JOURNAL OF CLINICAL MICROBIOLOGY, Feb. 1976, p. 105-109
Vol. 3, No. 2

Printed in U.S.A.

Computer-Assisted Bacterial Identification Utilizing Antimicrobial Susceptibility Profiles Generated by Autobac 1

BRUCE H. SIELAFF,[1] * EUGENE A. JOHNSON, AND JOHN M. MATSEN[2]

Division of Biometry and Health Information Systems, School of Public Health and Departments of Laboratory Medicine and Pathology, Pediatrics, and Microbiology, University of Minnesota, Minneapolis, Minnesota 55455

Received for publication 5 September 1975

A computer program was developed to identify bacteria solely on the basis of their relative susceptibility to various antimicrobial agents. A sample of 481 clinical isolates from nine of the most commonly isolated gram-negative groups was identified by the quadratic discriminant function technique. Various combinations of antimicrobials were tried, and one set of 18 resulted in a more than 97% correlation with conventional identification procedures. The antimicrobial set could be decreased to 14, while a better than 95% correlation with the conventional procedures was maintained.

Human error and delay in reporting results, in any laboratory science, are undesirable features of the diagnostic process. Human error has been minimized in other scientific and administrative endeavors through the use of both automation and computerization. Delay in reporting results in clinical microbiology has been addressed by mechanization and by rapid methods. This study outlines the use of a semiautomated mechanized device (Autobac 1) for antimicrobial susceptibility testing, coupled with a computerized program for utilizing the data from the susceptibility profile for organism identification. Since these methodologies are accomplished in the span of 3 to 5 h, the potential usefulness of this system is evidenced by the reduction in human error and acceleration of the reporting process for both the identification of species and results of antimicrobial susceptibility tests.

The concept of using antibiotic susceptibilities for bacterial identification is not entirely new. In 1971, Gilardi (5) found that susceptibility profiles could be used to assist in the identification of lactose-nonfermenting, gram-negative bacteria. He used susceptibility information, for 16 antibiotics, obtained from the disk agar diffusion method. Sutter and Finegold (11) used susceptibility profiles to place gram-negative anaerobic bacilli into five different groups. Further testing was then required to complete the identification.

A Baysean mathematical model was used by Friedman and MacLowry (4) to classify bacteria. Their data base contained probability data on the susceptibility profiles of 31 species of bacteria. This data was collected over a period of several years. When 1,000 clinical isolates were classified by this method, there was an 86% agreement with the identification obtained by conventional biochemical procedures.

MATERIALS AND METHODS

Source of organisms. The bacteria used in this study were obtained, where possible, from fresh clinical isolates from the clinical microbiology laboratory at the University of Minnesota Hospitals. Within each bacterial classification, the cultures were selected at random from the fresh isolates. Where sufficient fresh isolates were not available, stock cultures were also used. Here again, the cultures were selected at random. The organisms were identified in the clinical laboratory by conventional biochemical procedures (2, 3).

Source of antimicrobial agents and media. The antimicrobial agents, in the form of elution disks and diffusion disks, were supplied by Pfizer, Inc. The tubes of phosphate-buffered saline and tubes of eugonic broth to be used in the Autobac were also provided by Pfizer, Inc.

Method used for determining susceptibility profiles. The susceptibility of the bacteria to the various antibiotics was determined through the use of an Autobac 1 (Pfizer, Inc.) (8, 10). The Autobac 1 is a semiautomated antimicrobial susceptibility testing system. It combines the speed of automation with the flexibility of manual procedures. It is designed to determine the susceptibility of a bacteria to a panel of up to 12 antimicrobial agents simultaneously. The result for each antimicrobial is an index of susceptibility called the "light scatter index"

[1] Present address: Pfizer, Inc., Groton, Conn. 06340.
[2] Present address: University of Utah, College of Medicine, Salt Lake City, Utah 84132.

(LSI). The index runs from 0.00 (resistant) to 1.00 (susceptible) in increments of 0.01. The test results, except for certain slow-growing organisms, are available within 3 to 5 h after the test is begun. It is the LSI values that are used in the following method.

Statistical technique used for identification. Identification is accomplished by means of a multivariate statistical technique known as the quadratic discriminant function. The quadratic discriminant function is based on the multivariate normal model. Two adjacent, intersecting multivariate normal distributions have a point of equal probability in the overlap region. This point of equal probability can be used as a boundary for classification. To classify an individual, all that is necessary is to determine on which side of the boundary the individual falls. This equi-probability boundary minimizes mis-classification, assuming that the size of the populations are both approximately equal. If the two populations are of greatly different size, then the proportion of each population that is mis-classified will be minimized, but the total number of mis-classifications will not be minimized. An adjustment for the difference in population size would have to be made.

