparing it to standing height will provide an estimate of spinal foreshortening due to prior spinal radiation.

Treatment of short stature/impaired linear growth in childhood cancer survivors

1.3 We suggest against using growth hormone in cancer survivors who do not have growth hormone deficiency to treat for short stature and/or poor linear growth following spinal irradiation. (2|⊕OOO)

1.4 We suggest against treatment with growth hormone in children with short stature and/or impaired linear growth who are being treated with tyrosine kinase inhibitors. (2|⊕OOO)

Growth hormone deficiency in childhood cancer survivors

Diagnosis of growth hormone deficiency in childhood cancer survivors

2.1 We recommend lifelong periodic clinical assessment for growth hormone deficiency in survivors treated for tumors in the region of the hypothalamic–pituitary axis and in those exposed to hypothalamic–pituitary axis radiation treatment ≥18 Gy (*e.g.*, various brain tumors, nasopharyngeal carcinoma, acute lymphoblastic leukemia, lymphoma). (1|⊕⊕⊕O)

Technical remark: The consensus of the writing committee is to assess height in children every 6 to 12 months.

2.2 We recommend against relying solely on serum insulin-like growth factor-1 levels in childhood cancer survivors exposed to hypothalamic–pituitary axis radiotherapy to make the diagnosis of growth hormone deficiency. (1|⊕⊕OO)

2.3 We advise using the same provocative testing to diagnose growth hormone deficiency in childhood cancer survivors as are used for diagnosing growth hormone deficiency in the noncancer population. (Ungraded Good Practice Statement)

2.4 We recommend against the use of growth hormone–releasing hormone alone or in combination with arginine in childhood cancer survivors to diagnose growth hormone deficiency after hypothalamic–pituitary axis radiation. (1|⊕⊕OO)

2.5 We suggest against using spontaneous growth hormone secretion (*e.g.*, 12-hour overnight sampling) as a diagnostic test in determining GH deficiency in childhood cancer survivors. (2|⊕OOO)

2.6 We recommend that formal testing to establish a diagnosis of growth hormone deficiency is not required in childhood cancer survivors with three other confirmed anterior pituitary hormone deficits. (1|⊕⊕OO)

2.7 We recommend retesting adult cancer survivors exposed to hypothalamic–pituitary axis radiation treatment and with a diagnosis of isolated growth hormone deficiency in childhood. (1|⊕⊕OO)

Treatment of growth hormone deficiency in childhood cancer survivors

2.8 We recommend offering growth hormone treatment in childhood cancer survivors with confirmed growth hormone deficiency based on the safety and efficacy demonstrated in that population. (1|⊕⊕OO)

2.9 We suggest waiting until the patient has been 1 year disease-free, following completion of therapy for malignant disease, before initiating growth hormone treatment. (2|⊕OOO)

2.10 In childhood cancer survivors who have chronic stable disease and thus may not ever be "disease-free" (particularly survivors treated for optic pathway tumors), we advise discussing the appropriateness of growth hormone treatment and its timing with their oncologist. (Ungraded Good Practice Statement)

2.11 We advise treating growth hormone-deficient childhood cancer survivors with similar growth hormone treatment regimens as are appropriate for individuals with growth hormone deficiency from the noncancer population. (Ungraded Good Practice Statement)

Central precocious puberty in childhood cancer survivors

Diagnosis of central precocious puberty in childhood cancer survivors

3.1 We recommend periodically assessing childhood cancer survivors for evidence of central precocious puberty if they have a history of hydrocephalus, tumors developing in or near the hypothalamic region, and/or have been exposed to hypothalamic–pituitary radiation. (1|⊕⊕⊕O)

3.2 We recommend against using testicular volume as the primary or sole indicator of degree of sexual development in male childhood cancer survivors previously treated with gonadotoxic agents, such as alkylating agents or testicular radiotherapy. (1|⊕⊕⊕O)

3.3 We recommend measuring serum testosterone, preferably using liquid chromatography–tandem mass spectroscopy, and luteinizing hormone levels prior to 10:00 AM to complement the clinical assessment of male childhood cancer survivors who are suspected of or are at risk for developing central precocious puberty and were exposed to gonadotoxic treatments. (1|⊕⊕OO)

Technical remark: Clinicians need to interpret plasma LH levels in patients exposed to gonadotoxic treatments in the context of their medical history and physical examination. Elevated LH levels in such patients may be due to primary gonadal injury rather than to the onset of central puberty.

Treatment of central precocious puberty in childhood cancer survivors

3.4 We advise that the indications and the type of treatment regimens for central precocious puberty in childhood cancer survivors should be similar to those used for central precocious puberty in the noncancer population. (Ungraded Good Practice Statement)

Hypogonadotropic hypogonadism in childhood cancer survivors

Diagnosis of luteinizing hormone/follicle-stimulating hormone deficiency in childhood cancer survivors

4.1 We recommend screening for luteinizing hormone/follicle-stimulating hormone deficiency in childhood cancer survivors exposed to hypothalamic–pituitary axis radiation at doses ≥30 Gy and in those with a history of tumors or surgery affecting the hypothalamic–pituitary axis region. (1|⊕⊕⊕O)

4.2 We advise using the same strategies to diagnose luteinizing hormone/follicle-stimulating hormone deficiency in childhood cancer survivors as are used in the noncancer population. (Ungraded Good Practice Statement)

Treatment of luteinizing hormone/follicle-stimulating hormone deficiency in childhood cancer survivors

4.3 We advise following the same treatment approach to luteinizing hormone/follicle-stimulating hormone deficiency in childhood cancer survivors as is appropriate in the noncancer population. (Ungraded Good Practice Statement)

Central hypothyroidism–thyroid-stimulating hormone deficiency in childhood cancer survivors

Diagnosis of central hypothyroidism in childhood cancer survivors

5.1 We recommend lifelong annual screening for thyroid-stimulating hormone deficiency in childhood cancer survivors treated for tumors in the region of the hypothalamic–pituitary axis and those exposed to ≥30 Gy hypothalamic–pituitary radiation. (1|⊕⊕⊕O)

5.2 We advise using the same biochemical tests to screen for thyroid-stimulating hormone deficiency in childhood cancer survivors as are used in the noncancer population. (Ungraded Good Practice Statement)

5.3 We recommend against using serum triiodothyronine, thyroid-stimulating hormone surge analysis, or thyrotropin-releasing hormone stimulation to diagnose thyroid-stimulating hormone deficiency. (1|⊕⊕OO)

Treatment of thyroid-stimulating hormone deficiency in childhood cancer survivors

5.4 We advise using the same approach to treat thyroid-stimulating hormone deficiency in childhood cancer survivors as is used in the noncancer population. (Ungraded Good Practice Statement)

Adrenocorticotropic hormone deficiency in childhood cancer survivors

Diagnosing adrenocorticotropic hormone deficiency in childhood cancer survivors

6.1 We recommend lifelong annual screening for adrenocorticotropic hormone deficiency in childhood cancer survivors treated for tumors in the hypothalamic–pituitary region and in those exposed to ≥30 Gy hypothalamic–pituitary radiation. (1|⊕⊕⊕O)

6.2 We suggest screening for adrenocorticotropic hormone deficiency in childhood cancer survivors exposed to between ≥24 Gy and 30 Gy hypothalamic–pituitary radiation who are >10 years postradiation or develop clinical symptoms suggestive of adrenocorticotropic hormone deficiency. (2|⊕OOO)

6.3 We advise using the same screening and dynamic testing procedures to diagnose adrenocorticotropic hormone deficiency in childhood cancer survivors as are used in the noncancer population. (Ungraded Good Practice Statement)
Technical remark: Clinicians should consider the influence of oral estrogen on total cortisol levels, as it can increase cortisol-binding globulin raising total, but not free, cortisol levels.

Treating adrenocorticotropic hormone deficiency in childhood cancer survivors

6.4 We advise that clinicians use the same glucocorticoid regimens as replacement therapy in childhood cancer survivors with adrenocorticotropic hormone deficiency as are used in the noncancer population with adrenocorticotropic hormone deficiency. (Ungraded Good Practice Statement)

6.5 We recommend that clinicians instruct all patients with adrenocorticotropic hormone deficiency regarding stress dose and emergency glucocorticoid administration and instruct them to obtain an emergency card/bracelet/necklace regarding adrenal insufficiency and an emergency kit containing injectable high-dose glucocorticoid. (1|⊕⊕⊕O)

Commissioned Systematic Review

The Guideline Writing Committee commissioned two systematic reviews to support this guideline. The first systematic review aimed to evaluate the effect of GH treatment (GHT) in childhood cancer survivors on adult height (AH), risk of type 2 diabetes mellitus (T2DM), abnormal lipids, metabolic syndrome, quality of life, secondary tumors, and disease recurrence. Studies showed that GHT vs no GHT in this patient group was associated with a significant gain in AH and no significant association with the occurrence of secondary tumors or tumor recurrence. Studies that compared childhood cancer survivors receiving GHT to normal age- or sex-matched controls or controls with idiopathic GH deficiency (GHD) or short stature showed that GHT was associated with either improved or unchanged risk of T2DM, lipid profile, or metabolic syndrome. GHT was also associated with improvement in quality of life.

The second systematic review aimed to determine the best screening and diagnostic tests for GHD in childhood cancer survivors exposed to hypothalamic–pituitary radiation. The major challenge in this review was the lack of a "gold" standard to diagnose GHD. There was high variability in the confirmatory testing each study used. The insulin tolerance test (ITT) seems to be the most accepted reference test in the reviewed studies, either alone or in combination with arginine; although standardization of the testing dose and strategy among different practice groups is lacking. Studies included in this report spanned 4 decades; therefore, changes in clinical practice and assay methods can account for some of this variability. IGF-I and IGF-binding protein 3 had a suboptimal diagnostic accuracy, and their results were correlated. The patterns of diagnostic accuracy of all the tests evaluated suggested a similar pattern to what we see in patients who are not childhood cancer survivors.

Introduction

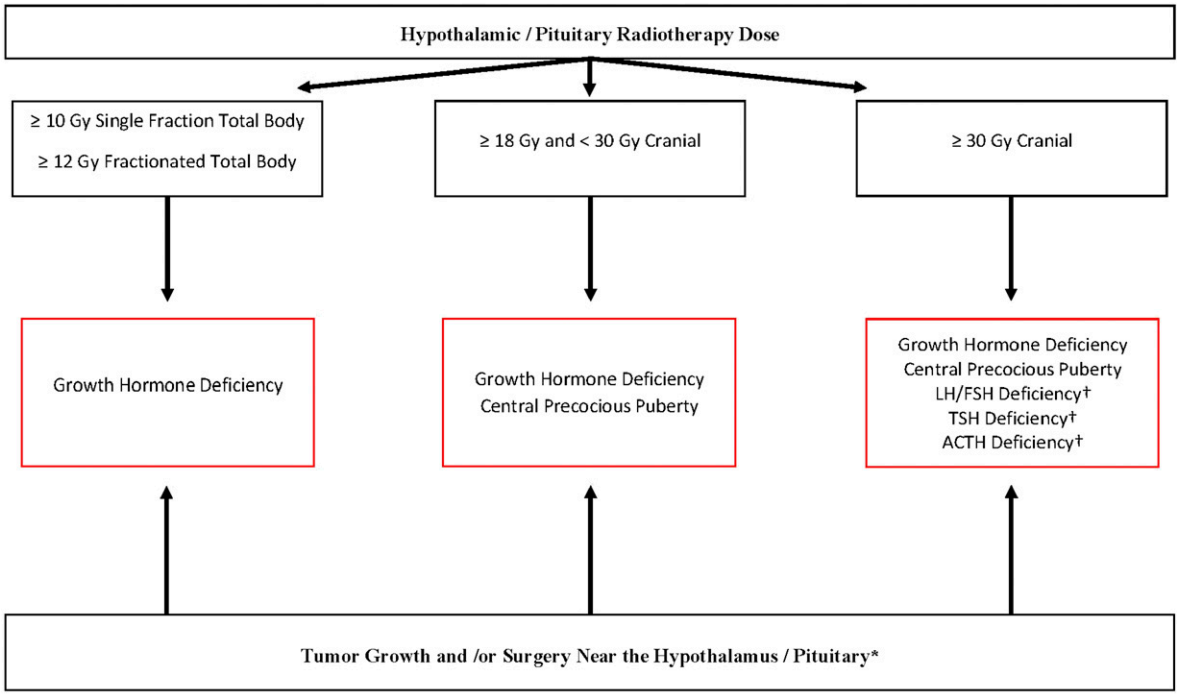
Cancers are relatively rare in the pediatric age group and account for only ~1% of the cancer burden in the entire population (1). Due to improvements in treatment and supportive care, current 5-year survival rates are >80% overall (2). The number of childhood cancers survivors is ever increasing and by the year 2020, it is estimated that there will be half a million survivors of childhood cancer residing in the United States. As the number of survivors has increased, there has been a growing awareness that survivors are at far greater risk of developing serious medical complications compared with noncancer controls (3). In particular, endocrine disorders are highly

prevalent among cancer survivors; recent data indicate that 40% to 50% of survivors will develop at least one endocrinopathy over the course of their lifetime (4, 5). Risk of developing endocrine disorders is dependent on a wide range of variables, including host (*e.g.*, age, sex, genetic background), disease (*e.g.*, diagnosis, tumor location), and treatment (*e.g.*, surgery, chemotherapy, radiation) factors. Radiation exposure to key endocrine organs (*e.g.*, hypothalamus, pituitary, thyroid and gonads) is the single most important risk factor and puts survivors at extremely high risk of developing an endocrine abnormality over time. Importantly, radiation-induced abnormalities are, in general, both dose- and time-dependent such that the higher the dose and the longer the interval following treatment, the greater the risk. Thus, endocrine disorders may not develop for decades after completion of cancer treatment, underscoring the critical importance of lifelong surveillance for those at risk (4, 5).

The current guideline focuses on the diagnosis and treatment of abnormalities of the hypothalamic–pituitary (HP) (Fig. 1) and management of growth disorders commonly encountered in childhood cancer survivors. Impaired linear growth and short AH can be due to both endocrine [*e.g.*, central precocious puberty (CPP), GHD] and nonendocrine (*e.g.*, medications, poor nutrition, radiation to the spine) factors. Not surprisingly, those diagnosed and treated for cancer at the youngest ages are the most affected. Currently, the only proven therapies for short stature are confined to the treatment of CPP and GHD, which follow the same general principles used in children without a cancer history.

Abnormalities of the HP are observed in survivors with tumors in the region of the HP, following surgery in the region of the HP, or, most commonly, following radiation to the HP (Fig. 1). Although HP dysfunction is generally observed acutely in individuals with tumors and/or surgery in the region of the HP, HP dysfunction is usually not observed for months to years following HP radiation. Whereas CPP and GHD occur following relatively low doses of HP radiation [*e.g.*, ≥18 Gy following standard fractionation, ≥12 Gy following hyperfractionation in the setting of total body irradiation (TBI)], deficits of the other anterior pituitary hormones [LH/FSH, thyroid-stimulating hormone (TSH), adrenocorticotropic hormone (ACTH)] develop almost exclusively in survivors previously exposed to HP doses ≥30 Gy. For the most part, the diagnosis and treatment of anterior pituitary deficits in cancer survivors follow the same general principles as are used in the noncancer population.

Although there have been a large number of excellent studies assessing disordered growth and HP abnormalities in this population that have informed this guideline, limitations exist. The focus of most studies has been on understanding prevalence and risk factors for the various outcomes, with fewer studies addressing diagnosis and management. Most of the data are descriptive and often limited to relatively small case series. Furthermore, much of the data relate to survivors treated in prior decades



†Deficiencies in LH/FSH, TSH and / or ACTH may appear after treatment with lower doses of radiotherapy with longer follow-up.
*Tumor and surgery –induced damage may acutely cause multiple hypothalamic-pituitary deficits in addition to central diabetes insipidus.

Figure 1. Common hypothalamic pituitary late effects in survivors of childhood cancer.

with therapies that may have been abandoned [*e.g.*, prophylactic cranial radiation therapy (CRT) for acute lymphoblastic leukemia] or modified in the current era [*e.g.*, reduced dose craniospinal irradiation (CSI) for medulloblastoma]. Data on the late effects of newer treatments such as targeted biologicals, immune modulators, and conformal radiation with protons are extremely limited. Additional areas requiring further research include: management of impaired growth following spinal radiation and in those receiving long-term therapy with tyrosine kinases such as imatinib; optimal frequency and duration of screening studies following HP radiation; and risks and benefits of GHT in adult survivors of childhood cancer.

1. Short Stature/Impaired Linear Growth in Childhood Cancer Survivors

Epidemiology, morbidity, and mortality

Impaired growth is defined by a loss in height SD over time and may be transient or progressive. Short stature is characterized by a standing AH of >2 SD below the mean for age and sex. Growth impairment and short stature in childhood cancer survivors may result from: alterations in HP hormone secretion due to tumors in the suprasellar/optic pathway region, surgery or CRT involving the HP axis, primary hypothyroidism (resulting from thyroidal radiation or high-dose alkylating agent chemotherapy) (6, 7), and radiation-induced impairment of spinal growth. The effects of both cranial and spinal radiation are dose- and time-dependent. Additional causes of growth impairment may include a malnourished state, growth suppressive effects of medications (*e.g.*, glucocorticoids, tyrosine kinase inhibitors [TKIs]), and medications associated with accelerated epiphyseal/physeal closure, such as retinoids.

The prevalence of adult short stature ranges from ~9% in studies of childhood acute leukemia survivors (8–10) to as high as 40% among survivors of childhood brain tumors (11).

Etiology and clinical manifestations

The major risk factors for impaired growth and short stature in cancer survivors are CRT, CSI, TBI, and younger age at the time of treatment (Table 1). Exposure to 18 to 30 Gy CRT may result in GHD and precocious puberty, whereas doses >30 Gy may result in multiple pituitary hormone deficiencies (12, 13). Exposure to CRT can also result in an earlier onset or altered tempo of puberty, including onset of breast development between ages 8 and 9 years, peak height velocity at age ≤10 years, and early menarche (14–18). Importantly, children who have both GHD and concomitant early or precocious puberty may not demonstrate a significant growth deceleration due to the stimulatory effects of sex hormones on linear growth, and the treating endocrinologist might miss a diagnosis of GHD unless he/she is knowledgeable in this regard.

Table 1. Established Risk Factors of Short Stature and HP Dysfunction in Childhood Cancer Survivors

Complication	Host Factors	High Risk Treatment Exposures
Short AH or impaired linear growth	• Younger age at cancer/tumor treatment • Precocious puberty • Genetic syndromes (such as neurofibromatosis type 1) • Fanconi anemia	• Radiation fields involving the HP region[a] • Spinal irradiation[b] • TBI • Glucocorticoids • TKIs • Retinoic acid
GHD	• Younger age at cancer/tumor treatment • Greater elapsed time since cancer/tumor treatment	• Tumor or surgery involving the HP region • HP radiation ≥18 Gy • TBI ≥10 Gy in single fraction • TBI ≥12 Gy fractionated
CPP	• Younger age at cancer/tumor treatment • Female sex • Increased body mass index • Neurofibromatosis type 1	• Tumor or surgery involving HP region or optic pathways • Hydrocephalus • HP radiotherapy ≥18 Gy
LH/FSHD	• Greater elapsed time since cancer/tumor treatment • Presence of other HP deficits	• Tumor or surgery involving the HP region • HP radiation ≥30 Gy
TSHD	• Greater elapsed time since cancer/tumor treatment • Presence of other HP deficits	• Tumor or surgery involving the HP region • HP radiation ≥30 Gy
ACTHD	• Greater elapsed time since cancer/tumor treatment • Presence of other HP deficits	• Tumor or surgery involving the HP region • HP radiation ≥30 Gy

[a]Cranial, infratemporal (ear), nasopharyngeal, orbital (eye), and Waldeyer's ring.

[b]Also includes fields involving the abdomen, chest, mediastinum, and pelvis.

Boney structures previously exposed to radiation may be at risk for poor growth; this effect is potentially greater with higher radiation doses and younger age at exposure. Exposure to spinal radiation can result in disproportionate short stature due to impaired spinal growth, which helps differentiate spinal radiation–related growth impairment from the symmetrical impairment caused by other etiologies, such as GHD.

