Comparison of Electrocardiogram Interpretations by Family Physicians, a Computer, and a Cardiology Service

Douglas Woolley, MD, Mark Henck, MD, Jerry Luck, MD
Greenville, North Carolina; Bryn Mawr and Hershey, Pennsylvania

Background. Some family physicians may be under pressure to relinquish the interpretation of outpatient electrocardiograms to cardiologists. The purpose of this study was to determine whether the quality of electrocardiogram (ECG) interpretations by family physicians justifies this pressure, and whether an immediately available computerized ECG interpretation program could serve as an appropriate backup for the family physician.

Methods. Family practice faculty and residents at a university-based residency program provided written interpretations of 301 ECGs ordered over an 11-month period. Their ECG findings were compared with those from a computerized interpretation program and the readings of the cardiology service. All interpretations were then compared with those of a fellowship-trained electrocardiographer, whose readings served as the reference standard.

Results. Discrepancy was found between the family physician and the electrocardiographer on 33% of those items that had any potential clinical significance. The computer interpretation and the cardiologist's interpretation agreed with that of the electrocardiographer on 63% and 71% of these discrepancies, respectively (not statistically different).

Conclusions. Family physicians reached a level of agreement with the reference standard in ECG interpretation that was comparable to previously published reports for expert interrater agreement. In this study, however, the quality of ECG readings by family physicians was further improved by expert review. The quality of computer-assisted ECG interpretation was comparable to that of review provided by a cardiology service. Furthermore, computerized interpretation may be clinically more useful because it is immediately available.

Key words. Electrocardiogram; cardiology; family practice; quality assurance. *J Fam Pract 1992; 34:428-432.*

Among tests obtained and interpreted in the family physician's office, the resting standard 12-lead electrocardiogram (ECG) is one of the most common, with approximately two ECGs ordered for every 100 patient visits.[1,2] It is generally considered an important diagnostic and management tool by family physicians. More than two thirds of all family physicians routinely perform ECG studies in the office.[3] Most of these physicians read their own ECG studies, and base their clinical actions on their findings. In some settings, however, family physicians are under growing pressure to relinquish formal interpretation of routine office ECG studies to consulting cardiologists. Proponents of this position suggest that physicians who are not cardiologists are likely to have inadequate ECG reading skills, and that patient care based on their ECG interpretations may be inferior.[4] Data are not available, however, to support the view that a cardiologist's interpretation, routinely returned to the clinician up to several days after the study is done, is more useful than that of the family physician, particularly when the ECG is obtained to help make immediate clinical decisions. Froom and Froom[1] and Grauer and Curry[5] have demonstrated that family physicians can achieve proficiency in the interpretation of over 95% of all ECGs findings seen in the primary care setting.

The present study was designed to answer these questions: Is the quality of ECG interpretations provided by family physicians acceptable? Is it significantly improved by expert backup? Can a high-quality computer-based ECG interpretation program compare favorably with review by a cardiology service in providing this expert backup?

Submitted, revised, January 3, 1992.

From the Department of Family Medicine (Dr Woolley), East Carolina University School of Medicine, Greenville, North Carolina; and the Department of Radiology (Dr Henck), Bryn Mawr Hospital, Bryn Mawr, Pennsylvania, and the Division of Cardiology (Dr Luck), Hershey Medical Center, Pennsylvania State University, Hershey. Requests for reprints should be addressed to Douglas Woolley, MD, Department of Family Medicine, ECU School of Medicine, Greenville, NC 27858.

