Who Gets Missed: Coverage in a Provincial Newborn Screening Program for Metabolic Disease

Donald W. Spady, MD*; L. Duncan Saunders, MBChB‡; and Fiona Bamforth, MBBS§

ABSTRACT. *Objective.* **To determine coverage of the newborn screening program (NSP) for metabolic disease in Alberta, Canada, and to determine reasons for not being screened.**

Study Design. **Coverage was estimated by deterministic matching of live birth registration data with newborn screening data for the year 1992. Demographic characteristics of not-matched infants were compared with good-match infants using logistic regression.**

Results. **For 42 392 live births, there were 41 553 screening records, of which 40 593 infants were very good matches to NSP records. Another 960 were possible matches. A total of 839 infants were not matched at all, and coverage was estimated at 98.0%. Determinants of infant not-matched status were death in week 1 (adjusted odds ratio [OR]: 383); birth weight of <1500 g (adjusted OR: 18.9) or between 1500 and 2500 g (adjusted OR: 3.2); having a mother who was single (adjusted OR: 2.7) or formerly married (adjusted OR: 12.9); or being born out of hospital (OR: 19.2). The calculated 98% coverage is close to an estimate of 98.3% made by the NSP comparing total births with initial screenings.**

Conclusion. **The matched data give insight as to who was missed and point to the need for closer attention for infants at greater risk of not being screened for metabolic disease.** *Pediatrics* **1998;102(2). URL: http://www.pediatrics.org/cgi/content/full/102/2/e21;** ***infant, newborn; neonatal screening; metabolism, inborn errors.***

ABBREVIATIONS. PKU, phenylketonuria; NSP, newborn screening program; NM, not matched; GM, good match; PM, possible match; OR, odds ratio; OOH, out of hospital.

Since the introduction of the Guthrie test,[1] screening of newborns has become an effective method to detect certain metabolic diseases. Disorders commonly screened for include congenital hypothyroidism and phenylketonuria (PKU), but some programs include other disorders.[2] These diseases are rare, present at birth, hard to detect clinically in the newborn, and can impair the health and mental development of the child permanently if not diagnosed and treated promptly. Where newborn screening programs (NSPs) exist, it is commonly recommended that all newborns undergo screening at between 1 and 7 days of age.[3–5]

A primary concern in NSPs is the completeness of coverage; ie, are all newborns screened? This usually is estimated by comparing the number of unique screen samples analyzed with the number of births registered for the same time period and region. Rarely is a formal link made between the two, and even more rarely is any attempt made to determine which children are not screened.

In Alberta, Canada, all newborns, ~42 000 per year, are screened for PKU, congenital hypothyroidism, biotinidase deficiency, and tyrosinemia. The program is part of the Alberta Hereditary Diseases Program and is carried out by the Department of Laboratory Medicine and Pathology of the Walter Mackenzie Health Sciences Centre in Edmonton, Alberta. Although coverage is estimated to be close to 100%, it has never been evaluated formally.

This project had three objectives: 1) to determine the number of infants born alive in Alberta in 1992 who had blood obtained for newborn screening; 2) to define the characteristics of those infants who were not screened or who could not be identified with confidence as being screened; and 3) to determine the degree to which screened infants with results requiring follow-up were followed up appropriately and a final disposition made. This article addresses the first two objectives.

SUBJECTS AND METHODS

Data came from two sources: 1) the Government of Alberta Register of live births, and 2) the Newborn Screening Program of the Alberta Hereditary Diseases Program. Ethics approval for the project was obtained from the ethics committee of the University of Alberta Faculty of Medicine.

Coverage was estimated by deterministic matching of all registered live births in Alberta in 1992 with screening program data to determine the number of infants screened. Infants remaining not matched (NM) were considered to be unscreened. The demographic and biological characteristics of the unscreened infants were compared with screened infants. Variables used in the analysis were those available from the two datasets plus variables derived from the datasets. Table 1 shows the variables available from each of the datasets.

Data preparation consisted of 1) abstracting, cleaning, and verifying the screening data from laboratory files for the period from January 1, 1992, to January 31, 1993 into a working screen file; 2) obtaining from Alberta Registries the birth data file for all live births in Alberta in 1992 and a file for all infants born in 1992 who died in the first year of life and preparing these data for merging with the screen file; 3) estimating a birth date for the screen record based on the date of the report and infant age at that time; and 4) merging the screen file and birth file to make a file suitable for analysis (Fig 1).

Data were stored in Paradox[6] and Reflex[7] databases and were

From the *Departments of Pediatrics, ‡Public Health Sciences, Faculty of Medicine, and §Laboratory Medicine and Pathology, Walter Mackenzie Health Sciences Center, University of Alberta, Edmonton, Alberta, Canada.
This research was carried out as a project in partial completion of the requirements for an MSc (Public Health Sciences) for Dr Spady.
Received for publication Nov 20, 1997; accepted Apr 8, 1998.
Reprint requests to (D.W.S.) Department of Pediatrics, 2C3.62 WMC, University of Alberta, Edmonton, Alberta, Canada T6G 2R7.

TABLE 1. Variables Available From Screening and Vital Statistics Records and Used for Analysis of Coverage in an NSP in Alberta, Canada, in 1992

Variables Common to Both Datasets	Variables Recorded Only in Birth Registry	Variables Recorded Only on Neonatal Screening Form
Date of birth[a]	Last name of mother	Date of sample
Last name of child	First name of mother	Is baby >24 hours old?
First name of child	Birthweight	Age at screening
Sex	Gestation	Time of sample
Birth site/blood source[b]	Date of birth	TSH result
	Mother's date of birth	Biotinidase result
	Mother's marital status	Tyrosine result
	Normal city of residence	Phenylalanine result
	Postal code of residence	
	Mother's health plan number	
	Total infants live born	
	Total infants stillborn	

[a] For matching purposes, date of birth used estimated date for the screen file and registered date for the birth file.
[b] Birth site was considered equivalent to blood source site for matching purposes. In a city, the birth site and blood source site could be one of several hospitals; smaller communities had only one hospital.

