ABSTRACT—Several consultant cardiologists were invited to assess the accuracy of the Lifepak 200 automated external defibrillator (AED) (Fig. 1), from ECG records collected from pre-hospital cardiac arrest victims. They were asked to classify the ECG rhythms, and also give an opinion on whether or not a shock should have been given, and the potential harm of inappropriate treatment. As there was no absolute agreement between cardiologists in rhythm classification, sensitivity of the AED for ventricular fibrillation varied from 78% to 100%, and the specificity was between 92% and 100% according to each cardiologist. They agreed that all ventricular fibrillation should be shocked and failure to do so would reduce a patient's chances of survival; and that all other rhythms, and asystole, should not be shocked. Most experts believed shocking asystole would not be harmful, but opinions regarding the potential harm of administering shocks to patients with pulseless rhythm were mixed.

The first automated external defibrillator (AED) was described in 1979 [1]. Initially, the accuracy of the device was of prime concern and trials were considered imperative before use of AEDs could be recommended [2]. A consistent format for reporting accuracy of AEDs was established; ie 'sensitivity', meaning ability to recognise ventricular fibrillation and advise a shock; and 'specificity', or ability to recognise non-ventricular fibrillation. Accuracy is measured according to the response of the device to individual ECG segments rather than number of patients treated. Subsequent trials reported sensitivities of 89–92% and specificities of 98–100% in the field [3,4].

Although critics initially expressed concern over the possibility of inappropriate shock administration, supporters of automatic defibrillation have countered this by pointing out that the few errors in rhythm recognition made by AEDs result in failure to shock rhythms requiring defibrillation (usually low voltage ventricular fibrillation), rather than inappropriate defibrillation of organised rhythms.

The limitations of current technology do not allow 100% sensitivity and 100% specificity of AEDs. It is accepted that where machine specificity is increased, there is a proportionate loss in machine sensitivity, and vice versa.

Clinicians who are unfamiliar with AEDs may be confused over the issue of whether AEDs are 'safe' because of the inability to demonstrate 100% accuracy, and reluctant to sanction the use of these devices in their own area. Therefore definition of universally accepted optimum levels of sensitivity and specificity, and justification of the levels chosen, is required.

However, defining an optimum level of sensitivity and specificity for AEDs is complicated by other factors. First, there is no universal standard definition [5] for ventricular fibrillation. Second, there are 'transitional' cardiac arrest rhythms which cannot easily be classified, for which shock administration may not be inappropriate [6]. Therefore, the assessment of AED accuracy depends on arrhythmia definition and consistency in rhythm interpretation, and consistent views on which rhythms should be defibrillated and which should not.

In view of these difficulties, several experts were asked independently to assess the accuracy of an AED, to give an opinion on which ECG rhythms should or should not be defibrillated, and what might be the consequences of inappropriate shock administration or failure to shock.

Device used

The Lifepak 200 AED (Physio-Control Corporation) (Fig. 1) used in this study has already been subjected to field and laboratory testing, with sensitivities of 96% for coarse ventricular fibrillation, 69–75% for fine ventricular fibrillation, and 100% specificity, using the following definitions: coarse ventricular fibrillation amplitude > 0.2 mV (fine ventricular fibrillation 0.08–0.2 mV, asystole < 0.08; shocks are permitted for QRS widths of >/= 160msec, rates > 120/min, and no apparent P waves.

Method

Ambulance staff in South London used the Lifepak 200 AED to treat patients with cardiac arrest for 15
months. The results of this trial, in terms of patient survival, have already been reported [7]. ECG data from 267 patients were recorded onto a cassette tape, and retrieved from the machine and recorded onto paper after the event.

Altogether, 685 segments of ECG were collected. The strips were randomised and split into seven batches of 90 to 100 strips. Fourteen cardiologists were approached to assist with assessing the strips, in order that each batch of strips could be assessed by two cardiologists. It was anticipated that assessing 100 strips would take approximately 20 to 30 minutes, which was considered the maximum time which could be requested of an independent assessor.

Each assessor received an explanatory sheet outlining the purpose of the research and including guidelines for completing the assessment (Fig. 2).

For each ECG segment, they were asked to allocate the rhythm to one of five categories, and give their opinion on a number of issues (Fig. 2).

