Accuracy of Interpretation of Preparticipation Screening Electrocardiograms

Allison C. Hill, MD, Christina Y. Miyake, MD, MS, Stafford Grady, MD, and Anne M. Dubin, MD

Objective To evaluate the accuracy of pediatric cardiologists’ interpretations of electrocardiograms (ECGs).

Study design A series of 18 ECGs that represented conditions causing pediatric sudden cardiac death or normal hearts were interpreted by 53 members of the Western Society of Pediatric Cardiology. Gold-standard diagnoses and recommendations were determined by 2 electrophysiologists (100% concordance).

Results The average number of correct ECG interpretations per respondent was 12.4 ± 2.2 (69%, range 34%-98%). Respondents achieved a sensitivity of 68% and a specificity of 70% for recognition of any abnormality. The false-positive and false-negative rates were 30% and 32%, respectively. Based on actual ECG diagnosis, sports participation was accurately permitted in 74% of cases and accurately restricted in 81% of cases. Respondents gave correct sports guidance most commonly in cases of long QT syndrome and myocarditis (98% and 90%, respectively) and least commonly in cases of hypertrophic cardiomyopathy, Wolff-Parkinson-White syndrome, and pulmonary hypertension (80%, 64%, and 38%, respectively). Respondents ordered more follow-up tests than did experts.

Conclusions Preparticipation screening ECGs are difficult to interpret. Mistakes in ECG interpretation could lead to high rates of inappropriate sports guidance. A consequence of diagnostic error is overuse of ancillary diagnostic tests. (J Pediatr 2011; -

See editorial, p 

Sudden death in young athletes is shocking and devastating, especially because most of these seemingly healthy athletes have silent heart disease.1-7 The tragedy of sudden cardiac deaths (SCDs) and subsequent community outcries prompted some European countries to mandate preparticipation screening electrocardiograms (ECGs).8 Since the introduction of preparticipation screening ECGs, the annual rate of SCD in the Venuto region of Italy has decreased significantly, mostly because of the detection of arrhythmogenic right ventricular cardiomyopathy.9

In the United States, the American Heart Association (AHA) has redressed the issue and has not changed its position, which states that it is not prudent to recommend routine ECGs because of the large number of athletes in the United States, the low frequency of diseases leading to SCD, the low rate of SCD itself, and the frequent false-positives, which would lead to unnecessary anxiety among athletes and their families as well as to unjustified exclusion from sports or even life insurance coverage. Thus the AHA recommends a thorough history and physical exam every 2 years.10

Despite the AHA recommendations, there has been a push to have screening ECGs included in preparticipation examinations. Controversy exists in the literature regarding whether these ECGs should be mandated. Similarly, controversy exists regarding the recently released AHA guidelines concerning medications for attention deficit hyperactivity disorder; they conclude that “it is reasonable to consider adding an ECG, which is of reasonable cost, to the history and physical examination in the cardiovascular evaluation of children.”11

One unknown aspect of this controversy is the accuracy of interpretation of preparticipation ECGs. This study was designed to assess the accuracy of pediatric cardiologists’ interpretations of ECGs in demonstrating diseases commonly underlying SCD and in identifying factors associated with successful interpretations.

Methods

This study was an online questionnaire-based study investigating the accuracy of pediatric cardiologists’ interpretations of ECGs. We selected 18 12-lead ECGs from the ECG database at Lucile Packard Children’s Hospital. The 18 ECGs included 8

| AHA | American Heart Association |
| ECG | Electrocardiogram |
| HCM | Hypertrophic cardiomyopathy |
| LQTS | Long QT syndrome |
| PAH | Pulmonary arterial hypertension |
| SCD | Sudden cardiac death |
| WPW | Wolff-Parkinson-White syndrome |

From the Department of Pediatrics (A.H.), and Department of Pediatrics, Division of Cardiology, Stanford University, Stanford, CA (C.M., A.D.); and Pediatric Cardiology Associates, San Jose, CA (S.G.)

A.D. receives fellowship support from Medtronic. The other authors declare no conflicts of interest.

0322-3476/5 - see front matter. Copyright © 2011 Mosby Inc.
All rights reserved. 10.1016/j.jpeds.2011.05.014
from patients with normal hearts and 10 from patients with conditions that commonly underlie SCD (1 with long QT syndrome (LQTS); 4 with hypertrophic cardiomyopathy (HCM); 2 with Wolff-Parkinson-White syndrome (WPWs); 1 with pulmonary arterial hypertension (PAH); and 2 with myocarditis). The distribution of normal and abnormal ECGs was arbitrary. Typical ECG findings for each of these diagnoses were present on the ECGs. Some of the normal ECGs did demonstrate common findings for athletic hearts, including sinus arrhythmia, low atrial rhythm, and sinus bradycardia. Each ECG had only 1 correct primary diagnosis and a short list of secondary findings.