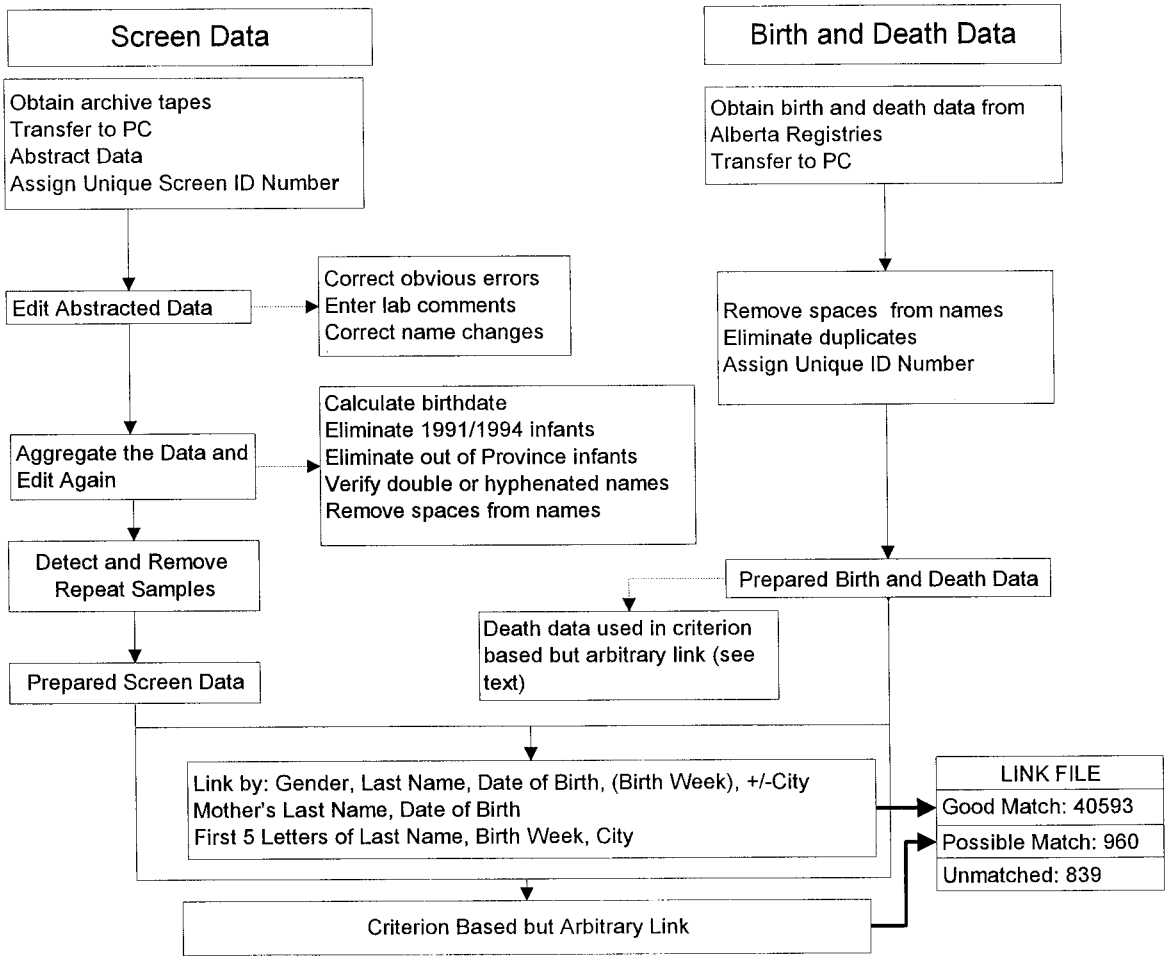

Fig 1. Screen data were abstracted from laboratory archive files and, after cleaning and editing to make a file suitable for linking, were linked to vital statistics data provided by Alberta Registries using deterministic criteria. Records matched using these criteria were considered GM. After all possible records had been deterministically matched, records were matched using less stringent criteria. These records were considered PM. The remaining vital statistics data were NM and thus considered to be unscreened.

analyzed using SPSS.[8] Matching was performed with an iterative series of SPSS MATCH procedures and also by visual matching of screen and birth records. Except as noted below, matching was deterministic and based on the last name of infant and mother, sex, date, hospital and community of birth, and normal residence of the infant. In general, for matching purposes the place where blood was obtained was considered to be the place of birth, and the estimated date of birth on the screen record was compared with the registered date of birth. Because a child may take the mother's last (maiden) name or it may have a hyphenated name, visual matching was attempted as the last stage of the matching procedure.

Each matched record was assigned a level of confidence reflecting one author's (D.W.S.) subjective estimate of the quality of the match for that record. An infant had very good match (GM) status if confidence in the match was high and had a possible match (PM) status if confidence was low. The PM infants were chosen using the criteria that the birth record child had to have a birth

date closest to but before the estimated birth date of an existing unmatched screen result; had to be alive on the date the blood was recorded as obtained; and had to be from the same community, the next nearest community, or the most likely referral community as the existing unmatched screen result. Cases in which the criteria were equally met by more than one match were assigned a match at random. The children remaining in the birth file after all the screen records had been matched were considered NM; thus there were three categories of infants: GM, PM, and NM.

On completion of matching, unique personal identifiers were deleted and the age when blood was obtained was calculated, using date of birth from the birth certificate, and date of sample from the screen record. The variable native was created by considering that infants with homes on native reserves were very likely of native origin.

Statistical Analysis

The GM and PM groups were combined to obtain an estimate of coverage only. For much of the remaining analysis, the PM group was not included. Descriptive summary statistics were obtained for all variables. The statistical significance of analysis of categorical data was assessed using contingency table analysis and χ^2 statistics, and of continuous data using analysis of variance with post hoc comparisons using Scheffé's procedure with a probability criterion of $P < .05$.

The variables birth weight and maternal age also were divided into categories for use in contingency table analysis. Birth weight was divided into <1.5 kg, 1.5 to 2.5 kg, and >2.5 kg categories, and maternal age was divided into <20 years, 20 to 35 years, and >35 years categories.

The choice of variables to explore as determinants of nonmatching was based on availability of data, clinical intuition, and the results of Streetly.[9] After initial exploration comparing single variables with match status, logistic regression was used to account for several factors simultaneously, using matching status (GM vs NM) as the dependent variable. Those variables significantly associated with match status in the exploratory analysis were included as independent variables.

RESULTS

Data were available for 43 155 screen records: 41 553 were initial screen results, and 1602 were repeat screen results from 1485 infants. Of the 42 392 live births registered, 40 593 were GM infants, 960 were PM, and 839 were NM. The sum of GM and PM infants (41 553) is the best estimate of coverage and is 98.0% of the eligible population. Almost 75% of all infants were born to married women, 22% were born to women who had never married, and the remaining 3% were born to women who were divorced, widowed, separated, or whose marital status was unknown. For analytic purposes, the mothers in these last four groups were combined to form one category called formerly married. For 1554 infants (3.7%), the probability of native status was high. Of the 167 infants who died in the first week of life, 13 were GM, 14 were PM, and 140 were NM infants.

