Use of computers in clinical electrocardiography: an evaluation

H.K. WOLF, PH D; R.D. GREGOR, MD, FRCP[C]; B.M. CHANDLER, MD, FRCP[C]

The use of computers in clinical electrocardiography is increasing rapidly; however, the role of computers with respect to the electrocardiographer has not been established. At present all electrocardiograms (ECGs) processed by computer are also interpreted by electrocardiographers; hence effort is duplicated.

In an investigation of whether conditions can be defined under which the electrocardiographer can use the computer more profitably by eliminating some of the duplication, ECGs recorded in a university teaching hospital were processed by a computer program and and subsequently reviewed by 1 of 10 electrocardiographers. For ECGs interpreted as showing normal sinus rhythm the rate of agreement between computer and human reviewer was 99%. For those showing a normal ECG pattern (contour) the rate of direct agreement was only 88%. However, the rate of occurrence of clinically significant differences was only 1.64%; hence the rate of essential agreement for this classification was 98.36%. Other classifications with good agreement were myocardial infarction, sinus bradycardia and sinus tachycardia.

Therefore, in circumstances comparable to those of this investigation it is feasible for electrocardiographers to use computers to reduce greatly their workload without compromising the quality of the service provided.

L'emploi de l'ordinateur en électrocardiographie clinique prend rapidement de l'ampleur; toutefois, le rôle de l'ordinateur face au spécialiste en électrocardiographie n'a pas encore été établi. En ce moment tous les électrocardiogrammes (ECG) soumis à l'ordinateur sont aussi interprétés par un spécialiste en électrocardiographie; il en résulte donc une duplication d'effort.

Dans une étude visant à déterminer si l'on peut définir les conditions selon lesquelles le spécialiste en électrocardiographie pourrait utiliser l'ordinateur à meilleur profit en éliminant certaines duplications, les ECG enregistrés dans un hôpital universitaire ont été analysés à l'aide d'un programme d'ordinateur, puis subséquemment revus par 1 ou l'autre de 10 spécialistes en électrocardiographie. Pour les ECG interprétés comme présentant un rythme sinusal normal le taux de concordance entre l'ordinateur et le réviseur humain a été de 99%. Pour ceux montrant un tracé électrocardiographique normal le taux de concordance directe n'a été que de 88%. Toutefois, le taux des différences cliniquement significatives n'a été que de 1.64%; le taux de concordance essentielle pour cette classification a donc été de 98.36%. Les autres classifications présentant une bonne concordance ont été l'infarctus du myocarde, la bradycardie sinusale et la tachycardie sinusale.

Dans des conditions comparables à celles de cette étude il est donc possible pour le spécialiste en électrocardiographie d'employer l'ordinateur dans le but de réduire sensiblement sa tâche sans pour autant compromettre la qualité du service pourvu.

From the department of physiology and biophysics and the department of medicine, Dalhousie University, Halifax

Reprint requests to: Dr. H.K. Wolf, Department of physiology and biophysics, Rm. 4G1, Sir Charles Tupper Medical Building, Dalhousie University, Halifax, NS B3H 4H7

Computer programs for the interpretation of electrocardiograms (ECGs) have been developed for the past 15 years and several have gained national or even international prominence. The report of a recent study conducted for the Department of Health, Education, and Welfare in the United States lists nine programs available to the medical community in that country, some of them in the public domain and others surrounded by secrecy to protect commercial interests.[1]

One feature appears to be common to all of them — lack of advice from the developer on how the programs are to be used clinically. The program descriptions that appear in the literature do not include a clear development objective but are limited to generalities like "contribution to improvement of health care".[2-6] It is not even clear for which user the interpretation reports are formulated — the electrocardiographer, who is thoroughly familiar with electrocardiographic terminology, or the surgeon or general practitioner, who has to extract from the report information relevant to patient management. Judging by the prevailing practice the program developers seem to assume that it is the former. Hence there is one more reason for the electrocardiographer to review and interpret for the user every computer ECG interpretation — a task that often is more time-consuming than interpretation without the aid of the computer.

Although claims have been made that the presence of a computer report reduces the time needed by an electrocardiographer for ECG interpretation,[7] no convincing data have been presented to substantiate this claim; our experience does not allow us to support such a claim.

