A Long-term ST Database for Development and Evaluation of Ischemia Detectors

F Jager, GB Moody[1,2], A Taddei[3], G Antolič[4], M Emdin[3], A Smrdel, B Glavić, C Marchesi[3], RG Mark[1,2]

Faculty of Computer and Information Science, Ljubljana, Slovenia
[1]Harvard-M.I.T., Division of Health Sciences and Technology, Cambridge, USA
[2]Cardiology Division, Beth Israel Deaconess Medical Center, Boston, USA
[3]CNR Institute of Clinical Physiology, Pisa, Italy
[4]Department of Cardiology, University Medical Center, Ljubljana, Slovenia

Abstract

In this paper we report the status of an ongoing international collaborative research effort to produce a new long term ST database (LTST DB), a collection of seventy annotated ambulatory records containing transient ischemic and non-ischemic ST changes. We present the selection criteria for records, an annotation protocol with definitions of transient ST events, interactive graphic tools for manual and automatic annotating, and the annotation procedure.

1. Introduction

One of the tasks during ambulatory electrocardiographic (AECG) monitoring is analysis of transient ST-segment and T-wave changes compatible with ischemia. Most AECG instruments do not distinguish between ischemic and non-ischemic ST and T changes. The European Society of Cardiology ST-T database (ESC DB) [1], developed by the ICP group in Pisa, is a generally available set of 90 representative 2-hour 2-lead ambulatory excerpts with well-characterized and annotated ischemic and non-ischemic ST segment and T wave changes. It is an invaluable tool, and represents a standard set of material for developing algorithms to detect ST episodes, assessing their quality, and quantifying ST deviations. The ESC DB contains small numbers of non-ischemic ST episodes, "mixed" episodes (non-ischemic episodes containing ischemic episodes within), cases of slow ST level drift, as well as varied temporal patterns of transient ischemic ST changes likely to be connected to underlying mechanisms responsible for ischemia, which led us to develop a new long-term database (LTST DB) [2] as a complement to the ESC DB. The LTST DB is developed by a joint collaborative effort of our research groups from Ljubljana, Boston and Pisa. It is planned to contain up to 70 annotated 2-channel ambulatory records, each 24 hours in duration. The goals of the LTST DB are to better represent the wide variety of "real-world" data with many examples of mixed and non-ischemic ST events, to include quasi-periodic and other temporal patterns in ST level to enable researchers to study these phenomena in a broader context, to stimulate development of algorithms for differentiating ischemic from non-ischemic ST events, and to permit more reliable prediction of clinical performance. The aim of this paper is to report the current status of the LTST DB.

2. Methods

We are selecting new records from those obtained in routine clinical practice at the BIDMC in Boston and at the ICP in Pisa. The records are being collected from subjects with and without known coronary disease (to model real-world clinical conditions as far as possible) while documenting significant numbers of ischemic and non-ischemic ST events. Each selected record must contain ***significant*** ($> 100\mu$V) transient ST segment episodes corresponding to known ischemia (ischemic ST episodes), non-ischemic ST episodes (axis-related non-ischemic ST episodes, episodes of slow ST level drift, and heart-rate related ST episodes), or mixed episodes (significant episodes containing other significant "sub-episodes"). Each record is accompanied by a report including results of other clinical investigations (if performed) such as ECG stress tests, echocardiograms, coronary angiography, ventriculography, or thallium scintigraphy; relevant information on patient conditions, i.e., on electrolytes, medications and diagnosis; and technical information about the record. We are also including a number of 24-hour records with proven acute myocardial ischemia such

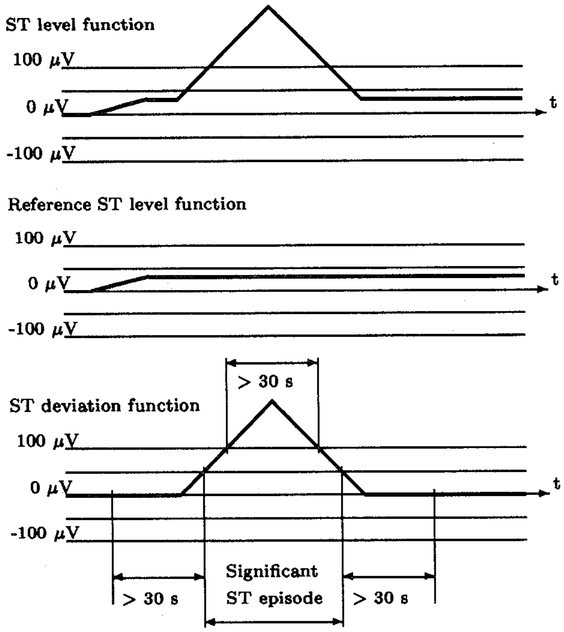

Figure 1. Definition of significant ST episode. The ST deviation is defined as the difference between the ST level and the reference ST level.