ISSN 0094-3509

Methods

All ECGs ordered during an 11-month period by physicians in the family practice center on the campus of a university hospital were included in the study. When a physician requested an ECG, it was immediately obtained on a Hewlett-Packard 4750-A recorder by one of the trained family practice nurses. The physician was asked to read the study immediately and complete a questionnaire, noting whether the study was normal, describing all specific abnormal findings, stating the reason for the study, whether interpretations of previous studies were available for comparison, and whether the information from this study influenced immediate patient care. Within minutes after completing the questionnaire, the physician was provided with a computer-generated interpretation from the Hewlett-Packard ECG diagnostic program,[6] available as backup in the family practice center for the previous 7 years. All computer interpretations were added to the pool of ECGs from the university hospital and ambulatory clinics for review by general cardiologists of the routine ECG interpretation service, who were unaware of the present study. Unlike the family physician who provided the initial interpretation, the cardiologist referred to the computer printout while interpreting the ECG study. This procedure was consistent with the standard practice for interpretations in this and many other hospitals and outpatient sites.[7] The standard for comparison was the interpretation by one of the authors (J.L.), a cardiologist with fellowship training and clinical expertise in cardiac electrophysiology, referred to here as the electrocardiographer. The standard for interpretation was validated in two ways. First, another fellowship-trained electrocardiographer interpreted a subset of the study ECGs. Then the first electrocardiographer unknowingly reinterpreted a subset of the study ECGs he had seen a few months before. Thus, interobserver and intraobserver reliability checks were conducted.

The findings of the family physician, computer, cardiologist, and electrocardiographer were coded according to the categories of Hancock et al.[8] Based on criteria established by consensus of the authors prior to data collection, each coded finding was assigned one of three levels of probable clinical significance: no clinical significance, minor clinical significance, or major clinical significance. We obtained a measure of agreement between the family physician and the electrocardiographer, and the family physician and cardiologist by comparing interpretations on all findings so that if error had been committed, its significance would be categorized as either "minor" or "major." We determined how often the cardiologist and the electrocardiographer concurred whenever either specialist differed from the family physician interpretation on these "clinically significant" items. Then, reviewing the computer interpretations, we determined how many of the clinically significant errors (as determined by our standard, the electrocardiographer) the family physician could have caught immediately by comparing his or her interpretation with that of the computer. In this four-tiered ECG interpretation sequence, involving the family physician, computer, cardiologist, and electrocardiographer, we determined whether subsequent levels of interpretation provided statistically significant improvement over previous levels.

Results

During the 11 months from April 1988 through February 1989, a total of 357 ECG studies were obtained in 14,198 visits to the family practice center, or 2.5 ECGs per 100 patient visits. The rate is similar to those reported by Froom and Froom[1] (2.1/100), Nissan[2] (2.0/100), and Rosenblatt et al[3] (2.3/100). All 357 ECGs were submitted to the cardiology service for official interpretation. Of these, 301 tracings (84%) had been interpreted by the family physician according to study protocol and were included in the study. The ECGs not properly handled according to the study protocol did not represent interpretations of greater complexity or diagnostic difficulty than those submitted. Of the 301 ECGs, only 36% were subsequently interpreted as "normal study" by the cardiologist; of the 56 ECGs ordered but not included in the study, 42% were interpreted as normal. There were 2.1 abnormal findings per abnormal ECG among those properly handled, and 2.0 among those that were not a part of the study. Neither difference is statistically significant by chi-square analysis ($P > .10$).

We found a level of agreement of 0.67 (95% CI = 0.62 to 0.72) between the family physician and the electrocardiographer on diagnoses of potential clinical significance. A total of 112 discrepancies of potential clinical significance were found on 85 ECGs, or 28% of the ECGs in the study. Of these, 42 tracings were thought by the electrocardiographer to have discrepancies of potentially minor significance, whereas 43 were thought to have at least one discrepancy of potentially major significance. Errors by family physicians were grouped into three categories: "errors of omission," where abnormalities that were probably significant were not seen; "errors of misinterpretation," where abnormalities were seen but misdiagnosed; and "errors of commission," where abnormalities were noted that likely did not exist (Table 1). The most common error was overlooking probable infarcts. The second most common

Table 1. Types and Numbers of Significant Errors by Family Physicians in Interpreting 301 ECGs, as Determined by an Electrocardiographer

Diagnostic Category	Number of Each Error Type*			Subtotal for Each Diagnostic Category
	O	MI	C	
Infarct	27	2	10	39
Conduction abnormality	8	5	10	23
Rhythm	9	6	—	15
Ischemia/strain/hypertrophy	1	9	3	13
Axis	5	1	3	9
Lead placement	5	1	—	6
Other	4	1	2	7
Total	59	25	28	112

**O denotes omission; MI, misinterpretation; C, commission.*

error was noting infarcts and conduction abnormalities where they likely did not exist.