Accuracy of Interpretation of Preparticipation Screening Electrocardiograms

Members of the Western Society for Pediatric Cardiology were e-mailed a link to an online survey. For each ECG, respondents were asked whether the ECG was normal; to describe their interpretation of the ECG; what (if any) heart disease they thought the patient may have; what additional tests they would order; and whether they would recommend allowing the patient to participate in athletic events on the basis of the ECG. All questions were multiple choice except the interpretations of ECGs and diagnoses, which were presented in a fill-in-the-blank format. Respondents were provided with the age and sex of each patient. No further clinical history or computer ECG interpretation was given so as to avoid affecting the accuracy of the ECG interpretation. No time limit was enforced. Participants took the survey online, so they could reference outside sources if desired at the time of the survey.

The online survey was taken by 2 pediatric electrophysiologists who were blinded to any clinical information or previous ECGs. The 2 electrophysiologists had 100% concordance for all diagnoses. Treatment recommendations were based on current AHA guidelines.

All scoring was completed at Stanford by the lead author. Respondents were awarded 1 point for each correct diagnosis, for a possible total of 18. A point was also given for each time a respondent correctly allowed or restricted athletic participation, for a possible total of 18 points.

The participants completed a questionnaire regarding their board certification, number of years practicing pediatric cardiology, practice type, and number of ECGs read per month.

The study was approved by the Institutional Review Board of Stanford University. Informed consent was obtained.

Each survey was scored as a percentage of correct diagnoses and correct permission for sports participation. Comparisons were made among all respondents and general cardiologists. Summary statistics included mean score and standard deviation. Comparisons were performed using Chi-square test. Sensitivity, specificity, false-positives and false-negatives were determined. Expert responses were compared using the kappa coefficient. A P value of <.05 was considered significant.

Table. Survey respondents

<table>
<thead>
<tr>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queried</td>
</tr>
<tr>
<td>Responded</td>
</tr>
<tr>
<td>Board certified in pediatric cardiology</td>
</tr>
<tr>
<td>Number of years practicing pediatric cardiology</td>
</tr>
<tr>
<td>Fewer than 5</td>
</tr>
<tr>
<td>5-15</td>
</tr>
<tr>
<td>16-25</td>
</tr>
<tr>
<td>More than 25</td>
</tr>
<tr>
<td>Practice location</td>
</tr>
<tr>
<td>University-based</td>
</tr>
<tr>
<td>Private</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Practice type</td>
</tr>
<tr>
<td>General cardiology</td>
</tr>
<tr>
<td>Echocardiography</td>
</tr>
<tr>
<td>Catheterization</td>
</tr>
<tr>
<td>Electrophysiology</td>
</tr>
<tr>
<td>Cardiac intensive care</td>
</tr>
<tr>
<td>Heart failure/transplant</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Number of ECGs read per month</td>
</tr>
<tr>
<td>Fewer than 30</td>
</tr>
<tr>
<td>30-60</td>
</tr>
<tr>
<td>61-100</td>
</tr>
<tr>
<td>More than 100</td>
</tr>
</tbody>
</table>

Results

We emailed the survey link to 212 pediatric cardiologists. Although 97 respondents began the survey, only 53 finished it. Incomplete surveys were excluded. The rate of participation was 25%. The participants included 53 pediatric cardiologists, most of whom had been practicing for 5 to 15 years (43%) in university-based (47%) practices where 43% read more than 100 ECGs per month (Table).

The mean score for correct ECG interpretation was 12.4 (69%) ± 2.2 of the total possible 18 (range, 7-17), significantly lower than expert responses (P = .005). The respondents’ scores for accurately identifying ECG findings are shown in Figure 1. Dependent on electrocardiographic findings, the range of scores for correct interpretation was 34% to 98%. In the ECGs that demonstrated findings consistent with HCM, 59% of all respondents correctly identified them (P = .01). Normal ECGs were also significantly misinterpreted; only 71% of all respondents read them correctly (P = .02). Other diagnoses (myocarditis, LQTS, WPW, and PAH) were misinterpreted by all respondents and by general cardiologists more often than by experts. However, the difference between all respondents and experts as well as between general cardiologists and experts did not reach statistical significance (P > .05), either because the power or the difference in score was not large. Respondents achieved a sensitivity of 68% and a specificity of 70% for recognition of any abnormality in ECGs. The false-positive rate was 30%, and the false-negative rate was 32%.