**Analysis of results**

A comparison of the rhythm categories assigned by each of the cardiologists was performed using Cohen’s kappa [8]. This test gives a measure of diagnostic consistency between independent examiners. All rhythms assigned by each cardiologist were reclassified into two categories, ‘ventricular fibrillation’ or ‘all other rhythms’. The results for each pair of cardiologists were cross-tabulated in a four-way table and the test applied. Rhythms assigned by each pair of cardiologists were then similarly classified to the categories of ‘fine ventricular fibrillation’/‘other rhythms’, asystole/‘other rhythms’ and ‘EMD’/‘other rhythms’, and the test applied for each comparison.

A kappa score of one would indicate perfect agreement, a zero score that the classification was done as if at random, and a negative score that there would have been greater consistency if rhythms had been randomly assigned. A score of 0.8 or more indicates ‘good’ agreement, over 0.6 ‘substantial’ agreement, and over 0.4 ‘moderate’ agreement [8].

**Sensitivity and specificity**

For each batch of strips, the strips each cardiologist categorised as ‘ventricular fibrillation’ or ‘fine ventricular fibrillation’ were counted and the percentage of strips in these two categories that triggered an AED

---

**Fig. 1. Lifepak 200 automated external defibrillator (AED)**

**Fig. 2. Guidelines for assessors**

1. Traces are calibrated to 1Mv
2. Include ventricular standstill in asystole category
3. Include in ‘artefact’ section if the rhythm cannot be determined due to the presence of artefact.
4. Since there is no universally accepted criteria for ‘fine VF’, define ‘fine VF’ according to your own clinical practise
5. EMD stands for ‘electro-mechanical dissociation’, or an organised electrical rhythm with no output.
6. Assume no pulse is present
shock (i.e. sensitivity), was calculated. For all other categories (asystole, EMD, and artefact), the percentage of these segments not shocked by the AED was calculated (specificity).

Results

Seven batches of electrocardiograms were each assessed by two cardiologists. For each batch, Table 1 shows rhythms assigned by both cardiologists, the kappa score for each rhythm category, AED sensitivity for ventricular fibrillation and 'fine' ventricular fibrillation, and specificity for non-ventricular fibrillation according to rhythms assigned by each person.

Treatment of rhythms categorised and the implications of inappropriate treatment

In all cases where ventricular fibrillation or 'fine' ventricular fibrillation were identified, the responses of the cardiologists to the questions, 'should a shock have been given?' and 'is inappropriate shock administration or failure to shock detrimental to the patient?', are shown in Figure 3. In all cases where EMD, asystole or artefact were identified, their responses are shown in Fig. 4.

Discussion

Consistency of rhythm interpretation between the cardiologists

There was no absolute agreement between any of the pairs of cardiologists in any ECG rhythm category, but consistency was 'good' or 'substantial', according to kappa score, in most categories. Agreement between assessors for rhythms categorised as 'fine ventricular fibrillation was generally poorer than for other rhythms, particularly in the distinction between fine ventricular fibrillation and ventricular fibrillation, and fine ventricular fibrillation and asystole.

Sensitivity and specificity

Consistency in classification of arrhythmias between the cardiologists was good according to the analysis using kappa, but the differences that did occur resulted in a wide variation of accuracy (particularly sensitivity) attributed to the AED. Differences in sensitivity to ventricular fibrillation within batches according to different rhythm classification ranged from 0–14%. Specificity was above 90% and comparable between and within all batches (between 95% and 99%).

Table 1. Comparison of ECG categories assigned by each assessor in each batch (C = cardiologist)