We found 115 discrepancies of potential clinical significance when comparing the interpretations of the family physician and the general cardiologist. The electrocardiographer agreed with the cardiologist that a reading error had been committed on 86 of these findings but disagreed on 29. The electrocardiographer cited an additional 26 errors not cited by the cardiologist. Using Dice's "proportion of specific agreement,"[9] this gives a value of 0.76 (95% CI = 0.70 to 0.82) for "expert" interrater agreement between the cardiologist and electrocardiographer on errors by family physicians. A second electrocardiographer read a subset (51) of the study ECGs. Comparing his interpretations with those of the first electrocardiographer by the same method gives us a second, and comparable, "expert" interrater agreement on probable family physician reading errors of 0.75 (95% CI = 0.63 to 0.87). The first electrocardiographer unknowingly reread a subset (40) of the study's ECGs. Analysis of these interpretations gives an intrarater agreement on probable family practice reading errors of 0.91 (95% CI = 0.83 to 0.99). The repeated high level of interrater agreement supports the accuracy of our standard, and the high intrarater agreement supports its reliability.

The initial Hewlett-Packard computer interpretation of the ECG, before being reviewed and changed by the cardiology service, was available on a random subset of 59 of the 85 studies in which the electrocardiographer had noted potentially significant errors in interpretation by family physicians. The computer interpretation agreed with the interpretation by the electrocardiographer concerning potentially significant errors in interpretation of 42 of the 67 items in the subset of ECGs reviewed. The cardiologist's interpretation agreed with 48 of the 67 items. By chi-square analysis, this difference was not statistically significant ($P > .10$). The sequential improvement in the accuracy of ECG interpretation by the family physician, the computer, and the cardiologist as defined by agreement with the interpretations of the electrocardiographer on all findings of potential clinical significance is illustrated in Figure 1. For the family physician the level of agreement was 0.67 (95% CI = 0.61 to 0.74), for the computer it was 0.88 (95% CI = 0.83 to 0.93) and for the cardiologist it was 0.91 (95% CI = 0.87 to 0.95). The only statistically significant improvement occurred with the computer interpretation.

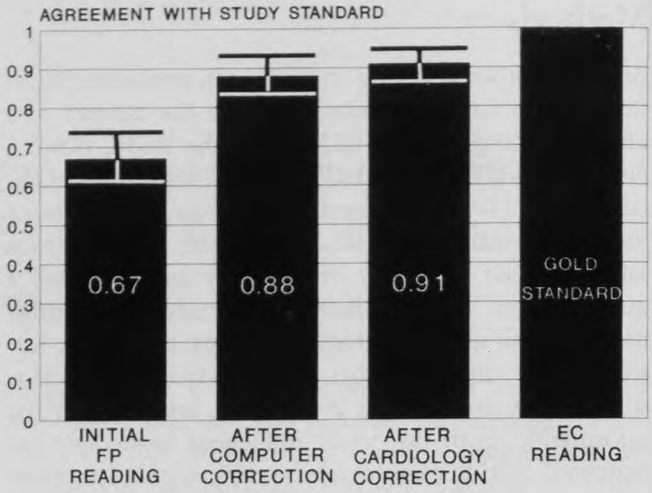

Figure 1. Accuracy of electrocardiogram readings done by the family physician, the computer, the cardiologist, compared with the electrocardiographer (the "gold" standard). Error bars represent 95% confidence intervals.

According to the electrocardiographer only 108 of the 301 study ECGs were normal, or 35.9% (95% CI = 33.2 to 38.6%), compared with 62.4% reported by Froom and Froom[1] and 69.7% reported by Nissan.[2] Our patient population was significantly older than the population studied by these two authors, with 54% (95% CI = 49.0 to 59.0%) older than 60 years (Nissan: 38.8%) and 47.0% (95% CI = 42.0 to 52.0%) older than 65 years (Froom and Froom: 41.8%). The most common diagnoses corresponded generally with those of Froom and Froom,[1] with discrepancies between the previous study and our study primarily found in the frequency of accurately detecting axis deviation and conduction abnormalities (Table 2).