Scores were given for correct allowing of sports participation based on the actual underlying ECG diagnosis, and were
independent of the findings respondents reported for each ECG. The mean score for correct sports-participation guidance for all respondents was 14 (78%) of the total possible 18 (range, 7-17), with a SD of 1.9 (P = .003). Respondents correctly restricted sports participation most often in cases of LQTS and myocarditis and less often in findings of HCM, WPW, or PAH. All respondents correctly allowed sports participation in 74% of patients with normal ECGs (P = .005). Of the respondents, 20% would allow sports participation by patients with evidence of HCM, 11% for those with myocarditis, 36% for those with WPW, 2% for those with LQTS, and 62% for those with PAH. In patients with normal ECGs, 26% of respondents would restrict sports participation (Figure 2). Respondents gave the wrong sports guidance usually because of inaccurate interpretations of the ECG (72%), not because they did not know the AHA guidelines for sports guidance.

Because general cardiologists interpret preparticipation ECGs routinely in the clinical setting, data from their responses were analyzed. They performed slightly better than all respondents in correct diagnoses, although the differences did not reach statistical significance. The computer-derived interpretations accurately identified ECG findings in 12 of the 18 ECGs in the survey, which was not statistically different from those of the respondents. No associations were found between survey performance and length of time in practice, number of ECGs read per month, or practice type. Electrophysiologists’ performance was better than that of the rest of the respondents, although the difference was insignificant. No associations were found among survey performance and practice type, length of time in practice, or number of ECGs read per month.

Respondents were asked whether they would order a follow-up echocardiogram, Holter monitor, exercise test, or no follow-up tests at all for each ECG. The 53 respondents ordered 508 follow-up tests for the 18 ECGs that the experts also would have ordered. However, they ordered an additional 380 tests that were deemed unnecessary by the experts, and they would have missed ordering 340 tests that were necessary for appropriate diagnosis.

Figure 1. Percentage of correct ECG interpretations: respondents’ scores of accurate findings in ECG interpretation broken down by underlying disease.

Figure 2. Percentage of correct sports participation guidance: respondents’ scores compared with correct guidance for actual ECG diagnosis.
There were inconsistencies in accurate diagnosis, additional testing, and sports guidance. Figure 3 shows a schematic of how the four ECGs showing HCM were interpreted. Respondents were able to identify correctly that an ECG showing HCM was abnormal and warranted an echocardiogram 85% of the time and required exclusion from sports 80% of the time, even though only 60% were able to diagnose HCM correctly on the basis of the ECG. This meant that 25% of respondents recognized an abnormality but could not identify it. Similarly, in ECGs showing myocarditis, 50% correctly diagnosed it, but 88% would order an echocardiogram and restrict sports participation. Thus, 38% recognized an abnormality but could not identify it.

**Discussion**

ECGs demonstrating diseases that underlie SCD are difficult to interpret, even for pediatric cardiologists; their accuracy rate in this study was 69%. We found that the accuracy rate did not differ significantly in respondents, despite variations in length of time practicing pediatric cardiology, number of ECGs read per month, or practice type. These findings are in agreement with other studies showing that the accuracy of ECG interpretation ranges from 53% to 96%. Others have also shown that ECG interpretation is fraught with difficulties and that interpretations of the same ECG by multiple cardiologists vary greatly. Even when an individual cardiologist reads the same ECG on separate occasions, the interpretation varies substantially.

One factor that influences accuracy of ECG interpretation is the amount of clinical history offered to the cardiologist. Less-experienced cardiologists interpret ECGs more accurately when given clinical history than do more experienced cardiologists, whose accuracy is not affected by the presence of clinical history. Another factor contributing to accuracy is underlying diagnosis. The respondents in our study scored lowest when interpreting ECGs showing ventricular hypertrophy (PAH, HCM) and WPW, probably because of the similarities in these pathologies and the normal changes in athletes’ hearts. Others have also found that pediatric cardiologists disagree substantially in diagnosing ventricular hypertrophy on ECG. Given these difficulties with ECG interpretation, it may come as little surprise to some that in certain scenarios, the computer interpretation is more accurate than physicians’ interpretations. It may be suggested that expert pediatric electrophysiologists read all the preparticipation screening ECGs, but this would be impractical. With 10.7 million young athletes in the United States, the approximately 200 electrophysiologists would have to read 223 ECGs per day in addition to their regular workload.

Another major issue concerning the institution of mass-screening ECGs is the challenge of identifying very few abnormal ECGs in a multitude of normal ECGs. The experimental-psychology literature has thoroughly described the prevalence effect, a phenomenon in which people commonly miss rare things on visual searches. When volunteers were given baggage-screening tasks, errors in detecting targets increased as the prevalence decreased and as the variety of targets increased. This phenomenon is easily applied to the ECG screening dilemma: the likelihood of identifying pathology on ECG is low as the prevalence of pathology is low and the variety of pathology is high.

