■ *A cooperative phenylketonuria screening program involving private nongovernmental laboratories, individual physicians and local and state health departments has been in operation for two years. The system has evolved to the point where practically all newborns are tested. The accuracy of laboratory work has been verified by an ongoing evaluation program which has resulted in continual improvement in level of performance. There are two areas in which some beneficial changes might be considered. One is the reduction of costs of the testing and follow-up by increasing volume and centralization of work. The other is greater cooperation of the medical community in collecting the data necessary to evaluate the program and expedite the final diagnosis.*

Two Years of PKU Testing in California

The Role of the Laboratory

GEORGE C. CUNNINGHAM, M.D., *Berkeley*

THE SCREENING OF NEWBORN INFANTS for phenylketonuria has become a standard procedure and in many cases is required by law. An important aspect of the effectiveness of such a program is the ability of the screening laboratory to perform the test reliably and expedite follow-up of suspicious results. This report will review the PKU program in California with emphasis on the role of the testing laboratory.

The recommendations of many authorities, including the American Academy of Pediatrics[1] and the U.S. Children's Bureau,[2] indicated that such screening should be performed in central laboratories. It was believed that such laboratories could more easily monitor their precision and accuracy. The volume of testing would reduce the cost of testing and a central reference point for information would allow for immediate and direct follow-up of positive screening results.

From the Hereditary Defects Unit, California State Department of Public Health, Berkeley.

Supported in part by a grant from the United States Children's Bureau.

Submitted 2 April 1968.

Reprint requests to: Hereditary Defects Unit, California State Department of Public Health, 2151 Berkeley Way, Berkeley 94704.

In California, the screening program has been implemented in a unique way. Rather than attempting a centralization of testing, the resources of hospital and private laboratories have been used. Any licensed clinical laboratory may apply to the State Department of Public Health for approval to participate as a screening laboratory. On the basis of available information, the Department has approved the microbiological inhibition assay (MIA)[3] and the fluorometric method[4] as the only techniques applicable to newborn screening. Laboratories report results of all positive and follow-up tests to the Department as a part of a comprehensive reporting system. The laboratories establish their own fee for the test, which is paid by the parents. This report summarizes the first two years' experience with regard to participation of the hospital, the laboratory and the physician.

Results of Program

At the end of the second year of the program, 39 cases of PKU had been detected and diagnosed as a result of screening 624,998 newborns. This is an incidence of one case per 16,000 infants tested. In 1966, 307,704 infants were tested and 16 cases detected. In 1967, 317,294 infants were tested and an additional 23 cases discovered.

This incidence is below that predicted from screening programs conducted in other states and foreign countries (1 in 10,000). It is not clear from the statistics relating to the other studies whether the higher incidence reported represented cases of high serum phenylalanine or true phenylketonuria. All diagnosed cases are carefully analyzed to exclude other causes of high serum phenylalanine. The incidence also reflects random occurrence of cases from birth cohort to birth cohort.

Hospital Participation

Information was requested from 436 hospitals and at completion of our tally all had supplied the requested information.

During the period from 1 January 1966 to 31 December 1966, the hospitals reported 312,032 live infants discharged, of which 307,704 had had the test, approximately 98.6 percent of the total reported. From 1 January 1967 to 31 December 1967, the hospitals reported 320,805 live infants discharged, of which 317,294 had had the test, approximately 98.9 percent of the total reported. There are no definitive statistics available on the number of live births in 1966, but the Bureau of Vital Statistics estimates there were 315,265 births in these hospitals. This represents reporting of 99 percent of the births.

The State Department of Public Health is taking appropriate action to keep the number of infants not tested to a minimum. The reasons infants were not tested are summarized in Table 1.

The law allows exception in the case of parental objection and medical contraindication to testing. Most of the infants listed in the table as "transfered to another hospital" were subsequently tested. Infants in categories 4, 5, 6 and 7 in Table 1—that is, those with inadequate specimens or with the test omitted in error before discharge from the hospital—have been referred to the local health departments and many of them may have received the test after discharge. While there is some degree of under-reporting of births and a few infants are not being tested, it is apparent that the program has been highly successful in insuring that the vast majority of infants receive the test.