If the computer saves little or no

time for the reviewing electrocardiographer the costs of producing a computer ECG interpretation have to be justified some other way. Greater accuracy of a combined reviewer–computer ECG interpretation in comparison with an electrocardiographer's report would be such a justification; however, to the best of our knowledge this has not been shown, though studies have demonstrated that the computer interprets ECGs almost as well as a competent electrocardiographer.[8,9]

When the cardiology group at Dalhousie University medical school was confronted with the question whether and under what terms it should get involved with computer interpretation of ECGs it was decided to conduct a study to clarify the roles of computer and electrocardiographer in ECG interpretation in clinical situations and to establish the benefits an electrocardiographer can obtain from using computer interpretation of ECGs.

Methods

Study design

The degree of accuracy with which electrocardiographers can diagnose cardiovascular disease on the basis of ECGs has been found to average between 70% and 80%, with a rate of false-positive reports of 10% to 20%.[10,11] Similar tests with computer programs have yielded similar results.[8,11] The extent of correlation in findings between computer and electrocardiographer determines what effect a combination of the two reports might have.

In the case of perfect correlation in findings between computer and electrocardiographer, accuracy is not improved by the combining of the reports. The only benefit for the electrocardiographer in this instance is that he can use the computer as an "understudy".

When the correlation in findings between computer and electrocardiographer is less than perfect, the possibility of improvement in accuracy of interpretation depends on the willingness of the reviewing electrocardiographer to consider the dissenting opinion of the computer. In a hypothetical situation in which computer and electrocardiographer each has a 75% accuracy rate but their performance is maximally uncorrelated, the combination of computer and electrocardiographer will be correct in interpreting 50% of the ECGs since the two will arrive at the same interpretation. Of the remaining 50% half will be correctly interpreted by the computer and half by the electrocardiographer. The electrocardiographer may become sufficiently distracted by the computer's interpretation that he accepts more incorrect than correct computer interpretations; as a result his accuracy decreases — at the extreme to 50%. On the other hand the difference in interpretation may lead him to review particularly carefully; as a result he may choose more of the correct computer interpretations and thus improve his accuracy score, possibly even to 100%.

Hence, maximally uncorrelated performance of computer and electrocardiographer carries the risk of reduction in accuracy for the combined report, but it also holds the promise of an improvement over the individual performance of a computer or an electrocardiographer. The latter possibility prompted us to investigate the combined performance of computer and electrocardiographer under conditions that would minimize the correlation between their findings. This was achieved in two ways:

1. A computer program was selected that analyses the Frank-lead vectorcardiogram, whereas the electrocardiographers based their interpretation on the standard 12-lead ECG.
2. A computer program was selected that uses statistical methods of classification and had been trained on the basis of clinical information.[12] In contrast, the reviewing electrocardiographers used more or less standard criteria to interpret the ECGs through logical decisions.

The study was conducted in two phases. During the first phase, contingency tables were produced for all major diagnostic entities to describe the amount of agreement and disagreement between computer and electrocardiographer without giving any indication of which interpretation was right, an inherent limitation for many diagnostic categories that are defined solely on the basis of the ECG. The results reported in this paper are primarily those of phase 1.

The attempt of the second phase is to establish the diagnostic accuracy of computer and electrocardiographer for diagnostic categories that permit comparison of ECG interpretation with independent clinical evidence. This part of the study is still in progress and is being conducted in cooperation with medical centres in Kingston and Winnipeg.

Data collection

The data for phase 1 were collected over a 2-year period during which two of the teaching hospitals in Halifax recorded all ECGs in a format that permitted interpretation by computer and electrocardiographer.[13] Ten electrocardiographers reviewed in rotation all computer interpretations and made corrections by adding or deleting statements as they deemed necessary. The amended computer interpretations became the official ECG reports for the patients' charts. To avoid ambiguity in the terms they used, the electrocardiographers agreed on definitions for the important diagnostic statements prior to the study, and these statements were used exclusively for ECG report amendments.