| Rhythm          | 1     | 2     | 3     | 4     | 5     | 6     | 7     | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | C13 | C14 |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Total VF No. agreed Kappa | 32    | 23    | 16    | 16    | 18    | 22    | 15    | 19  | 11 | 7  | 16 | 16 | 17 | 18 |
| Sensitivity of AED (%) to VF | 78    | 92    | 94    | 81    | 94    | 81    | 100   | 100 | 92 | 100 | 88 | 94 | 100 | 100 |
| Total 'fine' VF No. agreed Kappa | 6     | 15    | 9     | 10    | 7     | 5     | 4     | 6   | 18 | 23 | 22 | 6  | 15 | 10 |
| Sensitivity of AED (%) to 'fine' VF | 44    | 57    | 11    | 30    | 100   | 100   | 100   | 83  | 83 | 78 | 31 | 50 | 47 | 60 |
| Total asystole No. agreed Kappa | 37    | 39    | 42    | 53    | 42    | 36    | 42    | 36  | 45 | 43 | 32 | 50 | 47 | 34 |
| Total 'EMD' No. agreed Kappa | 19    | 21    | 31    | 19    | 33    | 37    | 27    | 26  | 21 | 22 | 24 | 20 | 21 | 38 |
| Artefact | 3     | 2     | 2     | 2     | 0     | 0     | 2     | 3   | 0  | 1  | 5  | 7  | 0  | 0  |
| 'Specificity' | 98    | 97    | 100   | 100   | 95    | 96    | 92    | 99  | 99 | 97 | 97 | 94 | 96 | 96 |
Implications of incorrect treatment by the AED

In most cases, the cardiologists agreed that all ventricular fibrillation should be shocked, and failure to do so might reduce the chances of survival; asystole should not be shocked, but all but one of the cardiologists believed that administering a shock would not be harmful.

In most cases, the cardiologists thought that EMD or artefact should not be shocked. Opinions were divided regarding the potential harm of shocking artefact or organised rhythms without a pulse, but each person answered consistently. Some believed shocking would always be harmful, while others believed shocking would not be harmful in any segment. Others said that shock administration may be harmful in cases where the ECG rhythm was organised, with a rate of above about 40, but not in segments where there were isolated complexes or regular complexes with a very slow rate.

Implications for developing AED algorithms

These findings have implications for developing AED algorithms. The difficulty in achieving 100% accuracy in AEDs is largely due to the inability of the device to differentiate between fine ventricular fibrillation and asystole. It has previously been suggested that the possibility of increasing sensitivity to 'fine' ventricular fibrillation at the expense of reducing specificity to asystole should be examined if it is 'medically acceptable' [4]. Since a different component of the algorithm is responsible for detecting QRS complexes, reducing specificity for asystole should not alter specificity for rhythms that show distinct QRS complexes with rates above 10 per minute.

Thus, if it is believed that shock administration to 'fine' VF is always indicated, and that shock administration to asystole, although not indicated, is unlikely to harm the patient; there is little benefit in meticulously adjusting the AED algorithm to achieve an acceptable level of sensitivity for 'fine' ventricular fibrillation in order to maintain 100% specificity for asystole. Also, as far as rhythm classification is concerned, most inconsistency occurred in the classification.
tion of fine ventricular fibrillation and ventricular fibrillation, and fine ventricular fibrillation and asystole. This implies a subjective element in ECG rhythm classification, and a broader margin of safe practice where treatment of asystole and ‘fine’ VF is concerned.

Disadvantages of reducing the specificity of AEDs

The views expressed were based on the assumption that for each of the ECG segments assessed, no pulse was present. Many of the ECG rhythms were bizarre and difficult to categorise, and the cardiologists were aware that the rhythms had originated from pre-hospital arrest victims. It was their informal opinion that the fate of such victims was unlikely to be made worse, whatever treatment was administered. These views are therefore tailored to the use of the AED by operators who can be relied upon to recognise cardiac arrest in a situation where the fate of the patient is unlikely to be worsened. These conditions may not always apply, particularly in the case of less-experienced rescuers for whom AEDs are also intended (for example, lay first-aiders, or family members).

Making sure that patients are pulseless before applying the AED, and that the device is not used if the patient is moving or being moved (since a reduction in specificity for asystole may also result in some loss of specificity for artefact [9]), are already important aspects in the training of an AED operator: these aspects will become crucial if the safety limits of AEDs are narrowed. If AEDs become less specific, the consequential increase in the length of the training programme or level of ability necessary to become an operator are likely to reduce the pool of potential AED users.

An alternative would be to develop AEDs of variable sensitivity and specificity. For users who can be relied on accurately to diagnose cardiac arrest, a highly sensitive device likely to detect ‘fine’ VF is the best option, even if this results in occasional defibrillation of asystole. For inexperienced users the safety aspect is paramount, and a highly specific device which may fail to recognise some examples of ‘fine’ VF would be the best choice. Since the survival rate from low amplitude VF is only about 6% [5], this would not have major consequences in terms of patient survival.

