Screening for Phenylketonuria in New York City
Threshold Values Reconsidered

H. HANSEN, MD, DrPH, A. SHAHIDI, PhD, and Z. A. STEIN, MA, MB, BCh

THE PREDICTIVE VALUE of positive test results for phenylketonuria (PKU) in the New York City screening program is examined in this report. This value is expressed by the proportion of PKU cases diagnosed in the followup of children whose screening test results are designated as presumptive positive. The New York City experience suggests that the threshold value for presumptive positive screening test results can be raised without compromising sensitivity. Such a modification would reduce the followup load considerably and improve the efficiency of determining which children are in need of preventive treatment.

Screening for PKU at birth is one of the largest and most successful of recent preventive public health programs. In the United States, 20 million infants have been screened—including almost all children born in New York after 1964—since the program was started in the mid-1960s. Previously, PKU cases accounted for about 1 percent of residents in institutions for the mentally retarded. Since systematic screening was introduced, no PKU cases have been reported among patients admitted to institutions that were surveyed in the United States and Britain (*1, 2*).

In the United States, screening is inexpensive and efficient as a result of blood sample collection in the newborn nursery and of automation of the test procedure. An analysis of the New York State program for 1965-70 (*3*) estimated costs per live birth of $1.61 for sample collection in hospitals and of 89 cents for laboratory testing. By contrast, followup of positive screening results is expensive and inefficient. Cost estimates for collecting a followup blood sample in New York State ranged from $7.50 to more than $40, depending on the region. Of approximately 2,190 infants with a presumptive positive screening test result in New York State from 1965 to 1970, only 119 or 5.4 percent were confirmed as having PKU. We believe that it is possible to reduce the high false positive rate without increasing the risk of missing cases.

Data Sources and Methods

Sources. The observations reported here, which cover all children who had presumptive positive screening test results and were born in New York City from 1966 through 1970, are based on two sources. The Bureau of Laboratories of the New York City Health Department provided the results of screening and followup tests. The New York City PKU center at Bellevue Hospital provided diagnostic classifications for the children tested and examined.

Screening and followup. Until the end of 1966, in New York City a screening test result of a phenylalanine (PA) concentration of 6 mg/100 ml blood or more was considered presumptive positive. In 1967, the threshold value requiring followup was lowered to 4 mg/100 ml or more. A presumptive positive infant was judged to be normal if a followup test result indicated 5.6 mg/100 ml or less. The bacterial inhibition assay for phenylalanine in blood spots dried on filter paper was the primary screening procedure, and followup was done by fluorimetry for phenylalanine on microcapillary blood samples.

Screening and followup test results were provided for 1,094 presumptive positive newborns from 1966

☐ *Dr. Hansen is a research scientist, Epidemiology of Brain Disorders Research Unit, New York State Psychiatric Institute, and assistant professor of public health (epidemiology), Columbia University School of Public Health. Dr. Stein is director of the Epidemiology of Brain Disorders Research Unit and professor of public health (epidemiology), Columbia University School of Public Health. Dr. Shahidi is director of the Inborn Metabolic Errors Screening Laboratory, New York City Department of Health. Dr. Selma Snyderman, professor of pediatrics, New York University of Medicine, provided results of diagnostic evaluations, and Dr. Mervyn Susser, professor and head of the Division of Epidemiology, Columbia University School of Public Health, constructively criticized the argument and the manuscript.* *The work reported here was partially supported by a grant (U-2233) from the Health Research Council of the City of New York.*

Tearsheet requests to H. Hansen, MD, Division of Epidemiology, Columbia University School of Public Health, 600 W. 168th St., New York, N.Y. 10032.

through 1970. Followup was classified as (*a*) complete when the child had a normal followup test result or was referred for diagnostic evaluation, (*b*) incomplete when the last followup test result was abnormal and there was no record of referral, and (*c*) unknown when no followup or referral was recorded.

