of the patients were less than 3 years of age, 57.7 percent were under age 10, and 42.3 percent over age 10.

REFERENCE

(*1*) U. S. Public Health Service: Manual of serologic tests for syphilis. Supplement No. 22, to J. Ven. Dis. Inform., 1949, pp. 109–119.

A Statistical Evaluation Of the FPM Test

By AD HARRIS
SIDNEY OLANSKY, M.D.
HENRY MILLER, B.S.

A further evaluation of the efficiency of the filter paper microscopic (FPM) test (*1*) as a "detector" test for syphilitic infection, based on specimens from 276 donors, including the specimens from the 266 donors analyzed in the preliminary report (*2*) is presented in this paper. As noted in that report, collection and distribution of blood specimens were carried out by the Eastern Medical Center, Durham, N. C., and by the staff of the Venereal Disease Research Laboratory at the Alto Medical Center, Alto, Ga.

Each Tuesday, blood specimens from 10 to 30 donors were collected and five filter paper strips and five tubes of whole blood were prepared from each specimen. One filter paper strip and one tube of whole blood from each donor were sent to each of the following laboratories: Dr. Kahn, Dr. Kline, Mr. Mazzini, the Eastern Medical Center, and the Venereal Disease Research Laboratory. On the following Friday each laboratory performed the FPM test on the filter paper in accordance with the test protocol and, in addition, any modification of the FPM test that they might devise. The tube of whole blood was used to perform tests commonly used in the various laboratories, hereafter referred to as standard tests. Antigen used for the FPM and VDRL tests, if performed, was distributed by the Venereal Disease Research Laboratory.

At the time of collection and distribution of specimens, the Eastern Medical Center and the Alto Medical Center established a diagnosis by clinical and serologic findings of all donors from whom the specimens for this study were taken.

For purposes of evaluation, two methods of comparison of the FPM test are presented. The first method discussed is that of comparing the test results of each participating laboratory with established diagnostic results. The second method disregards diagnostic findings and compares the results of the FPM test findings in a particular laboratory with the other tests performed in that laboratory. The second method of evaluation has been used for the following reasons: (*a*) A diagnosis by clinical means is not always obtainable in actual practice; (*b*) an error in diagnosis is possible. Certain specimens collected from donors diagnosed as positive or doubtful have been found to be negative by the standard tests. The agreement by all tests in a laboratory as to the negativity of these particular (diagnosed positive) specimens ranges from 2 specimens in the Venereal Disease Research Laboratory to 18 specimens in the Kahn laboratory. This type of disagreement (diagnosed positive, tested negative) may be due to laboratory technique or to an error in diagnosis. In either case a comparison of tests within laboratories seems justifiable.

In both methods, comparison of specimens as to agreement or disagreement is made only when tests have been performed. Tests giving doubtful reactions are considered positive, since it is not the purpose of this study to analyze the quantitative results produced by the FPM test.

Method 1

The various diagnoses of syphilis, based on clinical and serologic findings, established by the Eastern and Alto Medical Centers have been classified as follows:

Mr. Harris is a serologist and the assistant director of the Venereal Disease Research Laboratory; Dr. Olansky is the director; Mr. Miller is a statistician in the Division of Venereal Disease, Public Health Service.

1. Positive—secondary; early latent; late latent; asymptomatic central nervous system (CNS); symptomatic CNS (paresis, tabes, etc); cardiovascular; CNS plus cardiovascular syphilis (CVS); congenital; serorelapse; infectious relapse.

2. Negative—normal; other venereal diseases, not syphilis.

Since clinical manifestations were not considered by the testing laboratories in determining results, and negative blood tests are highly probable in patients with primary syphilis, the following diagnoses of the centers were eliminated from comparison: (*a*) primary; (*b*) reinfection regardless of stage of syphilis; and (*c*) suspect only or late syphilis.

Table 1 shows the percent of agreement and disagreement of the serologic test results obtained in each of the five testing laboratories as compared with the diagnoses of the Alto and Eastern Medical Centers. Tables giving the serologic test results as compared with the diagnoses made in these laboratories may be obtained from the authors upon request.