Laboratory Participation

The laboratory is responsible for obtaining a reliable initial test and communicating this information to the physician and the Department. The California regulations provide that the physician be notified immediately by telephone and that he arrange for a repeat test within seven days. Assistance in this regard is available from the local health departments.

In 1966 the laboratories performed 311,953 initial tests, 1,162 of which (0.37 percent) were

TABLE 1.—*Distribution of Infants Not Tested*

Reason For Not Testing	*1966*		*1967*	
	Number	*Percent*	*Number*	*Percent*
1. Parental objection on religious grounds	252	(5.8%)	202	(5.8%)
2. Medical contraindication to testing	91	(2.1%)	197	(5.6%)
3. Transferred to another hospital for testing	773	(17.9%)	977	(27.8%)
4. Inadequate specimen obtained in hospital	200	(4.6%)	187	(5.3%)
5. Omission of test by error or no reason given	2528	(58.4%)	774	(22.1%)
6. Discharged "too early" for reliable test	283	(6.5%)	81	(2.3%)
7. Miscellaneous	201	(4.6%)	1093	(31.1%)
Total Infants Not Tested	4328	(100.0%)	3511	(100.0%)

positive. According to reports received from the laboratories, a repeat test was obtained on 861, or 74 percent of the infants with initial positive tests; in 134 cases the test was again positive. With the cooperation of local health departments, the Department obtained information on the disposition of these 134. All were determined to be temporary elevations and not phenylketonuric except for the 16 diagnosed PKU and five lost to follow-up because of death or inability to locate.

The large number of initial positive tests which were not repeated (299) is a product of inaccurate reporting by the laboratories and their failure to report repeats of initial positive tests to the local health departments in the first and second quarters. Health departments reported that repeat tests were done in all but 74 of the 905 cases in which initial positive tests were reported to them.

In 1967 the laboratories performed 334,831 initial tests, of which 905 or 0.27 percent were positive. The follow-up figures for this year are more accurate, as every initial positive reported by the laboratories was verified before being included in the count. At the close of the reporting period a repeat test had been obtained on 799 or 87 percent of the 905 initial positives, 117 of which were again positive. Twenty-four had been lost to follow-up and in 82 cases a repeat test was pending. All but one of the 117 infants with doubly positive tests were diagnostically evaluated. Twenty were confirmed as having PKU, eight are still being studied, and the rest were found to be not phenylketonuric. By the time these cases are closed, we anticipate follow-up tests will have been obtained on 90 percent or more. In order to estimate the effectiveness of follow-up, a thorough study of follow-up activity in the third quarter of 1966 was carried out. There were 302 initial positives reported in this period. Of these, 193 were repeat tested in the third quarter and 77 in the fourth quarter. There were 32, approximately 10 percent of the initial positive, that had not had repeat test results reported to the Department six months after the initial positive test. In ten of these instances the cases were closed because the parents moved without forwarding addresses, and one infant died in the interim. Considering the logistic difficulties involved, this study was interpreted as indicating a reasonably complete, if at times undesirably slow, follow-up of initial positives. More data is being collected to document the exact time lag between the first positive result and the follow-up test and results.

It is not possible to answer in a definitive fashion why follow-up of initial positives is sometimes so slow. In some instances the laboratory has not pursued its obligation to inform the physician and Department promptly, but in many cases the physicians are at fault. It requires special effort to arrange for retesting and interpretation of results to the family. Some physicians are not aware of their legal obligations and see no necessity of repeating a test when the positive level is low. Only if the physicians are conscientious and cooperative can the early detection of PKU be assured. It is important in evaluation of this kind of effort to have documented results on all babies with positive tests so that implementation can be made more effective and practical with a minimum of effort for all concerned.

With the cooperation of participating laboratories, the reporting forms have been improved and are currently being converted to a punched card data collection system to simplify monitoring and follow-up reports in this program.