Diagnostic referral. Children with followup test results of 5.7 mg/100 ml or more were referred to the designated PKU center for New York City. A very high screening test value usually resulted in immediate referral. Of the 1,094 presumptive positive children, 59 had a record of referral. The PKU center gave us information on 56 of these children and the diagnoses for the remaining 3 were obtained from other sources. The three categories of diagnostic outcome were classic PKU, hyperphenylalaninemic variants (HPA), and normal. The concentration of blood PA was only one of several diagnostic parameters. PKU children had persisting PA levels above 20 mg/100 ml; HPA children with levels between 15 and 20 mg/100 ml were usually treated, and HPA children with levels below 15 mg were usually not treated.

Results

Over the 5 years, 736,469 children were registered as born alive in New York City. Table 1 shows the annual incidence of presumptive positive screening test results, table 2 the attainment of followup, and table 3 the diagnoses that were established.

Table 1 demonstrates the instability of the rates for PA values of 4 and 6 mg/100 ml over the 5 years. In contrast, the rate for values of 8 mg/100 ml or more altered little during these years.

The program aimed to retest all infants having values of 6 mg/100 ml or more during 1966 and 4 mg/100 ml or more from 1967 through 1970. Table 2 shows the accomplishment of followup by PA value on screening. Followup was complete for 70 percent of the 1,094 presumptive positive children; 25 percent had no record of followup, and for the remainder, followup was incomplete. However, the table shows

Table 1. Presumptive positive screening test results for 1,094 infants, by phenylalanine (PA) value and annual number and rate per 100,000 live births, New York City PKU screening program, 1966–70

PA value (mg/100 ml)	Year of birth											
	1966		1967		1968		1969		1970		Total	
	Number	Rate	Number	Rate	Number	Rate	Number	Rate	Number	Rate	Number	Rate
4[1]			95	65.2	148	104.3	184	125.8	146	96.9	573	98.3
6	205	133.7	74	50.8	59	41.6	48	33.8	24	16.1	410	55.7
8	11		15		11		5		8		50	
12	5	14.3	3	17.1	1	16.2	10	15.7	7	12.1	26	14.8
20	6		7		11		8		3		35	
Total	227	148.0	194	133.1	230	162.1	255	179.7	188	128.6	1,094	148.5

[1] Designated as negative in 1966.

Table 2. Followup categories for 1,094 infants, by results of screening test (mg/100 ml blood phenylalanine), New York City PKU screening program, 1966–70

Followup category	Screening test result (PA mg/100 ml)											
	4		6		8		12		20		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Complete	379	66.1	290	70.7	37	74.0	22	84.6	35	100.0	763	69.7
Not complete	20	3.5	36	8.8	5	10.0	1	3.8	..	....	62	5.7
Unknown	174	30.4	84	20.5	8	16.0	3	11.5	..	....	269	24.6
Total[1]	573	100.0	410	100.0	50	100.0	26	99.9	35	100.0	1,094	100.0

[1] Some percentages do not total 100 because of rounding.

Table 3. Diagnostic classifications for 763 presumptive positive infants with complete followup, by screening test result, New York City PKU screening program, 1966–70

Diagnostic category	Screening test result (PA mg/100 ml)										Total	
	4		6		8		12		20			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
PKU	...	....	1	0.3	1	2.7	8	36.4	29	82.9	39	5.1
HPA,[1] treated	...	....	...	....	..	....	4	18.2	3	8.6	7	0.9
HPA, not treated	1	0.3	5	1.7	2	5.4	2	9.2	..	....	10	1.3
Normal	378	99.7	284	97.9	34	91.9	8	36.4	3	8.6	707	92.7
Total[2]	379	100.0	290	99.9	37	100.0	22	100.2	35	100.1	763	100.0

[1] Hyperphenylalaninemic variants.
[2] Some percentages do not total 100 because of rounding.

that the omissions in followup decreased consistently with rising PA values at screening, and there were no omissions among the 35 children with screening test results of 20 mg/100 ml.