The procedure begins by computing the covariance matrices for the different groups. The formula for calculating the elements of the covariance matrix is the standard covariance formula as follows:

$$s_{x_i x_j} = \frac{n \sum_{k=1}^{n} x_{ik}x_{jk} - \sum_{k=1}^{n} x_{ik} \sum_{k=1}^{n} x_{jk}}{n^2 - n} \quad (1)$$

where $s_{x_i x_j}$ is the covariance between variables x_i and x_j (when $i = j$, the formula reduces to that of the variance of variable x_i) and n is the number of observations on variables x_i and x_j (this number is the same for both x_i and x_j; it represents the number of individuals in the data set). These matrices contain the variances and covariances for the different variables in each group.

Also calculated are mean vectors for each group. They contain the means for the different variables in each group. If the set of variables used for classification is changed, then a new set of matrices must be constructed. These matrices will be constructed using only the LSI values for those variables in the new variable set.

For classifying an individual into one of *NG* groups, the following function is calculated for each group:

$$f(X)_i = p_i (2\pi)^{-\frac{NV}{2}} |S_i|^{-1/2} e^{-\frac{q_i}{2}} \quad (2)$$

where p_i can be either the proportion of the sample in group i or the prior probability of group i; NV is the number of variables; $|S_i|$ is the determinant of the covariance matrix for group i; $q_i = (X - \bar{x}_i)' S_i^{-1} (X - \bar{x}_i)$, where X is the vector of LSI values for the organism to be classified; $\bar{x}_i$ is the mean vector for the i^{th} group; $'$ means the matrix transpose; and S_i^{-1} is the inverse of the covariance matrix for the i^{th} group. It can be seen that equation 2 without the p_i is just the probability density function for the multivariate normal model.

When this function has been computed for all groups, the group with the greatest value for $f(X)$ is selected as the group identification for the unknown organism. In actual practice, since the relative magnitude, used for comparison between groups, is important rather than the actual value, the constant factor involving π is eliminated from the calculations.

In the real world environment, most multivariate distributions are not normal. Fortunately, the quadratic discriminant function procedure is very robust. Very good classification rules can be generated for populations that are very nonnormal. For further discussion of multivariate normality and the quadratic discriminant functions, see Anderson (1), Grams et al. (6, 7), and Michaelis (9).

RESULTS

A total of 31 antimicrobial agents were investigated. Because multiple concentrations of certain agents were used, a total of 48 possible variables (Table 1) were examined in the course of these studies. Many of these variables were eliminated because they gave no information useful for differentiation. Most of the strains of all groups were either all susceptible or all resistant to the antimicrobial at that concentration. Other variables provided redundant information (the same information as another variable) and could, therefore, be eliminated. Table 2 shows the subset of the total variable set that showed the most promise.

The antimicrobials were tested against a sample of 481 bacterial isolates. The composition of this sample is shown in Table 3. These results were used to construct the matrices used for classification. When the 481 isolates in Table 3 were classified according to the 18 antimicrobials in Table 2, there was greater than 97% agreement with the identification arrived at by the clinical laboratory by conventional identification procedures. These results can be seen in Table 4. Nearly 80% of the disagreements consisted of organisms identified as *Citrobacter* or *Enterobacter* by the clinical lab and something else by the susceptibility profile. These two genera were consistently found to be the most difficult to identify.

An attempt was made to decrease the number of variables to a minimum without a large sacrifice in percentage of agreement. Various subsets of different sizes of the 18 variables were tried. Table 5 shows the results of a particular subset of 14 variables. There is a loss in percentage of agreement of less than 2%. As the number of variables was decreased below 14, the percentage began to drop rapidly. For prac-

TABLE 1. *Antibiotics investigated*

Agent	Disk mass (μg)	Agent	Disk mass (μg)	Agent	Disk mass (μg)
Ampicillin	3.6	Kanamycin	5.0	Penicillin G	0.2[a]
Bacitracin	18.0[a]	Kanamycin	18.0	Penicillin G	2.0[a]
Carbenicillin	50.0	Lincomycin	2.0	Penicillin G	10.0[a]
Carbenicillin	120.0	Methenamine mandelate	1.0[b]	Polymyxin B	12.5[a]
Cephalothin	15.0	Methacycline	30.0	Polymyxin B	50.0[a]
Chloramphenicol	5.0	Methicillin	5.0	Polymyxin B	300.0[a]
Clindamycin	2.0	Nafcillin	1.0	Streptomycin	2.0
Cloxacillin	1.0	Nalidixic acid	5.0	Streptomycin	10.0
Colistin	2.0	Nalidixic acid	15.0	Streptomycin	20.0
Colistin	13.0	Neomycin	5.0	Tetracycline	0.5
Doxycycline	0.5	Neomycin	20.0	Tetracycline	1.5
Doxycycline	1.6	Nitrofurantoin	15.0	Trimethoprim/sulfamethoxazole	1.25 23.75
Erythromycin	2.5	Novobiocin	5.0	Vancomycin	10.0
Erythromycin	15.0	Novobiocin	30.0	Vancomycin	30.0
Furizolidone	100.0	Oleandomycin	6.0	Viomycin	2.0
Gentamicin	9.0	Oleandomycin	15.0	Viomycin	10.0

[a] Mass measured in units.
[b] Mass measured in milligrams.