Two communities, each with >13 000 births per year, accounted for 63% of all births. Communities with between 100 and 1000 births per year accounted for ~24% of births, and smaller communities accounted for the remainder. There were 226 infants who were born out of hospital (OOH); these were considered home births. A large proportion of infants were born in communities different from their home community.

Data describing the GM, PM, and NM groups are shown in Table 2. Compared with the GM group, NM infants more often died in infancy, were of low birth weight, and were born after a shorter gestation.

Nearly 5.4% of NM infants were born OOH, compared with only 0.4% of GM infants. Only 167 (73.9%) of the 226 infants born OOH were matched. As well, NM infants were slightly more likely than were GM infants to have come from a community with a high native population and were more likely to live in a community different from the community or administrative health region in which they were born (data not shown).

Overall, 6.2% of infants had a birth weight <2500 g, and 1.2% had a birth weight <1500 g. However, 32.1% in the NM group and 14.9% in the PM group were <2500 g, but only 5.5% of the GM group was <2500 g. For infants with a birth weight <1500 g, the differences between the NM and GM groups were even more striking (18.6% vs 0.7%). Because so many NM infants died in the perinatal period, birth weights excluding deaths in the first week of life were compared. With first-week deaths excluded, infants <2500 g constituted 24.2% of NM infants compared with 5.5% for the GM group.

Only 50.8% of NM infants had married mothers, compared with 76.0% of GM infants. The infants of formerly married mothers were much more likely to be NM than were those whose mothers marital status was recorded as single. Of the 839 women in the NM group, 14.5% were younger than 21 years, compared with 9.1% of the GM group. Older mothers were generally evenly distributed between the two groups (9.4% and 8%, respectively).

Table 3 summarizes data describing the relationship of match status to maternal age, birth weight, and marital status. Data for infants whose birth weight was missing are excluded. Except for the formerly married, younger than 21-year group, of whom there were only 20 cases, matching was most likely to occur among infants with a birth weight >2500 g born to married mothers 21 to 35 years of age. At every age, as birth weight rises, the proportion of infants matched also rises. The proportion of GM infants rises also as marital status changes from

TABLE 2. Characteristics of NM, PM, and GM Groups of Infants Born in Alberta in 1992 Who Should Have Had Screening for Metabolic Disease in the Newborn Period

Variable	NM			PM			GM			Post hoc Contrasts (Scheffé's Test) $P < 0.05$
	Mean	SD	*N*	Mean	SD	*N*	Mean	SD	*N*	
Birth weight	2740	1129	781	3164	780	954	3375	560	40 593	NM < PM < GM
Gestation	35.2	6.5	839	38.2	3.7	959	39.1	1.9	40 591	NM < PM < GM
Total live born	2.1	1.6	839	2.1	1.4	960	2.0	1.2	40 593	NM > GM
Mother's age	27.4	5.9	838	27.2	6.3	958	28.0	5.3	40 590	NM, PM < GM
Age at death	3.4	18.0	146	4.7	5.4	16	72.3	69.1	86	NM, PM < GM

TABLE 3. Relationship of Match Status to Marital Status, Maternal Age, and Infant Birth Weight for Infants Born in Alberta in 1992 Who Should Have Been Screened for Metabolic Disease in the Newborn Period

Marital Status	Age (y)	Weight (g)	NM	PM	GM	Total Infants	Percent GM
	<20	<2500	24	9	123	156	78.8
		2500+	60	110	2035	2205	92.3
Single	20–35	<2500	54	31	458	543	84.3
		2500+	136	201	6070	6407	94.7
	35+	<2500	4	5	22	31	71.0
		2500+	5	4	207	216	95.8
	<20	<2500	1	0	15	16	93.8
		2500+	1	5	295	301	98.0
Married	20–35	<2500	140	61	1411	1612	87.5
		2500+	205	260	26 251	26 716	98.3
	35+	<2500	13	13	145	171	84.8
		2500+	22	32	2718	2772	98.1
	<20	<2500	0	0	3	3	100.0
		2500+	0	0	8	8	100.0
Formerly married	20–35	<2500	11	16	60	87	68.9
		2500+	85	165	633	883	71.7
	35+	<2500	4	7	11	22	50.0
		2500+	15	32	124	171	72.5

formerly married to single to married. These data suggest that match status was influenced by maternal age, marital status, and birth weight.

Initial exploration of the data using logistic regression eliminated the variables describing probability of native status, maternal age, and mobility as important determinants of match status. Variables included in the final regression were death in week 1, birth weight (three categories), marital status (three categories), and birth OOH. The results of this regression are generally consistent with the stratified analysis performed previously (Table 4) but demonstrate the independent effects of the significant variables.

In summary, 98% of infants were screened; however, 839 infants were not screened. Those at greater risk of not being screened were those born OOH, of low birth weight, to unmarried or formerly married mothers, or dying in the perinatal period.

DISCUSSION

Data Quality

The credibility of this study rests on the assumption that all initial screen records were obtained; that the match was complete and accurate; and that all repeat samples were detected, recorded as such, and eliminated from estimates of coverage. All the screening data were on archive data tapes of the clinical laboratories of the Walter Mackenzie Health Sciences Centre of the University of Alberta. Data abstraction was exhaustive and involved at least four separate scans of the screening data files. The quality and accuracy of the birth data were assumed to be very high, because they were obtained from vital statistics data provided by the Government of Alberta.

Because both the computer and visual match processes were based on specific identifiers such as last name and date and place of birth, it is likely that the accuracy of the match is high. With the assignment to each record of a level of confidence of a match, the subsequent analysis, which is based primarily on records with strong confidence of a very good match, is conservative in its conclusions.

A sample was considered a repeat sample if it was labeled as such or if records with equivalent identifiers were found during data preparation or matching. The method of matching for repeat analyses was almost always a visual match. The search for repeat analyses ended after finding the first repeat, although on occasion several sequential analyses were found for an infant.

TABLE 4. OR and 95% Confidence Interval (CI) for NM (Unscreened) Infants Born in Alberta in 1992

Variable	OR	95% CI	
Birth weight (g)			
>2500	1		
1500–2499	3.2	2.5	4.1
<1500	18.9	14.2	25.2
Marital status			
Married	1		
Single	2.7	2.3	3.2
Formerly married	12.9	10.1	16.4
Born OOH	19.2	13.1	28.2
Died in week 1	383	206	710

Program Performance

Guidelines have been developed by various agencies that describe the institution and management of an NSP.[3–5] Although these guidelines differ in specifics, they all emphasize the concepts of 1) centralization; 2) rigorous coverage of all infants at an age >24 hours and <7 days; 3) the use of accepted screening tests for the respective metabolic diseases; 4) a clear line of responsibility for transmission of all screen results to the physician or hospital identified as the sender; 5) a more direct and immediate method of informing physicians of abnormal results; 6) follow-up of suspect results until a diagnosis is made and definitive therapy is instituted; 7) some form of ongoing evaluation, usually participation in national or international quality control programs

that allow an assessment of sample processing; and, 8) keeping accurate records.