The diagnoses established among those followed are shown in table 3 by screening test result. Of the 763 presumptive positive infants who were retested, 93 percent were considered normal. No child with a screening test result of 4 mg/100 ml was found to require treatment, and only one each of the infants with a screening test result of 6 or 8 mg/100 ml needed treatment. On the other hand, of the 57 children with values of 12 mg/100 ml or more, 77 percent were treated.

Conclusions

Reliability of the screening test results. Over the 5 years, the frequency of 4 mg/100 ml values ranged from 65 to 126 per 100,000. The frequency of 6 mg/100 ml values declined consistently from 134 to 16 per 100,000 live births. Variation in demographic factors that might have influenced PA test values, for instance ethnicity and birthweight, was not of a magnitude that could have produced such great annual fluctuations of incidence. We suggest that changing test procedures account for the instability of 4 and 6 mg/100 ml values. From the stable frequency of PA values of 8 mg/100 ml or more, ranging from 12 to 17 per 100,000 live births, we infer that the procedure is reasonably reliable at these levels.

Cases missed because of incomplete followup. Although diagnostic classifications were not made for nearly one-third of the 1,094 presumptive positive children, only 3 of them may have required treatment. This estimate derives from the frequency of treated cases among those followed up, applied to those not followed up, at the 6, 8, and 12 mg/100 ml level ($0.003 \times 120 + 0.027 \times 13 + 0.546 \times 4$). Two of the three cases would have been classified as PKU during the 5 years. Thus, under the current threshold value for followup, one phenylketonuric child might be missed every 4 years in New York City, given the current number of 100,000 births a year.

Predictive value of screening test results. The threshold value was lowered in 1967 on the assumption that the inclusion of infants with a value of 4 mg/100 ml would yield PKU cases that might otherwise be missed. Yet no child in this group was subsequently found to require treatment. Therefore, for the New York City experience, the predictive value of the 4 mg/100 ml level for the detection of PKU was zero. Followup at this level, although completed for only two-thirds of the children, accounted for about half of the total completed followups (379 of 763).

Some uncertainty must remain regarding the predictive value of 6 mg/100 ml results. Among the 410 infants with a 6 mg/100 ml value, PKU was diagnosed for 1 of the 290 followed up. This case occurred in 1966 when the 6 mg/100 ml value was excessively frequent and presumably less reliable than in later years. Among the 583,135 live births from 1967 to 1970, 90 percent of the presumptive positive infants (778 of 867) had 4 or 6 mg/100 ml values; 513 of the 778 children were followed up, and none required treatment. During the 5 years, 985 children had screening values of 4 or 6 mg/100 ml; 669 (68 percent) of these children were followed up and only 1 (0.15 percent) required treatment.

At screening values of 8 mg/100 ml or more, however, the program can be efficient. Of the 111 children with these initial values, 94 (85 percent) were followed up and among the 94, 45 children (48 percent) were diagnosed and treated.

Choice of thresholds for followup. If the program's experience with 736,000 newborn children is taken to be representative of the current New York City screening program, then the program outcome for different threshold values of presumptive positive screening test results can be estimated. These projections are shown in table 4 for 100,000 live births, the current annual number of children born in New York City. Maintaining current threshold level and attainment of followup would yield an annual average of 148.6 children who would require followup; 103.8 would have completed followup, and 6.2 of these would be treated; and 0.4 children in need of treatment would be missed. If the threshold value were raised to 6 mg/100 ml, with attainment of followup unchanged, the followup load would be reduced by 50 percent and no additional cases would be missed. Thus, effectiveness of case detection would be identical, but efficiency would rise twofold.

A realistic goal for improving the rate of followup would be to reduce the number of incomplete followups by 50 percent at each screening test level. With this rate of followup, raising the threshold value of 6 mg/100 ml would reduce the followup load by only 40 percent, but case detection would increase; both effectiveness and efficiency would be greater than under the current program. Even raising the threshold value to 8 mg/100 ml, with incomplete followups reduced by 50 percent, would result in superior effectiveness and efficiency.

