

Screening for Congenital Adrenal Hyperplasia: Adjustment of 17-Hydroxyprogesterone Cut-Off Values to Both Age and Birth Weight Markedly Improves the Predictive Value

BERNHARD OLGEMÖLLER, ADELBERT A. ROSCHER, BERNHARD LIEBL, AND RALPH FINGERHUT

Labor Becker, Olgemöller and Colleagues (B.O., R.F.), D-81671 Munich, Germany; Department of Clinical Chemistry and Biochemical Genetics (A.A.R.), Dr. von Hauner Children's Hospital, Ludwig-Maximilians-University, D-80337 Munich, Germany; and Public Health Newborn Screening Center of the State of Bavaria (B.L.), D-85764 Oberschleißheim, Germany

Newborn screening procedures for congenital adrenal hyperplasia (CAH) are still suboptimal because of low specificity, particularly in premature infants. This study evaluated a multitiered strategy for reporting abnormal 17-hydroxyprogesterone screening values that simultaneously takes into account not only birth weight but also age at sampling. A cautious three-tiered cut-off scheme was used during the initial 24 months of CAH screening in Bavaria. Data were then reanalyzed using five birth weight classes to reflect more precisely the markedly higher values in low-birth-weight newborns. Because 17-hydroxyprogesterone values apparently decline with increasing age, these classes were then further

subdivided into a total of 21 groups according to the age at sampling. Based on this reanalysis, we defined new multitiered cut-off levels and used them for the next 18 months. A total of 538,466 newborns were screened from January 1999 to June 2002; 51 CAH cases were detected. Application of the new threshold values resulted in a 35% reduction of the total recall rate (from 1.13% to 0.74%) and an increase in the positive predictive value from 0.84% to 1.29% without reducing diagnostic sensitivity. This improvement of CAH screening can be achieved by simply using request forms that ask for both age and birth weight at the time of sampling. (*J Clin Endocrinol Metab* 88: 5790–5794, 2003)

CONGENITAL ADRENAL HYPERPLASIA (CAH) is a group of disorders characterized by deficient cortisol biosynthesis; in most cases steroid-21-hydroxylase deficiency is the underlying cause of CAH (1–4). The estimated worldwide incidence of classic CAH is around 1:14,000 (5). The spectrum of clinical presentations ranges from forms with neonatal symptoms, *i.e.* salt wasting (SW) and simple virilizing (SV) forms, to nonclassical forms that might not manifest until adulthood (6, 7).

To detect boys at risk of salt-losing crisis and prevent incorrect gender assignment in females, many neonatal screening programs for CAH have been implemented worldwide. Such screening is more likely to detect CAH in neonates than diagnosis by clinical signs alone (7–11).

CAH screening, however, is afflicted with significant problems. Diagnostic sensitivity for SV-CAH is not optimal when screening, because of cost constraints, is restricted to a single sample taken shortly after birth. Although taking a second sample would solve this problem, it would also lead to increased detection of nonclassical-CAH, which is not the primary goal of CAH screening (12).

Other problems arise from low specificity of 17 α -hydroxyprogesterone (17-OHP) screening methods, which are all based on immunological methods. Cross-reactivity, especially with steroid sulfates (13) and stress caused by illness contribute to a high recall rate, especially in premature

infants (8, 9, 14, 15). The possibility of mildly elevated newborn 17-OHP concentrations because of heterozygosity for CYP21 mutations might potentially add to this problem.

It has been shown that adjusting the cut-off levels to either gestational age (15) or birth weight (BW) (14) can lower the recall rate significantly. However, neither of these schemes takes into account that physiological 17-OHP values also depend heavily on age at sampling (16).

This variability is especially important when 17-OHP screening values from early sampling ages (d 3 or even earlier) need to be interpreted. Such unfavorable sampling times are becoming increasingly common because of early discharge policies and the widespread introduction of expanded newborn screening for metabolic disorders by tandem mass spectrometry (17). To improve efficiency of CAH screening, we evaluated a multitiered strategy for defining abnormal screening values that accounts for both BW and age at sampling.