Results

Rhythm interpretation

The results of 12 977 rhythm interpretations by the computer and indications of the electrocardiographers' responses are summarized in Table I. Interpretations in the dominant category, normal sinus rhythm, had an almost perfect rate of acceptance by the electrocardiographers — 99%. Acceptance in the other categories was lower and in one case only 33.1%. Analysis of the rejected interpretations revealed that in many instances the disagreement was due to the difference in duration of the ECG signal used for interpretation. Although the 10-second vectorcardiogram for the computer was recorded contiguously with the 20-second standard lead tracing for the electrocardiographer, some change in features could occur during the additional 10 seconds. This is especially true for isolated extrasystolic complexes, which might be found in one signal but not in the other. The be-

Table I—Results of electrocardiographers' review of major rhythm classifications produced by computer for 12 977 unselected electrocardiograms (ECGs)

Computer classification	Relative frequency (%)	% accepted unaltered by reviewer
Normal sinus rhythm	59.0	99.0
Sinus bradycardia	8.0	98.1
Sinus arrhythmia	8.0	75.1
First-degree atrioventricular block	5.0	83.5
Sinus tachycardia	4.0	98.0
Supraventricular premature beats	3.8	38.9
Ventricular premature beats	3.2	50.3
Electronic pacemaker (normal function)	1.8	92.9
Electronic pacemaker (demand function)	1.3	80.0
Other abnormal rhythms	5.9	33.1

haviour of the demand pacemakers might also be different in the two signals.

Disagreement also occurred for borderline cases of bradycardia and tachycardia. Differences between computer and electrocardiographer in the method of measuring heart rate, or minor fluctuations in heart rate, can explain most of the disagreements. Other rejections of rhythm interpretation were atttributable to the missing of P waves by the computer, which resulted in normal rhythms being misclassified as abnormal. In a few instances flutter waves were misinterpreted as P waves (in 7 of 7627 ECGs classified as having normal sinus rhythm).

The changes made by the electrocardiographers in the interpretations for electronic pacemakers all had to do with the type of pacemaker — fixed-rate versus demand. Although the computer reliably detected pacemaker spikes when present, it sometimes erred in the type of pacemaker. This error also can be attributed to the short signal available for analysis.

ECG contour interpretation

The computer program used for contour interpretation provides a differential diagnosis, giving relative probabilities for normal, left ventricular hypertrophy, right ventricular hypertrophy, anterior myocardial infarction, lateral myocardial infarction, pulmonary emphysema, and right and left bundle branch block.

The acceptance of the computer diagnoses by the electrocardiographers is summarized in Table II. We considered only ECGs for which the computer's most likely diagnosis had a probability of more than 60%. The 21.5% of ECGs thus eliminated usually had similar probabilities for several diagnoses; in the absence of a clear diagnostic decision by the computer no acceptance/rejection evaluation can be performed.

An average of 84% of the computer's definite classifications were accepted by the electrocardiographers. For several other categories there was substantial disagreement between computer and electrocardiographer, which can only be resolved in phase 2 of the study, when independent evidence will be used to establish accuracy of interpretation for computer and electrocardiographer. The three categories "normal", "myocardial infarction" and "conduction defects", however, had a high rate of acceptance and warrant close inspection of the rejections.

Reviewer differences: There were significant differences between the 10 electrocardiographers in their acceptance of computer interpretation. The rate of acceptance of the computer classification "normal" ranged between 75% and 97%. This difference reflects a true difference in the characteristics of the electrocardiographers since the performance of the computer was consistent. They differed almost as much in their rate of acceptance of the classification "myocardial infarction", the rate ranging from 77% to 92%.

Performance differences between the electrocardiographers existed despite the fact that they had agreed prior to the study to use a common set of interpretation criteria. Our results therefore confirm the earlier findings by Simonson and colleagues[10] about the variability of human interpreters of ECGs.

Significance of interpretation changes: The changes the electrocardiographers made to the computer classifications "normal" and "myocardial infarction" are summarized in Table III. In both categories most of the changes were of questionable significance. The definite statement of the computer was modified to a borderline statement such as "possible mild left ventricular hypertrophy" or "slight axis deviation", reflecting the uncertainty of the electrocardiographer in reviewing ECGs for which the computer offered a second, less probable, interpretation. In some instances the electrocardiographer considered the second interpretation as more significant than the computer did and made it the principal interpretation, and in other instances the electrocardiographer disagreed with the computer's interpretation, yet was so uncertain about the diagnosis that he did not offer his own interpretation but instead commented only on the rhythm or repolarization pattern.