Methods of Laboratory Evaluation

The specimens used to evaluate laboratory performance were prepared in the following manner: Outdated human blood was obtained from the blood bank, thoroughly mixed for homogeneity and divided into a suitable number of portions. One portion was used without further treatment. To the other portions, phenylalanine solution (200 mg per 100 ml) in isotonic saline solution was added in varying amounts to obtain concentrations desired, usually a high normal, approximately 3 mg per 100 ml, a low abnormal in the range 4 to 6 mg per 100 ml, and a high abnormal level, 7 to 10 mg per 100 ml, with respect to the screening level of 4 mg per 100 ml. The exact values used for each evaluation are given in Table 2.

All specimens were prepared, assigned random numbers, and mailed on a Monday to allow for analysis that week. Specimens were to be done on arrival at the laboratory and results returned by the following Monday. This was to minimize the effects of change of concentration with time, and to speed up the process of returning results to participants.

Laboratories approved for the microbiological inhibition assay method received filter paper soaked with whole blood. Laboratories approved

for the fluorometric test received liquid whole blood. The unknowns were supplied in limited amounts so that only one or two determinations could be run.

Evaluation Criteria

For the fluorometric test, where a numerical value is available, the best estimate of the true value was taken as the mean of all values reported by the participating laboratories after excluding very extreme values. The latter were defined as values beyond three standard deviations of all reported values around the mean. On a few occasions the distribution of values was such that, even after such correction, the mean was displaced by abnormal collections of values. In these instances the mean of 20 determinations by the Department's laboratory was used. An acceptable range was established by taking three standard deviations as calculated for the day-to-day test variability in the Department's laboratory. Each value reported by a laboratory was then classified as acceptable or unacceptable. Table 2 gives the values used to determine acceptability in each evaluation. Values outside of this range were regarded as unacceptable values.

Three standard deviations, as determined in our laboratory, were used instead of the conventional two standard deviations, since it was felt that until the laboratories had had some experience with the program, and some opportunity to accurately standardize their techniques, it would be unfair to prescribe the customary limits. On the other hand, use of the standard deviation between laboratories would have given an unacceptably high tolerance of error. It is anticipated that the limits of tolerance can now be narrowed as indicated by the results of our surveys.

All laboratories were given five unknowns on the first evaluation (March to May 1966). Laboratories with at least four values in the acceptable range were classified as satisfactory. On the second evaluation (5 December 1966), 13 unknowns were mailed, ten of them being from the same specimen. This gave four levels. In addition, a standard deviation was calculated from the ten identical specimens to give an estimate of precision. Acceptable values on at least four of these five criteria were classified as satisfactory. The July 1967 mailing consisted of six specimens, two from each of three levels.

Acceptable limits for the MIA method were more difficult to establish. Reporting was frequently in terms of ranges—for example, between 2 and 4 mg per 100 ml was reported "less than four." A single numerical value was assigned each report according to the following scheme:

Reading Reported	*Value Assigned (mg per 100 ml)*
0	Repeat
Less than 2	1
Equal to 2	2
Greater than 2 but less than 4	3
Equal to 4	4
Greater than 4 but less than 6	5
Equal to 6	6
Greater than 6 but less than 8	7
Equal to 8	8
Greater than 8 but less than 10	9
Equal to 10	10
Greater than 10 but less than 12	11
Equal to 12	12
Greater than 12 but less than 20	16
Equal to 20	20
Greater than 20	20+

All reported values were used to calculate the mean for the MIA technique. These means were found to correspond well with those obtained in the Department by the fluorometric technique and by laboratories performing fluorometric determinations on the same specimens, although the variations of the MIA results were much larger.