Estimates of coverage vary among countries and over time, but in established programs coverage often approaches 100%.[10–13] In the United States, coverage for 1992 varied from ~76% for Nevada to 100% for several other states. However, Puerto Rico, a US trust territory, screened only 8% of its newborn population.[2]

Although acceptable for administrative purposes, most estimates of coverage are crude, commonly comparing only total unique screens and total births for the same area and time. With this method, it was estimated that in 1992, 98.3% of infants in Alberta were screened. This study extended the usual estimate of coverage by linking screen records with birth data and determining actual coverage and showed that 98.0% of children were screened at least once. This is in good agreement. However, the relatively low value of 2% nonscreened infants represents 839 infants. It is possible that some of these infants were born in Alberta but moved out of province before being screened; however, this situation was determined in only 27 infants. This does not include an additional 531 infants who were first screened only after 7 days of age, nor does it include 47 infants who had inadequate or unsatisfactory samples but who did not undergo repeat screening (data not shown). The figure of 98% is less than that of 99.1% reported recently by Gray and colleagues,[14] who matched birth and screen records for 8751 consecutive births at two large hospitals for part of 1993. However, the data reported here reflect all infants born in one province for 1 year.

Total coverage remains a goal of all NSPs. In Alberta, with 43 000 births year and screen coverage of 98%, on average about once every 5 years, a child with congenital hypothyroidism (incidence, 1:4000)[15] will be missed because of a failure to screen. Apart from the personal tragedy that these missed cases represent, they also constitute a significant legal concern. In the United States, up to 1986, of 76 missed cases, 29% resulted in legal action, and settlements ranged from $1 million to $20 million.[16]

Who Got Missed

A strength of this report is the description of the characteristics of infants who were not screened. The results obtained agree in part with those of Streetly et al,[9] who suggested that infants in special care units, infants born in one area but normally lived in another, and infants of high-risk ethnic groups also were more likely to be missed. Gray[14] also reported that low birth weight infants were at particular risk of being missed. Although the current data verify that low birth weight infants are at significant risk of being missed, they extend these findings to show that risk of being missed is also associated with single parenthood, home birth, and death in the first week of life. Our results differ from Streetly in that neither mobility nor ethnic (native) status was independently associated with being missed. This may be in part attributable to our inability to identify accurately ethnic status or mobility.

A current concern regarding nonscreening is the practice of discharge from hospital before 24 hours of age. Gray[14] showed that infants discharged before 24 hours were 25 times more likely not to be screened than those discharged after 24 hours. We cannot determine with accuracy which children were discharged before 24 hours and who were not screened, thus, we cannot comment about this specific, and important, concern.

Death in week 1 of life is the strongest predictor of nonscreening, with 140 of 167 deaths by age 7 days in the NM group. Because most deaths occurred in the first day of life, this finding is not surprising, although it could be argued that even very sick infants should be screened. The logistic regression showed that independent of death in week 1, very low birth weight infants (<1500 g) and infants with birth weight between 1500 and 2500 g were 19 times and 3 times, respectively, less likely to be screened compared with infants with birth weights >2500 g. No specific reason for this observation is apparent; however, the problem of small preterm babies being missed requires resolution.

Although fewer in absolute number, infants born OOH also were at great risk of not being screened. Of the 226 infants born OOH, 58 were not screened, resulting in a nonscreen rate 19 times that for children born in hospital. Many of those screened were >7 days of age (data not shown), potentially prolonging the time to diagnosis and treatment.

Finally, infants of formerly married mothers or of mothers whose marital status was unknown also were at high risk of not being screened. Infants of single mothers were at lesser risk, and infants of mothers who were married were at least risk.

An unexpected result is the marked difference in the odds ratios (ORs) for being NM for infants of mothers with formerly married marital status (OR: 12.9) compared with infants of single mothers (OR: 2.7). It seems reasonable that infants of mothers in these two groups might have a greater frequency of being NM than do infants of married mothers, but it is not obvious why infants of formerly married mothers are at such increased risk. A possible reason is confusion in surnames of the infants as recorded when screened and on the birth registration form. Although visual matching was performed as part of the matching procedure and every effort was made to detect alternate names, it is possible that some were missed.

CONCLUSION

Some children with disease will be missed in screening programs. This may be attributable to the vagaries of the disease or because the screening tests are not perfect.[17] However, the primary determinants of an affected child not being detected are that the child was never screened; the blood sample obtained was misinterpreted, unsatisfactory, or obtained too early in the child's life; or follow-up for repeat blood analysis was incomplete.[18–20] This report has addressed the first reason. Of the 781 infants not matched and for whom all the data in Table 4 were available, 609 had one or more risk factors as deter-

mined by logistic regression. Among 11 772 infants with one risk factor, 434 (3.7%) were missed, and among 971 infants with two or more risk factors, 175 (18%) were missed. However, among 29 585 infants with no risk factor, only 172 (0.58%) were missed.

ACKNOWLEDGMENTS

This work was supported by the Department of Pediatrics of the Faculty of Medicine of the University of Alberta.

We thank Kathy Mazurek of the Department of Clinical Laboratories and Mr Harry Calhoun of the EPICORE project for facilitating the access to data, and David Spady who wrote the necessary computer programs to extract the screening data from the laboratory data archive.