The significant alterations of computer interpretation were therefore limited to a change from normal to definite myocardial infarction, definite right ventricular hypertrophy or significant ventricular conduction defect. Of lesser clinical importance were the changes from myocardial infarction to definite left or right ventricular hypertrophy or normal.

Table II—Relative frequency of classifications based on contour analysis of 10 647 ECGs

Computer classification	Relative frequency (%)	% accepted unaltered by reviewer
Normal	49.7	88
Myocardial infarction	18.9	84
Left ventricular hypertrophy	6.3	70
Ventricular conduction defects	1.6	85
Pulmonary emphysema	1.2	28
Right ventricular hypertrophy	0.8	29
Borderline*	21.5	50

*Computer interpretations for which all statements had probabilities of less than 60%.

Table III—Relative frequency of electrocardiographers' substitute interpretations for the computer classifications "normal" and "myocardial infarction" that were rejected

Electrocardiographers' substitute interpretations	Relative frequency of substitute (%); computer classification	
	Normal	Myocardial infarction
Definite myocardial infarction	0.1	—
Definite left ventricular hypertrophy	—	0.14
Definite right ventricular hypertrophy	0.04	0.07
Significant ventricular conduction defect	1.5	2.87
Normal	—	0.7
Computer's second, less probable, classification accepted	1.3	4.1
Borderline abnormal	7.86	2.3
No substitute	1.0	5.4
Total	11.8	15.6

Discussion

The results of our investigation have shown that the computer is by no means a replacement for the electrocardiographer. For most rhythm interpretations and many contour interpretations there was frequently a difference of opinion between computer and electrocardiographer. This is not surprising since the electrocardiographers were not a homogeneous group and there is as much difference between electrocardiographers as between computer and electrocardiographers. The question who is right in all the instances of significant

difference can only be answered when objective evaluation techniques are available to test computer programs and electrocardiographers. Such methods are being developed in our department as well as in several others outside Canada.

However, cardiologists do not have to wait until these issues are resolved to make profitable use of the computer. We have identified two classifications, normal sinus rhythm and normal ECG contour, for which there was general agreement between computer and all the electrocardiographers. Such classifications may apply to nearly 30% of our hospital population, though the percentage may well be higher in centres not as heavily oriented towards tertiary care as is our centre. Our results indicate that ECGs classified by the computer as normal with respect to rhythm and contour do not require review by an electrocardiographer since significant alterations by the latter are rare. Instead, review by a technician to safeguard against obvious computer errors in pattern recognition should be sufficient.

Under these circumstances it is economically feasible for electrocardiographers to use the computer for ECG interpretation, provided a sufficiently large number of ECGs are being processed. This is true especially if the benefits the computer will provide in storage and retrieval of old ECG interpretations are considered. An improvement in the accuracy of ECG interpretation by use of the computer will then be an additional benefit, once it has been demonstrated convincingly.

We thank the following electrocardiographers of Dalhousie University medical school for their participation in this study: R.N. Anderson, C. Felderhof, M.M. Gorelick, J.H. Haldane, L.K. Harris, A.R. MacNeil, J.R. Rae, K.A. Sample, E.R. Smith, L.C. Steeves and J.W. Stewart.

This study was supported in part by the Health Service and Insurance Commission of Nova Scotia.