Ranges for satisfactory performance using the MIA technique are also detailed in Table 2. These were selected empirically, as no data were available on the reproducibility, in terms of standard deviation, for the method at the time of these evaluations. The standard deviation obviously increases with increasing concentration, and no valid figure

TABLE 2.—*Means and Acceptable Ranges of Phenylalanine Unknowns (mg per 100 ml)*

Evaluation Date	*MIA* Mean*	*Accepted Range*	*Evaluation Date*	*Fluorometric Mean*	*Accepted Range*
3-28-66	2.0	0- 3.9	4-18-66	1.0	0.7- 1.3
	5.3	4.0- 7.9		4.4	3.5- 5.3
	9.7	6.1-12.0		8.0	6.4- 9.6
4- 4-66	1.2	0- 3.4	4-25-66	1.8	1.5- 2.1
	5.1	4.0- 7.9		6.6	5.6- 7.6
	9.3	6.0-12.0		10.8	9.3-12.3
5- 2-66	1.1	0- 2.0	5-16-66	1.5	1.1- 2.1
	4.5	2.1- 7.9		4.6	3.6- 5.6
	8.7	6.1-12.0		7.2	5.2- 9.2
12- 5-66	1.4	0- 4.0	11-28-66	1.8	0.2- 3.4
	5.6	3.0- 8.0		3.6	2.2- 5.0
				5.2	3.1- 7.2
	10.6	6.0-20.0		9.7	6.6-12.9
6-26-67	1.4	0- 3.0	7-17-67	2.2	1.5- 2.7
	4.7	3.0- 7.0		6.0	4.2- 7.8
	5.6	4.0- 8.0		8.4	6.0-10.8
	8.1	5.0-11.0			

*MIA=microbiological inhibition assay.

could be used to calculate reasonable confidence limits. Laboratory performance is classified as unsatisfactory if unacceptable values are reported for more than one concentration level.

Results of Laboratory Evaluations

The first evaluation of fluorometric technique combined the results of March, April and May 1966. Of the 93 laboratories returning results, 19 (20.4 percent) were unsatisfactory. On the second evaluation (December 1966), of 85 reporting laboratories, 15 (17.7 percent) were unsatisfactory. The most recent evaluation (June 1967) had four (5.9 percent) of the 68 laboratories evaluated in the unsatisfactory category.

With the MIA technique, the figures were: Of 144 laboratories responding, 14 (9.7 percent) were unsatisfactory on the first evaluation. On the second evaluation, eight of the 134 laboratories reporting (6 percent) were unsatisfactory, and the latest results indicate only six (4.8 percent) of the 126 participating laboratories were in this category. These results, which do not include laboratories discontinuing participation or failing to return reports, are summarized in Table 3.

The first point that must be made is that unsatisfactory performance as defined in this report does not imply that overall operation or performance of the laboratory was unsatisfactory. Most laboratories experiencing difficulties were able to correct their procedure and perform acceptably on subsequent testing. Moreover, this was the first year of operation of this program and improvement has been seen with additional experience. Initially, several laboratories with minimal experience or limited demand for the determination requested approval. Many of these have discontinued participation. At the inception of the program (March 1966) there were 245 participating laboratories. This number has progressively decreased and the number of approved laboratories is now 174. The proportion using the fluorometric technique has decreased from approximately 40 percent to 30 percent. At first, roughly two-thirds of approved laboratories were performing fewer than ten tests per quarter. This proportion dropped to 3 percent last year.

TABLE 3.—*Data on Unsatisfactory Performance by Laboratories*

Evaluation Date	*Laboratories Participating*		*Unsatisfactory*	
			Number	*Percent*
March-April-May 1966	Fluor.	93	19	20.4
	MIA	144	14	9.7
	Total	237	33	13.9
November-December 1966	Fluor.	85	15	17.7
	MIA	134	8	6.0
	Total	219	23	10.4
June-July 1967	Fluor.	68	4	5.9
	MIA	126	6	4.8
	Total	194	10	5.3

In comparing results of the fluorometric and MIA methods, it must be borne in mind that the criteria for acceptable performance were broader for the latter group because of the difficulties of setting limits with a semi-quantitative technique. As data is accumulated on the MIA technique, it is anticipated that acceptable limits will be narrowed. While both methods give reasonably accurate results as means, the reproducility of the fluorometric method is superior. Attempts to measure intralaboratory variation yielded a coefficient of 20 percent for the MIA technique and 10 percent for the fluorometric technique. A detailed analysis of each method is now being prepared and standards will be established in the near future.