REFERENCES

1. Guthrie R, Susi A. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics*. 1963; 32:338–343
2. Newborn Screening Committee. *The Council of Regional Networks for Genetic Services (CORN) National Newborn Screening Report—1992*. Atlanta, GA: CORN: 1995
3. American Academy of Pediatrics, Committee on Genetics. Issues in newborn screening. *Pediatrics*. 1992;89:345–349
4. Gruters A, Delange F, Giovannelli G, et al. Guidelines for neonatal screening programs for congenital hypothyroidism. ESPE Working Group on Congenital Hypothyroidism. *Horm Res*. 1994;41:1–2
5. LaFranchi S, Dussault JH, Fisher DA, et al. Newborn screening for congenital hypothyroidism: recommended Guidelines. AAP Section on Endocrinology and Committee on Genetics and American Thyroid Association Committee on Public Health. *Pediatrics*. 1993;91:1203–1209
6. Paradox. *Borland Paradox for Windows*. Version 5.0. Scotts Valley, CA: Borland International, Inc; 1994
7. Reflex: *Borland Reflex*. V2.0. Scotts Valley, CA: Borland International, Inc; 1989
8. SPSS. *Statistical Package for the Social Sciences*. Release 6.0.1. Chicago, IL: SPSS Inc; 1993
9. Streetly A, Grant C, Bickler G, Eldridge P, Bird S, Griffiths W. Variation in coverage by ethnic group of neonatal (Guthrie) screening programme in south London. *Br Med J*. 1994;309:372–374
10. Schoenberg D, Klett M. Screening for congenital hypothyroidism in the Federal Republic of Germany. Past, present, and future. In: Therrel BL, ed. *Advances in Neonatal Screening*. New York, NY: Elsevier Science Publishers; 1987:25–29
11. Connelly J. Australian experience in neonatal thyroid screening, 1977–1985. In: Therrel BL, ed. *Advances in Neonatal Screening*. New York, NY: Elsevier Science Publishers; 1987:31–34
12. Irie M, Nakajima H, Inomata H, Noruse H, Suwa S, Takasugi N. Screening of neonatal hypothyroidism in Japan. In: Therrel BL, ed. *Advances in Neonatal Screening*. New York, NY: Elsevier Science Publishers; 1987:41–47
13. Italian Committee of Experts to Neonatal Mass Screening. Neonatal mass screening for metabolic disorders in Italy. Technical report of Italian Society of Pediatrics. In: Therrel BL, ed. *Advances in Neonatal Screening*. New York, NY: Elsevier Science Publishers; 1987:473–474
14. Gray JE, Sorrentino JE, Matheson GA, et al. Failure to screen newborns for inborn disorders: a potential consequence of changes in newborn care. *Early Hum Dev*. 1997;48:279–285
15. Allen DB, Hendricks A, Sieger J, et al. Screening programs for congenital hypothyroidism. *Am J Dis Child*. 1988;142:232–236
16. Holtzman C, Stazyk W, Cordero J, Hannon H. Descriptive epidemiology of missed cases of phenylketonuria and congenital hypothyroidism. *Pediatrics*. 1986;78:553–558
17. MacMillan DR, Mabry CC. Differences in screening and confirmatory thyroid profiles in congenital hypothyroid subtypes: implications for missed cases. *J Perinatol*. 1995;15:126–130
18. Fisher DA. Effectiveness of newborn screening programs for congenital hypothyroidism: prevalence of missed cases. *Pediatr Clin North Am*. 1987;34:881–890
19. Leger J. Screening for congenital hypothyroidism in France. Misdiagnosed cases: collaborative study of screening centres in France. *Eur J Pediatr*. 1990;149:605–607
20. Sinai LN, Kim SC, Casey R, Pinto-Martin JA. Phenylketonuria screening: effect of early newborn discharge. *Pediatrics*. 1995;96:605–608

Who Gets Lost: Follow-up of Suspect Results in a Newborn Screening Program

Donald W. Spady, MD*; L. Duncan Saunders, MBChB‡; and Fiona Bamforth, MBBS§

ABSTRACT. ***Objective.*** **To determine reasons for and adequacy of follow-up testing of suspect results of metabolic screening in infants born in Alberta in 1992.**

Study Design. **Of 42 392 live births, 41 553 infants were deterministically matched using birth registry data. Infants requiring repeat analyses were determined from notes made on the screening report. Characteristics of infants needing repeat screening, and obtaining a repeat screen results, were determined by logistic regression using variables from the birth registry and the screening record.**

Results. **A total of 1375 infants required repeat screening. Infants with unsatisfactory samples were more likely to be born in a smaller community, of low birth weight, and to have the sample obtained after 7 days of age. Infants with biologically suspect results were more likely to be of low birth weight, to die in week 1 of life, and to be born in a large hospital. Repeat analyses were found for 663 infants. Boys, infants from smaller communities, and low birth weight infants were more likely to have the required repeat screening. Infants of single mothers were less likely to undergo repeat screening.**

Conclusions. **The results of this study demonstrate the need for a clear, time-oriented protocol of follow-up of newborn metabolic screening results.** ***Pediatrics*** **1998;102(2). URL: http://www.pediatrics.org/cgi/content/full/102/2/e21;** ***neonatal screening; metabolism, inborn errors; phenylketonuria; congenital hypothyroidism; infant, newborn.***

From the *Departments of Pediatrics, ‡Public Health Sciences, Faculty of Medicine, and §Laboratory Medicine and Pathology, Walter Mackenzie Health Sciences Center, University of Alberta, Edmonton, Alberta, Canada.

Received for publication Nov 20, 1997; accepted Apr 8, 1998.

Reprint requests to (D.W.S.) Department of Pediatrics, 2C3.62 WMC, University of Alberta, Edmonton, Alberta, Canada T6G 2R7.

ABBREVIATIONS. PKU, phenylketonuria; WMC, Walter Mackenzie Centre; GM, good match; PM, possible match; NM, not matched; NSP, newborn screening program; ACH, Alberta Children's Hospital; NSQ/US, nonsufficient quantity or unsatisfactory sample; TSH, thyroid stimulating hormone; OR, odds ratio.

Screening of newborns between 1 and 7 days of age detects certain metabolic diseases such as congenital hypothyroidism and phenylketonuria (PKU), which can impair the health and mental development of the child permanently if not diagnosed and treated promptly.[1] Newborn screening in Alberta, Canada, started in 1967 and is administered through the Department of Laboratory Medicine of the Walter Mackenzie Health Sciences Centre (WMC) in Edmonton, Alberta. Infants are screened for congenital hypothyroidism, PKU, biotinidase deficiency, and tyrosinemia.

A companion publication[2] addressed the issue of coverage of the Alberta program in 1992 and described the characteristics of children who were never screened. It showed that of the 42 392 infants born alive in Alberta in 1992, 40 593 were good matches (GM), 960 were possible matches (PM), and 839 infants were not matched (NM) and thus not screened. Factors found to be related independently to absence of screening were death in week 1 of life, low birth weight, birth out of hospital, and birth to a mother who was single or formerly married. This article addresses the timing of the initial screen, the reasons for follow-up testing, and the adequacy of follow-up of suspect results.

SUBJECTS AND METHODS

Normal Screening Protocol

Although guidelines exist for the provincial institution and management of a newborn screening program (NSP),[3–5] the current program lacks a well-defined protocol that ensures that all children will be screened and suspect results followed to resolution. In 1992 the Alberta program screened infants between 24 hours and 7 days of age. Infants screened before 24 hours underwent repeat analysis, preferably before 7 days of age. All samples were analyzed at the WMC. Normal results were mailed to physicians. Abnormal results usually were conveyed to the physician by telephone, with a request for a repeat sample. Atypical or less significant abnormal results were mailed to the physician with a request for a repeat sample. The names of infants needing repeat samples were logged and the name checked off when the repeat sample was obtained. The repeat log was, and is, a static document in that there was no time frame after which a reminder was sent to the physician to obtain a repeat sample.