There were 227 responses about the use of a previous ECG. For ECGs of this group demonstrating abnormal findings, there was a rate of reading errors of 31% when a previous ECG interpretation was available, and a rate of 32% when no previous interpretation was available. There was no apparent increase in accuracy of interpretation when previous ECG studies were available.

Table 2. A Comparison of Rank Order of the Most Common Electrocardiogram (ECG) Findings in Two Ambulatory Family Practice Studies

ECG Finding	Rank Order of Most Common Findings*	
	Hershey Medical Center	Froom and Froom[1]
Normal ECG	1	1
Abnormal ST-T wave	2	2
Old myocardial infarct	3	3
Left axis deviation or left anterior fascicular block	4	6
Premature ventricular beat	5	4
Left ventricular hypertrophy with or without strain	6	10
First degree anterior-ventricular block	8	9
Right bundle branch block	8	15
Atrial fibrillation	10	8
Atrial premature complex	10	5
Low voltage frontal leads	11	7
Left atrial abnormality	12	12

**Hershey Medical Center: 301 ECGs, 454 findings; Froom and Froom: 370 ECGs, 462 finding.*

Discussion

Did the family physicians in this study provide ECG interpretations of acceptable quality? There were potentially significant discrepancies between the family physician and electrocardiographer on 28% of study ECGs and 33% of abnormal findings. The family physicians correctly identified normal tracings (less than 3% error rate), but the ECGs of greater complexity and abnormality were more likely to be misread. However, the 67% agreement rate between family physicians and electrocardiographer on findings of potential clinical significance is still within the range of published values for "expert interrater" agreement of 60% to 70%.[10] Although they may be performing acceptably as measured against published standards, our study demonstrates that the quality of their interpretations may be significantly improved by expert review.

Can a high-quality computer ECG interpretation program routinely fill this expert review role? This study demonstrates that the computer program was effective in correcting most of the potentially significant ECG reading errors made by our family practice physicians. The degree of improvement in accuracy between the family physician's interpretation and the computer interpretation was statistically significant, whereas that between the computer and cardiology service was not ($P > .10$). The value of the computer backup is enhanced by its immediate availability, whereas routine interpretations from the cardiology service are not available for at least 1 day, and more typically, several days. The difference in availability may result in important differences in clinical outcome, thus making the use of computerized ECG interpretations more favorable. However, this is an issue best investigated by prospective study.

The patients studied here had a relatively high rate of abnormal ECGs compared with those in other studies.[1,2] The higher rate probably reflects the older and more impaired patient population. It is unlikely that the physicians in this study were overinterpreting the ECG findings compared with physicians in previous studies: (1) the standard diagnostic categories and criteria of Hancock et al[8] were used, and (2) all groups of ECG interpreters in this study had essentially the same number of abnormal findings on each ECG. In our family practice center we have emphasized that the routine ECG has little use in screening for coronary artery disease.[11] Therefore, the family practice physicians may have ordered significantly fewer screening ECGs on apparently normal patients than physicians in previous studies.[1,2]

Reading errors that were called "significant" for the purposes of this study often were not clinically significant. For example, the error of missing a left bundle branch block in an asymptomatic patient, assigned in our study a value of "major significance," would not typically be clinically significant.[12] This means that our study probably overestimates the number of clinically significant ECG reading errors made in the family physician's office. To understand the impact of such errors on quality of care, however, would require a controlled, prospective study of clinical outcomes.