Table 4. Projected followup load and case detection per 100,000 live births by levels of followup attainment and selected threshold values, based on New York City PKU screening program, 1966–70

Followup levels	Proportion of incompletes as in 1966–70		Proportion of incompletes reduced by 50 percent	
	≥ 4 mg/100 ml	≥ 6 mg/100 ml	≥ 6 mg/100 ml	≥ 8 mg/100 ml
Presumptive positives	148.6	70.8	70.8	15.2
Completed followups	103.8	52.2	61.5	13.9
Treated cases	6.2	6.2	6.4	6.3
Missed cases	0.4	0.4	0.2	0.3

Comment

Threshold values have been lowered in some PKU screening programs (*4*) and raised in others (*5*), ranging from more than 2 mg/100 ml to 6 mg/100 ml or more. Raising the threshold value carries the potential risk of generating false negative results, that is, of reducing sensitivity. False negative results may occur because the test is unreliable and fails to identify a true PA elevation or because the PA concentration is too low to give a presumptive positive result even though the test is reliable. Reliability of screening test results varies considerably within programs and between programs (*6*). To avoid false negatives resulting from low reliabilty, threshold values may be set well below those at which a reliable test would be sensitive enough to detect these cases.

False negative results seem to be extremely rare when the threshold value is less than 6 mg/100 ml. Data provided by a number of State programs that report careful diagnostic confirmation are summarized in table 5. These data corroborate the New York City experience. Among 3 million newborns screened elsewhere in New York State (1972 unpublished final report to the New York State Health Planning Commission, "Analysis of the New York State PKU Screening Program Using a Health Status Index," by Dr. J. W. Bush, Dr. M. M. Chen, and Dr. D. L. Patrick of the Division of Health Policy and Management, Department of Community Medicine, University of California, San Diego), Massachusetts (*4*), Pennsylvania (*5*), Michigan (*7*), and Oregon (1975 personal communication from G. R. Brandon of the Oregon Public Health Laboratory) none of the infants with confirmed PKU had an initial value of 4 mg/100 ml or less.

Holtzman and co-workers (*8*) have cautioned against raising the threshold value in PKU screening. Their survey of cases known to State health departments and PKU clinics in the United States and Canada revealed 23 cases that had been missed at newborn screening. The problem of laboratory error was compounded by the lack of centralized screening or laboratory control in several of the eight States reporting false negative results (*9*). However, Holtzman and associates did not base their argument against raising the threshold value on the lack of test reliability, but rather on the possibility that at early screening on discharge from the newborn nursery PA levels are minimally elevated in some PKU children (*8*).

A number of observations suggest that the risk of missing cases because of low PA levels in the first days of life may be negligible. These include the failure of repeat screening later in infancy to yield additional cases, the lack of a continuous distribution

Table 5. Screening test results (bacterial inhibition assay for blood phenylalanine) of infants verified as typical PKU, in screening programs of selected areas, 1962–74

Area and years	Screening laboratory and number of tests	Threshold value for followup (mg/100 ml)	Initial PA mg/100 ml of confirmed PKU cases
New York City, 1966–70	City laboratory, about 736,000	1966, 6 or more; 1967–70, 4 or more	4, none; 6, 1; 8, 1; 12 or more, 37
New York, excluding New York City, 1965–70	State laboratory and 2 regional laboratories, about 1 million	1965–66, 6 or more; 1967–70, 4 or more	4 or 6, none; 8 or more, all cases
Massachusetts, 1962–70	State laboratory, about 800,000	1962–63, 6 or more; 1964–67, 4 or more; 1968–70, more than 2	4 or 6, none; 8 or more, all cases
Pennsylvania, 1964–73	State laboratory, 554,972	1964, 4 or more; 1965–73, 6 or more	4, none; 6, 1; 8 or more, 49
Michigan, 1962–68	State laboratory, 409,554	1962–68, 4 or more	4, 6, or 8, none; 12 or more, 13
Oregon, 1966–74	State laboratory, 284,911	1966–74, 4 or more	4, none; more than 4, 20

from minimal to high PA elevations in true positives, and the absence of PA levels below 6 mg/100 ml after the first day of life in PKU children who had serial PA determinations.