TABLE 2. *Antibiotic subset*

Agent	Disk mass (μg)	Agent	Disk mass (μg)
Ampicillin	3.6	Kanamycin	5.0
Bacitracin	18.0[a]	Methenamine mandelate	1.0[b]
Carbenicillin	50.0	Nalidixic acid	5.0
Carbenicillin	120.0	Neomycin	5.0
Cephalothin	15.0	Nitrofurantoin	15.0
Colistin	2.0	Novobiocin	30.0
Colistin	13.0	Polymyxin B	50.0[a]
Erythromycin	15.0	Streptomycin	10.0
Furizolidone	100.0	Tetracycline	0.5

[a] Mass measured in units.
[b] Mass measured in milligrams.

TABLE 3. *Composition of sample*

Organism	No. of organisms
CITROB[a]	50
ENTEROB	48
ECOLI	75
HEREL	35
KLEB	59
PROTMIR	49
PROTOTH	51
P. morganii	(19)[b]
P. rettgeri	(17)
P. vulgaris	(15)
PSEUDO	62
P. aeruginosa	(35)
P. fluorescens	(15)
P. maltophilia	(12)
SERRAT	52

[a] CITROB, *Citrobacter;* ENTEROB, *Enterobacter;* ECOLI, *Escherichia coli;* HEREL, *Herellea;* KLEB, *Klebsiella;* PROTMIR, *Proteus mirabilis;* PROTOTH, indole-positive *Proteus;* PSEUDO, *Pseudomonas;* SERRAT, Serratia.
[b] Numbers in parentheses not included in totals of organisms tested.

tical purposes, therefore, that set of 14 variables shown in Table 5 appears to represent the minimum subset of the antimicrobials studied to date that is capable of providing a high level of agreement with conventional identification procedures.

DISCUSSION

This study was undertaken to determine the feasibility of using susceptibility profiles for the identification of bacteria. Within the scope of the study, this objective was successfully accomplished; the feasibility of this method was proved. To prevent dilution of effort in a feasibility study, it was necessary to limit the number of groups studied. Gram-negative bacteria were focused on for several reasons. First, since the Gram stain is fairly quick and easy to perform, differentiation of this group can be accomplished with relative ease. Second, since the members of this group are generally the most difficult to identify, they would provide the greatest test for the proposed identification system. Lastly, gram-negative bacteria comprise the vast majority of organisms currently

TABLE 4. *Group affiliation by susceptibility profile and conventional procedures – 18 variables*[a]

Group affiliation by conventional procedures	Group affiliation by susceptibility profile								
	CITROB[b]	ENTEROB	ECOLI	HEREL	KLEB	PROTMER	PROTOTH	PSEUDO	SERRAT
CITROB	45	2	2	0	1	0	0	0	0
ENTEROB	4	43	1	0	0	0	0	0	0
ECOLI	0	1	73	0	0	0	0	0	1
HEREL	0	0	0	35	0	0	0	0	0
KLEB	0	0	0	0	59	0	0	0	0
PROTMIR	0	0	0	0	0	51	0	0	0
PROTOTH	0	0	0	0	0	1	48	0	0
PSEUDO	0	0	0	0	0	0	0	62	0
SERRAT	0	0	0	0	0	0	0	0	52

[a] See Table 2. Percentage of agreement between susceptibility profile and conventional procedures was 97.3%.

[b] For abbreviation, see Table 3.

TABLE 5. *Group affiliation by susceptibility profile and conventional procedures – 14 variables*[a]

Group affiliation by conventional procedures	Group affiliation by susceptibility profile								
	CITROB[b]	ENTEROB	ECOLI	HEREL	KLEB	PROTMIR	PROTOTH	PSEUDO	SERRAT
CITROB	42	4	2	0	1	0	1	0	0
ENTEROB	2	44	2	0	0	0	0	0	0
ECOLI	1	1	72	0	0	0	0	0	0
HEREL	0	0	0	35	0	0	0	0	0
KLEB	0	1	1	0	57	0	0	0	0
PROTMIR	0	0	0	0	0	51	0	0	0
PROTOTH	0	0	0	0	0	1	48	0	0
PSEUDO	0	0	0	0	1	0	0	61	0
SERRAT	0	1	0	0	0	0	1	0	50

[a] Variables included ampicillin, bacitracin, carbenicillin 50 and 120, cephalothin, colistin 2 and 13, erythromycin 15, kanamycin 5, methenamine mandelate, neomycin 5, nitrofurantoin, novobiocin 30, and tetracycline 0.5. Percentage of agreement with conventional procedures was 95.6%.