Systemic therapy with retinoic acid and its derivatives is associated with premature epiphyseal closure in both animal models and human studies of noncancer populations (19). Studies of survivors of high-risk neuroblastoma reveal a significantly greater incidence of advanced bone ages in those treated with *cis*-retinoic acid (20, 21). This premature advancement and earlier closure of growth plates may explain, at least in part, the short AH seen in survivors treated with multimodality therapy that includes systemic *cis*-retinoic acid.

TKIs are targeted cancer therapies designed to disrupt specific signaling pathways involved in cellular growth and proliferation. Despite their intended specificity, nonselective, off-target effects on various protein kinases involved in chondrocyte accrual, as well as the GH/IGF-I signaling pathway, may result in growth deceleration and the potential for subsequent short AH (22).

Diagnosis and monitoring of short stature/impaired linear growth in childhood cancer survivors

1.1 We recommend prospective follow-up of linear growth for childhood cancer survivors at high risk for short AH, namely those exposed to CRT, CSI, or TBI at a young age and those with a history of inadequate weight gain or prolonged steroid requirement. (1|⊕⊕⊕O)

1.2 We recommend measuring standing height and sitting height in childhood cancer survivors treated with radiation that included the spine (*i.e.*, TBI, CSI, as well as radiation to the chest, abdomen, or pelvis). (1|⊕⊕OO)

Technical remark: Sitting height is measured directly using a sitting height stadiometer, and the lower segment can be determined by subtracting sitting height from standing height. Alternatively, the lower segment can be determined by measuring from the pubic symphysis to the floor, and the upper segment can be determined by subtracting leg length from height. The upper to lower segment ratio can then be calculated but differs depending on the method used and ethnicity. In situations where clinicians are unable to measure sitting height, measuring arm span and comparing it to standing height will provide an estimate of spinal foreshortening due to prior spinal radiation.

Evidence

The risk of growth impairment and adult short stature (height SD < -2 SD) in survivors of childhood leukemia is significantly higher among survivors treated before puberty, at younger ages, and at CRT doses >20 Gy (10, 23, 24). Among studies of survivors of leukemia, lymphoma, and a broad group of pediatric cancers (*e.g.*, osteosarcoma, Wilms' tumor, neuroblastoma, and soft tissue tumors of the head and neck), younger age at diagnosis and higher doses of CRT remained significant risk factors for adult short stature (9, 18). In a large study of 921 brain tumor survivors exposed to high-dose CRT, Gurney *et al.* (11) found that a significant number of adults diagnosed at younger ages had an AH <10th percentile, including 53% of those diagnosed before age 5 years, 46% of those diagnosed between 5 and 9 years, and 26% of those diagnosed between 10 and 20 years. Independent of age, those exposed to higher doses of CRT were more likely to have adult short stature than those not treated with CRT, with a threefold increased risk among those treated with >20 Gy and a fivefold increased risk among those treated with >60 Gy. These findings may be due to the development of multiple hormone aberrations, as detailed in subsequent sections of these guidelines.

Spinal radiation is an independent risk factor for short AH (10) and is associated with progressive growth impairment (25, 26). Survivors treated with higher doses of spinal radiation (>20 Gy) at younger ages, and to a larger volume of the spine, are at increased risk of short AH (11, 27). Compared with the proportionate short stature seen in GHD children resulting from CRT only, short AH associated with spinal irradiation results in disproportionate short stature, which is evident in the greater loss of spinal height SD relative to lower leg length SD (28–30). This disproportionate growth may be evident as early as 1 year following spinal radiation and becomes progressively more evident during puberty (26). Survivors treated with high-dose CSI (*e.g.*, >30 Gy for medulloblastoma) demonstrate the most significant losses in seated and standing AH (17, 25, 31).

Treatment of short stature/impaired linear growth in childhood cancer survivors

1.3 We suggest against using GH in cancer survivors who do not have GHD to treat for short stature and/or poor linear growth following spinal irradiation. (2|⊕OOO)

1.4 We suggest against treatment with GH in children with short stature and/or impaired linear growth who are being treated with TKIs. (2|⊕OOO)

Evidence

Studies on using GH to treat cancer survivors who do not have GHD are limited to a few small case series. In a study of 13 survivors of acute leukemia treated with cyclophosphamide and TBI, three of whom were not GH deficient, there was a progressive decline in height SD and impaired spine and leg growth despite GHT during a 3-year period (32). In a report of 51 high-risk neuroblastoma survivors treated with multimodal therapy (including TBI), Cohen *et al.* (33) described GHT in seven of these survivors. One had GHD, and six were initially thought to have GH neurosecretory dysfunction. Although short-term response to GH was good, the long-term response was not; of the two that achieved AH, even the patient with GHD remained >2 SD below midparental height.

Studies of GHD childhood cancer survivors treated with GH who were exposed to spinal irradiation also suggest a reduced benefit of GHT after spinal irradiation. Ciaccio *et al.* (34) found that among GH-deficient medulloblastoma survivors treated with 26 to 38 Gy CSI, the mean adult standing height decreased from −1.38 SD to −1.9 SD at AH in those treated with GH, whereas the standing height of those not treated with GH decreased from −1.55 SD to −3.4 SD. However, spinal heights in both groups were similar at AH, that is, −4.56 SD and −4.85 SD, respectively. In another study of 100 survivors with GHD treated with GH, those exposed to CSI had significantly lower growth responses to GHT (4.2 cm/y vs 6.7 cm/y) and significantly greater height loss from time of radiation to AH (−3.6 SD vs −1.6 SD) than those not exposed to spinal radiation (35). Any benefit to AH of GHT may be at the expense of further disproportionate growth (17, 31). Thus, full disclosure of this risk should be made and individual preferences considered when counseling survivors and their families about GHT in patients who have received spinal radiation treatment (RT).

Childhood cancer survivors exposed to TBI are also at risk for reduced spinal height, with the greatest risk among those treated at a younger age and with unfractionated TBI (29, 36, 37). Members of this patient subgroup who receive GHT for GHD may improve AH by preventing further height loss (38, 39); however, they may experience a worsening disproportion due to lack of spinal height gain (37, 40).

TKIs (such as imatinib and dasatinib) are mainstay treatments for chronic myelogenous leukemia and other malignancies that possess the BCR-ABL1 fusion protein; patients generally are treated with them for the long term to maintain molecular remission (41). Most studies report decreased growth in children who are using TKIs, with greater effects observed in prepubertal children and conflicting evidence of catch-up growth in pubertal children (42–44). Although the precise mechanisms underlying the growth deceleration associated with TKI therapy are unknown, reports of low serum IGF-I levels in children on TKI therapy suggest a possible disruption of the tyrosine kinases involved in the GH signaling cascade (44). Additional proposed mechanisms for growth failure include disrupted platelet-derived growth factor receptor-β, leading to altered recruitment and activity of chondrocytes.

We consider patients on continuous TKI therapy as having an active malignancy, as many will develop molecular evidence of persistent disease when TKI therapy is discontinued. Data on the safety and efficacy of GH use in these patients are quite limited (41); therefore, we cannot generally recommend GHT in this setting.

2. GHD in Childhood Cancer Survivors

GHD is characterized by inadequate GH secretion from the pituitary and is defined using different diagnostic tests. GHD can result from damage to the HP area due to tumors, surgery, and/or HP axis RT (8, 45–47). Additionally, researchers have described a few cases of GHD associated with the TKI imatinib (48, 49), and the immune modulator ipilimumab can cause hypophysitis (50).

GHD is the most common endocrine late effect in childhood cancer survivors treated with CRT (46, 51, 52). The prevalence of GHD varies depending on the type of tumor and treatment and is most frequent in survivors of suprasellar tumors and after high-dose HP axis RT (45–47). Adults with hypopituitarism on conventional hormone therapy that does not include GHT have an increased cardiovascular mortality in comparison with the general population (53, 54).

Clinical manifestations

GHD that develops in childhood usually affects linear growth (8, 55, 56). GHD in the cancer population has similar symptoms as we see in the noncancer population and may be associated (particularly in adults) with reduced lean body mass and increased fat mass, an adverse lipid profile, increased cardiovascular morbidity, impaired bone mineral density, impaired quality of life, and psychosocial problems (57–61). In a large cohort of 695 survivors of childhood cancer enrolled in the St. Jude Lifetime Cohort study, survivors with untreated GHD were more likely to have an increased weight-to-height ratio, decreased lean muscle mass, low energy expenditure, muscle weakness, and poor exercise tolerance compared with individuals without GHD (47).

Diagnosis of GHD in childhood cancer survivors

2.1 We recommend lifelong periodic clinical assessment for GHD in survivors treated for tumors in the region of the HP axis and in those exposed to HP axis RT ≥18 Gy (*e.g.*, various brain tumors, nasopharyngeal carcinoma, acute lymphoblastic leukemia, lymphoma). (1|⊕⊕⊕O)
Technical remark: The consensus of the writing committee is to assess height in children every 6 to 12 months.

2.2 We recommend against relying solely on serum IGF-I levels in childhood cancer survivors exposed to HP axis radiotherapy to make the diagnosis of GHD. (1|⊕⊕OO)

2.3 We advise using the same provocative testing to diagnose GHD in childhood cancer survivors as are used for diagnosing GHD in the noncancer population (Table 2). (Ungraded Good Practice Statement)

2.4 We recommend against the use of GHRH alone or in combination with arginine in childhood cancer survivors to diagnose GHD after HP axis radiation. (1|⊕⊕OO)

2.5 We suggest against using spontaneous GH secretion (*e.g.*, 12-hour overnight sampling) as a diagnostic test in determining GHD in childhood cancer survivors. (2|⊕OOO)

2.6 We recommend against formal testing to establish a diagnosis of GHD in childhood cancer survivors with three other confirmed anterior pituitary hormone deficits (Table 2). (1|⊕⊕OO)

2.7 We recommend retesting adult cancer survivors exposed to HP axis RT and with a diagnosis of isolated GHD in childhood (Table 2). (1|⊕⊕OO)

Evidence

HP axis radiation is a potent cause of GHD and the risk is directly related to the total dose delivered, the dose

Table 2. Related Guidelines Content

Recommendation Number	Guideline Title	Organization	Publication Year
Guidelines relevant to the diagnosis content of this guideline			
2.3, 2.6, 2.7	Guidelines for Growth Hormone and Insulin-Like Growth Factor-I Treatment in Children and Adolescents: Growth Hormone Deficiency, Idiopathic Short Stature, and Primary Insulin-Like Growth Factor-I Deficiency	Pediatric Endocrine Society	2017
4.2, 5.2, 6.3	Hormonal Replacement in Hypopituitarism in Adults: An Endocrine Society Clinical Practice Guideline	Endocrine Society	2016
Guidelines relevant to the treatment content of this guideline			
2.11	Evaluation and Treatment of Adult Growth Hormone Deficiency: An Endocrine Society Clinical Practice Guideline	Endocrine Society	2011
	Hormonal Replacement in Hypopituitarism in Adults: An Endocrine Society Clinical Practice Guideline	Endocrine Society	2016
	Guidelines for Growth Hormone and Insulin-Like Growth Factor-I Treatment in Children and Adolescents: Growth Hormone Deficiency, Idiopathic Short Stature, and Primary Insulin-Like Growth Factor-I Deficiency	Pediatric Endocrine Society	2017
4.3	Testosterone Therapy in Men with Hypogonadism: An Endocrine Society Clinical Practice Guideline	Endocrine Society	2018
	Treatment of Symptoms of the Menopause: An Endocrine Society Clinical Practice Guideline	Endocrine Society	2015
	Hormonal Replacement in Hypopituitarism in Adults: An Endocrine Society Clinical Practice Guideline	Endocrine Society	2016
5.4	Hormonal Replacement in Hypopituitarism in Adults: An Endocrine Society Clinical Practice Guideline	Endocrine Society	2016
6.4	Hormonal Replacement in Hypopituitarism in Adults: An Endocrine Society Clinical Practice Guideline	Endocrine Society	2016
	Diagnosis and Treatment of Primary Adrenal Insufficiency: An Endocrine Society Clinical Practice Guideline	Endocrine Society	2016

per fraction, and the time interval postirradiation (62–64). HP axis RT in children frequently causes abnormal HP function later in life (65). HP axis RT initially affects the hypothalamus, which is more sensitive to irradiation than is the anterior pituitary, based on responses to anterior pituitary–releasing hormone stimulation in patients with anterior pituitary hormone deficiencies (65, 66). GHD is usually the first established endocrine sequela of HP axis RT (51, 63). The prevalence varies depending on the population studied, follow-up time, type of stimulation test used, and peak GH cut-off levels (67).

GHD is a frequent consequence in childhood cancer survivors treated for tumors in the region of the HP (45, 68) and in brain tumor and nasopharyngeal carcinoma survivors exposed to HP axis radiotherapy ≥30 Gy (47, 62, 69). Furthermore, GHD is also reported in some acute lymphoblastic leukemia and lymphoma survivors exposed to 18 to 24 Gy to the HP axis (23, 70–72). In the even lower doses used for hematopoietic stem cell transplantation, GHD may occur after a single TBI dose of 10 Gy or fractionated doses of 12 Gy (73); however, with repeated assessments over time, there can be recovery (40).

In children, auxologic data collected every 6 to 12 months should be considered as the initial screen for GHD. Clinicians should further investigate for GHD when there is either linear growth deceleration or no linear growth acceleration during puberty and when they have ruled out other potential etiologies of growth failure (*e.g.*, undernutrition, spinal compromise, hypothyroidism, hypogonadism). Age-adjusted IGF-I levels measured in an accredited laboratory may be useful in screening for severe GHD but have limited utility when using a cut-off of −2 SD (52, 59, 74, 75). For example, Weinzimer *et al.* (74) found a sensitivity of 73% for IGF-I for the prediction of GHD in children with brain tumors. Additionally, clinicians need to interpret IGF-I levels in the context of sex steroid exposure (precocious puberty and hypogonadism). In adults, symptoms of GHD are nonspecific. GH testing should be considered in at-risk individuals with fatigue, increased abdominal fat mass, weight gain, low energy level, or hyperlipidemia. As for children, IGF-I levels may be useful in screening for severe GHD.

Clinicians who suspect GHD in childhood cancer survivors should perform provocative testing unless there are three other pituitary hormone deficiencies (76). In the general population, GHD is established via stimulation testing using the ITT, glucagon, arginine, levodopa, clonidine, or GHRH-arginine (if available) (77–79). Clinicians should not administer GHRH alone when the damage is primarily hypothalamic, as after radiation therapy, because it may give false-negative results (*i.e.*, normal GH responses despite true GHD). Likewise, the GHRH-arginine test can give a falsely normal GH response (71, 78, 80). Both ITT and glucagon testing allow evaluation of the complete hypothalamic–somatotroph axis (81). Based on a recent meta-analysis, the ability to diagnose GHD by different provocative tests after CRT is similar to the general population, with the ITT being most reliable; however, data are limited (82).

The same GH cut-off levels to stimuli are used in childhood cancer survivors as in the general population. Comparing studies to assess prevalence or incidence of GHD is problematic due to the use of different GH antibodies, GH standards, and GH assays (83), as well as the poor reproducibility of these tests (84). Data in noncancer pediatric populations suggest that clinicians often misdiagnose children as having GHD, especially when using peak GH values of 5 to 10 μg/L (83).

Older studies suggested that some children have GH neurosecretory dysfunction after cranial radiation, especially after low doses, where there is subnormal spontaneous GH secretion despite normal GH responses to stimulation testing (85). However, Darzy *et al.* (86) demonstrated normal physiologic GH secretion in adults who received cranial radiation in childhood, suggesting that this particular entity (radiation-induced GH neurosecretory dysfunction) either does not exist or is very rare. Additionally, due to the poor reproducibility of 12-hour overnight GH sampling and the overlap of responses in normal children and nonchildhood cancer survivors with GHD (76), as well as the impracticality of this test in clinical practice, we suggest against using spontaneous GH sampling, for example, 12-hour overnight sampling, as a diagnostic test in determining GHD in the childhood cancer survivors. In adults, the peak GH cut-off levels to diagnose GHD range from 3 to 5 μg/L for ITT and 3 μg/L for glucagon (52, 87, 88). Obesity, sex-steroid deficiency, and hypothyroidism can blunt GH secretion and yield a false-positive result (*e.g.*, falsely low GH levels) (83); for example, in an obese individual, the cut-off of 1 μg/L is used for glucagon (89, 90).

Treatment of GHD in childhood cancer survivors

2.8 We recommend offering GHT in childhood cancer survivors with confirmed GHD based on the safety and efficacy demonstrated in that population. (1|⊕⊕OO)

2.9 In childhood cancer survivors, we suggest waiting until the patient has been 1 year disease-free following completion of therapy for malignant disease, before initiating GHT. (2|⊕OOO)

2.10 In childhood cancer survivors who have chronic stable disease and thus may not ever be "disease-free" (particularly survivors treated for optic

pathway tumors), we advise discussing the appropriateness of GHT and its timing with their oncologist. (Ungraded Good Practice Statement)

2.11 We advise treating GH-deficient childhood cancer survivors with similar GHT regimens as are appropriate for individuals with GHD from the noncancer population (Table 2). (Ungraded Good Practice Statement)

Evidence

GHT is approved for both children and adults with confirmed GHD. GH dosing guidelines in the transition period after growth cessation are not well established. However, as GH secretion and IGF-I levels peak in puberty and decline overtime thereafter (91, 92), the effective GH dose needs to be higher in the transition period after growth cessation than in adulthood.

Childhood cancer survivors with GHD who receive GHT have a significant gain in height as compared with those who are not treated [see accompanying meta-analysis (82)]. However, those childhood cancer survivors with GHD who were also treated with CSI or TBI may have impaired spinal growth and not achieve target height. Higher spine radiation doses and radiation at a younger age are associated with impaired spinal growth (see section 1. "Short Stature/Impaired Linear Growth in Childhood Cancer Survivors"). GHT results in either an improvement or no difference in the risk of T2DM, dyslipidemia, and metabolic syndrome [see accompanying meta-analysis (82)]. The discrepant results between studies may be due to metabolic improvements being offset by the effect of GH on increasing insulin resistance. Likewise, quality of life after GHT is either improved or unchanged (58, 60, 93).

Concerns have been raised regarding the long-term safety of GHT in childhood cancer survivors, as GH and the target hormone, IGF-I, have *in vitro* proliferative effects, and IGF-I also has proangiogenic and antiapoptotic properties (94). Available data on the safety of GH in childhood cancer survivors are based on observational studies with selection bias and a lack of randomized placebo-controlled studies. Childhood cancer survivors have an increased risk of developing meningioma and glioma due to radiation exposure (95, 96); they also are at risk for GHD and will be potential candidates for GHT (67). Previous data on GH-treated childhood cancer survivors indicated that GHT might potentially induce a small increase in the relative risk of developing second neoplasms compared with survivors not receiving GHT (97, 98), with research indicating that meningiomas are the most common second neoplasm (97). However, the elevated risk decreased over time (99). Although the reason for this decrease is unknown, it has been speculated that GH-treated individuals may have been subjected to earlier and increased surveillance (100). Recent studies have shown no significant association between GHT and the development of a second neoplasm of the central nervous system (CNS) in childhood cancer survivors (100–102). In the systematic review and meta-analysis conducted for this guideline, there was no statistically significant difference in the occurrence of secondary tumors in survivors treated with GH compared with those not treated (OR, 1.34; 95% CI, 0.92 to 1.96). Similarly, studies show no significant change in the risk of tumor recurrence in survivors treated with GH, compared with those not treated (overall OR, 0.57; 95% CI, 0.31 to 1.02) [see accompanying meta-analysis (82)]. At a recent workshop, the Growth Hormone Research Society concluded that there are no indications of an increased risk of recurrence of primary cancers after GHT in children, and the association between GHT and risk of second tumors is insufficient to make recommendations against GHT (103).

However, few data are available to provide recommendations when to initiate GHT after cancer treatment (103). Traditionally, clinicians start GHT for survivors of malignant tumors at least 1 year after a childhood cancer survivor is disease-free; thus, the safety of GHT prior to that time is not clear. An exception is craniopharyngiomas (which are considered benign tumors); in these cases, GHT has been safely initiated earlier (as early as 0.7 year from diagnosis) (104). Additionally, there are patients who may have stable disease, rather than being disease-free. This is often the case in subjects with optic pathway gliomas (low-grade tumors frequently found in association with neurofibromatosis type 1). In these cases, disease can remain stable for prolonged periods, despite radiation and/or chemotherapy treatments (105). For these patients, clinicians should discuss whether to initiate GHT and its timing with the patient's oncologist.