Subjects and Methods

Subjects

Beginning in January 1999, 17-OHP was routinely measured in all newborns as part of a comprehensive new screening program in Bavaria that uses tandem mass spectrometry to screen for metabolic disorders. The recommended age for sampling was set to the third day of life to ensure timely therapeutic intervention. Over 42 months, 538,466 samples were analyzed and the documented participation rate was 98.8% (18, 19).

Methods

Whole blood was drawn by heel prick or venous puncture and dried on filter paper (S&S 2992, Schleicher & Schüll, Dussel, Germany) and

Abbreviations: BW, Birth weight; CAH, congenital adrenal hyperplasia; 17-OHP, 17 α -hydroxyprogesterone; ppv, positive predictive value; SV, simple virilizing; SW, salt wasting.

TABLE 1. 17-OHP threshold values (nmol/liter) before and after adjustment for both age and BW^a

	BW (g)	Age (d)	17-OHP values in nmol/liter		
			Normal	Elevated CAH possible	Markedly elevated CAH probable
Initial 3 cut-off values	<2000		<90	90–125	>125
	≥2000	0–3	<40	40–90	>90
Redefined cut-off values, adjusted for age and BW	<1000	≥4	<30	30–90	>90
		0–19	<200	200–300	>300
	1000–1500	20–29	<100	100–200	>200
		30–59	<60	60–150	>150
		≥60	<30	30–90	>90
		0–3	<150	150–200	>200
	1500–2000	4–13	<120	120–200	>200
		14–19	<80	80–200	>200
		20–29	<60	60–200	>200
		30–59	<40	40–125	>125
	2000–2500	≥60	<30	30–90	>90
		0–3	<80	80–150	>150
		4–13	<60	60–150	>150
		14–29	<40	40–150	>150
	>2500	≥30	<30	30–90	>90
		0–1	<60	≥60	^b
2–3		<50	50–125	>125	
4–13		<40	40–125	>125	
>2500	≥14	<30	30–90	>90	
	0–1	<60	≥60	^b	
	2–3	<40	40–90	>90	
		≥4	<30	30–90	>90

^a Conversion factor, 1 nmol/liter = 0.33 μg/liter.

^b For the first day of life, 17-OHP concentrations are so variable that a subdivision between possible and probable CAH is not possible.

17-OHP was then measured from punched 3.2-mm circles by time-resolved immunofluorescence with the AutoDelfia Neonatal 17-OHP test kit and the model 1235 automatic immunoassay system (Wallac, Turku, Finland).

Intraassay variation, established at two different concentrations [6.93 μg/liter (21 nmol/liter) and 19.47 μg/liter (59 nmol/liter)] was 8.5 and 5.3%, respectively; interassay variation was 13.3 and 10.5%, respectively (n = 20).

Initial cut-off values, established in an earlier pilot study, were set at 29.7 μg/liter (90 nmol/liter) for preterm babies (BW < 2000 g) and 13.2 μg/liter (40 nmol/liter) for term babies (BW > 2000 g when sampled during the first through fourth day of life). For term babies sampled on or after the fifth day of life, 9.9 μg/liter (30 nmol/liter) was used as the cut-off.

Abnormal 17-OHP values were internally classified as “elevated” or “markedly elevated” (Table 1). If a newborn had an elevated 17-OHP value, we made a low-urgency request for a second sample. If the 17-OHP value was markedly elevated, we issued a high-urgency request for a second sample and also immediately referred the child with the abnormal value to a pediatric endocrinologist.

By definition every result in a screening procedure that is above the applicable cut-off value is considered a recall. Therefore, in this study the recall rate represents the percentage of newborns in which a second test card was required.

Other programs have used less stringent definitions for a recall, specifically for the purpose of CAH screening (20). Our reported recall rates and positive predictive values can thus not be directly compared with such studies (21, 22).

Results

A cautious three-tiered cut-off scheme was used during the initial 24 months of CAH screening. Data from this period were then reanalyzed to subdivide 17-OHP concentrations according to both sampling age and BW. These reanalyzed data provided new multitiered cut-off levels that were used for the next 18 months. Initial and adjusted cut-off values are summarized in Table 1.