Eight different tests were performed in the Kline laboratory, three of which were modifications of the FPM test. Of 231 specimens used for comparison between the FPM test and the diagnosis, 161 were found to be positive and 13 negative by both the FPM test and the diagnosis. The percent of agreement is 75.3 percent. Fifty-six specimens were considered positive by the diagnoses, negative by the FPM test. The rate of disagreement is 24.7 percent. The FPM test modifications agree rather closely with those of the FPM test. The highest agreement found between the FPM test or any of its modifications is 81.8 percent for the FPM modification test. The difference between this rate of agreement and the lowest rate of agreement between the diagnosis and a standard test (88.7 percent for the Kline diagnostic) is statistically significant at the 5-percent level. This difference may be attributed to the inability of the FPM test and its modifications to identify specimens diagnosed as positive.

In the Venereal Disease Research Laboratory, the FPM test agreed with the diagnosis on 168

Table 1. Percent agreement and disagreement of serologic test results of five laboratories with the established diagnostic results [1]

Serologic test	Kline laboratory		Venereal Disease Research Laboratory		Eastern Medical Center		Mazzini laboratory		Kahn laboratory	
	Relation to diagnosis		Relation to diagnosis		Relation to diagnosis		Relation to diagnosis		Relation to diagnosis	
	Agreement	Disagreement	Agreement	Disagreement	Agreement	Disagreement	Agreement	Disagreement	Agreement	Disagreement
FPM	75.3	24.7	72.7	27.3	71.2	28.8	91.3	8.7	77.9	22.1
FPM Dri-Rite	------	------	------	------	77.3	22.7	------	------	------	------
FPM Mazzini cardiolipin	------	------	------	------	------	------	93.5	6.5	------	------
FPM modification	81.8	18.2	------	------	------	------	------	------	------	------
FPM modification standard	81.4	18.6	------	------	------	------	------	------	------	------
FPM standard	74.9	25.1	------	------	------	------	------	------	------	------
Kahn presumptive	------	------	------	------	------	------	------	------	90.2	9.8
Kahn standard	------	------	91.4	8.6	90.0	10.0	------	------	88.3	11.7
Kline diagnostic	88.7	11.3	------	------	------	------	------	------	------	------
Kline exclusion	94.8	5.2	------	------	------	------	------	------	------	------
Kline standard	95.7	4.3	97.2	2.8	------	------	------	------	------	------
Kolmer simplified	------	------	91.0	9.0	87.7	12.3	------	------	------	------
Mazzini cardiolipin	------	------	------	------	------	------	96.9	3.1	------	------
Mazzini flocculation	------	------	94.4	5.6	------	------	------	------	------	------
Rein-Bossak	------	------	96.3	3.7	------	------	------	------	------	------
VDRL slide	92.6	7.4	93.9	6.1	89.1	10.9	96.1	3.9	------	------

[1] Based on clinical and serologic findings at Eastern and Alto Medical Centers.

NOTE.—More detailed data may be obtained from authors.

of the 231 specimens used for comparison, a rate of agreement of 72.7 percent. With the exception of one comparison, disagreement between FPM and diagnostic results was due to 62 specimens diagnosed positive but negative by the FPM test. The highest degree of agreement in this laboratory is 97.2 percent and is found under the results of the Kline standard test. Only 2.8 percent were found to be in disagreement. All differences in the rate of agreement between the FPM test and the standard tests are significant beyond the 5-percent level, again showing the difference in the ability of the FPM and the standard tests to recognize positive diagnosed specimens.

The results of tests in the Eastern Medical Center laboratory and the Kahn laboratory are in accord with results in the laboratories already discussed. However, special attention is called to the comparison of diagnostic results and results of tests performed in the Mazzini laboratory. The rate of agreement between diagnostic results and the FPM test and a modification, the FPM-Mazzini-cardiolipin test (*2*) using the Mazzini antigen instead of the VDRL antigen, is approximately 20 percent greater than any other FPM test or FPM test modifications found in the other participating laboratories. The percent of agreement in the FPM test as compared to the agreement in the VDRL slide and the Mazzini-cardiolipin tests is significantly different at the 5-percent level. However, there is no significant difference between the FPM-Mazzini-cardiolipin test and the VDRL slide and Mazzini-cardiolipin tests. This high degree of agreement between diagnostic results and the FPM and FPM-Mazzini-cardiolipin test results was also noted in the previous report (*2*). In answer to an inquiry by the authors of that article, Mr. Mazzini reported that the FPM test performed in his laboratory was being carried out in accordance with the testing procedure.

