Published in final edited form as:

J Pediatr Endocrinol Metab. 2015 March ; 28(3-4): 387–392. doi:10.1515/jpem-2014-0135.

Evaluation of puberty by verifying spontaneous and stimulated gonadotropin values in girls

Vivian L. Chin

Department of Pediatrics, Children's Hospital at Montefiore, Bronx, NY, USA

Ziyong Cai

American International Group, NY, USA

Leslie Lam

Division of Endocrinology, Department of Pediatrics, Children's Hospital at Montefiore, Bronx, NY, USA

Bina Shah

Endocrinology Division, Department of Pediatrics, NYU Langone Medical Center, New York, NY, USA

Ping Zhou*

Division of Endocrinology, Department of Pediatrics, Children's Hospital at Montefiore, Bronx, NY, USA

Abstract

Background—Changes in pharmacological agents and advancements in laboratory assays have changed the gonadotropin-releasing hormone analog stimulation test.

Objective—To determine the best predictive model for detecting puberty in girls.

Subjects—Thirty-five girls, aged 2 years 7 months to 9 years 3 months, with central precocious puberty (CPP) (n = 20) or premature thelarche/premature adrenarche (n = 15).

Methods—Diagnoses were based on clinical information, baseline hormones, bone age, and pelvic sonogram. Gonadotropins and E2 were analyzed using immunochemiluminometric assay. Logistic regression for CPP was performed.

Results—The best predictor of CPP is the E2-change model based on 3- to 24-h values, providing 80% sensitivity and 87% specificity. Three-hour luteinizing hormone (LH) provided 75% sensitivity and 87% specificity. Basal LH lowered sensitivity to 65% and specificity to 53%.

Conclusions—The E2-change model provided the best predictive power; however, 3-h LH was more practical and convenient when evaluating puberty in girls.

*Corresponding author: Ping Zhou, MD, Division of Pediatric Endocrinology, Children's Hospital at Montefiore, Albert Einstein College of Medicine, 111 East 210th Street, Bronx, NY 10467, USA, Fax: +1 7184055609, pczhou@montefiore.org.

Authors' conflict of interest disclosure: The authors have no disclosures to report.

Keywords

central precocious puberty; girls; GnRHa stimulation testing; leuprolide stimulation test

Introduction

Traditionally, gonadotropin-releasing hormone (GnRH) stimulation test is considered the gold standard for the evaluation of puberty. Determination of central precocious puberty (CPP) is mostly clinical; however, the diagnosis can be verified with basal and stimulated serum gonadotropin values demonstrating elevated levels of luteinizing hormone (LH), follicle-stimulating hormone (FSH) and estradiol (E2) in girls.

Much has changed in recent years. First, laboratory assay improvements affected the cutoff values for LH, FSH, and E2 indicative of puberty. The sensitivity of the assays improved from the first-generation radioimmunoassays (RIAs), to the second-generation immunoradiometric assays (IRMAs), and to the most sensitive third-generation immunofluorometric assays (IFMAs) or immunochemiluminometric assays (ICMAs) (1). Second, the medication used in stimulation testing has changed. Factrel (Wyeth Ayerst Pharmaceuticals Collegeville, PA, USA), the medication previously used for GnRH stimulation tests, is no longer available in the United States. Most stimulation tests are now obtained with leuprolide acetate, an aqueous form of the GnRH analog (GnRHa), which is a synthetic nonapeptide with much greater potency. Due to the pharmacodynamic differences between these two medications, the timing and peak values of FSH and LH levels are different. After Factrel administration, LH levels peak after 20–40 min followed by a decline in values compared with leuprolide acetate, which can similarly stimulate LH to peak as early as 30 min to 4 h followed by sustained LH elevation (2–5). With these changes, the threshold values of GnRHa stimulation testing should be reevaluated.

The purpose of this study was to search for the best model predictive of puberty and to determine the accuracy of threshold values of LH, FSH, and E2.