After 42 months, the entire data set was reanalyzed to

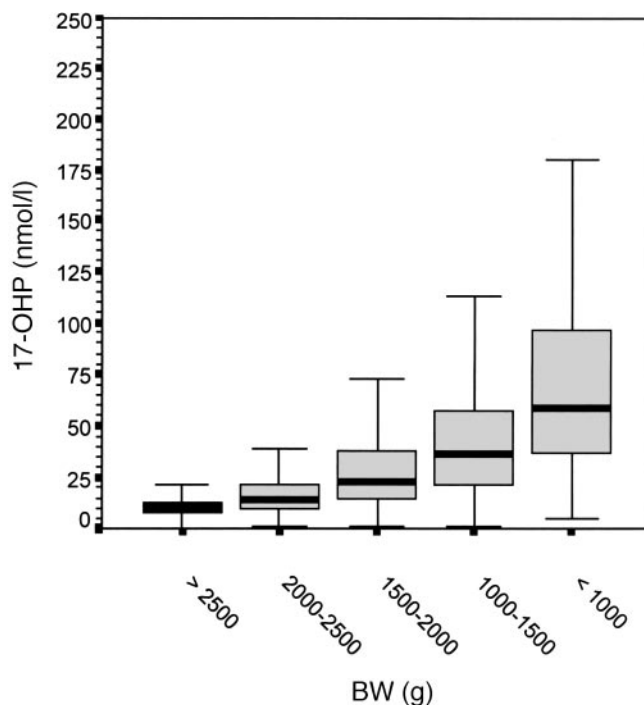


FIG. 1. 17-OHP values for babies aged 2–3 d classified in five birth weight groups. The boxes represent the interquartile range (25th–75th centiles); the horizontal black bars represent the median; the error bars represent the 95% confidence intervals of the mean (conversion factor: 1 nmol/liter = 0.33 μg/liter).

compare the two strategies. As expected median 17-OHP values were inversely related to BW (Fig. 1).

For the numerous babies with birth weight greater than

2000 g, the new cut-offs reflect the distribution of the 17-OHP values rather precisely. In low BW (<2000 g) babies, which represent 2.83% of our newborn population, 17-OHP values were markedly higher and also showed a high variance, presumably because of the high incidence of stressful illness.

In all BW classes, values declined with increasing sampling age, with the time course being faster in mature babies than premature ones (Fig. 2). Particularly in premature babies less than 1000 g, the onset of the decline was much later.