Under this first method of comparison, the laboratory results of the FPM test and the standard tests have shown a significant difference in agreement with the diagnoses. Although a few specimens diagnosed as negative are positive by the FPM test, the greatest difference may be attributed to the inability of the FPM test to identify specimens from donors in positive diagnostic categories.

Method 2

Results of method 2 are presented in table 2. By this procedure the results of the FPM test are compared with respect to agreement or disagreement with the various tests performed within each laboratory. To determine what variation could be expected within a laboratory, two standard testing methods whose efficiency is generally known are compared and will be referred to as control groups. Data from these groups are presented in table 3.

In the Kline laboratory, the results of comparison of tests by this method show a high degree of agreement between the FPM test and its modifications. This may be due to the similarity of reactivity between these tests. The FPM-FPM standard comparison reveals almost perfect agreement—99.6 percent. However, the agreement between the FPM test and the standard tests in this laboratory is quite low. The greatest amount of agreement in these comparisons is 84.2 percent in the FPM-Kline diagnostic comparison. The lowest agreement (78.0 percent) exists in the FPM-Kline standard and the FPM-Kline exclusion comparisons, and is attributable in each case to 60 specimens diagnosed negative by the FPM test but positive by the two Kline tests.

The tests used as a control group for this laboratory (VDRL slide-Kline standard) show an evident similarity, with an agreement rate of 96.3 percent (table 3). While perfect agreement cannot be expected, this does indicate rather definitely that the low rate of agreement between the FPM–VDRL slide and the FPM-Kline standard tests is not due to the fact that the VDRL slide and Kline tests lack specificity, but rather that the FPM test is inefficient in the identification of positive serologic specimens.

In the Venereal Disease Research Laboratory, no modifications of the FPM test were performed but, as in the Kline laboratory, the rate of agreement between the FPM test and the standard tests is rather low, lower than one might expect if these tests are comparable. The greatest amount of disagreement occurs in the

Table 2. Percent agreement and disagreement of blood specimen results obtained by the FPM test [1] and by other tests performed in five laboratories

Serologic test	Kline laboratory		Venereal Disease Research Laboratory		Eastern Medical Center		Mazzini laboratory		Kahn laboratory	
	Relation to FPM test		Relation to FPM test		Relation to FPM test		Relation to FPM test		Relation to FPM test	
	Agreement	Disagreement	Agreement	Disagreement	Agreement	Disagreement	Agreement	Disagreement	Agreement	Disagreement
FPM Dri-Rite					93. 3	6. 7				
FPM-Mazzini-cardiolipin							96. 3	3. 7		
FPM modification	93. 1	6. 9								
FPM modification standard	93. 4	6. 6								
FPM standard	99. 6	0. 4								
Kahn presumptive									85. 0	15. 0
Kahn standard			79. 2	20. 8	78. 3	21. 7			88. 2	11. 8
Kline diagnostic	84. 2	15. 8								
Kline exclusion	78. 0	22. 0								
Kline standard	78. 0	22. 0	72. 6	27. 4						
Kolmer simplified			79. 8	20. 2	79. 1	20. 9				
Mazzini-cardiolipin							92. 6	7. 4		
Mazzini flocculation			74. 1	25. 9						
Rein-Bossak			74. 7	25. 3						
VDRL slide	81. 0	19. 0	77. 6	22. 4	79. 8	20. 2	93. 3	6. 7		

[1] As performed in accordance with the test protocol.

NOTE. More detailed data may be obtained from authors.

FPM-Kline standard comparison, indicating that the agreement between the two tests was 72.6 percent; disagreement, 27.4 percent. Disagreement is due to 71 specimens indicated positive by the Kline test and negative by the FPM test. The control group (Kahn standard-Kline standard) indicates quite a difference in agreement in contrast to the FPM comparisons. Here the total agreement is 93.3 percent. Statistical tests carried out between the FPM-Kahn standard and the FPM-Kline standard comparisons showed no statistical difference. However, the difference between the Kahn standard-Kline standard comparison and the other two comparisons mentioned is highly significant, due to the large number of positive specimens called negative by the FPM test.