Repeat screening at 4 to 6 weeks of life should detect cases in infants who had been negative at discharge from the newborn nursery, if early screening encounters PA levels too low to be detected. The two programs with the largest experience in repeat screening now consider a second screening test unnecessary. The Massachusetts program did not detect an additional case among 435,000 repeats (*4*), and the Oregon program found two initially negative cases among 20,000 repeats in 1965 when the procedure was still being standardized, but none among the subsequent 217,000 repeats (*10*). One of the two "cases" was later reclassified as a hyperphenylalaninemic variant (G. R. Brandon). The Massachusetts and Oregon programs rely on a single State laboratory which participates in an external quality control program.

The distribution of screening test results of true positives is not compatible with the hypothesis of minimal PA levels. Under this hypothesis, a continuous distribution of PA levels from minimal to moderate to high elevations would be expected among the true positive results. The screening outcomes for almost 4 million infants (table 5) include no PKU case with an initial result of 4 mg/100 ml. Confirmed cases with an initial result of 6 mg/100 ml were extremely rare: 1 in New York City among 736,000 screened, none in upstate New York among 1 million, none in Massachusetts among 800,000, 1 in Pennsylvania among 555,000, and none in Michigan among 410,000.

Newborn infants with PKU who had serial PA determinations have not shown PA levels too low to be detected at discharge from the nursery. A review of 22 cases reported up to 1966 demonstrated a minimal PA level of 5.5 mg/100 ml, which was observed during the first 24 hours of life (*11*). An additional 15 cases with serial PA determinations observed in the California program between 1966 and 1970 showed a minimal PA level of 4.8 mg/100 ml at 2 hours after birth (*12*). The PA level rose rapidly in all children. These data appear to be consistent with the observations on 46 infants with 150 PA determinations between birth and 8 days of age, analyzed by Holtzman and associates (*13*). Although "20 of the 150 determinations were 6 mg/100 ml or less," and 15 of these were obtained "on or before the third day of age," regression analysis yielded PA values of 7.1 and 8.2 mg/100 ml, in males and females respectively, at 24 hours after birth.

False negative results have not been documented thoroughly. However, two observations conflict with the hypothesis of minimal PA elevations. False negative results would be expected to show an excess of early screening tests compared with true positive results. The survey by Holtzman and associates reported age at screening for 23 infants with false negative results (*9*) and 302 with true positive results (*8*). The modal age was 3 days for each of the two groups. None of the false negative infants was screened at 1 day of age, and 4 were screened at 2 days of age. Of the true positive infants, 13 were screened at 1 day of age and 45 at 2 days. Thus, 17 percent of the false negative infants and 19 percent of the true positive infants were screened before they were 3 days old.

The hypothesis of minimal PA elevations in infants with false negative results is based on the assumption of insufficient food intake before screening. The California experience (*12*) did not support this assumption; one-fourth of the true positive infants had low PA intake (0–199 mg) before they were tested, but none of the false negative infants was in this category.

What recommendations can be made from these observations? It is clear that priority must be given to optimizing and monitoring the reliability of testing. It is not economical or rational to allow the choice of a threshold value for presumptive positive results to be determined by less reliable procedures when high reliabilty can be attained. Further, to set a low threshold value is to engender a large volume of fruitless and costly followup work that may be poorly executed simply because of the volume. Followup of a false positive result raises concern in the parents about the health of the child, and this concern should not be disregarded. Choices must be determined in part by such local circumstances as the size and accessibility of the population and professional and monetary resources.