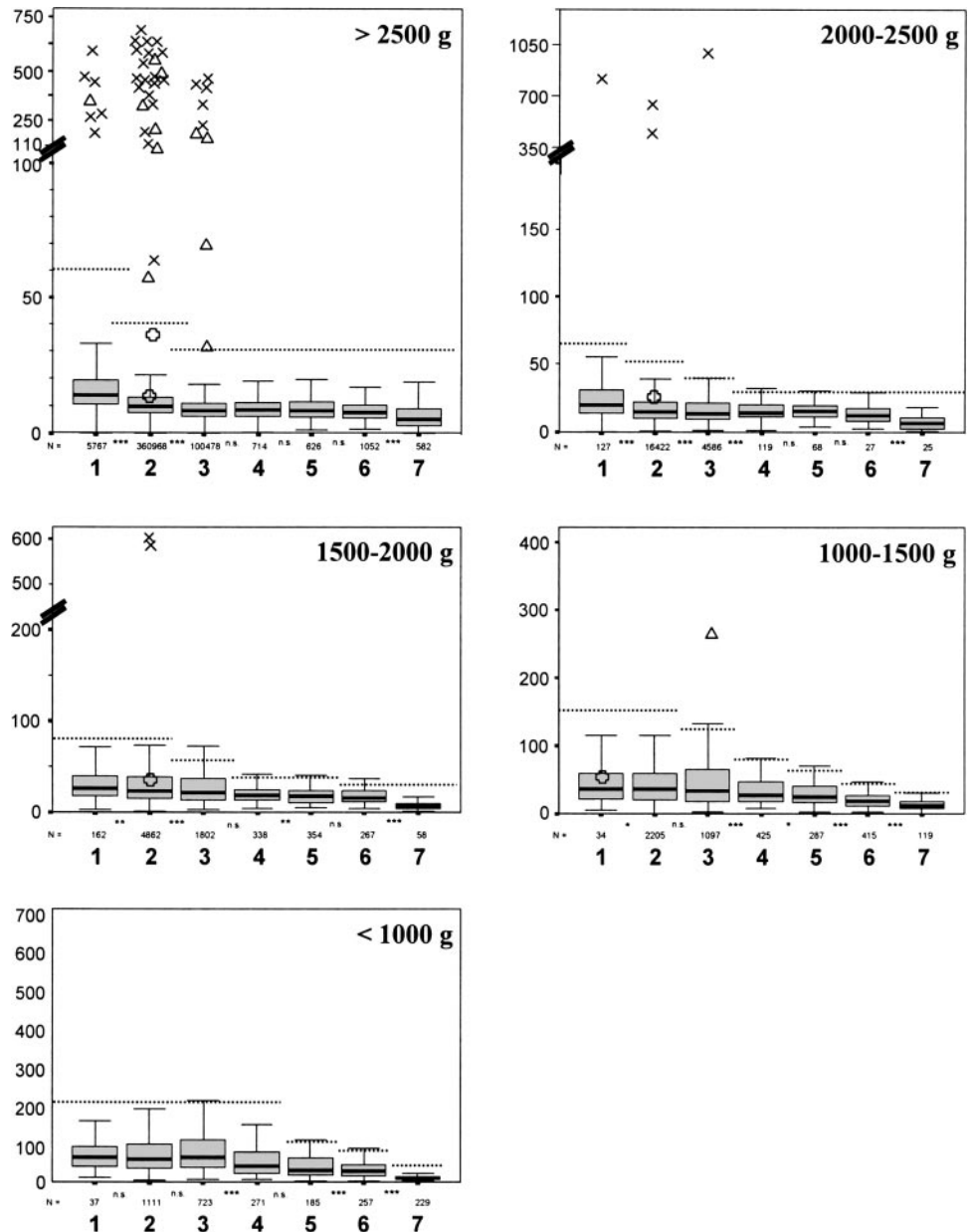
Using previously published criteria (23), we classified CAH into SW- and SV-CAH based on the referring physicians' written reports of the confirmatory procedures. We found 39 cases of SW-CAH among a total of 51 CAH subjects of 538,466 newborns. Nine of the 51 subjects were designated SV-CAH and one subjects was a putative nonclassical-CAH.

Thus, screening detected classic CAH with an incidence of 1:11,218 (95% confidence interval 1:8461 to 1:15,215). The SW-CAH incidence was 1:13,806 (95% confidence interval 1:10,100 to 1:19,416) with a sex ratio of 1.17 (21 males/18 females).

One baby showed elevated 17-OHP levels on three consecutive samples [84.15, 197.67, and 189.42 $\mu\text{g}/\text{liter}$ (255, 599 and 574 nmol/liter, respectively)]. This baby was later determined to be 3- β -hydroxysteroid dehydrogenase deficient on the basis of urinary metabolite analysis, which showed very high concentrations of 3- β -hydroxy-5-ene steroids (24). The high 17-OHP levels we observed most likely were due to the conversion of steroid precursors by extraadrenal enzyme activity (25).