Summary

The findings, and the views expressed by the consultant cardiologists, are reassuring regarding AED accuracy and function. It is unlikely that 100% accuracy of AEDs could ever be achieved since classification of ECGs associated with cardiac arrest appears to involve an element of subjectivity, although this does not result in unsafe practice.

AEDs can be programmed to shock precise electrocardiographic criteria, and it is suggested that AEDs should be either predominantly sensitive or specific, dependent on the type of user for which they are intended. Neither option would have major limitations in terms of patient survival or safety.

Acknowledgements

We are grateful to Physio-Control Corporation for loan of the defibrillators, to British Telecom plc for assistance with funding, and also to the cardiologists who took part in the study: Dr Cliff Bucknell, Dr Nigel Buller, Dr John Clelland, Dr Wyn Davies, Dr Rex Daw-
Automated external defibrillators

References

1 Daick AW, Welborn WS, Rullman RG, et al. An automatic cardiac resuscitator for emergency treatment of cardiac arrest. Medical Instrumentation 1979;13:78.

2 Cummins RO, Eisenberg MS, Moore JE, et al. Automatic external defibrillators: clinical, training, psychological, and public health issues. Ann Emerg Med 1985b;14:735-41.

3 Stults HR, Brown DD, Kerber RE. Efficacy of an automated external defibrillator in the management of out-of-hospital cardiac arrest: validation of the diagnostic algorithm and initial clinical experience in a rural environment. Circulation 1986;73:701-9.

4 Edwards DG. Development of a decision algorithm for a semi-automatic defibrillator. Ann Emerg Med 1989;18:1264-8.

5 Weaver WD, Cobb LA, Dennis D, et al. Amplitude of ventricular fibrillation waveform and outcome after cardiac arrest. Ann Intern Med 1985;102:53-5.

6 Cummins RO, Stults KR, Haggar B, et al. A new rhythm library for testing automatic external defibrillators: performance of three devices. J Am Coll Cardiol 1988;11:597-602.

7 Walters G, D'Auria D, Glucksman EE. Controlled trial of automated external defibrillators in the London Ambulance Service. J Roy Soc Med 1990;83:565-5.

8 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

9 Bocka JJ. Automatic external defibrillators. Ann Emerg Med 1989;18:1264-8.

Address for correspondence: Dr E. Glucksman, Accident and Emergency Department, King's College Hospital, London SE5 9RS.

Notes for contributors

The Journal seeks articles of merit, whether they be reports of original research or critical reviews. The Editor is assisted in his choice of articles by expert referees. The following notes are for the guidance of potential contributors:

The title of the article should be as succinct as possible and the text should be interspersed with subheadings, as appropriate.

A summary of not more than 200 words should be included and will be printed at the start of each article.

The text should be typewritten in double spacing on one side of A4 paper and pages should be numbered consecutively. Please write as concisely as possible. Amendments should be made in the text and not in the margins. A separate title page is requested; this should carry the title, the name of the author(s), medical degrees and principal appointment(s), and, in addition, a full postal address for correspondence should be given. Two copies of each article should be supplied.

SI units should be used in the text and the author is responsible for their correct usage.

References

References should conform to the Vancouver style. They should give names of all authors, followed by initials (unless there are more than six authors in which case the first six only should be given, followed by et al). The title of the article should be followed by the title of the journal (abbreviated to the style of Index Medicus), the year of publication, volume number, and first and last pages. References to books should give the names of authors, (and/or editors), title, place of publication and publisher, followed by the year of publication.

Please note that the accuracy of references will be the responsibility of the author.

Each Figure (illustration) must be submitted on a separate sheet and given a number which accords with the order in which it appears in the text. Tables also should be presented in this way. Captions to Figures should be as brief as possible and listed separately. Each Figure should be identified with the name of the author as well as a number. The type area is 17.5 cm × 22.5 cm deep (single column width 8.4 cm). Figures will be reduced, where necessary, to fit this format.

Line Figures should be presented on plain white paper or drawing paper. SI units should be used throughout. Photographs (including x-rays and micrographs) should be printed on high-quality glossy paper and any lettering should be presented on an overlay.

Fillers on medical and related topics (historical, anecdotal or critical) are welcomed. They should not exceed 700 words, and are published at the discretion of the editor.

This Journal will accept disks. Details are printed elsewhere.