Despite difficulties in ECG interpretation, many proponents of preparticipation ECGs look to the Italian experience. In Italy, mandatory preparticipation screening including ECGs for all athletes has been correlated with an 89% decline in SCD. However, the sudden-death rate in Italy was 2.5-fold higher when this law was first implemented in 1982, compared with similar US populations. By 2004, the rate of SCD in Italy was not significantly different from that in the United States. This was partially the result of the detection of arrhythmogenic right ventricular cardiomyopathy in Italian athletes. Perhaps more important, Italy has a different structure for the preparticipation screening exam. The history, physical examination, and ECG (including interpretation) are performed by a sports cardiologist who has completed a 4-year training program, works in a sports medicine facility, and sees athletes for these examinations on a regular basis. This system is in contrast to the American system, in which a general pediatrician usually does the history and physical examination and then faxes the ECG to a pediatric cardiologist for interpretation.

Although arrhythmogenic right ventricular cardiomyopathy is a major contributor to SCD in Italy, HCM is the most common cause of SCD in the United States. Of note, the Italian data showed no significant difference in the number of disqualified athletes due to detection of HCM before versus after ECG screening was mandated. The incidence of death due to HCM before and after ECG screening is not clear in the Italian data because the incidences of all cardiomyopathies are reported together.

In our study, electrocardiographic evidence of HCM was commonly misinterpreted. Although this is disappointing, some respondents were able to identify that an abnormality was present, and they ordered additional testing more often.
than they were able to report diagnoses. The spectrum of abnormalities seen on an ECG of a healthy athlete can be indistinguishable from that of some patients with HCM. If pediatric cardiologists were to read an ECG as evidencing HCM even though the heart was normal, many young people would be inappropriately excluded from sports. The implications of restricting healthy patients from participating in sports are particularly worrisome in light of the childhood obesity epidemic. On the other hand, we also found a considerable number of respondents who would not restrict patients’ athletic participation despite ECGs with evidence of HCM, myocarditis, WPW, or LQTS. The consequences could be devastating because patients and families could have a false sense of security after knowing the results of an ECG.

In addition to the consequences of false-positives and false-negatives, a preparticipation ECG screening program brings enormous cost. In this study, the proportion of unnecessary tests ordered was high. Based on reimbursements approved by the Center for Medicare Services, the approximate reimbursement is $104 for an exercise test, $431 for an echocardiogram, and $160 for a Holter monitor.29,10 If there are 10.7 million participants younger than 40 years old in all amateur and competitive sports, the cost of unnecessary tests would be over $1 billion.1

The difficulty in ECG interpretation identifies a need for reevaluation of pediatric cardiology training. According to the American College of Cardiology/AHA recommendations, all physicians must interpret 500 supervised ECGs to attain initial competency.30 Beyond general training, cardiology fellows should interpret 3500 supervised ECGs to obtain competency, according to the Accreditation Council for Graduate Medical Education Residency Review Committee for Cardiovascular Diseases.31 Despite these recommendations, pediatric cardiology fellowship programs do not have a standard curriculum for teaching ECG interpretation or formal testing of interpretation competency.32

Several limitations are present in this study. It is not a population-based study and the sample is small. Thus the utility of our cost estimates are limited. The respondents were self-selected because they were members of an academic society and voluntarily chose to complete the survey. The computer interpretation and clinical information were not included in the ECGs in our survey. Reading ECGs on screen in a survey is clearly different from working in a clinical setting. Last, the gold standard used was the interpretations of 2 pediatric electrophysiologists at our institution. The difference between the scores of our electrophysiologists used as the gold standard and those of other electrophysiologists may be attributable to increased attention and focus during the survey. Although our electrophysiologists are not immune to the difficulties of ECG interpretation, they can provide the most expert interpretations available for pediatric ECGs in our institution. Given these limitations, we caution against drawing unsupported conclusions from these data.

SCD is an emotional issue. An ideal screening test would have high sensitivity, high specificity, and easy interpretation. However, preparticipation ECGs do not have perfect sensitivity or specificity, and the interpretations are difficult, even for pediatric cardiologists. Preparticipation screening ECGs could lead to high rates of inappropriate inclusion and exclusion from sports. The potential cost of incorrect ECG interpretations is enormous. If preparticipation screening ECGs are found to be beneficial, it is important to consider carefully who should read these ECGs, how to train the people reading them, which criteria should be used for interpretation, and how to do further education concerning sports restriction.

References

18. Mased D, Dawdy JA, Melendez LJ. Strict reliance on a computer algorithm or measurable ST segment criteria may lead to errors in thrombolytic therapy eligibility. Am Heart J 2000;140:221-6.