It is important to note that GHT in children with GHD who had been treated with CSI and TBI may result in improvement in leg length but not spinal height (see section 1. "Short Stature/Impaired Linear Growth in Childhood Cancer Survivors"). Additionally, GHT may exacerbate an existing scoliosis, a condition not uncommon following either spinal surgery and/or spinal RT. GHT in the noncancer population is associated with an increased incidence of slipped capital femoral epiphysis (SCFE) (76). GHT in the noncancer population also commonly results in a decrease in insulin sensitivity and a compensatory increase in insulin secretion (76). As childhood cancer survivors are at increased risk for both SCFE (106) and metabolic

syndrome (107, 108) (particularly after TBI), the potential risk for SCFE and T2DM should be a factor when clinicians consider GHT in survivors. Therefore, when considering GHT, clinicians need to carefully weigh the potential risks against the potential benefits. Similar to the Pediatric Endocrine Society's recommendations for the noncancer population (76), we recommend that clinicians monitor serum IGF-I concentrations in patients on GHT and ensure they are kept within the normal range for sex, age, and pubertal status.

3. Central Precocious Puberty in Childhood Cancer Survivors

Epidemiology, morbidity, and mortality

The prevalence of CPP in childhood cancer survivors is estimated at 11.9% to 15.2% (109, 110). CPP in the context of a CNS insult may be associated with the accelerated fusion of the growth plates and potentially significant losses in AH (109, 111). Early sexual development may also result in challenges regarding the psychosocial adjustment of patients (112).

Etiology and clinical manifestations

CPP is defined by the onset of pubertal development before the age of 8 years in girls and 9 years in boys as a result of the premature activation of the HP gonadal axis (109, 113). Table 1 summarizes the risk factors of CPP in childhood cancer survivors.

Diagnosis of central precocious puberty in childhood cancer survivors

3.1 We recommend periodically assessing childhood cancer survivors for evidence of CPP if they have a history of hydrocephalus, tumors developing in or near the hypothalamic region, and/or have been exposed to HP radiation. (1|⊕⊕⊕O)

3.2 We recommend against using testicular volume as the primary or sole indicator of degree of sexual development in male childhood cancer survivors previously treated with gonadotoxic agents, such as alkylating agents or testicular radiotherapy. (1|⊕⊕⊕O)

3.3 We recommend measuring serum testosterone, preferably using liquid chromatography–tandem mass spectroscopy, and LH levels prior to 10:00 AM to complement the clinical assessment of male childhood cancer survivors who are suspected of or are at risk for developing CPP and were exposed to gonadotoxic treatments. (1|⊕⊕OO)

Technical remark: Clinicians need to interpret plasma LH levels in patients exposed to gonadotoxic treatments in the context of their medical history and physical examination. Elevated LH levels in such patients may be due to primary gonadal injury rather than to the onset of central puberty.

Evidence

CPP is among the most common endocrine complications in children with tumors arising near the hypothalamus or optic pathways (such as low-grade gliomas), and it is often associated with neurofibromatosis type 1 (109, 111, 113–115). The prevalence of CPP in patients with these tumors is between 26% and 39% (109, 111, 115). Exposure to HP axis radiation at a wide range of doses (20 to 50 Gy) has also been associated with CPP, albeit with a lower frequency (6.6%) (109). Additional risk factors for CPP include hydrocephalus (113), young age at HP axis RT, and (in patients exposed to HP axis RT) female sex and increased body mass index (110, 116).

The clinical diagnosis of pubertal onset in female childhood cancer survivors primarily relies on the observation of breast development (as in the noncancer population). The diagnosis of pubertal onset in male childhood cancer survivors previously exposed to gonadotoxic treatments (*e.g.*, alkylating agents, testicular radiation) requires a different approach than in males in the noncancer population, in whom increasing testicular volume is an early clinical indicator of the onset of puberty. Testicular volume is known to be affected by gonadotoxic cancer treatments because of germ cell and Sertoli cell injury (117–119). Research has indicated that testicular size plots below the 10th percentile for chronologic age in up to 50% of male survivors of acute lymphoblastic leukemia treated with CRT and extended abdominal radiotherapy (117), and it averages −2.0 SD in pediatric hematopoietic stem cell transplant recipients exposed to TBI (119). Studies have shown that the testes of adult childhood cancer survivors are significantly smaller than controls (118) and correlate with impaired germ cell function (120).

Given the limited reliability of testicular volume as a means of pubertal staging in male childhood cancer survivors treated with gonadotoxic modalities, laboratory markers such as AM serum testosterone and LH plasma levels may allow an earlier and more accurate assessment in this subset of patients whose Leydig cell function is less frequently affected than their germ cell function (117, 119). Liquid chromatography–tandem mass spectroscopy should preferably be used to measure serum testosterone levels (79). Medical providers should be aware that serum LH elevations could be due to primary gonadal injury rather than to the central onset of

puberty in patients exposed to gonadotoxic treatments, and clinicians should interpret laboratory data in the context of the patient's cancer history and clinical presentation (121). Measuring serum testosterone is especially helpful in boys exposed to gonadotoxic therapies and who are at risk for CPP given the challenges in the interpretation of clinical parameters (*e.g.*, testicular volume) and LH values. In girls with elevated gonadotropins, the assessment of pubertal progression based on Tanner staging (breast development), the measurement of estradiol levels, and the assessment of uterine length and shape via ultrasound can help distinguish between CPP and primary gonadal insufficiency (122).

As discussed in section 1 ("Short Stature/Impaired Linear Growth in Childhood Cancer Survivors"), the interpretation of growth velocity in childhood cancer survivors should be based not only on chronological age, but also on pubertal stage because of the frequent association between CPP and GHD (109, 111). Patients with genetic syndromes (such as neurofibromatosis type 1) and those exposed to craniospinal radiotherapy may also experience CPP (26, 123). GHD may compromise a patient's ability to experience linear growth acceleration during puberty in general and CPP in particular. Otherwise, the diagnostic work-up of childhood cancer survivors suspected of CPP follows the general steps followed in the general pediatric population (Table 2) (121, 124–126).

Treatment of central precocious puberty in childhood cancer survivors

3.4 We advise that the indications and the type of treatment regimens for CPP in childhood cancer survivors should be similar to those used for CPP in the noncancer population (Table 2). (Ungraded Good Practice Statement)

Evidence

Historical data on patients with tumor-related CPP who were not treated with pubertal suppression are scarce but suggest poor AH outcomes; although these patients may not have received treatment for other complications, including GHD (127). Studies have shown that pubertal suppression with gonadotropin-releasing hormone agonist (GnRHa) improves the AH of patients with CPP (not necessarily childhood cancer survivors) in comparison with their predicted AH at baseline (128–131). One study compared a cohort of 26 patients with CPP (31% related to a CNS insult) diagnosed at a young age (median 5 years) and treated with GnRHa to historical controls matched for demographic factors and etiology (131). The report showed a significant improvement in final or near AH in the treated group (-0.9 ± 0.3 SD in females, -1.7 ± 1.6 in males) in comparison with nontreated historical controls (-1.9 ± 0.2 in females, -3.2 ± 6.4 SD in males; $P = 0.01$ for both) (131). These data allow speculation that childhood cancer survivors with CPP most likely benefit from pubertal suppression with GnRHa (109, 113). Available AH data have nevertheless indicated that patients may not experience a complete recovery of their growth, and patients and families should be informed of the multifactorial nature of growth impairment in childhood cancer survivors (109, 129). Children with a history of hydrocephalus, HP tumors, and/or radiotherapy may experience nonprecocious but early onset puberty (8 to 9 years in girls or 9 to 10 years in boys) or rapid tempo of puberty (132, 133). Data are limited regarding the benefits of treatment with GnRHa on these forms of puberty in childhood cancer survivors, except in patients who also have GHD and in whom pubertal suppression, in association with GH therapy, seems to result in improved height outcomes (17, 132). There are no data supporting the use of GnRHa to augment the AH prospects of childhood cancer survivors experiencing normal pubertal development. The overall course of treatment of CPP in childhood cancer survivors can follow the advice in place for noncancer populations (121, 134).

4. Hypogonadotropic Hypogonadism in Childhood Cancer Survivors

Epidemiology, morbidity, and mortality

The estimated prevalence of LH/FSH deficiency (LH/FSHD) in childhood cancer survivors is 10.8% (47). Depending on the age of onset, LH/FSHD may manifest as delayed puberty (absence of signs of puberty after the ages of 13 years in girls and 14 years in boys) (135) or interrupted puberty, or LH/FSHD may manifest during adulthood as amenorrhea (females) or symptoms related to low testosterone (males). Untreated LH/FSHD in older childhood cancer survivors may be associated with adverse cardiovascular and bone health outcomes (47). Confounders related to the interpretation of low testosterone levels in obese men have complicated the understanding of the true impact of this problem (47, 136).

Etiology

LH/FSHD may occur within the context of panhypopituitarism following the direct anatomical injury of the HP area due to tumor growth or surgery. LH/FSHD may also occur as a late effect of HP axis radiation, especially at doses ≥30 Gy (47, 65, 137).

Diagnosis of LH/FSHD in childhood cancer survivors

4.1 We recommend screening for LH/FSHD in childhood cancer survivors exposed to HP axis radiation at doses ≥30 Gy and in those with a history of tumors or surgery affecting the HP axis region. (1|⊕⊕⊕O)

4.2 We advise using the same strategies to diagnose LH/FSHD in childhood cancer survivors as are used in the noncancer population (Table 2). (Ungraded Good Practice Statement)

Evidence

Childhood cancer survivors who are at risk for developing LH/FSHD because of their tumor or treatment history require periodic evaluation of their HP gonadal function (47, 65, 137–139). These assessments are important given the potential occurrence of LH/FSHD as a late effect and the rather nonspecific symptoms associated with this deficiency, especially in males (47, 79, 136).

The diagnosis of LH/FSHD in the general pediatric and adolescent population is complicated by the high prevalence of constitutional delay of growth in boys and by the presence of several congenital causes for LH/FSHD such as Kallmann syndrome, pituitary stalk interruption, and midline defects (140). Although such considerations may pertain to childhood cancer survivors experiencing pubertal delay, exposure to HP axis radiation and the presence of other pituitary deficiencies are generally robust indicators of LH/FSHD (135). Undetectable, low, or declining serum testosterone levels (males) or undetectable or low estradiol levels (females) in the setting of low or inappropriately normal levels of gonadotropins past 13 years of age in girls and 14 years of age in boys are suggestive of LH/FSHD during adolescence (135, 140). The diagnosis of LH/FSHD in adult childhood cancer survivors follows the same steps as those outlined by the Endocrine Society for the general population (79). Medical providers should cautiously interpret gonadotropin and sex hormone levels in obese or underweight individuals. It is important to note that childhood cancer survivors exposed to high-dose HP axis radiation, especially in the range ≥50 Gy, are at risk for developing hyperprolactinemia (65), which may also occur as a side effect of various drugs, especially antipsychotics. Hyperprolactinemia should be ruled out in patients with suspected LH/FSHD, as is the case in the general population (79).

Treatment of LH/FSHD in childhood cancer survivors

4.3 We advise following the same treatment approach to LH/FSHD in childhood cancer survivors as is appropriate in the noncancer population (Table 2). (Ungraded Good Practice Statement)

Evidence

Pubertal induction in adolescent female childhood cancer survivors with LH/FSHD can follow guidelines available for the general pediatric population (135, 140). The use of testosterone to induce puberty in boys during adolescence does not seem to adversely impact future fertility prospects in patients with LH/FSHD (141); however, data specific to childhood cancer survivors are limited (142).

Clinicians can use the same guidelines for the diagnosis and management of LH/FSHD in adult childhood cancer survivors as they do for adults with LH/FSHD (Table 2). Other measures potentially improving bone health, such as adequate dietary calcium intake and vitamin D supplementation, should be offered, along with sex hormone replacement (79). The benefits of estrogen replacement have been deemed to outweigh the risk of breast cancer in women 40 to 49 years in the general population (79). There are no data to support the need for a different approach in female childhood cancer survivors requiring estrogen/progesterone replacement for LH/FSHD. Preliminary studies do not support an increased risk of secondary breast cancer following spinal radiotherapy (143). In a recent study of hormone replacement therapy in women with premature ovarian insufficiency and a history of exposure to chest RT, the risk of breast cancer remained significantly lower than in childhood cancer survivors who retained normal ovarian function (144). Those with premature ovarian insufficiency who received hormone replacement therapy had a modestly increased risk of breast cancer, but not to the same degree as those with endogenous hormone production.

Medical providers should be aware of known drug interactions between antiepileptic medications and estrogen replacement, with potential repercussions on the efficacy of either treatment when the other is added or doses are changed (79). Antiepileptic drugs with enzyme-inducing properties (such as carbamazepine, oxcarbazepine, phenobarbital, phenytoin, and topiramate) may decrease the efficacy of sex hormones by interfering with their metabolism or by increasing the secretion of sex hormone–binding globulin. Other, nonenzyme-inducing, antiepileptic drugs such as lamotrigine, valproate, and levetiracetam have been shown to cause changes in plasma sex hormone concentrations but the mechanisms are unknown. Conversely, estrogen replacement may increase seizure risk in patients treated for epilepsy because of increased neuron excitability and/or interference with drug metabolism (as with lamotrigine) (145).

Estrogen replacement therapy increases the production of thyroid-binding globulin and may decrease the production of IGF-I; these interactions may require increasing the doses of levothyroxine and human recombinant GH in patients on treatment of hypothyroidism and GHD, respectively (79). The effect of estrogen replacement on circulating IGF-I may be avoided if transdermal formulations are used in lieu of oral forms (79). Clinicians should discuss the risks and benefits of estrogen replacement with patients and base treatment decisions on relevant guidelines and patient preferences.

5. Central Hypothyroidism–TSH Deficiency in Childhood Cancer Survivors

Epidemiology, Morbidity, and Mortality

The diagnosis of central hypothyroidism or TSH deficiency (TSHD) is complicated, and there is no uniform definition by which to make the diagnosis. Most endocrinologists diagnose TSHD in a patient recognized to be at risk for hypothalamic damage based on a low-normal or below-normal free T4 (fT4) level with a TSH level in the normal, below normal, or mildly elevated range (146). The diagnosis is more likely when based on progressively declining fT4 levels over time (147, 148). In the context of this broad definition, studies have reported that the prevalence of TSHD in childhood cancer survivors who have CNS tumors or were treated with HP axis RT is 2.6% to 14.9% (47, 111, 149–157). Symptoms are subtle, often delaying diagnosis.

Etiology

TSHD and ACTH deficiency (ACTHD) are among the least common anterior pituitary hormone deficits. TSHD is most often present after high-dose HP RT (47, 151); however, neither TSHD nor ACTHD is commonly present with doses <24 Gy or after TBI. A study of adult survivors of childhood cancer reported TSHD in 7.5% of participants with HP RT dose ≥30 Gy as an independent risk factor (47). Clinicians should evaluate patients with tumors in the HP region for TSHD if they have had suprasellar surgery or other hypothalamic deficiencies.

Factors associated with TSHD include hypothalamic involvement, radiation site and dose, and time elapsed since radiation exposure (111, 150, 152, 156). In general, chemotherapeutic agents have not been associated with TSHD (150, 155).

Diagnosis of central hypothyroidism in childhood cancer survivors

5.1 We recommend lifelong annual screening for TSHD in childhood cancer survivors treated for tumors in the region of the HP axis and those exposed to ≥30 Gy HP radiation. (1|⊕⊕⊕O)

5.2 We advise using the same biochemical tests to screen for TSHD in childhood cancer survivors as are used in the noncancer population (Table 2). (Ungraded Good Practice Statement)

5.3 We recommend against using serum triiodothyronine, TSH surge analysis, or thyrotropin-releasing hormone stimulation to diagnose TSHD. (1|⊕⊕OO)

Evidence

Clinicians should obtain fT4 and TSH levels at least annually (47, 147, 149, 151, 152). An fT4 level at the lower limits of normal or below the reference range in conjunction with a low, normal, or mildly elevated TSH level that does not appear appropriate for the fT4 level in the setting of disruption to the hypothalamus or pituitary is evidence of TSHD. The case for TSHD is made stronger with progressively decreasing fT4 levels (148). TSHD can develop many years after radiation exposure, and for this reason we recommend lifelong surveillance (158).

Previous data indicated that hidden central hypothyroidism was an early, subtle hypothalamic abnormality that clinicians could detect via the TSH surge pattern or TSH response to stimulation testing (159). Subsequent research indicated that the abnormalities of TSH dynamics uncovered by TSH surge analysis and thyrotropin-releasing hormone stimulation testing represent subtle variations that are not indicative or predictive of TSHD (160).

Treatment of TSHD in childhood cancer survivors

5.4 We advise using the same approach to treat TSHD in childhood cancer survivors as is used in the noncancer population (Table 2). (Ungraded Good Practice Statement)

Evidence

The treatment of central hypothyroidism in childhood cancer survivors is no different than in other children/adolescents with TSHD. The thyroid axis is one of the more resilient axes, and there is an increased risk of damage to other HP endocrine axes, which should be addressed. Clinicians should therefore perform regular, careful surveillance for GHD, full and partial ACTHD, and abnormalities of LH/FSH secretion in childhood cancer survivors who had RT, tumors, or surgery in the area around the hypothalamus and pituitary and who are diagnosed with TSHD (13, 147, 149). Clinicians should confirm the existence of an intact adrenal axis before beginning thyroid hormone replacement, recheck fT4 levels 4 to 6 weeks after dose adjustment or starting GH replacement, and maintain fT4 levels in the middle to upper half of the normal range. TSH levels are not useful

in monitoring the adequacy of thyroid hormone replacement in subjects with TSHD (79, 161). Many survivors at risk for TSHD are also at risk for seizures, and treatment with antiepileptic medications such as phenytoin, carbamazepine, oxcarbazepine, and topiramate can accelerate the metabolism of thyroid hormones. Consequently, clinicians should monitor thyroid hormone levels after starting or changing the dose of antiepileptics (145, 162). Commonly used fT4 assays, which use competitive binding methods, may give artifactually low fT4 levels in patients who are treated with antiepileptics (*e.g.*, phenytoin, carbamazepine, oxcarbazepine) due to displacement of thyroid hormone from binding proteins (163). Confirmation of the low fT4 level by a direct method, such as equilibrium dialysis or ultrafiltration, may be indicated in these patients.

6. ACTHD in Childhood Cancer Survivors

Epidemiology, morbidity, and mortality

ACTHD is characterized by inadequate cortisol secretion due to impaired production/secretion of ACTH. It can result from damage to the hypothalamus and/or pituitary gland due to tumors and/or surgery in the HP region (*e.g.*, craniopharyngiomas, suprasellar germinomas, optic pathway gliomas) (45, 164, 165) or to HP injury following high-dose (>30 Gy) HP radiation (47).

The prevalence of ACTHD (excluding exogenous steroid use) varies by tumor type and treatment (164–166). ACTHD has been associated with increased morbidity and mortality in pediatric survivors (167, 168).

Etiology and clinical manifestations

We list the major risk factors for ACTHD in Table 1. Although transient ACTHD secondary to exogenous glucocorticoids is very common in this population, particularly during active cancer treatment, this guideline focuses on permanent forms of ACTHD. The clinical symptoms most commonly associated with ACTHD in cancer survivors are similar to those described in the noncancer population (79). Given the nonspecific nature of these symptoms, it may be very difficult to distinguish between symptoms related to the underlying cancer, comorbidities from the disease and its treatment, or the presence of ACTHD. Partial ACTHD may be asymptomatic; thus, clinicians might not diagnose it unless they have a high degree of suspicion.

Diagnosing ACTHD in childhood cancer survivors

6.1 We recommend lifelong annual screening for ACTHD in childhood cancer survivors treated for tumors in the HP region and in those exposed to ≥30 Gy HP radiation. (1|⊕⊕⊕O)

6.2 We suggest screening for ACTHD in childhood cancer survivors exposed to between ≥24 Gy and 30 Gy HP radiation who are >10 years post-radiation or develop clinical symptoms suggestive of ACTHD. (2|⊕OOO)

6.3 We advise using the same screening and dynamic testing procedures to diagnose ACTHD in childhood cancer survivors as are used in the noncancer population (Table 2). (Ungraded Good Practice Statement)

Technical remark: Clinicians should consider the influence of oral estrogen on total cortisol levels, as it can increase cortisol-binding globulin raising total, but not free, cortisol levels.