Subjects and methods

This is a retrospective study approved by the institutional review board at Montefiore Medical Center. Inclusion criteria are girls who had been evaluated in Montefiore's pediatric endocrinology clinic during the period from 2006 to 2010 for precocious puberty defined by breast and/or pubic hair development before 7 years for White girls and before 6 years for Black girls (6). The definition is based on a 1999 recommendation by the Drug and Therapeutics and Executive Committees of the Lawson Wilkins Pediatric Endocrine Society (6). All the subjects had GnRHa testing done between 8:00 am and 10:00 am, using leuprolide acetate subcutaneous (20 µg/kg). Serum LH, FSH, and E2 levels at baseline, 3 h, and 24 h were measured based on a previously devised protocol (2). Patients who had been diagnosed with peripheral precocious puberty or who had been treated with any medications for pubertal problems were excluded from this study.

Serum LH, FSH, and E2 were measured by Montefiore's medical laboratory using the IMMULITE 2000 System by Siemens Medical Solutions Diagnostics. All the samples were measured in duplicate. LH and FSH were determined with a solid-phase, two-site chemiluminescent immunometric assay. The analytical sensitivity for LH and FSH was 0.05 and 0.1 IU/L (or mIU/mL), respectively. E2 was determined with a solid-phase, competitive chemiluminescent enzyme immunoassay with a sensitivity of 55 pmol/L (15 pg/mL).

With the use of the aforementioned inclusion and exclusion criteria, a total of 35 females, ages 2 years 7 months to 9 years 3 months, of mainly African-American or Hispanic race, were included in the study. Data, taken at the initial visit at the approximate time of the GnRHa stimulation test and at follow-up visits 6–12 months later, were retrospectively collected for this study. These data fields included the following: Tanner stage of breasts and pubic hair, height, weight, body mass index (BMI), growth velocity, skeletal age, pelvic sonogram, brain magnetic resonance imaging (MRI), and serum levels of LH, FSH, and E2 at baseline, 3 h, and 24 h post leuprolide stimulation.

Based on the aforementioned clinical data except the stimulated levels of LH, FSH, and E2, clinical diagnoses were retrospectively determined by the consensus of the majority of three pediatric endocrinologists independently. From these clinical diagnoses, these 35 patients were divided into two groups: group 1 for subjects with CPP and group 2 for subjects without CPP, with premature thelarche (PT) or premature adrenarche (PA).

Data analysis

Means comparisons were performed using Student's t-tests with significance set at $\alpha = 0.05$. Logistic regression modeling using the diagnosis of CPP as the dependent variable and LH, FSH, and E2 at various time points as independent variables was carried out to select the best predictive model for CPP. Data analysis was performed using SAS 9.1 statistical software (SAS Institute, Inc, Cary, NC, USA).

Results

Baseline clinical characteristics are summarized in Table 1. When compared to PT/PA girls, girls with CPP were older, with mean age of 7 years 8 months vs. 5 years 11 months ($p = 0.002$); more skeletally advanced from chronological age by 29.9 months vs. 17.4 months; and had larger ovarian volumes on the right (2 vs. 0.8 cm³, $p = 0.02$) and the left (2.4 vs. 0.9 cm³, $p = 0.03$). Uterine volumes and endometrial stripe size were not consistently reported for all girls and were excluded from statistical analysis. Clinical advancement in Tanner breast staging occurred in nine of 20 girls in the CPP group vs. only four of 15 in the PT/PA group. Five girls in the CPP group had microadenomas incidentally noted on brain MRI, while none of the girls in the PT/PA group had any evidence of brain abnormality. Other parameters such as BMI z-score ($p = 0.12$), growth velocity ($p = 0.74$), height percentile ($p = 0.69$), and target height percentile ($p = 0.37$) were not statistically different between the two groups.

Data from the GnRHa stimulation testing for these two groups are summarized in Table 2 and Figures 1–3.

With the use of logistic regression modeling, the best univariate model uses as its predictor the percentage change in estradiol level at the 24th hour relative to the level at the 3rd hour:

$$\ln\left(\frac{P}{1-P}\right) = a + b \cdot x,$$

where $a = 1.7209$, $b = 6.1808$, and $x = \frac{\text{EST24H} - \text{EST3H}}{\text{EST3H}}$.

This model has a sensitivity of 80% and a specificity of 87%, with a threshold of 27.8% change.