In six subjects we did not receive sufficient information on confirmatory tests, but SW-CAH could probably be excluded

FIG. 2. Mean 17-OHP values (nmol/liter) for five BW (b.w. in figure) groups and seven age groups: 1) 0-1 d; 2) 2-3 d; 3) 4-13 d; 4) 14-19 d; 5) 20-29 d; 6) 30-59 d; and 7) 60 d or more. The boxes represent the interquartile range (25th-75th centile); the horizontal black bars represent the median; the error bars represent the 95% confidence interval of the mean; dotted line represent cut-off values. 17-OHP values of patients with CAH are depicted by these symbols: x, true positive (SW); \oplus , false negative, Δ true positive (other). Levels of significance relative to next age group are indicated: n.s., $P > 0.05$; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$ (conversion factor: 1 nmol/liter = 0.33 $\mu\text{g}/\text{liter}$).



on the basis of initial and follow-up values of 17-OHP, initial information on the clinical state of the patient, or molecular testing.

False negative CAH screening was documented in two subjects who had initial normal 17-OHP values; in those cases, either a second sample (taken accidentally) or appearance of clinical signs eventually led to a diagnosis of SV-CAH. In three other instances we encountered communication errors in which treatment with dexamethasone resulted in a normal screening value.

Table 2 shows the practical impact of using multitiered threshold values that are adjusted for not only BW but also age at sampling. If these new cut-off values were used, a total of 3961 samples would require follow-up; these included 3560 elevated and 401 markedly elevated 17-OHP values. If the initial three tiers of cut-off values were used, a much higher total of 6071 samples (4868 elevated and 1203 markedly elevated) would require follow-up. The recall and false positive rate for the three-tiered cut-off scheme would be particularly high (0.68%) for babies screened on or before 3 d of age.

Application of the new multitiered cut-off values dropped the total recall rate from 1.12% to 0.73%. Total recall for babies more than 2000 g was 0.316% when samples were taken at the age of 2 or 3 d. Overall specificity increased from 98.9% to 99.3% and the total positive predictive value (ppv) from 0.84% to 1.29%. The benefit was most evident in the group of high-urgency recalls (markedly elevated 17-OHP values) as shown by an increase in ppv from 3.74% to 9.23%.

Discussion

The economic and psychosocial costs associated with false positive CAH screening results are high. The pure laboratory costs of a false positive screening result (duplicates of internal repeats and of follow-up samples, separate sample processing) are about 10 times the cost of a normal screening sample. If the costs for (unnecessary) clinical follow-up are added, the figure becomes significantly higher.

Therefore, schemes that account for the marked physiological variance in newborn 17-OHP values have been in-

troduced and include adjusting 17-OHP cut-off values for either BW (14) or gestational age (15).

However, no study had so far systematically evaluated whether adjusting 17-OHP cut-off values for sampling age might further reduce recall rates, even though there were indications that it might. For instance, 17-OHP cut-off values have almost all been based on a fixed, recommended sample collection time; in practice, however, samples are often collected before the recommended time. In addition, age-adjusted 17-OHP reference ranges for premature babies, which frequently show elevated values, have not been available.

We have shown here that concurrent adjustment of cut-off values for both BW and sampling age leads to a major improvement of CAH screening specificity without any loss of sensitivity.

With the cut-off values we initially used, our false positive rate was 1.12%, which agrees well with Torresani *et al.* (15), who reported a 1.6% false positive rate when low-urgency controls in premature babies are included in the calculation. With the new, more stringent cut-offs, we were able to significantly reduce the false positive rate to 0.73% and increase the ppv for babies with BW below or above 2000 g.

The overall statistics from our screening program, when evaluated after 42 months, clearly demonstrated that all confirmed cases of CAH detected under the initial, less stringent guidelines would have also been identified with the new, much more stringent guidelines.

The total incidence of classic CAH (SW and SV) detected through newborn screening in Bavaria is 1:11,218. This figure is in agreement with old data from Bavaria derived from clinically detected cases (26). We suspect, however, that the share of SV-CAH not recognized by our early screening should be somewhat higher than the few cases that came to our attention (as false negative results).

In summary, this study clearly shows that efficiency of CAH screening can be substantially improved by using simply obtainable information on sampling age to adjust 17-OHP cut-off values. This approach allows CAH screening to proceed with reasonable follow-up effort, even when early blood sampling is necessary.