Evidence

Radiation-induced ACTHD is known to be both time- and dose-dependent (169). Following HP radiation, ACTHD appears to occur less commonly than GH and LH/FSH deficiencies and is present primarily in childhood cancer survivors treated with doses of HP radiation >30 Gy (47, 151, 166, 170, 171), although the precise prevalence varies depending on the population studied, length of follow-up, and the type of biochemical testing used (4% to 43%). The two largest studies to assess ACTHD risk following HP radiation reported most cases of ACTHD in survivors exposed to >30 to 40 Gy HP radiation (47, 166). Although ACTHD is uncommon in subjects treated with HP doses ≤24 Gy (47, 166, 170, 172), a recent study of adult survivors of acute leukemia followed for >10 years reported biochemical evidence of ACTHD in more than a third of survivors exposed to a mean HP dose of 24 Gy (173). Data indicate that new cases of ACTHD emerge as late as 25 or more years after HP radiation (47).

A variety of tests are available for diagnosing ACTHD, including the ITT, standard- and low-dose ACTH stimulation test, glucagon stimulation test, and the overnight oral metyrapone test. Controversy exists as to which modality is the most reliable in establishing a diagnosis of ACTHD, irrespective of the underlying cause (90, 166, 169, 174, 175). Additionally, several factors can affect the determination of cortisol levels in plasma, including changes in cortisol-binding globulin. Of note, females taking oral contraceptives have elevated cortisol-binding globulin levels, which can make the interpretation of cortisol levels difficult (174). Although many view the ITT as the "gold standard" to diagnose ACTHD, most clinicians use the ACTH stimulation test due to its convenience and safety profile (174). The Endocrine Society's guideline on hormonal replacement in hypopituitarism in adults (79) includes

recommendations for testing for ACTHD. Two recently published systematic reviews and meta-analyses on ACTH stimulation tests for diagnosing adrenal insufficiency—one performed by the Endocrine Society in both adults and children (176) and one confined only to ACTHD in children (177)—concluded that the standard- and low-dose ACTH stimulation tests had similar accuracy for diagnosing ACTHD. Moreover, the study by Ospina *et al.* (176) concluded that both standard- and low-dose ACTH stimulation tests are adequate to rule in, but not rule out, ACTHD. Depending on the dilution method used when performing the low-dose ACTH stimulation test, there may be considerable variation in the actual dose delivered, raising the risk of inaccurate dosing and invalid results (178). Appreciation of the pretest probability of ACTHD and the limitations of the assays for cortisol (as well as the limitations of the various dynamic tests) are critical in establishing a diagnosis of ACTHD.

Treating ACTHD in childhood cancer survivors

6.4 We advise that clinicians use the same glucocorticoid regimens as replacement therapy in childhood cancer survivors with ACTHD as are used in the noncancer population with ACTHD (Table 2). (Ungraded Good Practice Statement)

6.5 We recommend that clinicians instruct all patients with ACTHD regarding stress dose and emergency glucocorticoid administration and instruct them to obtain an emergency card/bracelet/necklace regarding adrenal insufficiency and an emergency kit containing injectable high-dose glucocorticoid. (1|⊕⊕⊕O)

Evidence

The Endocrine Society's guidelines on primary adrenal insufficiency (179) and treating ACTHD in adults (79) include recommendations for physiologic daily replacement and for treating suspected adrenal crisis. Separate studies exist for treating ACTHD during childhood and adolescence (174). However, there are no specific studies addressing the treatment of ACTHD in childhood cancer survivors. Glucocorticoid deficiency has been shown to impair free water clearance, which can mask the symptoms of polyuria in subjects with central diabetes insipidus (180). Thus, when initiating glucocorticoid replacement therapy, clinicians should monitor for the development of diabetes insipidus in at-risk patients and the exacerbation of symptoms in those with pre-existing partial diabetes insipidus. Some antiepileptics enhance hepatic CYP450 isoenzyme activity (*e.g.*, phenytoin, carbamazepine, oxcarbazepine, and topiramate), which can affect the metabolism of glucocorticoids, especially dexamethasone. A recent guideline reviews the management of glucocorticoids in subjects taking enzyme-inducing antiepileptics (79).

Method of Development of Evidence-Based Clinical Practice Guidelines

GRADE approach

The guideline writing committee followed the approach recommended by the Grading of Recommendations, Assessment, Development, and Evaluation Group, an international group with expertise in the development and implementation of evidence-based guidelines (181). A detailed description of the grading scheme has been published elsewhere (182). The writing committee used the best available research evidence to develop the recommendations. The writing committee also used consistent language and graphical descriptions of both the strength of a recommendation and the quality of evidence. In terms of the strength of a recommendation, strong recommendations use the phrase "we recommend" and the number 1, and conditional recommendations use the phrase "we suggest" and the number 2. Cross-filled circles indicate the quality of the evidence, such that ⊕OOO denotes very low–quality evidence; ⊕⊕OO, low quality; ⊕⊕⊕O, moderate quality; and ⊕⊕⊕⊕, high quality. The writing committee has confidence that persons who receive care according to the strong recommendations will derive, on average, more benefit than harm. Conditional recommendations require more careful consideration of the person's circumstances, values, and preferences to determine the best course of action. Linked to each recommendation is a description of the evidence and the values that the writing committee considered in making the recommendation. In some instances, there are remarks in which the writing committee offers technical suggestions for testing conditions, dosing, and monitoring. These technical comments reflect the best available evidence applied to a typical person being treated. Often this evidence comes from the unsystematic observations of the writing committee and their preferences; therefore, one should consider these remarks as suggestions.

In this guideline, the writing committee made several statements to emphasize the importance of shared decision-making, general preventive care measures, and basic principles of treatment of hypothalamic–pituitary and growth disorders in childhood cancer survivors. They labeled these "Ungraded Good Practice Statement." Direct evidence for these statements was either unavailable or not

systematically appraised and considered out of the scope of this guideline. The intention of these statements is to draw attention to these principles.

Conflict of interest

The Endocrine Society maintains a rigorous conflict-of-interest review process for developing clinical practice guidelines. All writing committee members must declare any potential conflicts of interest by completing a conflict-of-interest form. The Clinical Guidelines Subcommittee reviews all conflicts of interest before the Society's Council approves the members to participate on the writing committee and periodically during the development of the guideline. All others participating in the guideline's development must also disclose any conflicts of interest in the matter under study, and most of these participants must be without any conflicts of interest. The Clinical Guidelines Subcommittee and the writing committee have reviewed all disclosures for this guideline and resolved or managed all identified conflicts of interest.

Conflicts of interest are defined as remuneration in any amount from commercial interests; grants; research support; consulting fees; salary; ownership interests [*e.g.*, stocks and stock options (excluding diversified mutual funds)]; honoraria and other payments for participation in speakers' bureaus, advisory boards, or boards of directors; and all other financial benefits. Completed forms are available through the Endocrine Society office.

Appendix. Conflict of Interest of Endocrine Disorders in Survivors of Childhood Cancer Guideline Writing Committee

Writing Committee Member	Employment	Uncompensated Memberships	Uncompensated Leadership	Personal Financial	Organizational Financial	Spousal/ Family Information
Charles A. Sklar, MD (Chair)	Director, Long-Term Follow-Up Program, Memorial Sloan-Kettering Cancer Center	Pediatric Endocrine Society	None Declared	• Sandoz (conference honorarium) • St. Jude Children's Research Hospital (research consultant, compensated)	None declared	None declared
Zoltan Antal, MD	Chief of Pediatric Endocrinology, Weill Cornell Medicine and New York Presbyterian Hospital; Assistant Attending, Memorial Sloan-Kettering Cancer Center	None declared	None declared	• Novo Nordisk (speaker's bureau, compensated 12/2016)	None declared	None declared
Wassim Chemaitilly, MD	Assistant Member and Director, Division of Endocrinology, St. Jude Children's Research Hospital	None declared	None declared	• Novo Nordisk (compensated consultant 2/2014 and 9/2015) • Pfizer (compensated consultant 12/2016)	None declared	None declared
Laurie E. Cohen, MD	Clinical Chief of Division of Endocrinology, Director of Neuroendocrinology Program, Boston Children's Hospital; Associate Professor of Pediatrics, Harvard Medical School	Pediatric Endocrine Society, Program Committee, Growth Hormone Guideline Task Force, Society for Pediatric Research, Children's Oncology Group, American Academy of Pediatrics, Women in Endocrinology, Growth Hormone Research Society	None declared	• Scherer Clinical Communications, speaker (Grant from Novo Nordisk)	• Versartis Pharmaceuticals (site PI) • Ascendis Pharmaceuticals (site PI)	None declared
Cecilia Follin, PhD	Skåne University Hospital, Lund, Sweden	European Society of Endocrinology	None declared	None declared	None declared	None declared
Lillian R. Meacham, MD	Professor of Pediatrics, Medical Director Cancer Survivor Program, Emory University School of Medicine	• Pediatric Endocrine Society • Children's Oncology Group	National Children's Cancer Society (Medical Advisory Board)	None declared	None declared	None declared
M. Hassan Murad, MD	Professor of Medicine, The Mayo Clinic	None declared	None declared	None declared	None declared	None declared

Acknowledgments

Financial Support: This guideline was supported by the Endocrine Society. No other entity provided financial or other support.

Correspondence: Charles A. Sklar, MD, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, Room H1111, New York, New York 10065. E-mail: sklarc@mskcc.org.

Disclosure Summary: See the Appendix.

Disclaimer: The Endocrine Society's clinical practice guidelines are developed to be of assistance to endocrinologists by providing guidance and recommendations for particular areas of practice. The guidelines should not be considered inclusive of all proper approaches or methods, or exclusive of others. The guidelines cannot guarantee any specific outcome, nor do they establish a standard of care. The guidelines are not intended to dictate the treatment of a particular patient. Treatment decisions must be made based on the independent judgement of healthcare providers and each patient's individual circumstances.

The Endocrine Society makes no warranty, express or implied, regarding the guidelines and specifically excludes any warranties of merchantability and fitness for a particular use or purpose. The Society shall not be liable for direct, indirect, special, incidental, or consequential damages related to the use of the information contained herein.

References

1. Robison LL, Hudson MM. Survivors of childhood and adolescent cancer: life-long risks and responsibilities. *Nat Rev Cancer*. 2014;**14**(1):61–70.
2. National Cancer Institute. SEER incidence data, 1973–2014. Available at: seer.cancer.gov/data/. Accessed 7 November 2017.
3. Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, Friedman DL, Marina N, Hobbie W, Kadan-Lottick NS, Schwartz CL, Leisenring W, Robison LL; Childhood Cancer Survivor Study. Chronic health conditions in adult survivors of childhood cancer. *N Engl J Med*. 2006;**355**(15):1572–1582.
4. de Fine Licht S, Winther JF, Gudmundsdottir T, Holmqvist AS, Bonnesen TG, Asdahl PH, Tryggvadottir L, Anderson H, Wesenberg F, Malila N, Holm K, Hasle H, Olsen JH; ALiCCS Study Group. Hospital contacts for endocrine disorders in Adult Life after Childhood Cancer in Scandinavia (ALiCCS): a population-based cohort study. *Lancet*. 2014;**383**(9933):1981–1989.
5. Mostoufi-Moab S, Seidel K, Leisenring WM, Armstrong GT, Oeffinger KC, Stovall M, Meacham LR, Green DM, Weathers R, Ginsberg JP, Robison LL, Sklar CA. Endocrine abnormalities in aging survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *J Clin Oncol*. 2016;**34**(27):3240–3247.
6. Boulad F, Bromley M, Black P, Heller G, Sarafoglou K, Gillio A, Papadopoulos E, Sklar C. Thyroid dysfunction following bone marrow transplantation using hyperfractionated radiation. *Bone Marrow Transplant*. 1995;**15**(1):71–76.
7. Oudin C, Auquier P, Bertrand Y, Chastagner P, Kanold J, Poirée M, Thouvenin S, Ducassou S, Plantaz D, Tabone MD, Dalle JH, Gandemer V, Lutz P, Sirvent A, Villes V, Barlogis V, Baruchel A, Leverger G, Berbis J, Michel G. Late thyroid complications in survivors of childhood acute leukemia. An L.E.A. study. *Haematologica*. 2016;**101**(6):747–756.
8. Gurney JG, Kadan-Lottick NS, Packer RJ, Neglia JP, Sklar CA, Punyko JA, Stovall M, Yasui Y, Nicholson HS, Wolden S, McNeil DE, Mertens AC, Robison LL; Childhood Cancer Survivor Study. Endocrine and cardiovascular late effects among adult survivors of childhood brain tumors: Childhood Cancer Survivor Study. *Cancer*. 2003;**97**(3):663–673.
9. Knijnenburg SL, Raemaekers S, van den Berg H, van Dijk IWEM, Lieverst JA, van der Pal HJ, Jaspers MWM, Caron HN, Kremer LC, van Santen HM. Final height in survivors of childhood cancer compared with height standard deviation scores at diagnosis. *Ann Oncol*. 2013;**24**(4):1119–1126.
10. Chow EJ, Friedman DL, Yasui Y, Whitton JA, Stovall M, Robison LL, Sklar CA. Decreased adult height in survivors of childhood acute lymphoblastic leukemia: a report from the Childhood Cancer Survivor Study. *J Pediatr*. 2007;**150**(4):370–375.e1.
11. Gurney JG, Ness KK, Stovall M, Wolden S, Punyko JA, Neglia JP, Mertens AC, Packer RJ, Robison LL, Sklar CA. Final height and body mass index among adult survivors of childhood brain cancer: Childhood Cancer Survivor Study. *J Clin Endocrinol Metab*. 2003;**88**(10):4731–4739.
12. Darzy KH, Shalet SM. Hypopituitarism following radiotherapy. *Pituitary*. 2009;**12**(1):40–50.
13. Rose SR, Horne VE, Howell J, Lawson SA, Rutter MM, Trotman GE, Corathers SD. Late endocrine effects of childhood cancer. *Nat Rev Endocrinol*. 2016;**12**(6):319–336.
14. Moëll C, Marky I, Hovi L, Kristinsson J, Rix M, Moe PJ, Garwicz S. Cerebral irradiation causes blunted pubertal growth in girls treated for acute leukemia. *Med Pediatr Oncol*. 1994;**22**(6):375–379.
15. Cicognani A, Cacciari E, Rosito P, Mancini AF, Carlă G, Mandini M, Paolucci G. Longitudinal growth and final height in long-term survivors of childhood leukaemia. *Eur J Pediatr*. 1994;**153**(10):726–730.
16. Didcock E, Davies HA, Didi M, Ogilvy Stuart AL, Wales JK, Shalet SM. Pubertal growth in young adult survivors of childhood leukemia. *J Clin Oncol*. 1995;**13**(10):2503–2507.
17. Gleeson HK, Stoeter R, Ogilvy-Stuart AL, Gattamaneni HR, Brennan BM, Shalet SM. Improvements in final height over 25 years in growth hormone (GH)-deficient childhood survivors of brain tumors receiving GH replacement. *J Clin Endocrinol Metab*. 2003;**88**(8):3682–3689.
18. Noorda EM, Somers R, van Leeuwen FE, Vulsma T, Behrendt H; Dutch Late Effects Study Group. Adult height and age at menarche in childhood cancer survivors. *Eur J Cancer*. 2001;**37**(5):605–612.
19. DiGiovanna JJ. Isotretinoin effects on bone. *J Am Acad Dermatol*. 2001;**45**(5):S176–S182.
20. Hobbie WL, Mostoufi SM, Carlson CA, Gruccio D, Ginsberg JP. Prevalence of advanced bone age in a cohort of patients who received cis-retinoic acid for high-risk neuroblastoma. *Pediatr Blood Cancer*. 2011;**56**(3):474–476.
21. Steineck A, MacKenzie JD, Twist CJ. Premature physeal closure following 13-*cis*-retinoic acid and prolonged fenretinide administration in neuroblastoma. *Pediatr Blood Cancer*. 2016;**63**(11):2050–2053.
22. Lodish MB, Stratakis CA. Endocrine side effects of broad-acting kinase inhibitors. *Endocr Relat Cancer*. 2010;**17**(3):R233–R244.
23. Viana MB, Vilela MIOP. Height deficit during and many years after treatment for acute lymphoblastic leukemia in children: a review. *Pediatr Blood Cancer*. 2008;**50**(2, Suppl):509–516.
24. Dalton VK, Rue M, Silverman LB, Gelber RD, Asselin BL, Barr RD, Clavell LA, Hurwitz CA, Moghrabi A, Samson Y, Schorin M, Tarbell NJ, Sallan SE, Cohen LE. Height and weight in children treated for acute lymphoblastic leukemia: relationship to CNS treatment. *J Clin Oncol*. 2003;**21**(15):2953–2960.
25. Shalet SM, Gibson B, Swindell R, Pearson D. Effect of spinal irradiation on growth. *Arch Dis Child*. 1987;**62**(5):461–464.
26. Clayton PE, Shalet SM. The evolution of spinal growth after irradiation. *Clin Oncol (R Coll Radiol)*. 1991;**3**(4):220–222.
27. Brownstein CM, Mertens AC, Mitby PA, Stovall M, Qin J, Heller G, Robison LL, Sklar CA. Factors that affect final height and change in height standard deviation scores in survivors of