A second model using only one value, the 3-h LH value, can be specified as

$$\ln\left(\frac{P}{1-P}\right) = a + b \cdot x,$$

where $a = -3.4699$, $b = 0.6852$, and $x = \text{LH3H}$.

Compared to the previous model, this model has a lower sensitivity of 75% and the same specificity of 87% using a threshold of 5.061 IU/L at 3 h.

Alternatively, using the basal LH in a single predictor model lowers the sensitivity to 65% and the specificity to 53% using a threshold of 0.132 IU/L at baseline, with the following parameters: $a = -1.0266$, $b = 7.771$, and $x = \text{LH0H}$.

Discussion

Consistent with previously observed clinical evidence, our data on both basal and stimulated FSH levels showed a substantial overlap between the two groups, indicating that FSH is poor at predicting CPP (7).

Serum LH level has been suggested as the most valuable biochemical marker for predicting CPP. The basal cutoff level ranged from 0.1 to 0.6 IU/L, depending on the assay used. In the Neely et al. (8) study, using the most sensitive assay, ICMA, a level above 0.1 IU/L predicts true puberty, with 94% sensitivity and 88% specificity (9). But some studies suggest a sexual difference when using basal LH level to predict puberty. In another study by Resende et al. (7), a basal serum LH level < 0.2 IU/L by ICMA had a 100% sensitivity and specificity in distinguishing prepubertal from pubertal males. But the test only placed 50% of the girls with Tanner stage II breasts in the prepubertal range (7). Our data showed similar overlapping results for girls.

Most studies confirmed that peak LH level after stimulation with either GnRH or GnRHa has a better predictive power for detecting puberty over basal values (2, 7, 10–13). The old GnRH test with Factrel by IFMA set LH levels above 9.6 IU/L in boys and 6.9 IU/L in girls as positive pubertal responses (14). Previous studies have suggested that the leuprolide test is more potent. Based on different laboratory assays, the threshold value of LH varies

between 5.0 and 8.0 IU/L (15). The results of our study validate this, with a cutoff of 5.1 IU/L.

Only stimulated ovarian estradiol response to leuprolide was studied, possibly due to its potency and longer half-life than Factrel. Garibaldi et al. found that a GnRHa-stimulated serum estradiol level at 24 h may be the most useful indicator of early CPP (11). With RIA, a mean E2 level of 136 ± 11 pmol/L (range 92–176 pmol/L) indicates early puberty.

Generally, the basal E2 level used by most commercial laboratories is not a sensitive test, although some studies suggest that estradiol levels by tandem mass spectrometry allow for the detection of early puberty (7). Estradiol levels lower than 1 pg/mL (3.7 pmol/L) usually indicate the prepubertal stage (15, 16). However, this assay is not routinely used and is limited by cost and availability. The less expensive assay (ICMA) that was used in this study has a limited sensitivity of 15 pg/mL (55 pmol/L). Measurable estradiol in our assay only confirms relatively advanced puberty.

In our study, the best model for detecting CPP that we identified uses the percentage change in estradiol level at the 24th hour relative to the 3rd hour. This model provides 80% sensitivity and 87% specificity using 27.8% change as the threshold. Although this is the best performing model, it uses two test results, each from a separate day per patient; thus it is time consuming and impractical clinically. Alternatively, a model using a single value of LH at the 3rd hour is more practical and provides 75% sensitivity and 87% specificity with its threshold set at 5.1 IU/L. Lastly, the least sensitive model based on a single basal LH level was 65% sensitive and 53% specific using a threshold of 0.132 IU/L.

Our study has some limitations. One of them is that the diagnosis of CPP was made retrospectively by a majority of three pediatric endocrinologists' independent evaluations. The diagnosis of CPP was based mainly on the clinical signs of pubertal progression. There were three girls in the CPP group who had Tanner stage I at their initial visits, but all had clinically apparent rapid advancement of puberty within 6–12 months, with breast advancing to Tanner stage III, and even one who had menarche. In contrast, a few subjects who had advanced initial Tanner stage for breast were placed in the non-CPP group because their pubertal development showed no significant advancement during follow-up. Benign conditions such as premature thelarche with early presence of breast tissue or premature adrenarche with activation of the adrenal axis do not indicate any central activation of the hypothalamic-pituitary gonadal axis.