TABLE 2. Influence of different cut-off schemes on detection rates, false positive (fp) rates, and ppv

	BW (g)	Age (d)	fp (n)		tp (n)		n	fp rate (%)		ppv (%)		
			el.	m.el.	el.	m.el.		Rel. to the total population	Rel. to the resp. BW and age group	el.	m.el.	Total
Initial 3 cut-off values	<2,000		642	750	0	3	15,238	0.26	9.14		0.40	0.22
	≥2,000	0–3	3,324	353	3	34	414,453	0.68	0.89	0.09	8.79	1.00
		≥4	896	55	3	8	108,775	0.18	0.87	0.33	12.70	1.14
Total			4,862	1,158	6	45	538,466		1.12	0.12	3.74	0.84
Redefined cut-off values, adjusted for age and BW	<1,000		156	69	0	0	2,813	0.04	8.00			
	1,000–1,500		243	73	0	1	4,582	0.06	6.90		1.5	0.29 ^a
	1,500–2,000		442	55	0	2	7,843	0.09	6.34			
	2,000–2,500		817	35	0	1	21,374	0.16	3.99		2.78	0.47
	>2,500	0–1		494		7	5,767	0.09	8.57		1.40	1.40
		2–3	1,004	108	3	24	392,635	0.21	0.28	0.30	18.18	2.37
Total		≥4	390	24	3	7	103,452	0.08	0.40	0.76	22.58	2.36
			3,546	364	14	37	538,466		0.73	0.39	9.23	1.29

tp, True positive; el., elevated; m.el., markedly elevated; rel., relative.

^a Due to the low numbers of premature babies, the ppv was summarized for the BW groups <2000 g.

Acknowledgments

Received November 7, 2002. Accepted August 24, 2003.

Address all correspondence and requests for reprints to: Dr. rer. nat. Ralph Fingerhut, Ph.D., Labor Becker, Olgemöller and Kollegen, Führichstraße 70, D-81671 München, Germany. E-mail: r.fingerhut@labor-bo.de.