childhood cancer treated with growth hormone: a report from the childhood cancer survivor study. *J Clin Endocrinol Metab*. 2004; **89**(9):4422–4427.
28. Wallace WHB, Shalet SM, Morris-Jones PH, Swindell R, Gattamaneni HR. Effect of abdominal irradiation on growth in boys treated for a Wilms' tumor. *Med Pediatr Oncol*. 1990;**18**(6): 441–446.
29. Thomas BC, Stanhope R, Plowman PN, Leiper AD. Growth following single fraction and fractionated total body irradiation for bone marrow transplantation. *Eur J Pediatr*. 1993;**152**(11): 888–892.
30. Hovi L, Saarinen-Pihkala UM, Vettenranta K, Lipsanen M, Tapanainen P. Growth in children with poor-risk neuroblastoma after regimens with or without total body irradiation in preparation for autologous bone marrow transplantation. *Bone Marrow Transplant*. 1999;**24**(10):1131–1136.
31. Ogilvy-Stuart AL, Shalet SM. Growth and puberty after growth hormone treatment after irradiation for brain tumours. *Arch Dis Child*. 1995;**73**(2):141–146.
32. Papadimitriou A, Urena M, Hamill G, Stanhope R, Leiper AD. Growth hormone treatment of growth failure secondary to total body irradiation and bone marrow transplantation. *Arch Dis Child*. 1991;**66**(6):689–692.
33. Cohen LE, Gordon JH, Popovsky EY, Gunawardene S, Duffey-Lind E, Lehmann LE, Diller LR. Late effects in children treated with intensive multimodal therapy for high-risk neuroblastoma: high incidence of endocrine and growth problems. *Bone Marrow Transplant*. 2014;**49**(4):502–508.
34. Ciaccio M, Gil S, Guercio G, Vaiani E, Alderete D, Palladino M, Warman DM, Rivarola MA, Belgorosky A. Effectiveness of rhGH treatment on adult height in GH-deficient childhood survivors of medulloblastoma. *Horm Res Paediatr*. 2010;**73**(4):281–286.
35. Sulmont V, Brauner R, Fontoura M, Rappaport R. Response to growth hormone treatment and final height after cranial or craniospinal irradiation. *Acta Paediatr Scand*. 1990;**79**(5): 542–549.
36. Chow EJ, Liu W, Srivastava K, Leisenring WM, Hayashi RJ, Sklar CA, Stovall M, Robison LL, Baker KS. Differential effects of radiotherapy on growth and endocrine function among acute leukemia survivors: a childhood cancer survivor study report. *Pediatr Blood Cancer*. 2013;**60**(1):110–115.
37. Chemaitilly W, Boulad F, Heller G, Kernan NA, Small TN, O'Reilly RJ, Sklar CA. Final height in pediatric patients after hyperfractionated total body irradiation and stem cell transplantation. *Bone Marrow Transplant*. 2007;**40**(1): 29–35.
38. Sanders JE, Guthrie KA, Hoffmeister PA, Woolfrey AE, Carpenter PA, Appelbaum FR. Final adult height of patients who received hematopoietic cell transplantation in childhood. *Blood*. 2005; **105**(3):1348–1354.
39. Couto-Silva AC, Trivin C, Esperou H, Michon J, Baruchel A, Lemaire P, Brauner R. Final height and gonad function after total body irradiation during childhood. *Bone Marrow Transplant*. 2006;**38**(6):427–432.
40. Holm K, Nysom K, Rasmussen MH, Hertz H, Jacobsen N, Skakkebaek NE, Krabbe S, Müller J. Growth, growth hormone and final height after BMT. Possible recovery of irradiation-induced growth hormone insufficiency. *Bone Marrow Transplant*. 1996;**18**(1):163–170.
41. Samis J, Lee P, Zimmerman D, Arceci RJ, Suttorp M, Hijiya N. Recognizing endocrinopathies associated with tyrosine kinase inhibitor therapy in children with chronic myelogenous leukemia. *Pediatr Blood Cancer*. 2016;**63**(8):1332–1338.
42. Shima H, Tokuyama M, Tanizawa A, Tono C, Hamamoto K, Muramatsu H, Watanabe A, Hotta N, Ito M, Kurosawa H, Kato K, Tsurusawa M, Horibe K, Shimada H. Distinct impact of imatinib on growth at prepubertal and pubertal ages of children with chronic myeloid leukemia. *J Pediatr*. 2011;**159**(4):676–681.
43. Rastogi MV, Stork L, Druker B, Blasdel C, Nguyen T, Boston BA. Imatinib mesylate causes growth deceleration in pediatric patients with chronic myelogenous leukemia. *Pediatr Blood Cancer*. 2012; **59**(5):840–845.
44. Narayanan KR, Bansal D, Walia R, Sachdeva N, Bhansali A, Varma N, Marwaha RK. Growth failure in children with chronic myeloid leukemia receiving imatinib is due to disruption of GH/IGF-1 axis. *Pediatr Blood Cancer*. 2013;**60**(7):1148–1153.
45. Karavitaki N, Brufani C, Warner JT, Adams CBT, Richards P, Ansorge O, Shine B, Turner HE, Wass JAH. Craniopharyngiomas in children and adults: systematic analysis of 121 cases with long-term follow-up. *Clin Endocrinol (Oxf)*. 2005;**62**(4):397–409.
46. Schmiegelow M, Lassen S, Poulsen HS, Feldt-Rasmussen U, Schmiegelow K, Hertz H, Müller J. Cranial radiotherapy of childhood brain tumours: growth hormone deficiency and its relation to the biological effective dose of irradiation in a large population based study. *Clin Endocrinol (Oxf)*. 2000;**53**(2): 191–197.
47. Chemaitilly W, Li Z, Huang S, Ness KK, Clark KL, Green DM, Barnes N, Armstrong GT, Krasin MJ, Srivastava DK, Pui C-H, Merchant TE, Kun LE, Gajjar A, Hudson MM, Robison LL, Sklar CA. Anterior hypopituitarism in adult survivors of childhood cancers treated with cranial radiotherapy: a report from the St Jude Lifetime Cohort study. *J Clin Oncol*. 2015; **33**(5):492–500.
48. Kebapcilar L, Bilgir O, Alacacioglu I, Payzin B, Bilgir F, Oner P, Sari I, Calan M, Binicier O. Does imatinib mesylate therapy cause growth hormone deficiency? *Med Princ Pract*. 2009;**18**(5): 360–363.
49. Hobernicht SL, Schweiger B, Zeitler P, Wang M, Hunger SP. Acquired growth hormone deficiency in a girl with chronic myelogenous leukemia treated with tyrosine kinase inhibitor therapy. *Pediatr Blood Cancer*. 2011;**56**(4):671–673.
50. Joshi MN, Whitelaw BC, Palomar MTP, Wu Y, Carroll PV. Immune checkpoint inhibitor-related hypophysitis and endocrine dysfunction: clinical review. *Clin Endocrinol (Oxf)*. 2016;**85**(3): 331–339.
51. Shalet SM, Beardwell CG, Pearson D, Jones PHM. The effect of varying doses of cerebral irradiation on growth hormone production in childhood. *Clin Endocrinol (Oxf)*. 1976;**5**(3):287–290.
52. Darzy KH, Shalet SM. Radiation-induced growth hormone deficiency. *Horm Res*. 2003;**59**(Suppl 1):1–11.
53. Rosén T, Edén S, Larson G, Wilhelmsen L, Bengtsson BA. Cardiovascular risk factors in adult patients with growth hormone deficiency. *Acta Endocrinol (Copenh)*. 1993;**129**(3):195–200.
54. Bülow B, Hagmar L, Mikoczy Z, Nordström C-H, Erfurth EM. Increased cerebrovascular mortality in patients with hypopituitarism. *Clin Endocrinol (Oxf)*. 1997;**46**(1):75–81.
55. Kirk JA, Raghupathy P, Stevens MM, Cowell CT, Menser MA, Bergin M, Tink A, Vines RH, Silink M. Growth failure and growth-hormone deficiency after treatment for acute lymphoblastic leukaemia. *Lancet*. 1987;**329**(8526):190–193.
56. Clayton PE, Shalet SM, Morris-Jones PH, Price DA. Growth in children treated for acute lymphoblastic leukaemia. *Lancet*. 1988; **331**(8583):460–462.
57. Consensus guidelines for the diagnosis and treatment of adults with growth hormone deficiency: summary statement of the Growth Hormone Research Society Workshop on Adult Growth Hormone Deficiency. *J Clin Endocrinol Metab*. 1998;**83**(2): 379–381.
58. Mukherjee A, Tolhurst-Cleaver S, Ryder WDJ, Smethurst L, Shalet SM. The characteristics of quality of life impairment in adult growth hormone (GH)-deficient survivors of cancer and their response to GH replacement therapy. *J Clin Endocrinol Metab*. 2005;**90**(3):1542–1549.
59. Link K, Moëll C, Garwicz S, Cavallin-Ståhl E, Björk J, Thilén U, Ahrén B, Erfurth EM. Growth hormone deficiency predicts cardiovascular risk in young adults treated for acute lymphoblastic

leukemia in childhood. *J Clin Endocrinol Metab*. 2004;**89**(10): 5003–5012.
60. Follin C, Thilén U, Österberg K, Björk J, Erfurth EM. Cardiovascular risk, cardiac function, physical activity, and quality of life with and without long-term growth hormone therapy in adult survivors of childhood acute lymphoblastic leukemia. *J Clin Endocrinol Metab*. 2010;**95**(8):3726–3735.
61. van den Heijkant S, Hoorweg-Nijman G, Huisman J, Drent M, van der Pal H, Kaspers G-J, Delemarre-van de Waal H. Effects of growth hormone therapy on bone mass, metabolic balance, and well-being in young adult survivors of childhood acute lymphoblastic leukemia. *J Pediatr Hematol Oncol*. 2011;**33**(6): e231–e238.
62. Merchant TE, Rose SR, Bosley C, Wu S, Xiong X, Lustig RH. Growth hormone secretion after conformal radiation therapy in pediatric patients with localized brain tumors. *J Clin Oncol*. 2011; **29**(36):4776–4780.
63. Littley MD, Shalet SM, Beardwell CG, Ahmed SR, Applegate G, Sutton ML. Hypopituitarism following external radiotherapy for pituitary tumours in adults. *Q J Med*. 1989;**70**(262): 145–160.
64. Lam KSL, Tse VKC, Wang C, Yeung RTT, Ma JTC, Ho JHC. Early effects of cranial irradiation on hypothalamic-pituitary function. *J Clin Endocrinol Metab*. 1987;**64**(3):418–424.
65. Constine LS, Woolf PD, Cann D, Mick G, McCormick K, Raubertas RF, Rubin P. Hypothalamic-pituitary dysfunction after radiation for brain tumors. *N Engl J Med*. 1993;**328**(2):87–94.
66. Lannering B, Albertsson-Wikland K. Growth hormone release in children after cranial irradiation. *Horm Res*. 1987;**27**(1):13–22.
67. Mulder RL, Kremer LCM, van Santen HM, Ket JL, van Trotsenburg AS, Koning CC, Schouten-van Meeteren AY, Caron HN, Neggers SJ, van Dalen EC. Prevalence and risk factors of radiation-induced growth hormone deficiency in childhood cancer survivors: a systematic review. *Cancer Treat Rev*. 2009; **35**(7):616–632.
68. Daubenbüchel AM, Müller HL. Neuroendocrine disorders in pediatric craniopharyngioma patients. *J Clin Med*. 2015;**4**(3): 389–413.
69. Schmiegelow M, Lassen S, Poulsen HS, Feldt-Rasmussen U, Schmiegelow K, Hertz H, Müller J. Growth hormone response to a growth hormone-releasing hormone stimulation test in a population-based study following cranial irradiation of childhood brain tumors. *Horm Res*. 2000;**54**(2):53–59.
70. Brennan BM, Rahim A, Mackie EM, Eden OB, Shalet SM. Growth hormone status in adults treated for acute lymphoblastic leukaemia in childhood. *Clin Endocrinol (Oxf)*. 1998;**48**(6): 777–783.
71. Björk J, Link K, Erfurth EM. The utility of the growth hormone (GH) releasing hormone-arginine test for diagnosing GH deficiency in adults with childhood acute lymphoblastic leukemia treated with cranial irradiation. *J Clin Endocrinol Metab*. 2005; **90**(11):6048–6054.
72. Gurney JG, Ness KK, Sibley SD, O'Leary M, Dengel DR, Lee JM, Youngren NM, Glasser SP, Baker KS. Metabolic syndrome and growth hormone deficiency in adult survivors of childhood acute lymphoblastic leukemia. *Cancer*. 2006;**107**(6): 1303–1312.
73. Brennan BMD, Shalet SM. Endocrine late effects after bone marrow transplant. *Br J Haematol*. 2002;**118**(1):58–66.
74. Weinzimer SA, Homan SA, Ferry RJ, Moshang T. Serum IGF-I and IGFBP-3 concentrations do not accurately predict growth hormone deficiency in children with brain tumours. *Clin Endocrinol (Oxf)*. 1999;**51**(3):339–345.
75. Blijdorp K, van den Heuvel-Eibrink M, Pieters R, Boot A, Sluimer J, van der Lelij A-J, Neggers S. The limited screening value of insulin-like growth factor-I as a marker for alterations in body composition in very long-term adult survivors of childhood cancer. *Pediatr Blood Cancer*. 2012;**59**(4):711–716.
76. Grimberg A, DiVall SA, Polychronakos C, Allen DB, Cohen LE, Quintos JB, Rossi WC, Feudtner C, Murad MH; Drug and Therapeutics Committee and Ethics Committee of the Pediatric Endocrine Society. Guidelines for growth hormone and insulin-like growth factor-I treatment in children and adolescents: growth hormone deficiency, idiopathic short stature, and primary insulin-like growth factor-I deficiency. *Horm Res Paediatr*. 2016;**86**(6): 361–397.
77. Growth Hormone Research Society. Consensus guidelines for the diagnosis and treatment of growth hormone (GH) deficiency in childhood and adolescence: summary statement of the GH Research Society. *J Clin Endocrinol Metab*. 2000;**85**(11):3990–3993.
78. Darzy KH, Aimaretti G, Wieringa G, Gattamaneni HR, Ghigo E, Shalet SM. The usefulness of the combined growth hormone (GH)-releasing hormone and arginine stimulation test in the diagnosis of radiation-induced GH deficiency is dependent on the post-irradiation time interval. *J Clin Endocrinol Metab*. 2003; **88**(1):95–102.
79. Fleseriu M, Hashim IA, Karavitaki N, Melmed S, Murad MH, Salvatori R, Samuels MH. Hormonal replacement in hypopituitarism in adults: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*. 2016;**101**(11):3888–3921.
80. Lissett CA, Saleem S, Rahim A, Brennan BMD, Shalet SM. The impact of irradiation on growth hormone responsiveness to provocative agents is stimulus dependent: results in 161 individuals with radiation damage to the somatotropic axis. *J Clin Endocrinol Metab*. 2001;**86**(2):663–668.
81. Hoffman DM, O'Sullivan AJ, Baxter RC, Ho KK. Diagnosis of growth-hormone deficiency in adults. *Lancet*. 1994;**343**(8905): 1064–1068.
82. Tamhane S, Sfeir JG, Kittah NEN, Jasim S, Chemaitilly W, Cohen LE, Murad MH. GH therapy in childhood cancer survivors: a systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2018;**103**(8):2794–2801.
83. Polychronakos C. Laboratory diagnosis of growth hormone deficiency in children. In: Cohen LE, ed. *Growth Hormone Deficiency: Physiology and Clinical Management*. Cham, Switzerland: Springer International Publishing; 2016:95–107.
84. Carel JC, Tresca JP, Letrait M, Chaussain JL, Lebouc Y, Job JC, Coste J. Growth hormone testing for the diagnosis of growth hormone deficiency in childhood: a population register-based study. *J Clin Endocrinol Metab*. 1997;**82**(7):2117–2121.
85. Blatt J, Bercu BB, Gillin JC, Mendelson WB, Poplack DG. Reduced pulsatile growth hormone secretion in children after therapy for acute lymphoblastic leukemia. *J Pediatr*. 1984;**104**(2): 182–186.
86. Darzy KH, Pezzoli SS, Thorner MO, Shalet SM. Cranial irradiation and growth hormone neurosecretory dysfunction: a critical appraisal. *J Clin Endocrinol Metab*. 2007;**92**(5):1666–1672.
87. Corneli G, Di Somma C, Prodam F, Bellone J, Bellone S, Gasco V, Baldelli R, Rovere S, Schneider HJ, Gargantini L, Gastaldi R, Ghizzoni L, Valle D, Salerno M, Colao A, Bona G, Ghigo E, Maghnie M, Aimaretti G. Cut-off limits of the GH response to GHRH plus arginine test and IGF-I levels for the diagnosis of GH deficiency in late adolescents and young adults. *Eur J Endocrinol*. 2007;**157**(6):701–708.
88. Yuen KC. Laboratory diagnosis of growth hormone deficiency in adults. In: Cohen LE, ed. *Growth Hormone Deficiency: Physiology and Clinical Management*. Cham, Switzerland: Springer International Publishing; 2016:109–121.
89. Dichtel LE, Yuen KC, Bredella MA, Gerweck AV, Russell BM, Riccio AD, Gurel MH, Sluss PM, Biller BM, Miller KK. Overweight/obese adults with pituitary disorders require lower peak growth hormone cutoff values on glucagon stimulation testing to avoid overdiagnosis of growth hormone deficiency. *J Clin Endocrinol Metab*. 2014;**99**(12):4712–4719.
90. Hamrahian AH, Yuen KC, Gordon MB, Pulaski-Liebert KJ, Bena J, Biller BM. Revised GH and cortisol cut-points for the glucagon

stimulation test in the evaluation of GH and hypothalamic–pituitary–adrenal axes in adults: results from a prospective randomized multicenter study. *Pituitary*. 2016;**19**(3):332–341.
91. Bidlingmaier M, Friedrich N, Emeny RT, Spranger J, Wolthers OD, Roswall J, Körner A, Obermayer-Pietsch B, Hübener C, Dahlgren J, Frystyk J, Pfeiffer AF, Doering A, Bielohuby M, Wallaschofski H, Arafat AM. Reference intervals for insulin-like growth factor-1 (IGF-I) from birth to senescence: results from a multicenter study using a new automated chemiluminescence IGF-I immunoassay conforming to recent international recommendations. *J Clin Endocrinol Metab*. 2014;**99**(5):1712–1721.
92. Elmlinger MW, Kühnel W, Weber MM, Ranke MB. Reference ranges for two automated chemiluminescent assays for serum insulin-like growth factor I (IGF-I) and IGF-binding protein 3 (IGFBP-3). *Clin Chem Lab Med*. 2004;**42**(6):654–664.
93. Murray RD, Skillicorn CJ, Howell SJ, Lissett CA, Rahim A, Smethurst LE, Shalet SM. Influences on quality of life in GH deficient adults and their effect on response to treatment. *Clin Endocrinol (Oxf)*. 1999;**51**(5):565–573.
94. Chemaitilly W, Robison LL. Safety of growth hormone treatment in patients previously treated for cancer. *Endocrinol Metab Clin North Am*. 2012;**41**(4):785–792.
95. Neglia JP, Robison LL, Stovall M, Liu Y, Packer RJ, Hammond S, Yasui Y, Kasper CE, Mertens AC, Donaldson SS, Meadows AT, Inskip PD. New primary neoplasms of the central nervous system in survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *J Natl Cancer Inst*. 2006;**98**(21):1528–1537.
96. Strodtbeck K, Sloan A, Rogers L, Fisher PG, Stearns D, Campbell L, Barnholtz-Sloan J. Risk of subsequent cancer following a primary CNS tumor. *J Neurooncol*. 2013;**112**(2):285–295.
97. Sklar CA, Mertens AC, Mitby P, Occhiogrosso G, Qin J, Heller G, Yasui Y, Robison LL. Risk of disease recurrence and second neoplasms in survivors of childhood cancer treated with growth hormone: a report from the Childhood Cancer Survivor Study. *J Clin Endocrinol Metab*. 2002;**87**(7):3136–3141.
98. Woodmansee WW, Zimmermann AG, Child CJ, Rong Q, Erfurth EM, Beck-Peccoz P, Blum WF, Robison LL; GeNeSIS and HypoCCS International Advisory Boards. Incidence of second neoplasm in childhood cancer survivors treated with GH: an analysis of GeNeSIS and HypoCCS. *Eur J Endocrinol*. 2013;**168**(4):565–573.
99. Ergun-Longmire B, Mertens AC, Mitby P, Qin J, Heller G, Shi W, Yasui Y, Robison LL, Sklar CA. Growth hormone treatment and risk of second neoplasms in the childhood cancer survivor. *J Clin Endocrinol Metab*. 2006;**91**(9):3494–3498.
100. Patterson BC, Chen Y, Sklar CA, Neglia J, Yasui Y, Mertens A, Armstrong GT, Meadows A, Stovall M, Robison LL, Meacham LR. Growth hormone exposure as a risk factor for the development of subsequent neoplasms of the central nervous system: a report from the Childhood Cancer Survivor Study. *J Clin Endocrinol Metab*. 2014;**99**(6):2030–2037.
101. Mackenzie S, Craven T, Gattamaneni HR, Swindell R, Shalet SM, Brabant G. Long-term safety of growth hormone replacement after CNS irradiation. *J Clin Endocrinol Metab*. 2011;**96**(9):2756–2761.
102. Brignardello E, Felicetti F, Castiglione A, Fortunati N, Matarazzo P, Biasin E, Sacerdote C, Ricardi U, Fagioli F, Corrias A, Arvat E. GH replacement therapy and second neoplasms in adult survivors of childhood cancer: a retrospective study from a single institution. *J Endocrinol Invest*. 2015;**38**(2):171–176.
103. Allen DB, Backeljauw P, Bidlingmaier M, Biller BM, Boguszewski M, Burman P, Butler G, Chihara K, Christiansen J, Cianfarani S, Clayton P, Clemmons D, Cohen P, Darendeliler F, Deal C, Dunger D, Erfurth EM, Fuqua JS, Grimberg A, Haymond M, Higham C, Ho K, Hoffman AR, Hokken-Koelega A, Johannsson G, Juul A, Kopchick J, Lee P, Pollak M, Radovick S, Robison L, Rosenfeld R, Ross RJ, Savendahl L, Saenger P, Toft Sorensen H, Stochholm K, Strasburger C, Swerdlow A, Thorner M. GH safety workshop position paper: a critical appraisal of recombinant human GH therapy in children and adults. *Eur J Endocrinol*. 2016;**174**(2):1–9.
104. Müller HL, Gebhardt U, Schröder S, Pohl F, Kortmann RD, Faldum A, Zwiener I, Warmuth-Metz M, Pietsch T, Calaminus G, Kolb R, Wiegand C, Sörensen N; study committee of KRANIOPHARYNGEOM 2000/2007. Analyses of treatment variables for patients with childhood craniopharyngioma—results of the multicenter prospective trial KRANIOPHARYNGEOM 2000 after three years of follow-up. *Horm Res Paediatr*. 2010;**73**(3):175–180.
105. Bornhorst M, Frappaz D, Packer RJ. Pilocytic astrocytomas. *Handb Clin Neurol*. 2016;**134**:329–344.
106. Mostoufi-Moab S, Isaacoff EJ, Spiegel D, Gruccio D, Ginsberg JP, Hobbie W, Shults J, Leonard MB. Childhood cancer survivors exposed to total body irradiation are at significant risk for slipped capital femoral epiphysis during recombinant growth hormone therapy. *Pediatr Blood Cancer*. 2013;**60**(11):1766–1771.
107. Oudin C, Auquier P, Bertrand Y, Contet A, Kanold J, Sirvent N, Thouvenin S, Tabone MD, Lutz P, Ducassou S, Plantaz D, Dalle JH, Gandemer V, Beliard S, Berbis J, Vercasson C, Barlogis V, Baruchel A, Leverger G, Michel G. Metabolic syndrome in adults who received hematopoietic stem cell transplantation for acute childhood leukemia: an LEA study. *Bone Marrow Transplant*. 2015;**50**(11):1438–1444.
108. Faienza MF, Delvecchio M, Giordano P, Cavallo L, Grano M, Brunetti G, Ventura A. Metabolic syndrome in childhood leukemia survivors: a meta-analysis. *Endocrine*. 2015;**49**(2):353–360.
109. Chemaitilly W, Merchant TE, Li Z, Barnes N, Armstrong GT, Ness KK, Pui CH, Kun LE, Robison LL, Hudson MM, Sklar CA, Gajjar A. Central precocious puberty following the diagnosis and treatment of paediatric cancer and central nervous system tumours: presentation and long-term outcomes. *Clin Endocrinol (Oxf)*. 2016;**84**(3):361–371.
110. Armstrong GT, Whitton JA, Gajjar A, Kun LE, Chow EJ, Stovall M, Leisenring W, Robison LL, Sklar CA. Abnormal timing of menarche in survivors of central nervous system tumors: a report from the Childhood Cancer Survivor Study. *Cancer*. 2009;**115**(11):2562–2570.
111. Gan H-W, Phipps K, Aquilina K, Gaze MN, Hayward R, Spoudeas HA. Neuroendocrine morbidity after pediatric optic gliomas: a longitudinal analysis of 166 children over 30 years. *J Clin Endocrinol Metab*. 2015;**100**(10):3787–3799.
112. Mensah FK, Bayer JK, Wake M, Carlin JB, Allen NB, Patton GC. Early puberty and childhood social and behavioral adjustment. *J Adolesc Health*. 2013;**53**(1):118–124.
113. Trivin C, Couto-Silva AC, Sainte-Rose C, Chemaitilly W, Kalifa C, Doz F, Zerah M, Brauner R. Presentation and evolution of organic central precocious puberty according to the type of CNS lesion. *Clin Endocrinol (Oxf)*. 2006;**65**(2):239–245.
114. Rivarola MA, Belgorosky A, Mendilaharzu H, Vidal G. Precocious puberty in children with tumours of the suprasellar and pineal areas: organic central precocious puberty. *Acta Paediatr*. 2001;**90**(7):751–756.
115. Habiby R, Silverman B, Listernick R, Charrow J. Precocious puberty in children with neurofibromatosis type 1. *J Pediatr*. 1995;**126**(3):364–367.
116. Oberfield SE, Soranno D, Nirenberg A, Heller G, Allen JC, David R, Levine LS, Sklar CA. Age at onset of puberty following high-dose central nervous system radiation therapy. *Arch Pediatr Adolesc Med*. 1996;**150**(6):589–592.
117. Sklar CA, Robison LL, Nesbit ME, Sather HN, Meadows AT, Ortega JA, Kim TH, Hammond GD. Effects of radiation on testicular function in long-term survivors of childhood acute lymphoblastic leukemia: a report from the Children Cancer Study Group. *J Clin Oncol*. 1990;**8**(12):1981–1987.