At the Eastern Medical Center laboratory, the control group of Kahn standard and Kolmer simplified tests shows a total agreement of 95.2 percent, whereas the highest agreement between the FPM test and any standard test performed in this laboratory is only 79.8 percent (VDRL slide test). The FPM Dri-Rite test indicated 93.3-percent agreement with the FPM test. For further evidence of reactivity of a modification of the FPM test, the FPM Dri-Rite test and the Kahn test were compared and an agreement of 83.0 percent was found. This was not significantly different from the results

Table 3. Percent agreement and disagreement of control groups within laboratories

Laboratory	Control groups	Percent agreement	Percent disagreement
Kline	VDRL slide v. Kline standard	96. 3	3. 7
Venereal Disease Research Laboratory	Kahn standard v. Kline standard	93. 3	6. 7
Eastern Medical Center	Kahn standard v. Kolmer simplified	95. 2	4. 8
Mazzini	VDRL slide v. Mazzini-cardiolipin	99. 3	. 7
Kahn	Kahn standard v. Kahn presumptive	97. 0	3. 0

of comparisons of the FPM test with the other standard tests in this laboratory, indicating the similarity of the two FPM tests in their inability to agree with other tests in identifying positive specimens.

The results at the Mazzini laboratory show a higher percentage of agreement between the FPM test and standard tests than was obtained by any of the other laboratories. The agreement of FPM results with the diagnoses established by the medical centers was also much higher in this laboratory than in the other testing laboratories. However, the control group of VDRL slide-Mazzini-cardiolipin tests is also in high agreement (99.3 percent) which is significantly different (5-percent level) from the FPM-VDRL slide (93.3 percent) and the FPM-Mazzini-cardiolipin comparison (92.6 percent), again indicating the lack of reactivity in the FPM test (table 3).

The comparison of results from the Kahn laboratory shows somewhat higher agreements between standard tests (with the exception of the Mazzini laboratory) than those obtained in the other three laboratories. However, the conclusion reached from results obtained by all reporting laboratories appears to be the same—the failure of the FPM test to identify positive specimens. The degree of agreement between the control groups can be seen in table 3. These five groups of tests have an average agreement of 96.3 percent compared to 81.0 percent between the FPM and standard tests.

Since the FPM test is less reactive than the commonly employed laboratory tests, it would be less efficient as a "detector" test for field survey use. However, as stated in the previous report (*2*), this fact should not militate against the use of the FPM test under conditions in which another type of blood collection cannot be used.

Summary

1. Filter paper strips and tubes of whole blood from 276 donors were distributed to 5 laboratories.

2. Each laboratory performed the FPM test in accordance with the testing protocol and any modification of the FPM test if they so desired. The tube of whole blood was used to perform other testing procedures employed in the laboratories.

3. Two methods of analysis are presented. Both methods show that a significant difference (5-percent level) exists between the ability of commonly employed laboratory tests and the FPM test to identify positive and doubtful serologic specimens.

REFERENCES

(*1*) Hogan, Ralph B., and Busch, Shirley: Filter paper microscopic test for syphilis, or the FPM-test. A preliminary report. J. Ven. Dis. Inform. 31: 37–45 (1950).

(*2*) Harris, A., and Olansky, Sidney: A study of the filter paper microscopic (FPM) test for syphilis. Preliminary report. J. Ven. Dis. Inform. 32: 1–4 (1951).

Use of the FPM Test In a Control Program

By CHARLES R. FREEBLE, Jr., M.D.
BERTTINA ORSBURN, B.S.

When the filter paper microscopic (FPM) test for syphilis was reported by Hogan and Busch (*1*), it appeared to present a simple method of collecting, shipping, and testing blood. This method seemed especially desirable in screening children for congenital syphilis, since it obviated the necessity for jugular punctures. In Ohio, the test has been studied from the standpoint of ease of use under field conditions and by comparison with the standard serologic tests currently being performed by the Ohio Department of Health.

Use in Mass Testing Programs

The FPM technique is suitable for use in mass testing programs. Although more time is required to obtain a specimen of blood for the

Dr. Freeble and Mrs. Orsburn are with the communicable disease division of the Ohio Department of Health.