Additional limitations include overestimation of the accuracy measurements due to our small sample size of only 35 subjects. Furthermore, the E2 assay used was not the most sensitive assay available especially in the early pubertal ranges where there is a significant overlap with prepubertal values. Even so, E2 was still able to provide the best predictive model for the prediction of puberty using the percent change between the 3rd and the 24th hour in our study.

Conclusion

Basal gonadotropin levels are poor at predicting CPP. The percentage change in estradiol level at the 24th hour relative to the level at the 3rd hour in the GnRHa stimulation test is the most sensitive predictor. However, a 3rd-hour LH level provides the best clinical model for its practicality and convenience when evaluating puberty in girls.

References

1. Lee PA. Laboratory monitoring of children with precocious puberty. *Arch Pediatr Adolesc Med.* 1994; 148:369–76. [PubMed: 8148936]
2. Ibanez L, Potau N, Zampolli M, Virdis R, Gussinye M, et al. 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; 78:30–5. [PubMed: 7507123]
3. Chi CH, Durham E, Neely EK. Pharmacodynamics of aqueous leuprolide acetate stimulation testing in girls: correlation between clinical diagnosis and time of peak luteinizing hormone level. *J Pediatr.* 2012; 161:757–9.e1. [PubMed: 22809662]
4. Elsholz DD, Padmanabhan V, Rosenfield RL, Olton PR, Phillips DJ, et al. GnRH agonist stimulation of the pituitary-gonadal axis in children: age and sex differences in circulating inhibin-B and activin-A. *Hum Reprod.* 2004; 19:2748–58. [PubMed: 15513981]
5. Houk CP, Kunselman AR, Lee PA. The diagnostic value of a brief GnRH analogue stimulation test in girls with central precocious puberty: a single 30-minute post-stimulation LH sample is adequate. *J Pediatr Endocrinol Metab.* 2008; 21:1113–8. [PubMed: 19189683]
6. Kaplowitz PB, Oberfield SE. Reexamination of the age limit for defining when puberty is precocious in girls in the United States: implications for evaluation and treatment. Drug and Therapeutics and Executive Committees of the Lawson Wilkins Pediatric Endocrine Society. *Pediatrics.* 1999; 104:936–41. [PubMed: 10506238]
7. Resende EA, Lara BH, Reis JD, Ferreira BP, Pereira GA, et al. Assessment of basal and gonadotropin-releasing hormone-stimulated gonadotropins by immunochemiluminometric and immunofluorometric assays in normal children. *J Clin Endocrinol Metab.* 2007; 92:1424–9. [PubMed: 17284632]
8. Neely EK, Wilson DM, Lee PA, Stene M, Hintz RL. Spontaneous serum gonadotropin concentrations in the evaluation of precocious puberty. *J Pediatr.* 1995; 127:47–52. [PubMed: 7608810]
9. Houk CP, Kunselman AR, Lee PA. Adequacy of a single unstimulated luteinizing hormone level to diagnose central precocious puberty in girls. *Pediatrics.* 2009; 123:e1059–63. [PubMed: 19482738]
10. Eckert KL, Wilson DM, Bachrach LK, Anhalt H, Habiby RL, et al. A single-sample, subcutaneous gonadotropin-releasing hormone test for central precocious puberty. *Pediatrics.* 1996; 97:517–9. [PubMed: 8632938]
11. 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; 76:851–6. [PubMed: 8473395]
12. Roger, M.; Lahlou, N.; Chaussain, JL. Gonadotropin-releasing hormone testing in pediatrics. In: Ranke, MB., editor. *Diagnostics of endocrine function in children and adolescents.* Johann Ambrosius Barth; Heidelberg, Germany: 1996.
13. Sathasivam A, Garibaldi L, Shapiro S, Godbold J, Rapaport R. Leuprolide stimulation testing for the evaluation of early female sexual maturation. *Clin Endocrinol (Oxf).* 2010; 73:375–81. [PubMed: 20184599]
14. Brito VN, Batista MC, Borges MF, Latronico AC, Kohek MB, et al. Diagnostic value of fluorometric assays in the evaluation of precocious puberty. *J Clin Endocrinol Metab.* 1999; 84:3539–44. [PubMed: 10522992]