118. Siimes MA, Rautonen J. Small testicles with impaired production of sperm in adult male survivors of childhood malignancies. *Cancer*. 1990;**65**(6):1303–1306.
119. Ishiguro H, Yasuda Y, Tomita Y, Shinagawa T, Shimizu T, Morimoto T, Hattori K, Matsumoto M, Inoue H, Yabe H, Yabe M, Shinohara O, Kato S. Gonadal shielding to irradiation is effective in protecting testicular growth and function in long-term survivors of bone marrow transplantation during childhood or adolescence. *Bone Marrow Transplant*. 2007;**39**(8):483–490.
120. Romerius P, Ståhl O, Moëll C, Relander T, Cavallin-Ståhl E, Wiebe T, Giwercman YL, Giwercman A. Hypogonadism risk in men treated for childhood cancer. *J Clin Endocrinol Metab*. 2009; **94**(11):4180–4186.
121. Carel JC, Eugster EA, Rogol A, Ghizzoni L, Palmert MR, Antoniazzi F, Berenbaum S, Bourguignon JP, Chrousos GP, Coste J, Deal S, de Vries L, Foster C, Heger S, Holland J, Jahnukainen K, Juul A, Kaplowitz P, Lahlou N, Lee MM, Lee P, Merke DP, Neely EK, Oostdijk W, Phillip M, Rosenfield RL, Shulman D, Styne D, Tauber M, Wit JM; ESPE-LWPES GnRH Analogs Consensus Conference Group. Consensus statement on the use of gonadotropin-releasing hormone analogs in children. *Pediatrics*. 2009;**123**(4): e752–e762.
122. de Vries L, Phillip M. Role of pelvic ultrasound in girls with precocious puberty. *Horm Res Paediatr*. 2011;**75**(2):148–152.
123. Clementi M, Milani S, Mammi I, Boni S, Monciotti C, Tenconi R. Neurofibromatosis type 1 growth charts. *Am J Med Genet*. 1999; **87**(4):317–323.
124. Garibaldi LR, Aceto T Jr, Weber C, Pang S. The relationship between luteinizing hormone and estradiol secretion in female precocious puberty: evaluation by sensitive gonadotropin assays and the leuprolide stimulation test. *J Clin Endocrinol Metab*. 1993;**76**(4):851–856.
125. Ibáñez L, Potau N, Zampolli M, Virdis R, Gussinyé M, Carrascosa A, Saenger P, Vicens-Calvet E. Use of leuprolide acetate response patterns in the early diagnosis of pubertal disorders: comparison with the gonadotropin-releasing hormone test. *J Clin Endocrinol Metab*. 1994;**78**(1):30–35.
126. Neely EK, Wilson DM, Lee PA, Stene M, Hintz RL. Spontaneous serum gonadotropin concentrations in the evaluation of precocious puberty. *J Pediatr*. 1995;**127**(1):47–52.
127. Sigurjonsdottir TJ, Hayles AB. Precocious puberty. A report of 96 cases. *Am J Dis Child*. 1968;**115**(3):309–321.
128. Oostdijk W, Rikken B, Schreuder S, Otten B, Odink R, Rouwé C, Jansen M, Gerver WJ, Waelkens J, Drop S. Final height in central precocious puberty after long term treatment with a slow release GnRH agonist. *Arch Dis Child*. 1996;**75**(4):292–297.
129. Klein KO, Barnes KM, Jones JV, Feuillan PP, Cutler GB Jr. Increased final height in precocious puberty after long-term treatment with LHRH agonists: the National Institutes of Health experience. *J Clin Endocrinol Metab*. 2001;**86**(10):4711–4716.
130. Carel J-C, Roger M, Ispas S, Tondu F, Lahlou N, Blumberg J, Chaussain J-L. Final height after long-term treatment with triptorelin slow release for central precocious puberty: importance of statural growth after interruption of treatment. French study group of decapeptyl in precocious puberty. *J Clin Endocrinol Metab*. 1999;**84**(6):1973–1978.
131. Paul D, Conte FA, Grumbach MM, Kaplan SL. Long-term effect of gonadotropin-releasing hormone agonist therapy on final and near-final height in 26 children with true precocious puberty treated at a median age of less than 5 years. *J Clin Endocrinol Metab*. 1995;**80**(2):546–551.
132. Adan L, Souberbielle JC, Zucker JM, Pierre-Kahn A, Kalifa C, Brauner R. Adult height in 24 patients treated for growth hormone deficiency and early puberty. *J Clin Endocrinol Metab*. 1997;**82**(1):229–233.
133. Wendt S, Shelso J, Wright K, Furman W. Neoplastic causes of abnormal puberty. *Pediatr Blood Cancer*. 2014;**61**(4):664–671.
134. Lee PA, Klein K, Mauras N, Neely EK, Bloch CA, Larsen L, Mattia-Goldberg C, Chwalisz K. Efficacy and safety of leuprolide acetate 3-month depot 11.25 milligrams or 30 milligrams for the treatment of central precocious puberty. *J Clin Endocrinol Metab*. 2012;**97**(5):1572–1580.
135. Palmert MR, Dunkel L. Clinical practice. Delayed puberty. *N Engl J Med*. 2012;**366**(5):443–453.
136. Vigen R, O'Donnell CI, Barón AE, Grunwald GK, Maddox TM, Bradley SM, Barqawi A, Woning G, Wierman ME, Plomondon ME, Rumsfeld JS, Ho PM. Association of testosterone therapy with mortality, myocardial infarction, and stroke in men with low testosterone levels. *JAMA*. 2013;**310**(17):1829–1836.
137. Schmiegelow M, Lassen S, Poulsen HS, Schmiegelow K, Hertz H, Andersson A-M, Skakkebaek NE, Müller J. Gonadal status in male survivors following childhood brain tumors. *J Clin Endocrinol Metab*. 2001;**86**(6):2446–2452.
138. Metzger ML, Meacham LR, Patterson B, Casillas JS, Constine LS, Hijiya N, Kenney LB, Leonard M, Lockart BA, Likes W, Green DM. Female reproductive health after childhood, adolescent, and young adult cancers: guidelines for the assessment and management of female reproductive complications. *J Clin Oncol*. 2013; **31**(9):1239–1247.
139. Kenney LB, Cohen LE, Shnorhavorian M, Metzger ML, Lockart B, Hijiya N, Duffey-Lind E, Constine L, Green D, Meacham L. Male reproductive health after childhood, adolescent, and young adult cancers: a report from the Children's Oncology Group. *J Clin Oncol*. 2012;**30**(27):3408–3416.
140. Boehm U, Bouloux P-M, Dattani MT, de Roux N, Dodé C, Dunkel L, Dwyer AA, Giacobini P, Hardelin J-P, Juul A, Maghnie M, Pitteloud N, Prevot V, Raivio T, Tena-Sempere M, Quinton R, Young J. Expert consensus document: European Consensus Statement on congenital hypogonadotropic hypogonadism—pathogenesis, diagnosis and treatment. *Nat Rev Endocrinol*. 2015;**11**(9):547–564.
141. Rohayem J, Hauffa BP, Zacharin M, Kliesch S, Zitzmann M; "German Adolescent Hypogonadotropic Hypogonadism Study Group". Testicular growth and spermatogenesis: new goals for pubertal hormone replacement in boys with hypogonadotropic hypogonadism?—a multicentre prospective study of hCG/rFSH treatment outcomes during adolescence. *Clin Endocrinol (Oxf)*. 2017;**86**(1):75–87.
142. Sato N, Hasegawa T, Hasegawa Y, Arisaka O, Ozono K, Amemiya S, Kikuchi T, Tanaka H, Harada S, Miyata I, Tanaka T. Treatment situation of male hypogonadotropic hypogonadism in pediatrics and proposal of testosterone and gonadotropins replacement therapy protocols. *Clin Pediatr Endocrinol*. 2015;**24**(2):37–49.
143. Moskowitz CS, Malhotra J, Chou JF, Wolden SL, Weathers RE, Stovall M, Armstrong GT, Leisenring WM, Neglia JP, Robison LL, Oeffinger KC. Breast cancer following spinal irradiation for a childhood cancer: a report from the Childhood Cancer Survivor Study. *Radiother Oncol*. 2015;**117**(2):213–216.
144. Moskowitz CS, Chou JF, Sklar CA, Barnea D, Ronckers CM, Friedman DN, Neglia JP, Turcotte L, Howell RM, Henderson TO, Armstrong GT, Leisenring WM, Robison LL, van Leeuwen FE, Pike MC, Oeffinger KC. Radiation-associated breast cancer and gonadal hormone exposure: a report from the Childhood Cancer Survivor Study. *Br J Cancer*. 2017;**117**(2):290–299.
145. Paragliola RM, Prete A, Kaplan PW, Corsello SM, Salvatori R. Treatment of hypopituitarism in patients receiving antiepileptic drugs. *Lancet Diabetes Endocrinol*. 2015;**3**(2):132–140.
146. Hanley P, Lord K, Bauer AJ. Thyroid disorders in children and adolescents: a review. *JAMA Pediatr*. 2016;**170**(10):1008–1019.
147. Darzy KH. Radiation-induced hypopituitarism after cancer therapy: who, how and when to test. *Nat Clin Pract Endocrinol Metab*. 2009;**5**(2):88–99.
148. Gleeson HK, Shalet SM. The impact of cancer therapy on the endocrine system in survivors of childhood brain tumours. *Endocr Relat Cancer*. 2004;**11**(4):589–602.
149. Clement SC, Schoot RA, Slater O, Chisholm JC, Abela C, Balm AJM, van den Brekel MW, Breunis WB, Chang YC, Davila Fajardo R, Dunaway D, Gajdosova E, Gaze MN, Gupta S,

Hartley B, Kremer LCM, van Lennep M, Levitt GA, Mandeville HC, Pieters BR, Saeed P, Smeele LE, Strackee SD, Ronckers CM, Caron HN, van Santen HM, Merks JHM. Endocrine disorders among long-term survivors of childhood head and neck rhabdomyosarcoma. *Eur J Cancer*. 2016;**54**:1–10.
150. Ratnasingam J, Karim N, Paramasivam SS, Ibrahim L, Lim LL, Tan ATB, Vethakkan SR, Jalaludin A, Chan SP. Hypothalamic pituitary dysfunction amongst nasopharyngeal cancer survivors. *Pituitary*. 2015;**18**(4):448–455.
151. Brignardello E, Felicetti F, Castiglione A, Chiabotto P, Corrias A, Fagioli F, Ciccone G, Boccuzzi G. Endocrine health conditions in adult survivors of childhood cancer: the need for specialized adult-focused follow-up clinics. *Eur J Endocrinol*. 2013;**168**(3):465–472.
152. Hudson MM, Ness KK, Gurney JG, Mulrooney DA, Chemaitilly W, Krull KR, Green DM, Armstrong GT, Nottage KA, Jones KE, Sklar CA, Srivastava DK, Robison LL. Clinical ascertainment of health outcomes among adults treated for childhood cancer. *JAMA*. 2013;**309**(22):2371–2381.
153. Shalitin S, Gal M, Goshen Y, Cohen I, Yaniv I, Phillip M. Endocrine outcome in long-term survivors of childhood brain tumors. *Horm Res Paediatr*. 2011;**76**(2):113–122.
154. Bonato C, Severino RF, Elnecave RH. Reduced thyroid volume and hypothyroidism in survivors of childhood cancer treated with radiotherapy. *J Pediatr Endocrinol Metab*. 2008;**21**(10):943–949.
155. van Santen HM, Vulsma T, Dijkgraaf MG, Blumer RME, Heinen R, Jaspers MWM, Geenen MM, Offringa MO, de Vijlder JJM, van den Bos C. No damaging effect of chemotherapy in addition to radiotherapy on the thyroid axis in young adult survivors of childhood cancer. *J Clin Endocrinol Metab*. 2003;**88**(8):3657–3663.
156. Schmiegelow M, Feldt-Rasmussen U, Rasmussen AK, Poulsen HS, Müller J. A population-based study of thyroid function after radiotherapy and chemotherapy for a childhood brain tumor. *J Clin Endocrinol Metab*. 2003;**88**(1):136–140.
157. Sklar CA, Constine LS. Chronic neuroendocrinological sequelae of radiation therapy. *Int J Radiol Oncol Biol Phys*. 1995;**31**(5):1113–1121.
158. Merchant TE, Conklin HM, Wu S, Lustig RH, Xiong X. Late effects of conformal radiation therapy for pediatric patients with low-grade glioma: prospective evaluation of cognitive, endocrine, and hearing deficits. *J Clin Oncol*. 2009;**27**(22):3691–3697.
159. Rose SR, Lustig RH, Pitukcheewanont P, Broome DC, Burghen GA, Li H, Hudson MM, Kun LE, Heideman RL. Diagnosis of hidden central hypothyroidism in survivors of childhood cancer. *J Clin Endocrinol Metab*. 1999;**84**(12):4472–4479.
160. Darzy KH, Shalet SM. Circadian and stimulated thyrotropin secretion in cranially irradiated adult cancer survivors. *J Clin Endocrinol Metab*. 2005;**90**(12):6490–6497.
161. Ferretti E, Persani L, Jaffrain-Rea M-L, Giambona S, Tamburrano G, Beck-Peccoz P. Evaluation of the adequacy of levothyroxine replacement therapy in patients with central hypothyroidism. *J Clin Endocrinol Metab*. 1999;**84**(3):924–929.
162. Carl JS, Weaver SP, Tweed E, Edgerton L. Effect of antiepileptic drugs on oral contraceptives. *Am Fam Physician*. 2008;**78**(5):634–635.
163. Surks MI, DeFesi CR. Normal serum free thyroid hormone concentrations in patients treated with phenytoin or carbamazepine. A paradox resolved. *JAMA*. 1996;**275**(19):1495–1498.
164. Odagiri K, Omura M, Hata M, Aida N, Niwa T, Ogino I, Kigasawa H, Ito S, Adachi M, Inoue T. Treatment outcomes, growth height, and neuroendocrine functions in patients with intracranial germ cell tumors treated with chemoradiation therapy. Int J Radiol Oncol Biol Phys. 2012;84(3):632–638.
165. Collett-Solberg PF, Sernyak H, Satin-Smith M, Katz LL, Sutton L, Molloy P, Moshang T Jr. Endocrine outcome in long-term survivors of low-grade hypothalamic/chiasmatic glioma. *Clin Endocrinol (Oxf)*. 1997;**47**(1):79–85.
166. Rose SR, Danish RK, Kearney NS, Schreiber RE, Lustig RH, Burghen GA, Hudson MM. ACTH deficiency in childhood cancer survivors. *Pediatr Blood Cancer*. 2005;**45**(6):808–813.
167. Olsson DS, Andersson E, Bryngelsson I-L, Nilsson AG, Johannsson G. Excess mortality and morbidity in patients with craniopharyngioma, especially in patients with childhood onset: a population-based study in Sweden. *J Clin Endocrinol Metab*. 2015;**100**(2):467–474.
168. Mills JL, Schonberger LB, Wysowski DK, Brown P, Durako SJ, Cox C, Kong F, Fradkin JE. Long-term mortality in the United States cohort of pituitary-derived growth hormone recipients. *J Pediatr*. 2004;**144**(4):430–436.
169. Schmiegelow M, Feldt-Rasmussen U, Rasmussen AK, Lange M, Poulsen HS, Müller J. Assessment of the hypothalamo-pituitary-adrenal axis in patients treated with radiotherapy and chemotherapy for childhood brain tumor. *J Clin Endocrinol Metab*. 2003;**88**(7):3149–3154.
170. Patterson BC, Truxillo L, Wasilewski-Masker K, Mertens AC, Meacham LR. Adrenal function testing in pediatric cancer survivors. *Pediatr Blood Cancer*. 2009;**53**(7):1302–1307.
171. Spoudeas HA, Charmandari E, Brook CGD. Hypothalamo-pituitary-adrenal axis integrity after cranial irradiation for childhood posterior fossa tumours. *Med Pediatr Oncol*. 2003;**40**(4):224–229.
172. Crowne EC, Wallace WH, Gibson S, Moore CM, White A, Shalet SM. Adrenocorticotrophin and cortisol secretion in children after low dose cranial irradiation. *Clin Endocrinol (Oxf)*. 1993;**39**(3):297–305.
173. Follin C, Wiebe T, Moëll C, Erfurth EM. Moderate dose cranial radiotherapy causes central adrenal insufficiency in long-term survivors of childhood leukaemia. *Pituitary*. 2014;**17**(1):7–12.
174. Park J, Didi M, Blair J. The diagnosis and treatment of adrenal insufficiency during childhood and adolescence. *Arch Dis Child*. 2016;**101**(9):860–865.
175. Shankar RR, Jakacki RI, Haider A, Lee MW, Pescovitz OH. Testing the hypothalamic-pituitary-adrenal axis in survivors of childhood brain and skull-based tumors. *J Clin Endocrinol Metab*. 1997;**82**(6):1995–1998.
176. Ospina NS, Al Nofal A, Bancos I, Javed A, Benkhadra K, Kapoor E, Lteif AN, Natt N, Murad MH. ACTH stimulation tests for the diagnosis of adrenal insufficiency: systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2016;**101**(2):427–434.
177. Ng SM, Agwu JC, Dwan K. A systematic review and meta-analysis of Synacthen tests for assessing hypothalamic-pituitary-adrenal insufficiency in children. *Arch Dis Child*. 2016;**101**(9):847–853.
178. Cross AS, Helen Kemp E, White A, Walker L, Meredith S, Sachdev P, Krone NP, Ross RJ, Wright NP, Elder CJ. International survey on high- and low-dose Synacthen test and assessment of accuracy in preparing low-dose Synacthen. *Clin Endocrinol (Oxf)*. 2018;**88**(5):744–751.
179. Bornstein SR, Allolio B, Arlt W, Barthel A, Don-Wauchope A, Hammer GD, Husebye ES, Merke DP, Murad MH, Stratakis CA, Torpy DJ. Diagnosis and treatment of primary adrenal insufficiency: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*. 2016;**101**(2):364–389.
180. Linas SL, Berl T, Robertson GL, Aisenbrey GA, Schrier RW, Anderson RJ. Role of vasopressin in the impaired water excretion of glucocorticoid deficiency. *Kidney Int*. 1980;**18**(1):58–67.
181. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schünemann HJ, Edejer T, Varonen H, Vist GE, Williams JW Jr, Zaza S; GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;**328**(7454):1490.
182. Swiglo BA, Murad MH, Schünemann HJ, Kunz R, Vigersky RA, Guyatt GH, Montori VM. A case for clarity, consistency, and helpfulness: state-of-the-art clinical practice guidelines in endocrinology using the grading of recommendations, assessment, development, and evaluation system. *J Clin Endocrinol Metab*. 2008;**93**(3):666–673.