References

1. New MI 1998 Diagnosis and management of congenital adrenal hyperplasia. *Annu Rev Med* 49:311–328
2. Newfield RS, New MI 1997 21-Hydroxylase deficiency. *Ann NY Acad Sci* 816:219–229
3. Miller WL, Levine LS 1987 Molecular and clinical advances in congenital adrenal hyperplasia. *J Pediatr* 111:1–17
4. Mornet E, Crete P, Kuttann F 1991 Distribution of deletions and seven point mutations on CYP21 genes in the three clinical forms of steroid 21-hydroxylase deficiency. *Am J Hum Genet* 48:79–88
5. Speiser PW, Dupont J, Zhu D, Serrat J, Buegeleisen M, Tusie-Luna MT 1992 Disease expression and molecular genotype in congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *J Clin Invest* 90:584–595
6. Wedell A, Thilen A, Ritzen EM, Stengler B, Luthman H 1994 Mutational spectrum of the steroid 21-hydroxylase gene in Sweden: implications for genetic diagnosis and association with disease manifestation. *J Clin Endocrinol Metab* 71:1145–1152
7. Pang S, Wallace MA, Hofman L, Thuline HC, Dorche C, Lyon IC, Dobbins RH, Kling S, Fujieda K, Suwa S 1988 Worldwide experience in newborn screening for classical congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Pediatrics* 81:866–874
8. Al-Nuaim AR, Abdullah MA, Stevens B, Zain M 1995 Effect of gender, birth weight and gestational age on serum 17-hydroxyprogesterone concentration and distribution among neonates in Saudi Arabia. *Indian J Pediatr* 62:605–609
9. Balsamo A, Cacciari E, Piazzi S, Cassio A, Bozza D, Pirazzoli P, Zappulla F 1996 Congenital adrenal hyperplasia: neonatal mass screening compared with clinical diagnosis only in the Emilia-Romagna region of Italy, 1980–1995. *Pediatrics* 98:362–367
10. Suwa S 1994 Nationwide survey of neonatal mass-screening for congenital adrenal hyperplasia in Japan. *Screening* 3:141–151
11. Lebovitz RM, Pauli RM, Laxova R 1984 Delayed diagnosis in congenital adrenal hyperplasia. *Am J Dis Child* 138:571–573
12. Therrell Jr BL, Berenbaum SA, Manter-Kapanke V, Simmank J, Korman K, Prentice L, Gonzalez J, Gunn S 1998 Results of screening 1.9 million Texas newborns for 21-hydroxylase-deficient congenital adrenal hyperplasia. *Pediatrics* 101:583–590
13. Wong T, Shackleton CHL, Covey TR, Ellis G 1992 Identification of the steroids in neonatal plasma that interfere with 17 α -hydroxyprogesterone radioimmunoassays. *Clin Chem* 38:1830–1837
14. Allan DB, Hoffman GL, Fitzpatrick P, Laessig R, Maby S, Slyper A 1997 Improved precision of newborn screening for congenital adrenal hyperplasia using weight-adjusted criteria for 17-hydroxyprogesterone levels. *J Pediatr* 130:128–133
15. Torresani T, Grüters A, Scherz R, Burckhardt JJ, Harras A, Zachmann M 1994 Improving the efficacy of newborn screening for congenital adrenal hyperplasia by adjusting the cut-off level of 17 α -hydroxyprogesterone to gestational age. *Screening* 3:77–84
16. Gruñero de Papendieck L, Prieto L, Chiesa A, Bengolea S, Bergadá C 1998 Congenital adrenal hyperplasia and early newborn screening: 17 α -hydroxyprogesterone (17 α -OHP) during the first days of life. *J Med Screen* 5:24–26
17. Levy HL, Alber S 2000 Genetic screening of newborns. *Annu Rev Genomics Hum Genet* 1:139–177
18. Liebl B, Nennstiel-Ratzel U, von Kries R, Fingerhut R, Olgemöller B, Zapf A, Roscher AA 2002 Expanded newborn screening in Bavaria: tracking to achieve requested repeat testing. *Prev Med* 34:132–137
19. Liebl B, Nennstiel-Ratzel U, von Kries R, Fingerhut R, Olgemöller B, Zapf A, Roscher AA 2002 Very high compliance in an expanded MS-MS-based newborn screening program despite written parental consent. *Prev Med* 34:127–131
20. Steigert M, Schoenle EJ, Biason-Lauber A, Torresani T 2002 High reliability of neonatal screening for congenital adrenal hyperplasia in Switzerland. *J Clin Endocrinol Metab* 87:4106–4110
21. Working Group on Neonatal Screening of the European Society for Pediatric Endocrinology 2001 Procedures for neonatal screening for congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Horm Res* 55:201–205
22. Pang S, Kling S, Dobbins RH 1990 Illinois experience in newborn screening for congenital adrenal hyperplasia: a new guideline for follow-up approach. In: Bradford HB, Hannon WH, Therrell BL, eds. *Proceedings of the 7th National Neonatal Screening Symposium*, 1989, Washington, DC, 157–161
23. Krone N, Braun A, Roscher AA, Knorr D, Schwarz HP 2000 Predicting phenotype in 21-hydroxylase deficiency? Comprehensive genotyping in 155 unrelated, well defined patients from southern Germany. *J Clin Endocrinol Metab* 85:1059–1065
24. Caulfield MP, Lynn T, Gottschalk ME, Jones KL, Taylor NF, Malunowicz EM, Shackleton CH, Reitz RE, Fisher DA 2002 The diagnosis of congenital adrenal hyperplasia in the newborn by gas chromatography/mass spectrometry analysis of random urine specimens. *J Clin Endocrinol Metab* 87:3682–3690
25. Cara JF, Moshang Jr T, Bongiovanni AM, Marx BS 1985 Elevated 17-hydroxyprogesterone and testosterone in a newborn with 3 α -hydroxysteroid dehydrogenase deficiency. *N Engl J Med* 313:618–621
26. Mauthe I, Laspe H, Knorr D 1977 Zur Häufigkeit des kongenitalen adrenogenitalen syndroms (AGS): München 1963–1972. *Klin Padiatr* 189:172–176