[b] For abbreviations, see Table 3.

identified in the clinical microbiology laboratory.

As with most feasibility studies, questions arose during the course of the study. The principal question regarded the alteration of the susceptibility profile due to acquired resistance. This problem was not encountered during the course of the study, but nonetheless, the spector of mis-identifications due to this cause remains. Acquired resistance occurs because resistant mutants are selected for by widespread use of an antimicrobial agent. This resistance can then be transferred through r factors.

Obviously, to have a practical identification system, the number of groups included must be increased. Additional gram-negative genera, as well as gram-positive genera, will have to be added. It would also be extremely valuable if more speciation within the genera could also be accomplished. These are areas that need further investigation. Work is currently progressing in our laboratories to address the abovementioned points. A large number of non-antibiotic chemical compounds are being examined for their ability to differentiate bacterial groups by differential inhibition of growth. One of the criteria for selection is that the compound not be commonly used in the clinical setting to minimize the possibility of acquired resistance. If bacteria do not normally encounter an agent

in their environment, acquired resistance should not become a problem. We have also more than doubled the number of bacterial groups being studied, both by additional genera and increased speciation. It is our intention to increase the number of groups even further.

In summary, this study has shown that identification of bacteria using their relative susceptibility to various antimicrobial agents is a practicable approach. If the system is to prove usable in the clinical setting, though, there are questions that must be addressed and a large amount of additional work that must be performed.

ACKNOWLEDGMENTS

We wish to acknowledge the staff of the clinical microbiology laboratories at the University of Minnesota Hospitals for their cooperation and technical skill.

We wish to thank the Biotechnology Resources Branch of the National Institutes of Health who contributed support for the computer resource at the University of Minnesota where the data was processed.

This study was supported by a research grant from Pfizer, Inc.

LITERATURE CITED

1. **Anderson, T. W.** 1958. An introduction of multivariate statistical analysis. John Wiley & Sons, Inc., New York.
2. **Blair, J. E., E. H. Lennette, and J. P. Truant (ed.).** 1970. Manual of clinical microbiology. American Society for Microbiology, Bethesda, Md.
3. **Edwards, P. R., and W. H. Ewing.** 1955. Identification of enterobacteriaceae, 2nd ed. Burgess Publishing Co., Minneapolis, Minn.
4. **Friedman, R., and J. MacLowry.** 1973. Computer identification of bacteria on the basis of their antibiotic susceptibility patterns. Appl. Microbiol. 26:314-317.
5. **Gilardi, G. L.** 1971. Antimicrobial susceptibility as a diagnostic aid in the identification of non-fermenting gram-negative bacteria. Appl. Microbiol. 22:821-823.
6. **Grams, R. R., E. A. Johnson, and E. S. Benson.** 1972. Laboratory data analysis system: section III—multivariate normality. Am. J. Clin. Pathol. 58:188-200.
7. **Grams, R. R., E. A. Johnson, and E. S. Benson.** 1972. Laboratory data analysis system: section IV—multivariate diagnosis. Am. J. Clin. Pathol. 58:201-207.
8. **McKie, J. E., Jr., R. J. Borovoy, J. F. Dooley, G. R. Evanega, G. Mendoza, F. Meyer, M. Moody, D. E. Packer, J. Praglin, and H. Smith.** 1975. Autobac 1-A 3-hour automated antimicrobial susceptibility system. II. Microbiological studies. *In* C. Heden and F. Illeni (ed.), Automation in microbiology and immunology. John Wiley & Sons, Inc., New York.
9. **Michaelis, J.** 1973. Simulation experiments with multiple group linear and quadratic discriminant analysis. *In* T. Cacoullos (ed.), Discriminant analysis and applications. Academic Press, Inc., New York.
10. **Praglin, J., A. C. Curtiss, D. K. Longhenry, and J. E. McKie, Jr.** 1975. Autobac 1-A 3-hour automated antimicrobial susceptibility system. I. System description. *In* C. Heden and F. Illeni (ed.), Automation in microbiology and immunology. John Wiley & Sons, Inc., New York.
11. **Sutter, V. L., and S. M. Finegold.** 1971. Antibiotic disc susceptibility tests for rapid presumptive identification of gram-negative anaerobic bacilli. Appl. Microbiol. 21:13-20.