Diagnostic Accuracy of Ultrasound With Color Flow Doppler in Children With Thyroid Nodules

Anthony W. Gannon,[1] Jill E. Langer,[2] Richard Bellah,[3] Sarah Ratcliffe,[4] Jennifer Pizza,[5] Sogol Mostoufi-Moab,[5,6] Anne R. Cappola,[7] and Andrew J. Bauer[5]

[1]Division of Pediatric Endocrinology, Nemours/Alfred I DuPont Hospital for Children, Wilmington, Delaware 19803; [2]Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104; [3]Department of Radiology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104; [4]Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, Pennsylvania 19104; [5]Division of Endocrinology and Diabetes, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104; [6]Division of Oncology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104; and [7]Division of Endocrinology, Diabetes, and Metabolism, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Context: Thyroid nodules are increasingly recognized in children and are associated with a greater risk for thyroid cancer compared with adults. Thyroid ultrasound is the favored tool for evaluation of thyroid nodules; however, there are limited data regarding the accuracy of thyroid ultrasound to confirm features associated with a low risk of thyroid cancer in children.

Objectives: We examined whether thyroid ultrasound is capable of accurately identifying thyroid nodules at a low risk of malignancy in children.

Design and Setting: Using a retrospective cohort study design, we identified children age ≤18 years with thyroid nodules and adequate follow-up. Ultrasound images were reviewed independently by two blinded expert radiologists, and ultrasound characteristics were analyzed to determine optimal predictive value and reliability.

Patients and Results: A total of 417 subjects were found to have thyroid nodules, and 152 subjects had adequate follow-up; 59 (38.8%) of these were diagnosed with thyroid cancer. We evaluated 236 individual nodules. Features most consistent with benign nodules included small size, isoechoic echogenicity, partially cystic structure, sharp or noninfiltrative margins, absent Doppler flow, and absent calcifications. Significant variability was found between expert interpretations of ultrasound features. Thyroid nodule composition appears to be the most sensitive and reliable feature for stratifying the risk of thyroid cancer. Ultrasound accurately identified benign thyroid nodules in 80.9% of subjects (95% confidence interval, 74–86.6).

Conclusions: Ultrasonography is useful for the evaluation of thyroid nodules, but we found no combination of ultrasound features sufficient to exclude thyroid cancer without a biopsy. (***J Clin Endocrinol Metab*** **103: 1958–1965, 2018)**

The incidence of thyroid nodules and cancers in children and adolescents has increased over the past 30 years (1–3). The annual incidence of thyroid cancer has increased by 3.74% per year in all children, with the majority of cases being identified in adolescent female patients (4). Although thyroid nodules are generally associated with a low risk of malignancy in adults (~5% to 10%), the risk of malignancy is significantly higher in children (22% to 26%) (5, 6). Nonetheless, most thyroid nodules in children are benign.

Thyroid ultrasound is the principal tool for risk stratification of thyroid nodules. Although ultrasound

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received 13 November 2017. Accepted 7 March 2018.
First Published Online 12 March 2018

Abbreviations: ATA, American Thyroid Association; CI, confidence interval; OR, odds ratio.

features that suggest malignancy are well recognized, there are sparse data to confirm which nodules have a low risk of representing thyroid cancer for which biopsy and surgery could be reasonably avoided. Large studies in adults have demonstrated a low risk for thyroid cancer in nodules with partially cystic composition and isoechoic echogenicity combined with the absence of high-risk features (*e.g.*, calcifications, infiltrative borders, taller-than-wide shape, and extrathyroidal extension). The American Thyroid Association (ATA) guidelines for evaluation of thyroid nodules in adults recognize combinations of ultrasound features associated with low (<10%) risk of thyroid cancer and support withholding thyroid nodule biopsy until the nodule exceeds 1.5 cm in maximal dimension (7, 8). No such risk stratification strategy has been identified in children, so fine needle aspiration (FNA) biopsy is routinely recommended for children with solid or partially cystic thyroid nodules ≥1 cm in maximal diameter or any size nodule with any suspicious features (1).

Current data regarding ultrasound evaluation of thyroid nodules in children are limited by small numbers, variability in ultrasound protocols and technologies, referral bias, potential bias from nonblinded evaluation of ultrasounds, and a lack of assessment of reliability of ultrasound interpretation between radiologists (5, 9–14). These limitations undermine confidence in the predictive ability of thyroid ultrasonography in children.

This study was designed to examine the diagnostic accuracy of thyroid ultrasonography with color flow Doppler in a large cohort of children analyzed by experienced radiologists blinded to the subjects' clinical information. Our primary objective was to determine whether any ultrasound features alone or in combination could be associated with a low enough risk of malignancy (<10% risk) to allow for the same conservative approach currently practiced in adults. Additionally, we sought to determine the reliability of thyroid ultrasound features between expert radiologists.

Materials and Methods

We performed a retrospective cohort study including all children (defined as age ≤18 years) diagnosed with a thyroid nodule in the Children's Hospital of Philadelphia medical system between January 2009 and March 2013 who underwent adequate clinical follow-up to determine whether the nodule was malignant. The protocol was reviewed and approved by the Children's Hospital of Philadelphia institutional review board.

Subjects were identified through a system-wide search based on completion of a neck ultrasound (CPT 76536) for any reason and were considered for inclusion if chart review confirmed the presence of a thyroid nodule. Review of the ultrasound images and medical record was performed to confirm that subjects had a thyroid nodule ≥0.5 cm. Clinical follow-up was considered adequate if any of the following were achieved: (1) surgical pathology; (2) FNA biopsy consistent with a benign lesion plus follow-up including ultrasound imaging without evidence of substantial growth (>20% increase in maximal dimension) or new suspicious features for ≥12 months; or (3) clinical follow-up including ultrasound imaging without evidence of substantial growth (>20% increase in maximal dimension), no new suspicious sonographic features, and no clinical suspicion of malignancy for ≥12 months.

Study data were collected and managed using REDCap electronic capture tools hosted at the University of Pennsylvania. Ultrasound image sets (including dynamic images when available) were manually extracted and coded (ShowCase 5.0; Trilium Technology, Inc., Ann Arbor, MI) to remove identifiers. Two blinded radiologists with expertise in adult and pediatric thyroid ultrasonography evaluated the image sets independently. One radiologist reviewed all images for the study, and the acquired data were used for the primary analysis. Reliability of ultrasound interpretation was determined by duplicating a random selection of images (30%) for review by the primary radiologist and the second radiologist. Consistent with real-world experience, disagreements between reports were not reconciled.

Statistical analysis

Data analysis was performed using Stata 12.1 (StataCorp, College Station, TX). A P value of <0.05 was considered statistically significant, and two-sided tests were used. Group differences in subjects with thyroid carcinoma and subjects with benign lesions were determined using Wilcoxon rank sum or Kruskall-Wallis as indicated. Differences in proportions were assessed using Fisher's exact test or χ^2 test as indicated. Reliability was assessed through use of Cohen's κ. Based on our power calculations and on the expectation of a prevalence of malignancy of 30%, 149 subjects would be required to provide 90% power to detect an inferior negative predictive value of 90% with $\alpha = 0.05$.

Data were analyzed in a bivariate logistic regression model with adjustment for clustering within individuals. Variables that were clinically important or found to be statistically significant in bivariate modeling were included in a multivariable logistic regression model with adjustment for clustering; backward deletion was used to determine the variables that affected the final model. The final model was validated by a sensitivity analysis. Variables were further analyzed in isolation and in combination to determine their sensitivity, specificity, negative predictive value, and positive predictive value with 95% confidence intervals (CIs). Predictive modeling was accomplished with multivariable logistic regression modeling and receiver operator curve analysis.

Results

A total of 3069 children received neck ultrasounds during the study period. The most common indications were

cervical lymphadenopathy, cellulitis, salivary gland enlargement, vascular imaging, and neck masses. A total of 417 (13.6%) subjects were reported to have thyroid nodules, and 152 subjects with 241 unique thyroid nodules met criteria for inclusion in the study (Fig. 1). Of these subjects, 93 (61.2%) were diagnosed with benign thyroid nodules, and 59 (38.8%) were diagnosed with thyroid cancer (Table 1). Eighty-five subjects (55.9%) were diagnosed with a thyroid nodule in our hospital system, and the remainder of subjects were referred externally. The median age was 14.2 ± 3.8 years, and 121 (79%) were female. The majority (72.7%) of subjects were white and were asymptomatic at the time of diagnosis. Seventeen subjects (11.3%) had a history of cancer, and 16 (10.5%) had received total body irradiation or craniospinal or mediastinal radiation therapy prior to being diagnosed with thyroid nodules; no increased prevalence of thyroid cancer was seen in these subjects. Two subjects had received radioactive iodine ablation therapy for hyperthyroidism, and neither had thyroid cancer. The majority (59.2%) of subjects had a positive family history of thyroid disease, 18 (12.8%) had a positive family history of thyroid cancer, and 55 (39.3%) had a positive family history of nonthyroidal cancer; none of these historical features was significantly associated with final diagnosis of thyroid cancer in the subjects.

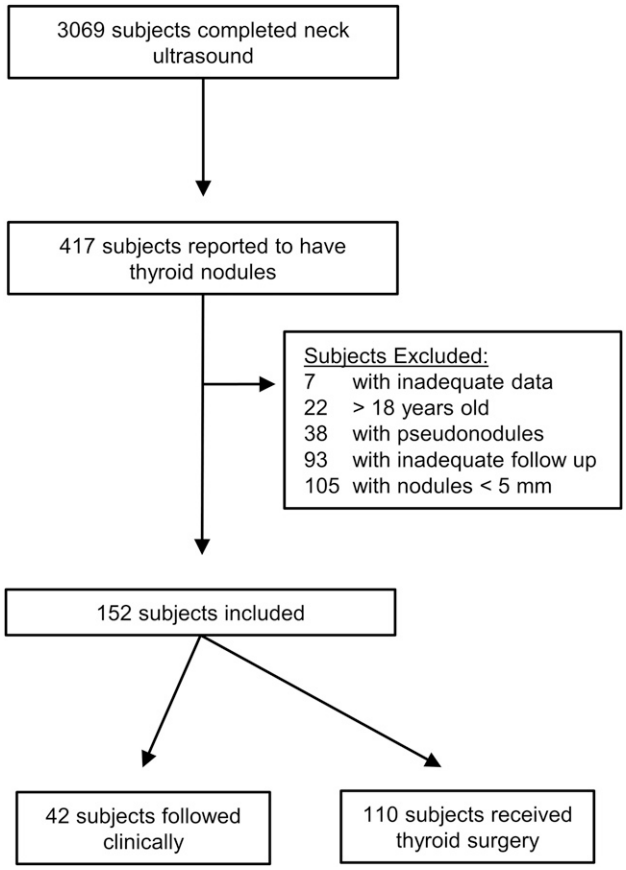


Figure 1. Consort diagram for subject selection and inclusion into the study.

Thyroid nodules were most commonly (64.5%) identified on palpation by the physician, a family member, or the subject; the remainder (35.5%) were found incidentally during unrelated imaging studies. Physical examination by a physician in our center identified a palpable nodule in 91 subjects (59.9%) and palpable cervical lymphadenopathy in 31 subjects (20.3%). The median thyroid-stimulating hormone value was 1.25 mIU/L (interquartile range, 1.55), and 36 subjects (35.3%) had positive antithyroid autoantibodies. There were no significant baseline differences in the characteristics between patients with benign and malignant thyroid nodules except for the distribution of race (P = 0.04) and the presence of palpable cervical lymphadenopathy (P = 0.01).

The endpoint of the study was achieved based on surgical pathology in 110 (72.4%) of subjects, by FNA biopsy followed by observation in 28 (12.3%) of subjects, and by observation without biopsy in 24 (15.3%) subjects.

Ultrasound findings

Ultrasound images identified 241 unique thyroid nodules. Data were complete for 236 nodules; 80 (33.8%) of these nodules were confirmed to be thyroid cancer (Table 2). A single nodule was identified in 94 (61.8%) subjects, 27 (17.8%) subjects had two nodules, and 31 (20.4%) subjects had three or more nodules. There was no association between the number of nodules and the risk of malignancy. The median maximal dimension of thyroid nodules was 1.6 ± 1.7 cm; benign thyroid nodules were smaller (1.5 ± 1.4 cm) than malignant thyroid nodules (2.0 ± 1.7 cm). Thyroid nodule size ≥1 cm was associated with an increased odds ratio (OR) for thyroid cancer (OR, 1.78; 95% CI, 1.15 to 2.75).

Thyroid nodule composition was ≤25% cystic in 152 (64%) nodules, and 68 (44.7%) of these nodules were malignant. Conversely, >25% cystic composition was observed in 84 nodules, and 12 (14.3%) of these nodules were malignant. Thyroid nodules with ≤25% cystic composition were found to harbor 91.4% of observations of nonisoechoic echogenicity, 96.7% of observations of infiltrative/microlobulated margins, and 91.5% of observations of calcifications in the population. Altogether, ≤25% cystic composition identified 85% of thyroid cancer cases in the study and was associated with a significant increase in the OR for thyroid cancer (OR, 3.46; 95% CI, 1.96 to 6.09).

Thyroid nodule echogenicity was isoechoic in 93 (39.4%) nodules, and 13 (13.9%) of these nodules were

Table 1. Baseline Characteristics of Subjects at Time of Diagnosis of a Thyroid Nodule

Feature	All Subjects (n = 152)	Benign Nodules (n = 93)	Malignant Nodules (n = 59)	*P* Value[a]
Historical findings				
Age, y (SD)	14.2 (3.8)	14.1 (3.8)	14.4 (4.1)	0.52
Female	121 (79.6)	74 (79.6)	47 (79.7)	0.99
White	104 (72.7)	70 (78.7)	34 (62.9)	0.04
Prepubertal	12 (9.2)	10 (12.5)	2 (3.9)	0.10
Incidental finding	54 (35.5)	36 (38.7)	18 (30.5)	0.30
Symptoms of tracheoesophageal pressure[b]	12 (7.9)	7 (11.9)	5 (5.38)	0.15
History of malignancy	17 (11.3)	12 (12.9)	5 (8.6)	0.59
History of radiation therapy[c]	16 (10.5)	11 (11.8)	5 (8.5)	0.60
Family history of thyroid disease[d]	84 (59.2)	52 (60.5)	32 (57.1)	0.70
Family history of thyroid nodules or thyroid cancer	45 (32.1)	29 (34.1)	15 (29.1)	
Family history of thyroid cancer	18 (12.8)	11 (12.9)	7 (12.5)	
Family history of nonthyroid malignancy	55 (39.3)	34 (40)	21 (38.2)	0.83
Clinical findings				
Diffuse goiter	60 (39.2)	35 (37.6)	25 (42.4)	0.56
Palpable nodule	91 (59.9)	50 (53.8)	41 (69.5)	0.05
Palpable cervical lymphadenopathy	31 (20.3)	13 (14)	18 (30.5)	0.01
TSH, mIU/L (SD)	1.25 (1.55)	1.14 (1.49)	1.31 (1.71)	0.30
Positive antithyroid antibodies	36 (35.3)	20 (32.3)	16 (40)	0.42

Values are n (%) unless noted otherwise.

Abbreviations: SD, standard deviation; TSH, thyroid-stimulating hormone.

[a] *P* values refer to statistical significance of the difference between subjects with benign and malignant nodules.

[b] Symptoms of dysphagia, dysphonia, or globus.

[c] Radiation therapy included total body, craniospinal, and/or mediastinal radiation therapy.

[d] Any disease causing abnormal thyroid function or thyroid injury, including thyroiditis, Graves disease, hypothyroidism, and hyperthyroidism.

malignant. Hypoechoic echogenicity was observed in 87 (36.9%) nodules, and 40 (45.9%) of these were malignant. Thirty-five nodules demonstrated hyperechoic echogenicity, and 16 nodules demonstrated mixed hypoechoic/hyperechoic echogenicity; 16 (45.7%) and 10 (62.5%) of these nodules, respectively, were malignant. The majority of high-risk ultrasound features were seen in nodules with nonisoechoic echogenicity, including 94.9% of infiltrative/microlobulated margin observations, 94.1% of extrathyroidal extension observations, and 84.5% of calcification observations. The presence of nonisoechoic echogenicity identified 83.8% of thyroid cancer cases in the study and was associated with a significantly increased OR for thyroid cancer (OR, 3.91; 95% CI, 2.18 to 7.01).

Thyroid nodule margins were sharply defined or noninfiltrative in 177 (75%) nodules, and 38 (21.4%) of these nodules were malignant. Infiltrative or microlobulated margins were seen in 58 (24.7%) nodules, and 41 (70.7%) of these were malignant. The presence of infiltrative or microlobulated margins was associated with a significantly increased OR for malignancy (OR, 5.58; 95% CI, 2.79 to 11.18).

Taller-than-wide structure on transverse imaging was seen in 33 (13.9%) nodules, and 16 (51.5%) of these nodules were malignant. Calcifications were present in 70 (29.8%) nodules, and 48 (68.6%) of these nodules were malignant. Extrathyroidal extension was described in 16 (6.8%) nodules, and 13 (81.3%) of these nodules were malignant. Each of these features was associated with a marked increase in OR for thyroid carcinoma (Table 2).

Intrarater reliability was found to be excellent for thyroid nodule size ($\kappa = 1.0$; $P < 0.001$), nodule composition ($\kappa = 0.8$; $P < 0.001$), and calcifications ($\kappa = 0.86$; $P < 0.001$), whereas it was moderate for nodule echogenicity ($\kappa = 0.57$; $P < 0.001$), nodule margins ($\kappa = 0.57$; $P < 0.001$), and Doppler flow ($\kappa = 0.58$; $P < 0.001$). Interrater reliability was variable, with excellent agreement for nodule size ($\kappa = 0.96$; $P < 0.001$), good agreement for nodule composition ($\kappa = 0.79$; $P < 0.001$) and Doppler flow ($\kappa = 0.77$; $P < 0.001$), moderate agreement for extrathyroidal extension ($\kappa = 0.4$, $P < 0.001$), fair agreement for nodule margins ($\kappa = 0.34$; $P < 0.001$) and calcifications ($\kappa = 0.39$; $P = 0.002$), and poor agreement for nodule echogenicity ($\kappa = 0.19$; $P = 0.003$).