In New York City, there is every reason for raising the PA threshold value for presumptive positives to 6 mg/100 ml. Further, the combined data for New York State, New York City, Massachusetts, Pennsylvania, and Michigan suggest that the case detection rate for screening test results of 6 mg/100 ml is approximately 1 in 1.5 million (table 5). There is strong evidence that case detection at this level may approach zero as the reliability of test results improves. Therefore, a strong case can be made for raising the threshold value to 8 mg/100 ml, maintaining high reliability, and instituting a procedure for the immediate referral and evaluation of the small number of positive screening tests that would be generated.

References

1. MacCready, R. A.: Admissions of phenylketonuria patients to residential institutions before and after screening programs of the newborn infant. J Pediatr 85: 383–385 (1974).
2. Spencer, D. A.: Untreated phenylketonuria. Lancet 1: 1351 (1974).
3. Bush, J. W., Chen, M. M., and Patrick, D. L.: Health status index in cost effectiveness: Analysis of PKU program. *In* Health status indexes. Hospital Research and Educational Trust, Chicago, 1973, pp. 173, 181, 182.
4. Levy, H. L., Shih, V. E., Karolkewicz, V., and MacCready, R. A.: Screening for phenylketonuria. Lancet 2: 522–523 (1970).
5. Wainer, S. C., and Sideman, L.: Nine years of PKU screening in Pennsylvania. Health Lab Sci 11: 306–311 (1974).
6. Hansen, H.: Prevention of mental retardation due to PKU: Selected aspects of program validity. Prev Med 4: 310–321 (1975).
7. Read, S. S., Allen, R. J., and Haddy, T. B.: Phenylketonuria in newborns. Mich Med 68: 691–697 (1969).
8. Holtzman, N. A., Mellits, E. D., and Kallman, C. H.: Neonatal screening for phenylketonuria. II. Age dependence of initial phenylalanine in infants with PKU. Pediatrics 53: 353–357 (1974).
9. Holtzman, N. A., Meek, A. G., and Mellits, E. D.: Neonatal screening for phenylketonuria. I. Effectiveness. JAMA 229: 667–670 (1974).
10. Buist, M. R. M., and Penn, R. L.: Follow up screening for phenylketonuria. N Engl J Med 290: 577–578 (1974).
11. Hsia, D. Y. Y.: Screening tests for the detection of phenylketonuria. *In* Phenylketonuria and allied metabolic diseases. Children's Bureau, Department of Health, Education, and Welfare, 1967, pp. 50–59.
12. Dontanville, V. K., and Cunningham, G. C.: Effect of feeding on screening for PKU in infants. Pediatrics 51: 531–538 (1973).
13. Holtzman, N. A., Meek, A. G., Mellits, E. D., and Kallman, C. H.: Neonatal screening for phenylketonuria. III. Altered sex ratio; extent and possible causes. J Pediatr 85: 175–181 (1974).

SYNOPSIS

HANSEN, H. (Columbia University School of Public Health), SHAHIDI, S. A., and STEIN, Z. A.: *Screening for phenylketonuria in New York City. Threshold values reconsidered. Public Health Reports, Vol. 93, May–June 1978, pp. 246–251.*

In New York City from 1966 to 1970, almost all 736,469 newborns were screened for phenylketonuria (PKU). Among 1,094 infants with presumptive positive test results, 763 were followed up and 46 of them were judged to require preventive treatment. The considerable annual variation observed in the frequency of values of 4 to 6 mg/100 ml blood phenylalanine suggested low reliability at this level. Screening test results of 4 mg/100 ml, with no cases detected among them, represented 53 percent of all false positive results; newborns with 6 mg/100 ml results yielded 1 infant in need of treatment and accounted for 40 percent of the false positive results. The large volume of presumptive positive results generated by these levels presumably contributed to incomplete followup. If the threshold value for followup were raised, the effectiveness and the efficiency of the screening program could be improved. The experiences of other large PKU programs in the United States support these observations.