All repeat analyses for PKU, biotinidase deficiency, or elevated tyrosine were performed at the WMC laboratory or at the Biochemical Genetics Laboratory at the Alberta Children's Hospital (ACH) in Calgary, Alberta. Repeat thyroid function tests were performed either as a repeat newborn screen or by serum analysis at a private laboratory. In the latter event, the screening program was not necessarily notified of the result. When samples analyzed at the ACH were identified as repeats or when an abnormal phenylalanine, biotinidase, or tyrosine result was obtained, a copy of the result was sent to the WMC. If the sample was not identified as a repeat and not otherwise identified as a newborn screen, no report would be sent to the WMC.

Data Preparation and Matching

Details of data preparation and the matching process have been reported.[2] An infant required repeat analysis if 1) the first sample was obtained before the infant was 24 hours of age, 2) it was an unsatisfactory specimen, or c) there was an abnormal result. The physician was requested to obtain a repeat specimen. Records were identified as a repeat if they were labeled as a repeat or if multiple records were found with matching personal identifiers. Repeat records were linked by a unique identifying number to the initial record and by the temporal sequence of sampling determined. The search for repeat records stopped after finding the first repeat, although sometimes several repeats were found for an infant. For infants needing follow-up, the primary data were those of the NSP. Vital statistics data from the Government of Alberta were used to describe demographic characteristics of the infants.

Statistical Analysis

Analysis was limited to the GM group. Data for the NM group are not applicable, and characteristics of the PM infants are unreliable. Summary statistics were obtained relating to time when blood samples were taken and number and type of results needing repeat analysis. Characteristics of infants requiring a repeat specimen and the way these characteristics varied with the reason for a repeat were compared with GM infants with normal results. Characteristics of infants who did not have a requested repeat screen obtained were compared with infants who did have a requested repeat screen. The statistical significance of categorical data was assessed using contingency table analysis and χ^2 statistics, and of continuous data using analysis of variance with post hoc comparisons using Scheffé's procedure and a probability criterion of $P < .05$. Birth weight and maternal age were stratified for use in contingency table analysis. After initial exploration of the data, logistic regression was used to control for several factors simultaneously, using need for repeat analysis versus no need as one dependent variable and repeat analysis done versus no repeat done as another dependent variable. The statistical packages SPSS[6] and STATA[7] were used for the analysis.

RESULTS

There were 43 155 records of screen data; 43 125 from the WMC and 30 from ACH. Of these, 1602 were repeat records from 1485 infants; however, 850 records were for infants for whom repeat analysis was not requested. A repeat sample was required for 1375 GM infants. The reasons for repeat screens are outlined in Table 1. Characteristics of infants needing repeat sampling varied with the reason for the repeat sample and are summarized in Table 2.

Logistic regression analysis of these data show that infants with nonsufficient or unsatisfactory specimens (NSQ/US) were more likely to be native, born in a smaller community, and of low birth weight; to have a mother who was formerly married; or to have the first sample taken after 7 days of age (Table 3).

TABLE 1. Reasons for Newborns Born in Alberta to Require a Repeat Sample for Metabolic Screening, 1992, and the Proportion Complying With a Request For a Repeat Sample

Reason	*n*	Reason Percent	Percent Repeated by Screening Program	<24 Hour Samples
<24 Hour sample	235	17.1	46.4	
Poor quality (NSQ/US)	140	10.2	65.5	2
Abnormal amino acid	149	10.8	50.0	1
High phenylalanine	8	0.1	100.0	—
High tyrosine	713	51.9	44.2	3
Low biotin	72	5.2	53.4	1
High TSH	58	4.2	24.1	10
Total	1375	100.0	47.9	17

TABLE 2. Characteristics of Infants Requiring Repeat Analysis Compared With Infants Not Requiring Repeat Analysis

Variable	Repeat Not Needed	Percent	Repeat Needed	Percent	χ^2
Birth weight (g)					***
<1500	245	83.9	47	16.1	
1500–2499	1752	89.6	204	10.4	
>2499	37 221	97.1	1123	2.9	
Marital status					***
Married	29 880	96.9	956	3.1	
Single	8530	95.7	384	4.3	
Formerly married	808	96.0	34	4.0	
Birth frequency					**
>10 Births/day	24 818	96.6	875	3.4	
1–10 Births/day	6827	97.2	197	2.8	
<1 Birth/day	7414	96.2	293	3.8	
Out of hospital	159	94.6	9	5.4	
Home community different from birth community					***
No	28 878	96.9	929	3.1	
Home different but in same health administrative region	5149	96.2	205	3.8	
Home different and out of health administrative region	5191	95.6	240	4.4	
Infant likely of native origin					***
No	37 870	96.8	1269	3.2	
Probable	1348	92.8	105	7.2	
Died in week 1					***
No	39 209	96.6	1370	3.4	
Yes	9	69.2	4	30.8	

* $P < .05$; ** $P < .01$; *** $P < .001$.

TABLE 3. Logistic Regression Analysis Describing Factors Accounting for Need to Repeat Newborn Metabolic Analyses in Infants Born in Alberta in 1992

Variable	NSQ/US Sample*		<24-h Sample		Abnormal Result	
	OR	95% CI	OR	95% CI	OR	95% CI
Birth weight (g)						
>2500	1.00		1.00		1.00	
1500–2499	1.70	0.82–3.51	1.21	0.67–2.17	4.92	4.16–5.83
<1500	6.35	2.11–19.03	5.65	2.76–11.57	6.32	4.37–9.13
Births in community						
>10 per day	1.00				1.00	
1–10 per day	1.32	0.72–2.43			0.85	0.71–1.01
<1 per day	6.47	4.33–9.67			0.81	0.68–0.98
Born out of hospital	8.79	2.80–27.62			0.67	0.21–2.15
Native status: 1 = yes	2.22	1.35–3.69			2.34	1.80–3.05
Marital status						
Married	1.00					
Single	1.19	0.80–1.75				
Formerly married	2.63	1.25–5.55				
Blood >7 days: 1 = yes	2.99	1.43–6.27	NA	NA		
Gender: 1 = male			1.39	1.07–1.81		
Died in week 1: 1 = yes					8.64	2.08–35.90

* In each instance, 0 indicates null; 1, positive response.
NA indicates not applicable.