Using a multivariable logistic regression model including all ultrasound features and accounting for clustering within subjects and statistical interaction between calcifications and Doppler flow patterns, we found that nonisoechoic echogenicity ($P = 0.019$) and extrathyroidal extension ($P = 0.01$) remained statistically significant predictors of thyroid cancer in children. A prespecified sensitivity analysis revealed that these features were

Table 2. Bivariable Logistic Regression Analysis of the Odds of Thyroid Cancer and Reliability of Ultrasound Features

	All Nodules (n = 236)	Benign Nodules (n = 156)	Malignant Nodules (n = 80)	OR for Thyroid Cancer Diagnosis (95% CI)	Intrarater Variability, κ (*P* Value)	Interrater Variability, κ (*P* Value)
No. of nodules						
One	94	57	37	0.77 (0.53–1.13)		
Two	54	37	17			
Three	88	62	26			
Size (greatest dimension), cm						
<1	65	50	15	1.78 (1.15–2.75)	1.0 (<0.001)	0.96 (<0.001)
1–4	155	97	58			
>4	16	9	7			
Nodule composition						
>25% cystic	84	72	12	3.46 (1.96–6.09)	0.8 (<0.001)	0.79 (<0.001)
≤25% cystic	152	84	68			
Echogenicity						
Isoechoic	93	80	13	1.0	0.57 (<0.001)	0.19 (0.003)
Hypoechoic	87	47	40	3.69 (1.97–6.93)		
Hyperechoic	35	19	16	3.99 (1.79–8.86)		
Mixed hypoechoic and hyperechoic	16	6	10	4.91 (1.98–12.18)		
Margins						
Sharply defined/ noninfiltrative	177	139	38	5.58 (2.79–11.18)	0.57 (<0.001)	0.34 (0.001)
Infiltrative/microlobular	58	17	41			
Calcifications						
Absent	165	133	32	6.85 (3.47–13.49)	0.86 (<0.001)	0.39 (0.002)
Present	70	22	48			
Doppler flow						
Absent	45	38	7	3.52 (1.24–10.01)	0.58 (<0.001)	0.77 (<0.001)
Increased	179	109	70			
Taller than wide						
Absent	203	140	63	2.22 (1.27–3.88)	0.58 (<0.001)	0.39 (0.002)
Present	33	16	17			
Extrathyroidal extension						
Absent	220	153	67	6.15 (2.39–15.86)	−0.03 (0.61)	0.4 (<0.001)
Present	16	3	13			

Logistic regression was performed with adjustment for clustering of features within individuals.

affected by clinical features such as subject age, pubertal status, and race/ethnicity but not by personal history of cancer or radiation exposure.

Receiver operator curve analysis

The clinical utility of ultrasound features for prediction of thyroid cancer was evaluated in bivariable and multivariable models (Figure 2; Table 3). Numerous patterns of features were observed in the sample, and this significantly affected the predictive value of each feature. Optimal sensitivity was found in ≤25% cystic thyroid nodule composition (85%; 95% CI, 75.3 to 92), increased Doppler flow (90.9%; 95% CI, 82.2 to 96.3), and nonisoechoic echogenicity (83.5%; 95% CI, 73.5 to 90.9). Optimal specificity was found in extrathyroidal extension (98.1%; 94.5 to 99.6), taller-than-wide shape (89.7%; 95% CI, 83.9 to 94), and infiltrative/microlobulated margins (89.1%; 95% CI, 83.1 to 93.5). The most powerful predictive values for thyroid cancer were extrathyroidal extension (81.2%; 95% CI, 54.4 to 96), infiltrative/microlobulated margins (70.7%; 95% CI, 57.3 to 81.9), and calcifications (68.6%; 95% CI, 56.4 to 79.1). The most powerful predictive values

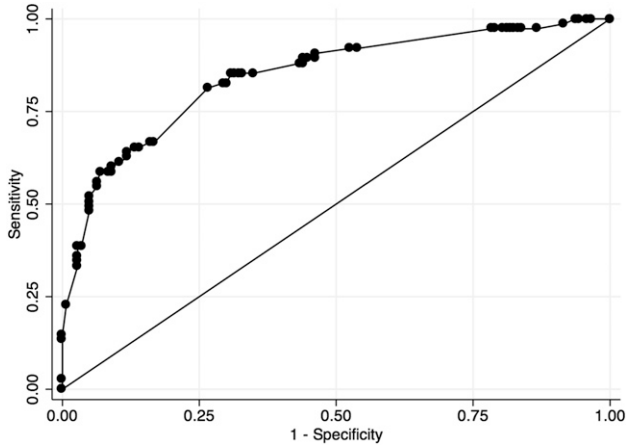


Figure 2. Receiver operator curve for performance of ultrasound features in correctly diagnosing thyroid cancer in children.

Table 3. Analysis of Predictive Statistics and Receiver Operator Curve for Ultrasound Features for Thyroid Carcinoma in Children

Variable	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	ROC (95% CI)
Nodule size >1 cm	81% (71–89.1)	32.1% (24.8–40)	38% (30.7–45.7)	76.9% (64.8–86.5)	0.57 (0.51–0.62)
≤25% solid nodule composition	85% (75.3–92)	46.2% (38.2–54.3)	44.7% (36.7–53)	85.7% (76.4–92.4)	0.66 (0.6–0.71)
Nonisoechoic echogenicity	83.5% (73.5–90.9)	52.6% (44.4–60.8)	47.8% (39.3–56.5)	86% (77.3–92.3)	0.68 (0.62–0.74)
Taller-than-wide	21.2% (12.9–31.8)	89.7% (83.9–94.0)	51.5% (33.5–69.2)	69% (62.1–75.3)	0.56 (0.5–0.61)
Increased Doppler flow	90.9% (82.2–96.3)	25.9% (19–33.7)	39.1% (31.9–46.7)	84.4% (70.5–93.5)	0.58 (0.54–0.63)
Presence of calcifications	60% (48.4–70.8)	85.8% (79.3–90.9)	68.6% (56.4–79.1)	80.6% (73.7–86.3)	0.73 (0.67–0.79)
Infiltrative/microlobular margin	51.9% (40.4–63.3)	89.1% (83.1–93.5)	70.7% (57.3–81.9)	78.5% (71.7–84.3)	0.71 (0.64–0.77)
Extrathyroidal extension	16.2% (8.9–26.2)	98.1% (94.5–99.6)	81.2% (54.4–96.0)	69.5% (63.0–75.6)	0.57 (0.53–0.61)
All features combined	58.7% (46.7–69.9)	91.6% (85.8–95.6)	78.6% (65.6–88.4)	80.9% (74–86.6)	0.8449

Abbreviations: NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operator curve.

for benign thyroid nodules were isoechoic echogenicity (86%; 95% CI, 77.3 to 92.3), >25% cystic composition (85.7%; 95% CI, 76.4 to 92.4), and absence of Doppler flow (84.4%; 95% CI, 70.5 to 93.5).

Multivariable analyses of these ultrasound features in combination demonstrated combined sensitivity of 58.7% (95% CI, 46.7 to 69.9), specificity of 91.6% (95% CI, 85.8 to 95.6), positive predictive value of 78.6% (95% CI, 65.6 to 88.4), and negative predictive value of 80.9% (95% CI, 74 to 86.6) for thyroid cancer.

Discussion

The diagnostic evaluation of thyroid nodules in children has thus far been limited to cohorts of children at increased risk of thyroid cancer due to exposure to radiation (10) or associated with referral to tertiary care centers for care (5, 9). Analysis of these data has been limited by small numbers of observations, a lack of a rigorous review of primary imaging data (9), a lack of assessment of reliability of radiology interpretation (5), and reconciliation of disagreements between radiologists (13, 14). This study was designed to address these limitations, and the principle question we sought to answer was whether thyroid ultrasound could reliably identify thyroid nodules in children associated with a low (<10%) risk of thyroid cancer so that biopsy may be avoided.

This study contains the largest and most rigorous primary evaluation of thyroid nodules in children thus far reported. By using a broad case-finding strategy, we recruited most of our subjects from the local population of the tertiary care center, and the majority of subjects had no known risk factors for thyroid nodules or thyroid cancer. We believe that this strategy was effective because of the diversity of subjects recruited and because the study met the prespecified recruitment goal by evaluating images obtained over the course of 4.25 years, compared with >10 years required by other large pediatric studies (5, 9). This short recruitment phase also allowed for greater consistency in ultrasound protocol and technology.

Once children were identified to have thyroid nodules, they received further evaluation at the direction of a dedicated pediatric thyroid center within the hospital system, which included either observation or diagnostic procedures. Although this was considered a strength during the design of the study, it introduced an important limitation in that children with low-risk ultrasound features were given appropriate reassurance and were less likely to complete follow-up, making them ineligible for inclusion in this study. Conversely, children with more concerning ultrasound features were more likely to complete the recommended follow-up and could be included in the study. This ascertainment bias resulted in a relatively high prevalence of thyroid cancer in our population, although the prevalence is similar to what has been reported by other groups (1, 5, 10, 15).

The greatest strength of this study is the rigorous evaluation of ultrasound image data in every subject. Although ultrasound features were evaluated by two expert radiologists, the analysis was based on features reported by a single radiologist, as would occur in a practical clinical setting. This strategy prevented a bias toward overestimation of the reliability and accuracy of the ultrasound when taken out of the research environment. Consistent with the findings of other groups, we observed substantial variation in the interpretation of ultrasound features between reviewers (6, 12, 13). In particular, we found fair-to-poor agreement in features such as nodule echogenicity, nodule margins, taller-than-wide shape, calcifications, and extrathyroidal extension. The reliability of imaging reporting was not assessed in other large studies of pediatric thyroid nodules, and this represents a practical challenge when relying on written

thyroid ultrasound reports for research or patient care. These challenges are further compounded considering the variation in the completeness of thyroid and neck ultrasound image acquisition and reports that are produced outside of dedicated thyroid centers.

Ultrasound findings such as infiltrative/microlobulated margins, calcifications, taller-than-wide appearance, and extrathyroidal extension were highly predictive of thyroid cancer, as has been reported (5, 6, 10, 11, 13). Although these features were quite specific for thyroid cancer, their utility was limited by a low number of observations and resultant poor sensitivity. Conversely, nodule size ≥1 cm, ≤25% cystic composition, nonisoechoic echogenicity, and increased Doppler flow were very sensitive for thyroid cancer but were also nonspecific.

Evaluation of combinations of features in a multivariable logistic regression model confirmed that all of the ultrasound features were important. However, we also found that the high degree of correlation between ultrasound features limited the predictive value of individual features. Combinations of ultrasound features, however, were quite useful in stratifying the odds of thyroid cancer in children. Thyroid nodule composition and echogenicity were closely associated with the presence of high-risk ultrasound features and were found to be favorable targets for risk stratification.

Thyroid nodule composition has been shown to be closely associated with risk of malignancy in adults (16, 17), but most studies in children have not explored this. Gupta *et al.* (5) reported an increased prevalence of thyroid cancer in thyroid nodules with entirely solid composition and <25% cystic composition (57% and 21%, respectively). In their study, ≤25% cystic composition identified 89.3% of thyroid cancers; this is in agreement with our findings. Thyroid nodule echogenicity has been closely associated with the risk of thyroid cancer in adults, but this pattern has not been consistently observed in pediatric studies. Lyshchik *et al.* (10) reported the prevalence of thyroid cancer to be 21% in children with isoechoic nodules, whereas the prevalence was 35% in children with nonisoechoic nodules. However, Mussa *et al.* (9) reported a 19% prevalence of thyroid cancer in nodules with isoechoic echogenicity, compared with 18.9% prevalence in nodules with nonisoechoic echogenicity. Our findings are in agreement with those of Lyshchik *et al.* (10), and it is unclear why the findings of Mussa *et al.* (9) differ, although we speculate that the lack of primary ultrasound image review in the study by Mussa *et al.* could limit the reliability of their findings.

Although we found that both thyroid nodule composition and nonisoechoic echogenicity were associated with thyroid cancer, we found that radiologist report of thyroid nodule composition was substantially more reliable than echogenicity; this finding was recently corroborated (12). Therefore, we conclude that thyroid nodule composition is the most important feature for initial risk stratification for a thyroid nodule in children.

The ATA guideline for evaluation of thyroid nodules in children recommends biopsy for thyroid nodules of any size with high-risk ultrasound features and in any solid or partially cystic thyroid nodule ≥1 cm in maximal dimension (1). This strategy is quite different from the ATA guideline for adults, which establishes size thresholds for biopsy based on the ultrasound features with an emphasis on three major components: the presence of high-risk ultrasound features (*e.g.*, irregular margins, microcalcifications, taller-than-wide shape, and evidence of extrathyroidal extension), thyroid nodule composition, and echogenicity (8). With this approach, nodules are assigned an estimated risk of malignancy, and those with lower risk would not undergo biopsy. Even nodules with high-risk features (>70% to 90% estimated risk of malignancy) would not undergo biopsy until the lesion was ≥1 cm in maximal dimension. Solid nodules with hypoechoic echogenicity (10% to 20% estimated risk of malignancy) would receive biopsy if they are ≥1 cm, whereas solid or partially cystic nodules of nonhypoechoic echogenicity (5% to 10% estimated risk of malignancy) would receive FNA biopsy if the nodule is ≥1.5 cm. Spongiform nodules and those with minimal solid content (<3% estimated risk of malignancy) could be monitored without biopsy. Comparison of our data with the ATA adult algorithm yields very different estimates of risk for thyroid cancer in children: 50% to 88% for high-suspicion patterns, 44.7% for solid or partially cystic hypoechoic nodules, 14% to 22% for isoechoic or hyperechoic nodules, and 14% for >25% cystic nodules. No ultrasound features alone or in combination were associated with a <10% risk for thyroid cancer in children.

This study was designed to evaluate the accuracy of thyroid ultrasound in discerning benign from malignant thyroid nodules in children. The relatively high prevalence of thyroid cancer that we identified in our study made this very difficult, even though the study was powered to withstand this. The low prevalence of low-risk features (*e.g.*, an entirely cystic structure) in our sample limit our ability to draw conclusions about their effectiveness in excluding thyroid cancer. Further prospective studies are needed to confirm the true risk associated with these findings.

Limitations

This study was limited by referral bias, which was minimized by our inclusive case-finding approach as described above. There was no significant difference between the prevalence of thyroid cancer between

children referred internally or externally ($P = 0.66$). We also performed sensitivity analyses for each statistical model, and there was no significant difference in findings based on the source of referral. Ascertainment bias also heavily affected our findings, resulting in a higher overall estimate of the prevalence of thyroid cancer, as discussed above. Additionally, the diagnosis of a benign thyroid nodule without surgical pathology is not conclusive, and it is possible that an indolent thyroid cancer may have been misclassified as benign.

Conclusion

Thyroid ultrasound is a powerful and important tool for the evaluation of thyroid nodules, but our analysis was unable to identify any combination of features associated with a "low suspicion" (<10% risk) of thyroid cancer in children. The lowest risk group for thyroid cancer appears to be children with thyroid nodules with >25% cystic composition, but even this group has a higher prevalence of malignancy than is observed in adults. This limitation is underscored by significant intra- and interobserver variability that can be expected even in expert hands. Current pediatric guidelines recommend biopsy of all solid or predominantly solid thyroid nodules ≥1 cm in diameter and nodules of any size with high-risk features; we were unable to identify any group for which reliance on ultrasound features alone was reliable to avoid biopsy.

Acknowledgments

Financial Support: This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases Grant T32DK063688-10 (to A.W.G.).

Correspondence and Reprint Requests: Anthony W. Gannon, MD, MSCE, Alfred I DuPont Hospital for Children, 1600 Rockland Road, Wilmington, Delaware 19801. E-mail: anthony.gannon@nemours.org.

Disclosure Summary: The authors have nothing to disclose.

References

1. Francis GL, Waguespack SG, Bauer AJ, Angelos P, Benvenga S, Cerutti JM, Dinauer CA, Hamilton J, Hay ID, Luster M, Parisi MT, Rachmiel M, Thompson GB, Yamashita S; American Thyroid Association Guidelines Task Force. Management guidelines for children with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2015;**25**:716–759.
2. Hogan AR, Zhuge Y, Perez EA, Koniaris LG, Lew JI, Sola JE. Pediatric thyroid carcinoma: incidence and outcomes in 1753 patients. *J Surg Res*. 2009;**156**(1):167–172.
3. Rivkees SA, Mazzaferri EL, Verburg FA, Reiners C, Luster M, Breuer CK, Dinauer CA, Udelsman R. The treatment of differentiated thyroid cancer in children: emphasis on surgical approach and radioactive iodine therapy. *Endocr Rev*. 2011;**32**(6):798–826.
4. Vergamini LB, Frazier AL, Abrantes FL, Ribeiro KB, Rodriguez-Galindo C. Increase in the incidence of differentiated thyroid carcinoma in children, adolescents, and young adults: a population-based study. *J Pediatr*. 2014;**164**(6):1481–1485.
5. Gupta A, Ly S, Castroneves LA, Frates MC, Benson CB, Feldman HA, Wassner AJ, Smith JR, Marqusee E, Alexander EK, Barletta J, Doubilet PM, Peters HE, Webb S, Modi BP, Paltiel HJ, Kozakewich H, Cibas ES, Moore FD Jr, Shamberger RC, Larsen PR, Huang SA. A standardized assessment of thyroid nodules in children confirms higher cancer prevalence than in adults. *J Clin Endocrinol Metab*. 2013;**98**(8):3238–3245.
6. Al Nofal A, Gionfriddo MR, Javed A, Haydour Q, Brito JP, Prokop LJ, Pittock ST, Murad MH. Accuracy of thyroid nodule sonography for the detection of thyroid cancer in children: systematic review and meta-analysis. *Clin Endocrinol (Oxf)*. 2016;**84**(3):423–430.
7. Cappelli C, Castellano M, Pirola I, Cumetti D, Agosti B, Gandossi E, Agabiti Rosei E. The predictive value of ultrasound findings in the management of thyroid nodules. *QJM*. 2007;**100**(1):29–35.
8. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 2015. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*. 2016;**26**(1):1–133.
9. Mussa A, De Andrea M, Motta M, Mormile A, Palestini N, Corrias A. Predictors of malignancy in children with thyroid nodules. *J Pediatr*. 2015;**167**:886–892.
10. Lyshchik A, Drozd V, Demidchik Y, Reiners C. Diagnosis of thyroid cancer in children: value of gray-scale and power doppler US. *Radiology*. 2005;**235**(2):604–613.
11. Corrias A, Mussa A, Baronio F, Arrigo T, Salerno M, Segni M, Vigone MC, Gastaldi R, Zirilli G, Tuli G, Beccaria L, Iughetti L, Einaudi S, Weber G, De Luca F, Cassio A; Study Group for Thyroid Diseases of Italian Society for Pediatric Endocrinology and Diabetology (SIEDP/ISPED). Diagnostic features of thyroid nodules in pediatrics. *Arch Pediatr Adolesc Med*. 2010;**164**(8):714–719.
12. Lim-Dunham JE, Erdem Toslak I, Alsabban K, Aziz A, Martin B, Okur G, Longo KC. Ultrasound risk stratification for malignancy using the 2015 American Thyroid Association Management Guidelines for Children with Thyroid Nodules and Differentiated Thyroid Cancer. *Pediatr Radiol*. 2017;**47**(4):429–436.
13. Koltin D, O'Gorman CS, Murphy A, Ngan B, Daneman A, Navarro OM, García C, Atenafu EG, Wasserman JD, Hamilton J, Rachmiel M. Pediatric thyroid nodules: ultrasonographic characteristics and inter-observer variability in prediction of malignancy. *J Pediatr Endocrinol Metab*. 2016;**29**(7):789–794.
14. Wienke JR, Chong WK, Fielding JR, Zou KH, Mittelstaedt CA. Sonographic features of benign thyroid nodules: interobserver reliability and overlap with malignancy. *J Ultrasound Med*. 2003;**22**(10):1027–1031.
15. Niedziela M. Pathogenesis, diagnosis and management of thyroid nodules in children. *Endocr Relat Cancer*. 2006;**13**(2):427–453.
16. Moon WJ, Jung SL, Lee JH, Na DG, Baek JH, Lee YH, Kim J, Kim HS, Byun JS, Lee DH; Thyroid Study Group, Korean Society of Neuro- and Head and Neck Radiology. Benign and malignant thyroid nodules: US differentiation–multicenter retrospective study. *Radiology*. 2008;**247**(3):762–770.
17. Frates MC, Benson CB, Doubilet PM, Kunreuther E, Contreras M, Cibas ES, Orcutt J, Moore FD Jr, Larsen PR, Marqusee E, Alexander EK. Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. *J Clin Endocrinol Metab*. 2006;**91**(9):3411–3417.