References

1. *Automated Electrocardiography in the United States*, prepared by Arthur D. Little, Inc., Cambridge, MA under contract 230-75-0212 for division of health care information systems and technology, health resources administration, National Centre for Health Services Research, Washington, 1976
2. BONNER RE, CREVASSE L, FERRER MI, et al: A new computer program for analysis of scalar electrocardiograms. *Comput Biomed Res* 5: 629, 1972
3. WEIHRER AL, WHITEMAN JR, ZIMMERMAN A, et al: Computer programs for an automated electrocardiographic system, in *Clinical Electrocardiography and Computers*, CACERES CA, DREIFUS LS (eds), New York, Acad Pr, 1970, p 83
4. CORNFIELD J, DUNN RA, BATCHLOR CD, et al: Multigroup diagnosis of electrocardiograms. *Comput Biomed Res* 6: 97, 1973
5. HU K, FRANCIS DB, GAU GT, et al: Development and performance of Mayo–IBM electrocardiographic computer analysis program (V70). *Mayo Clin Proc* 48: 260, 1973
6. WARTAK J, MILLIKEN JA, KARCHMAR J: Computer program for diagnostic evaluation of electrocardiograms. *Comput Biomed Res* 4: 225, 1971
7. *The Computer-Assisted ECG From Laboratory to Community — a Report of the Denver Demonstration*, DHEW publ no (HRA) 74-3104, US Public Health Service, health resources administration, Rockville, MD, 1973
8. PIPBERGER HV, MCCAUGHAN D, LITTMANN D, et al: Clinical application of a second generation electrocardiographic computer program. *Am J Cardiol* 35: 597, 1975
9. CREVASSE L, ARIET M: A new scalar electrocardiographic computer program. *JAMA* 226: 1089, 1973
10. SIMONSON E, TUKA N, OKAMOTO N, et al: Diagnostic accuracy of the vectorcardiogram and electrocardiogram. A cooperative study. *Am J Cardiol* 17: 829, 1966
11. GREGOR RD, WOLF HK, HARRIS LK: Effect of computer assistance on the performance of human electrocardiogram interpreters. *Ann R Coll Physician Surg Can* 10: 36, 1977
12. WOLF HK, MACINNIS PJ, STOCK S, et al: Computer analysis of rest and exercise electrocardiograms. *Comput Biomed Res* 5: 329, 1972
13. WOLF HK, GRAYSTONE P, LYWOOD DW, et al: Suggested minimum specifications for data acquisition instrumentation in computer-assisted ECG processing systems. *J Electrocardiol* 9: 239, 1976

Survival of patients treated for end-stage renal disease by dialysis and transplantation

MICHAEL R. HIGGINS,* MB, FRCP (EDIN), FRCP[C]; MICHAEL GRACE,† PH D, P ENG; JOHN B. DOSSETOR,‡ MB, FRCP (LOND), FRCP[C]

The results of treatment in 213 patients with end-stage renal disease who underwent hemodialysis, peritoneal dialysis or transplantation, or a combination, between 1962 and 1975 were analysed. Comparison by censored survival analysis showed significantly better ($P < 0.01$) patient survival with the integrated therapy of dialysis and transplantation than with either form of dialysis alone. There was no significant difference in survival of males and females but survival at the extremes of age was poorer. Analysis of survival by major cause of renal failure indicated best survival in patients with congenital renal disease. Graft and patient survival rates at 1 year after the first transplantation were 42% and 69%. The major cause of death in this series was vascular disease but infection was responsible for 50% of deaths after transplantation. While integration of dialysis with transplantation produces best patient survival, this course is possible only when sufficient cadaver kidneys are available.

On a analysé les résultats du traitement chez 213 patients souffrant de maladie rénale au stade terminal qui ont subi, entre 1962 et 1975, l'hémodialyse, la dialyse péritonéale ou la greffe, ou une combinaison. La comparaison, par analyse pondérée de la survie, a révélé une survie significativement meilleure ($P < 0.01$) pour le traitement comprenant dialyse et greffe que pour l'une ou l'autre des deux formes de dialyse employées seules. On n'a observé aucune différence significative entre les taux de survie des hommes et des femmes, mais la survie était moins bonne aux limites de l'intervalle d'âge. L'analyse de la survie pour les principales causes d'insuffisance rénale a indiqué une meilleure survie chez les patients souffrant de maladie rénale congénitale. Les taux de survie à 1 an pour les greffes et pour les patients après la première greffe ont été de 42% et de 69%, respectivement. La principale cause du décès dans ce groupe a été les maladies vasculaires mais l'infection compte pour 50% des causes de décès après une greffe. Alors que l'association de la dialyse et de la greffe donne la meilleure survie, ce mode de traitement n'est possible que quand on dispose d'une nombre suffisant de reins d'origine cadavérique.

From the department of medicine, University of Alberta, Edmonton

*Associate professor of medicine
†Director, research and development, the Dr. W.W. Cross Cancer Institute, Edmonton
‡Professor of medicine

Reprint requests to: Dr. Michael R. Higgins, Rm. 2125, University of Alberta Hospital, 112th St. and 83rd Ave., Edmonton, Alta. T6G 2B7

During the last 25 years dialysis and transplantation for end-stage renal disease have developed to the extent that rehabilitation is regarded as a more acceptable criterion of successful treatment than patient survival. While economic concerns may be responsible for this change in emphasis, patient sur-