Efforts are currently under way to improve performance. To avoid excessive variation due to too many minor modifications, the Department has circulated the technical details of the method as performed in our laboratories. We have also developed a standardized method of reading and reporting the MIA test. Increased support of the program would allow the Department to experiment with better methods of sample presentation and with the provision of certified reference standards. The Department will attempt to monitor variability of commercial standards used in the program and study variables in procedure used. Workshops in fluorometric and MIA techniques are being planned, and it is hoped that increased staff will allow on-site visits to laboratories unable to correct problems on their own.

Cost-Benefit Analysis of Program

If the 39 cases detected had gone undetected, eventual care of the patients in state hospitals would have cost $6,318,000 based on an estimated 30-year minimum life expectancy in the institution at a cost of $5,400 per patient per year. In addition these persons will be productive taxpayers and their families will have been spared the tragedy of mental retardation. The tax-supported costs of this program, per case, have amounted to less than $2,500 for screening and $8,000 for ten years' treatment services per patient, the latter provided

through Crippled Children Services. The total cost of detection and treatment thus far amounts to less than half a million dollars.

In March of 1966, participating laboratories were requested to indicate their charges for performance of the test. The charge for the microbiological inhibition assay test varies from $0.75 to $6.00. The mean charge was $2.30. For the fluorometric method, the range was $1 to $12 with a mean of $3.03. A repeat survey done in September 1967 reveals that the mean charge for the MIA technique has increased to $2.98 with a range of $0.90 to $7.50, and the mean for the fluorometric test is $4.91 with a range of $0.75 to $15.00. Analysis of these reports indicates that some reduction in costs could be achieved by increased volume of testing. The costs of test performance on the total newborn population of approximately 312,000 amounted to a little less than $900,000 in 1966. This amount was paid by the parents of newborns as part of their overall hospital maternity costs. The salaries of personnel administering this program in the state and local health departments, and operating expenses, the cost of forms and the like have been estimated at approximately $90,000 the first year and $60,000 in 1967. The administrative costs were necessarily higher the first year than would be expected once the program was well established. These costs were funded by federal, state and local taxes.

These costs might be compared with those reported from some other states that use centralized testing. Massachusetts reported a cost of $50,000 to test 114,000 infants; New York, for 500,000 tests, requested $227,343 for laboratory support. The District of Columbia requested $35,545 to screen 15,000 births. New Jersey appropriated $10,000 in 1965 for 60,000 births, and Wisconsin $27,000 for 56,000. Rhode Island reported a cost of $1.14 per test in the first year of its program. The cost of administration is apparently not included in these figures. Illinois, with 200,000 births, appropriated $400,000 to pay for a non-centralized program similar in many respects to the one in California.

REFERENCES

1. A.M.A.: Statement on treatment of phenylketonuria, Committee on the Handicapped Child, Edw. Davens, Chairman, American Academy of Pediatrics, Pediatrics, 35:501-503, March 1965.

2. Children's Bureau Publication, 1967: Recommended Guidelines for PKU Programs.

3. Guthrie, R., and Susi, A.: A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants, Pediatrics, 32:338-343, September 1963.

4. McCaman, M. W., and Robins, E.: Fluorometric method for the determination of phenylalanine in serum, J. Lab. Clin. Med., 59:885-890, May 1962.

THE "AMPICILLIN RASH" QUANDARY

How can you identify a rash caused by ampicillin?

"The rash that you get from ampicillin is very much like the rash you get with a viral disease, and this is one of the disturbing things. You treat a child with a respiratory disease; you give him a drug like ampicillin; and within a few days, he may be better but he now has a rash. Is it due to the respiratory disease, due to a viral disease, or due to ampicillin? This is one of the reasons I don't like to use ampicillin in respiratory disease; once you have that rash, you can't tell whether it's due to the disease or due to the ampicillin. Ampicillin is such a valuable drug that it would be much better to use something else. Save it until when you really need it."

— BENJAMIN M. KAGAN, M.D., Los Angeles

Extracted from *Audio-Digest Pediatrics*, Vol. 14, No. 9, in the Audio-Digest Foundation's subscription series of tape-recorded programs.