One of the shortcomings in performance of physicians at all levels in our study, particularly among family physicians, and also noted by Blake,[4] was the use of nonstandard, ambiguous, or hedging descriptive terms. Such comments as "small inferior Q waves noted" suggest hesitancy in the mind of the interpreter, and may contribute to confusion in clinical management. The availability of a reliable computerized interpretation program may have affected reader performance at several levels. Review of charts shows that the family physicians in this study, unless asked to do so, tended not to do a rigorous independent reading of the ECG before they received the computer interpretation. Such habits are likely to lead to a loss of interpretive skills in senior physicians and hinder skills development in resident physicians.[7] Moreover, the family physicians may not have been critical enough of the computer in its areas of greatest weakness (eg, arrhythmias). Likewise, the general cardiologist referred to the computer interpretation while reading the tracing, and in the process may not always have exercised appropriate independent critical judgment.

It is perhaps surprising to note that the family physicians' ECG reading accuracy was not related to the

availability of a prior ECG interpretation. Patients with more than one ECG in their chart were more likely to have cardiac disease and less stable tracings, so that prior interpretations often did not accurately describe current findings. Moreover, even where the tracings remained stable, disagreement in interpretation provided by the different experts reviewing the tracings over time was common.

The study results suggest that family physicians with ECG interpretation skills commensurate with those of our clinicians will benefit from expert backup. Where available, a high-quality computer ECG interpretation program can provide the majority of this backup. However, if computer backup is not available, family physicians should consider a program of ECG backup reading established directly with a cardiology colleague. Just how many and what kind of ECGs are reviewed will depend on the type of patients seen, the frequency with which ECGs are ordered, and the skill of the family physician. These factors, particularly the latter, should be monitored before such a program is instituted, and then monitored regularly thereafter.

What level of agreement with "expert" interpretation should be considered the minimal acceptable standard for a family physician? One might reasonably demand that family physicians perform essentially as well as the most skillful, readily available interpreter. In our setting this performance level could be defined by obtaining a measure of "blinded" expert interobserver agreement for general cardiologists within our own institution. This figure likely would be lower than the nonblinded expert interobserver figure obtained in this study and might approach previously published figures of 60% to 70%.[10] As shown, our family physicians, who were the only blinded ECG readers in this study, are already performing at this level.

Acknowledgments

The authors thank Jeanette Dolezal, PhD, for assistance in statistical analysis, Jeri Harris for editorial assistance, and Michael Markel, MD, for assistance in ECG interpretation.

References

1. Froom J, Froom P. Electrocardiogram abnormalities in primary care patients. J Fam Pract 1984; 189:223–5.
2. Nissan R, Encarnacion M. Clinical value of the electrocardiogram in ambulatory care. J Fam Pract 1987; 24:361–3.
3. Rosenblatt RA, Cherkin DC, Schneeweiss R, et al. The structure and content of family practice: current status and future trends. J Fam Pract 1982; 15:681–722.
4. Blake TM. Who is qualified to interpret electrocardiograms? South Med J 1985; 78:368–70.
5. Grauer K, Curry RW. Teaching ECG interpretation to residents. Fam Med 1984; 16:124–5.
6. Hodges M. A clinical evaluation of the H-P ECG analysis program: program accuracy and value of adjustable criteria. Comput Cardiol 1979; 6:167–70.
7. Grauer K, Kravitz L, Curry RW, Ariet M. Computerized electrocardiogram interpretations: are they useful for the family physician? J Fam Pract 1987; 24:39–43.
8. Hancock EW, Norcini JJ, Webster GD. A standardized examination in the interpretation of electrocardiograms. J Am Coll Cardiol 1987; 10:882–6.
9. Fleiss J. Statistical methods for rates and proportions. 2nd ed. New York: John Wiley & Sons, 1981:214.
10. Crevasse LE, Ariet M. Current status of computerized electrocardiography. In: Hurst JW, ed. Update III: The heart. New York: McGraw-Hill, 1980:101–20.
11. US Preventive Services Task Force. Screening for asymptomatic coronary artery disease. In: Guide to clinical preventive services. Baltimore: Williams & Wilkins, 1989: 3–10.
12. Rose G, Hamilton PS, Keen H, et al. Myocardial ischemia, risk factors and death from coronary heart disease. Lancet 1977; 1:105–9.