15. Carel JC, Eugster EA, Rogol A, Ghizzoni L, Group E-LGACC. et al. Consensus statement on the use of gonadotropin-releasing hormone analogs in children. *Pediatrics*. 2009; 123:e752–62. [PubMed: 19332438]
16. Fiers T, Casetta B, Bernaert B, Vandersypt E, Debock M, et al. Development of a highly sensitive method for the quantification of estrone and estradiol in serum by liquid chromatography tandem mass spectrometry without derivatization. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2012; 893–894:57–62.

Figure 1.
Graph of basal and stimulated LH values, expressed as mean $\pm$ 1 SE.

Figure 2.
Graph of basal and stimulated FSH values, expressed as mean $\pm$ 1 SE.

Figure 3.
Graph of basal and stimulated E2 values, expressed as mean $\pm$ 1 SE.

Table 1

Clinical characteristics and imaging results of CPP and PT/PA girls.

	CPP		Non-CPP		p-Value
	Mean (SE)	Range	Mean (SE)	Range	
n	20		15		
Age	7 years 8 months (2 months)	(6 years 4 months to 8 years 11 months)	5 years 11 months (6 months)	(2 years 7 months to 9 years 3 months)	0.002
Advanced BA, n	19		7		
BA adv., months	29.9 (2.4)	(8 to 48)	17.4 (5.6)	(-1 to 55)	0.02
Tanner I breast, n	3		3		
Tanner II breast, n	10		9		
Tanner III breast, n	5		2		
Tanner IV breast, n	1		0		
Breast not staged, n	1		1		
Tanner I PH, n	5		8		
Tanner II PH, n	7		4		
Tanner III PH, n	6		2		
Tanner IV PH, n	1		1		
PH not staged, n	1		0		
BMI (z-score)	1.53 (0.2)	(-0.9 to 2.8)	1 (0.3)	(-0.8 to 2.1)	0.12
Growth velocity, cm/year	6.3 (0.6)	(2.1 to 11.2)	6.7 (1.0)	(0 to 10.3)	0.74
Height, percentile	77.6 (5.7)	(17.5 to 95)	74.3 (6.5)	(37.5 to 95)	0.69
Target height, percentile	54.8 (6.7)	(5 to 95)	43.2 (9.4)	(7.5 to 95)	0.37
Brain MRI	5 with pituitary microadenoma		0 with pituitary microadenoma		
Right ovarian volume, cm ³	2 (0.4)	(0.5 to 6.8)	0.8 (0.2)	(0.1 to 1.8)	0.02
Left ovarian volume, cm ³	2.4 (0.5)	(0.07 to 8.9)	0.9 (0.1)	(0.5 to 1.3)	0.03

p-Values expressed are results of t-tests (PH, pubic hair). Advanced BA defined by number with BA advanced more than 1 year; BA adv. defined by BA advancement in months; BA, bone age.

Table 2

Biochemical data on LH, FSH, and E2 levels during GnRHa test.

	CPP		PT/PA	
	Mean±SE	(Min-max)	Mean±SE	(Min-max)
LH, mIU/mL				
0 h	0.32±0.08	(<0.1–1.67)	0.14±0.01	(<0.1–0.25)
3 h	9.06±1.41	(3.03–29)	3.92±0.44	(1.15–7.9)
24 h	2.95±0.95	(0.59–19.9)	0.67±0.06	(0.39–0.98)
FSH, mIU/mL				
0 h	2.80±0.33	(0.91–6.65)	2.08±0.27	(0.69–3.86)
3 h	21.21±1.54	(7.5–36.5)	32.16±5.6	(13–96.7)
24 h	12.61±1.54	(7.14–36.1)	8.02±1.04	(4.4–15.6)
E2, pg/mL				
0 h	27.45±3.29	(<2–60)	29.36±3.7	(<20–57)
3 h	35.5±2.97	(<20–63)	36.33±3.4	(<20–56)
24 h	128.4±34	(40–720)	51.2±7.7	(<20–110)