SUPPLEMENTARY INFORMATION

Interpretation performance of seven ECG programs in identifying arrhythmia and acute cardiovascular syndrome

J De Bie[1], I Diemberger[2], C. Martignani[2], G.Massaro[2],
[1]Mortara Instrument Europe s.r.l., Bologna, Italy
[2]Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Bologna, Italy

Supplementary Methods

Selection of ECGs

ECGs used in the study were acquired with equipment that complied with the requirements of International Electrotechnical Commission standard IEC 60601-2-51:2003 and were representative of those seen in hospital or acute-care settings. Anonymized ECGs had been obtained from consecutive patients in eight centers in the USA, Italy, and Australia, including three centers in Italy that provided pediatric ECGs. In addition, we used 300 + 300 ECGs randomly chosen from consecutive patients in two hospital databases outside Europe (one in the USA and one in Australia) and all ECGs from a European ambulance service database and a European university hospital database that had critical value statements flagging acute myocardial infarction from an automatic interpretation program (Mortara Instrument Veritas, version 7.3). To avoid selection bias, we excluded ECGs with statements already attached from the ACS analysis and included them only in the rhythm analysis. We excluded ECGs from patients with pacemakers because pacemaker spikes are usually detected in the analog front-end portion of electrocardiographs and are not faithfully recorded in stored digitized records.

Re-recording of ECGs

In order to use the ECGs, we had to convert them back into analog format and replay them into physical electrocardiographs. We required 10 s ECGs acquired at 1000 samples/s and, therefore, resampled those that were originally acquired at 500 samples/s by linear interpolation. When creating looped records, to avoid discontinuity we inserted a cubic spline into the samples between the first P wave and the last T wave, thus connecting the end and beginning of each record. The length of the spline was chosen to create an RR interval that was the average of the whole record. This process resulted in records that were not all exactly 10 s long. To enhance reproducibility of the capture, we stretched or compressed records to exactly 10 s in length. We excluded records that needed to be stretched or compressed by more than 10% or had a large offset or slope of the cubic spline.

To replay records, we used a Whaleteq MECG 2.0 Multichannel ECG Test System (WHALETEQ Co Ltd, Taipei City, Taiwan) connected to a laptop PC. The Whaleteq device is specifically constructed to reproduce ECG records faithfully and noise free. Up to four electrocardiographs were connected in parallel via the system's breakout box, with their original patient cables. The recording environment was as free as possible from other electric and magnetic disturbing sources, and care was taken to position the connecting wires to avoid interference between devices. Periodically, the printed ECGs were visually checked to ensure they were identical to the source record. The electrocardiographs were configured to provide a full ECG interpretation in English, and all filters

were set to a minimum or off, except for the AC-interference filter, which was set to the mains frequency of 50 Hz. Two operators, blinded to the aim of the study, entered record identifiers, age, and sex. If the patient's age and/or sex were unknown (~30% of ECGs, all in adults), age was set at 40 years and sex was randomly assigned as male or female. Recordings were saved digitally with paper backup copies. Any records with errors were corrected or the ECGs were re-recorded.

The most representative interpretation of each triplet for each device was used for analysis. If two of three interpretations were identical, one of these two was automatically chosen (occurred in 70–92% of cases). When all three interpretations were different, they were reviewed manually by one of the authors (JdB), without having access to the ECG waveform and the most representative interpretation was selected, based on that which had the most statements in common with the other two records. If that rule was not conclusive (which was rare), the interpretation that was most "normal" was selected. On average, only 15% of cases needed manual selection.

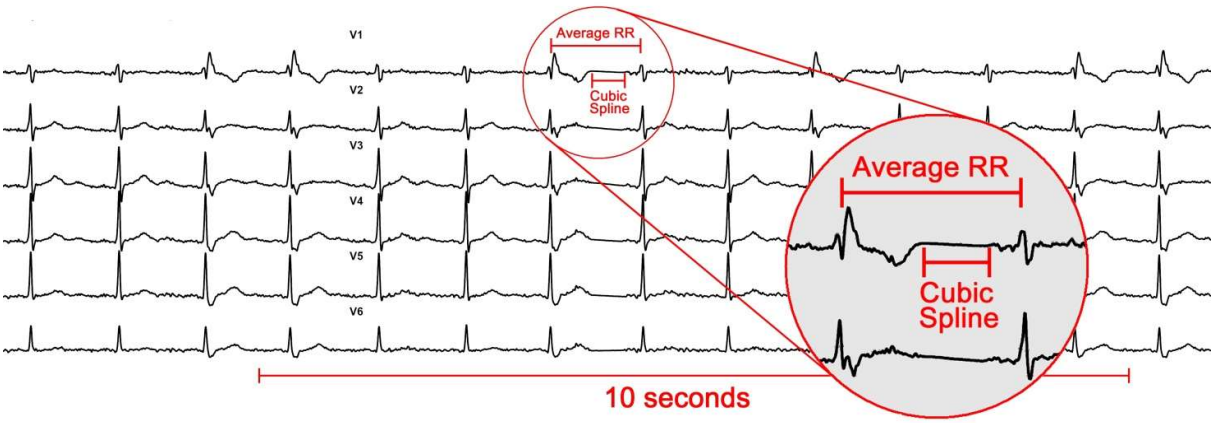
Manufacturer interpretation statements

Each manufacturer uses different wording for interpretation statements, including for probability and severity of conditions. For the rhythm analysis, we defined 52 clinical classes that could be combined to answer specific research questions. For example, the clinical class supraventricular rhythms included ectopic atrial rhythms, junctional rhythms, and the generic term supraventricular rhythm (including bradycardia and tachycardia for all these rhythms). Arrhythmias such as single or multiple ectopic beats and sinoatrial or atrioventricular blocks were not considered, and heart rate did not change rhythm classification (eg, sinus tachycardia was classified as sinus rhythm).
For the ACS analysis, we only considered the interpretation statements "critical value" or "critical test results" or other statements indicative of possible ACS (eg, "acute MI"). Text parsing rules were created to decide in which class the interpretations for infarction belonged. The final result of the parsing was a single TRUE/FALSE decision for each ECG for the ACS condition.

ECG Interpretation

ECGs presented for interpretation by experts were printed or displayed on a high-quality screen, with a 1 mm grid in a contrasting color, in a traditional 3×4 format with a 10 s rhythm lead II for the ACS readings and simultaneous 12 leads by 10 s for the rhythms, and with the patient's age and sex displayed but no interpretation statements.

For rhythm interpretations, if all seven programs agreed, the interpretation was accepted, and no human intervention was required. If any program disagreed, the rhythm shown in the ECG was established by an independent reviewer who was a highly experienced cardiologist and confirmed by author JdB. If the interpretations of the reviewers did not match, the case was discussed until consensus was reached.



Supplementary Figure S1: 14 s snapshot of a continuous record captured by replaying a 10 s ECG sample in a loop on an electrocardiograph

The start and end of the original record, where a cubic spline the length of the average RR interval was inserted, is enlarged. The aberrantly conducted beats show that the trace repeats every 10 s.