Infants with samples taken at <24 hours of age were more likely to be boys and of low birth weight. Infants with biologically suspect results were more likely to be of low birth weight, native, die in the first week of life, or to have been born in a hospital with >2000 births per year.

Blood samples were taken before 24 hours of age from 235 infants. Among these infants, 17 abnormal biochemical results, representing 7.2% of the records, were obtained. This figure is 2.5 times the proportion of abnormal results seen in infants first sampled after 24 hours of age. A high thyroid stimulating hormone (TSH) level was the most common abnormality; it was 30 times more common than that found among infants whose blood was taken after 24 hours.

Infants Obtaining the Requested Repeat

Repeat samples were found for only 663 of the 1375 GM infants requiring a repeat screen. Infants with birth weight <1500 g were more likely to have a repeat sample than were infants >2500 g (odds ratio [OR]: 3.11) (Table 4). Infants of single mothers were less likely than those of married mothers to get the necessary repeat specimen (OR: 0.64). Boys were more likely than girls to get the required repeat specimen (OR: 1.39), and infants born in communities where there were fewer than one birth per day were more likely to undergo the required repeat screen than were infants born in communities having more than 10 births per day (OR: 1.59).

The average age at repeat was 27.2 ± 22.4 (mean,

TABLE 4. Categorical Analysis and Logistic Regression Describing Factors Accounting for a Requested Repeat Metabolic Screen Sample From Infants Born in Alberta in 1992

Variable	Repeated		Not Repeated		χ^2 *P*	Logistic Regression	
	n	%	*n*	%		OR	95% CI
Birth weight (g)					**		
>2500	529	47.1	595	50.9		1.00	
1500–2500	101	49.5	103	50.5		1.23	0.99–1.66
<1500	33	70.2	14	29.8		3.11	1.63–5.94
Marital status					**		
Married	486	50.8	470	49.2		1.00	
Single	161	41.8	224	58.2		0.64	0.50–0.82
Formerly married	16	47.1	18	52.9		0.77	0.38–1.55
Births per day in community					*		
>10	410	46.8	466	53.2		1.00	
1–10	86	43.7	111	56.3		0.93	0.68–1.28
<1	161	54.9	132	45.1		1.59	1.21–2.10
Born out of hospital	6	66.7	3	33.3		2.94	0.72–12.07
Gender					**		
Female	288	44.2	354	55.8		1.00	
Male	375	51.8	348	48.2		1.39	1.12–1.73
Birth town same as home town							
Yes	442	47.5	487	52.4	NS		
No, but same health region	108	52.4	98	47.6			
No, and not same health region	113	47.1	127	52.9			
Native status							
No	619	48.7	651	51.3	NS		
Probable	44	41.9	61	58.1			
Died in week 1							
No	663	48.3	708	51.7	NS		
Yes	0	0.0	4	100.0			

* $P < .05$; ** $P < .01$.
NS indicates not significant.

SD) days, with a maximum of 270 days; 65% of the samples were obtained by 30 days of age (data not shown). All infants with high phenylalanine levels on initial screen underwent repeat screen at a mean age of 11.6 days and a maximum age of 23 days. Two had phenylketonuria and six had hyperphenylalaninemia. One child with hyperphenylalaninemia had two repeat samples obtained. Both results were abnormal, but the child was then lost to follow-up until tracked down to resolve inquiries related to this project. Review of the case shows that the child probably had a benign variant of hyperphenylalaninemia.

Biotinidase activity was reported initially as decreased in 72 infants. For 39 infants, repeat samples were obtained at an average age of 34.7 ± 18.6 days (range, 3 to 92 days). One infant was diagnosed with biotinidase deficiency.

For 713 GM infants, initial screens found high tyrosine levels. A repeat sample obtained after 1 month of age was requested, but only 315 (44.2%) repeat samples were found. The average age at repeat analysis was 34.5 ± 21.3 days (range, 2 to 236 days). Among the repeat screens, 14 samples showed a high tyrosine level; of these, 8 were repeated and found to be normal. A high tyrosine level often is seen in premature infants and, in this study, 163 (22.8%) of the GM infants with a high tyrosine level had birth weights <2500 g, whereas only 5.5% of all GM infants weighed <2500 g.

Initial TSH values were >25 mIU/L in 59 infants tested after 24 hours of age. Ten of 11 infants with a TSH >50 mIU/L were retested, some by private laboratories, and hypothyroidism was found in all. No record of repeat analysis was found for the 11th infant. Repeat analysis was found for only 12 of 47 infants whose TSH level was between 25 and 50 mIU/L.

There were 142 NSQ/US samples, of which 92 were rescreened at a mean infant age of 25.0 ± 28.6 days (range, 2 to 270 days). Relatively more of the NSQ/US samples came from centers where there was less than one birth a day or where the child was born out of hospital (probably a home birth).

In 235 instances, blood was obtained before 24 hours of age; however, only 109 had repeat analysis, after a mean interval of 9.1 ± 13.8 days (range, 1 to 91 days; median, 3 days). Of these 109 repeat screens, 7 showed abnormal results, and none of these have a record of additional repeat screening.

Infants Having a Repeat Sample That Was Not Indicated

For 884 infants, a repeat sample was obtained although the initial sample was normal. Among these 884, which were obtained at an average age of 20.5 days (median, 7 days), 33 had abnormal or NSQ/US results. Analysis of the data (not shown) suggests that these samples were taken from children who had neonatal problems. Nearly 28% had a birth weight <2500 g, but only 5% of GM children with no repeat had a birth weight <2500 g. These children may have been graduates of a newborn ICU who receive a second screen routinely before discharge.

DISCUSSION

One objective of this study was to determine the degree to which infants who underwent repeat

screening were followed appropriately until a final disposition could be made. Consistent with newborn screening guidelines, the assumption made in this report is that a newborn with an initial abnormal screen result, an unsatisfactory sample, or a sample obtained before 24 hours requires a repeat sample. The speed and completeness of this process reflect the effectiveness of the entire NSP because these children are at greatest risk for metabolic disease.

In our analysis, repeat samples were detected as such because they were labeled repeat or because records with equivalent identifiers were found during the editing, visual-matching, or computer-matching procedures. Although some records considered repeats could have been initial records, this is unlikely because the method of matching for repeat screens was always deterministic and almost always verified visually.

Some repeat samples may have been missed because they were not labeled as neonatal screens, but were requested specifically for phenylalanine, tyrosine, or biotinidase measures. This seems unlikely because these determinations are performed by only two laboratories in the province, and the files of both laboratories were searched specifically for all test terms that would report these values. A repeat thyroid screen may have been missed when a serum sample was analyzed at a private laboratory.

Follow-up of repeat screens is an area in which the NSP did not function well. Only 48% of the infants who required it actually underwent repeat analysis. Most of these were infants who had a high tyrosine level. All infants with high phenylalanine had repeat samples, but only 10 of 11 infants with a TSH value of >50 mIU/L and only 12 of 47 infants with a TSH value between 25 and 50 mIU/L were known to have had a repeat analysis. Finally, among those infants who did have the required repeat analysis, the analysis was often obtained after 30 days of age, diminishing the potential benefits of early intervention.

Ensuring that required follow-up samples are obtained is a common problem among all newborn screening systems. In the United States, the Council of Regional Networks for Genetic Services, National Newborn Screening Report, 1992, reported that 3.4% of abnormal phenylalanine results and 1.7% of abnormal thyroid function test results were lost to follow-up.[1] Data are not directly comparable with Alberta because of different diagnostic criteria, but given the very small sample, the Alberta data for these two disorders appear to be comparable. Ms J. Tuerck of the Oregon Screening Program and a member of the Maternal and Child Health Select Panel on Newborn Screening Systems, which has evaluated 14 states in the United States noted that "follow up problems seem to be common to all the states . . . follow up which works best seem to be those [sic] who have a designated person(s) at the program manager level, use medical consultants for infants with significant abnormal results, have written protocols which are followed and evaluated on a regular basis, and who have close effective communication with other parts of the program (lab and practitioners) and with other state or provincial programs" (J. Tuerck, April 11, 1996, personal communication).

Several explanations may account for Alberta's poor performance. There is no clear, time-oriented protocol to handle follow-up requests, and there are no personnel dedicated to this specific task. Although serious metabolic disease is being detected and clearly abnormal results are communicated by telephone to the physician, follow up is not well structured and the possibility of a "missed repeat and affected child" is real.

Repeat screens are not always labeled as such and thus may not be identified if patient demographics do not match the initial screen. Physicians may be unaware of what is good practice. For example, Sinai[8] documented that many infants who were discharged by 24 hours of age did not undergo repeat analysis for PKU, apparently because pediatricians and other physicians were unaware of the guidelines for newborn screening and thus for the need to obtain repeat analysis. Some physicians, on reading the message in the screening report "suggest clinical follow-up and serum thyroid function tests" or "suggest repeat screen in about 1 month or sooner if clinically indicated", may assume that if they think it is not clinically indicated, the screening need not be repeated. Unfortunately, these diseases cannot be excluded reliably on the basis of clinical impression only; nondisease requires laboratory verification.

Finally, analysis of some thyroid function samples could have been repeated in a private laboratory and no record of repeat analysis would be reported to the screening program. This is illustrated by the fact that the screening program had repeat records for only 24.1% of infants with high TSH values, whereas 37.3% of infants with high TSH values were known to have undergone repeat analysis.

In 142 cases, the sample was inadequate or unsatisfactory. Only 46.4% of these were repeated. Among the repeat analyses, there were several abnormal results that required additional verification. Where data are available, all eventually showed normal values. The higher rates of poor-quality samples from smaller centers may reflect inexperience on the part of personnel obtaining the sample. Very low birth weight infants were more likely than term infants to have NSQ/US samples.

A raised tyrosine level was the most common abnormal screen result, seen in 1.8% of those screened after 24 hours of age. Although the most common cause of a raised tyrosine level is transient tyrosinemia attributable to immaturity of the enzyme involved in its metabolism, tyrosinemia may have long-term effects for this small subset of infants.[9]

CONCLUSION

Among other requirements, an effective screening program must test all eligible subjects, ensure rapid follow-up and verification of suspect results, and promptly provide and institute definitive therapy.[10] Our earlier report showed that 98% of infants in Alberta were screened in 1992 and described the characteristics of nonscreened infants. This report evaluates further other aspects of the performance of

the Alberta Newborn Screening Program with explicit reference to GM infants who were screened. The most important finding is that nearly half of the required repeat screens were not performed. Review of other jurisdictions suggests that the best follow-up occurs when a protocol, detailing the procedures for the initial process of screening and the subsequent process of follow-up of abnormal or unsatisfactory results, is conceived and implemented. Such protocols appear to be most effective when there is an integrated system with designated responsibility and authority to ensure that the program runs effectively and efficiently.

REFERENCES

1. Newborn Screening Committee. *The Council of Regional Networks for Genetic Services (CORN), National Newborn Screening Report—1992*. Atlanta, GA: CORN; 1995
2. Spady DW, Saunders LD, Bamforth F. Who gets missed: coverage in a provincial newborn screening program. *Pediatrics*. 1998;102(2). URL: http://www.pediatrics.org/cgi/content/full/102/2/e21
3. American Academy of Pediatrics, Committee on Genetics. Issues in newborn screening. *Pediatrics*. 1992;89:345–349
4. Gruters A, Delange F, Giovannelli G, et al. Guidelines for neonatal screening programs for congenital hypothyroidism. ESPE Working Group on Congenital Hypothyroidism. *Horm Res*. 1994;1:1–2
5. LaFranchi S, Dussault JH, Fisher DA, et al. Newborn screening for congenital hypothyroidism: recommended guidelines. AAP Section on Endocrinology and Committee on Genetics, and American Thyroid Association Committee on Public Health. *Pediatrics*. 1993;91:1203–1209
6. SPSS. *SPSS for Windows*. Release 6.1.1. Chicago, IL: SPSS Inc; 1994
7. Statacorp. *Stata Statistical Software*. Release 5.0. College Station, TX: Stata Corporation; 1997
8. Sinai LN, Kim SC, Casey R, Pinto-Martin JA. Phenylketonuria screening: effect of early newborn discharge. *Pediatrics*, 1995;96:605–608
9. Stanbury JB, Wyngaarden JB, Fredrickson DS. Tyrosinemia and related disorders. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic Basis of Inherited Disease*. 6th ed. New York, NY: McGraw-Hill; 1989:556
10. Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. Geneva, Switzerland: World Health Organization; 1968

Who Gets Missed: Coverage in a Provincial Newborn Screening Program for Metabolic Disease

Donald W. Spady, L. Duncan Saunders and Fiona Bamforth

Pediatrics 1998;102;e21
DOI: 10.1542/peds.102.2.e21

PEDIATRICS®
OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Who Gets Missed: Coverage in a Provincial Newborn Screening Program for Metabolic Disease

Donald W. Spady, L. Duncan Saunders and Fiona Bamforth

Pediatrics 1998;102;e21
DOI: 10.1542/peds.102.2.e21