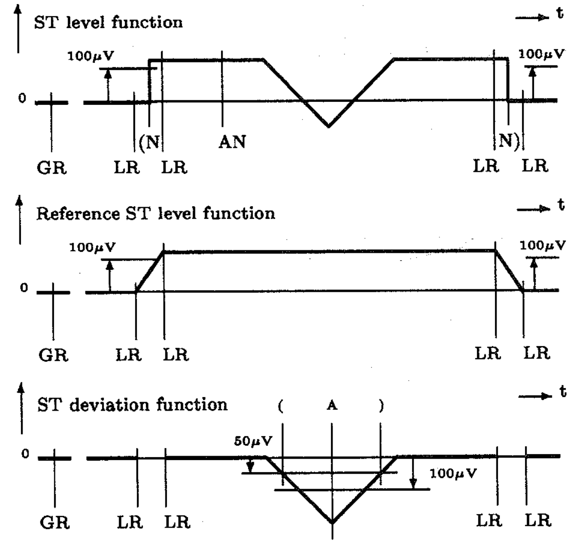

Figure 2. Schematic representation of annotating during a mixed ST episode (an axis-related non-ischemic ST episode containing an ischemic local episode). (*Legend:* "GR" - global reference; "LR" - local reference; "(N", "AN", "N)" - the beginning, extremum, and end of the mixed episode; "(", "A", ")" - the beginning, extremum, and end of the ischemic local episode.)

as Prinzmetal's, unstable, resting, or effort angina, selected from recordings made at the ICP from which the 2-hour excerpts in the ESC DB were obtained. The recordings are being made using standard AECG recorders. Since AECG recorders preserve frequency content in the signals typically up to 30 Hz (at best, 45 Hz), we digitize the records at 125 samples per second per channel with a resolution of 12 bits (5 μV).

The annotation protocol is compatible with that established for the MIT-BIH Arrhythmia and ESC database, and extended in order to permit more detailed description of non-ischemic and mixed ST events. ST events are defined and annotated in each ECG lead separately. We retained the ESC DB's definition of a significant ST episode:

- An episode begins when the magnitude of the ST deviation first exceeds 50μV;
- The deviation must reach 100μV or more throughout a continuous interval of at least 30 seconds;
- The episode ends when the deviation becomes smaller than 50μV, provided that it does not exceed 50μv in the following 30 seconds;

with the difference that the *ST deviation* is defined (see figure 1) as the difference between the *ST level* (measured 80 ms after the J point, or 60 ms after if heart rate exceeds 120 bpm) and the *reference ST level function*, and not as the difference between the ST level and a single reference value. The reference ST level function is defined as a piecewise linearly interpolated function between knot points in the ST level function that have been marked by *local reference annotations.* Any significant ST event must meet the significant episode criteria. *Each* significant ST episode is annotated by its beginning, extremum, and end, and by its type. *Ischemic ST episodes* typically exhibit a triangular pattern of ST deviation over time, with a characteristic change in ST segment morphology. *Axis-related non-ischemic ST episodes* are due to position-related (postural) changes in the cardiac electrical axis. Such an episode exhibits a stable ST deviation course and begins and/or ends with a significant concurrent axis shift. *Drift ST episodes* are due to slow (non-postural) changes in the cardiac electrical axis, or to effects of medication on repolarization. *Heart-rate related ST episodes* are similar to ischemic episodes accompanied by sharp increases in heart rate, but the change in ST segment morphology is due to shortening of the repolarization interval. Various significant *mixed ST episodes* could be of any type (ischemic or non-ischemic) containing subepisodes of any type within. Figure 2 schematically describes annotating during such a mixed episode. The boundaries of the mixed episode itself are determined by ST levels relative to the global-reference annotation. Those of the sub-episodes (local episodes) are defined by ST levels relative to the reference ST level function, which during the mixed episode approximates its general course.

The group at the FCIS in Ljubljana developed an interactive graphic editor (SEMIA, mainly programmed by the seventh author) for manual and automatic annotating. Preprocessing procedures [2] automatically derive time series of instantaneous heart rate, ST level and Karhunen-Loève (KL) transform-based QRS complex and ST segment morphology features. The interface was developed to provide detailed insight into the data, and to support and coordinate annotation of the records by cardiologists at different sites interacting via the Internet and without paper tracings. Figure 3 illustrates the appearance of the interface to the user. Annotations may be inserted, examined, moved, deleted, and modified. Events that satisfy the criteria of significant ST episodes can be identified and annotated automatically within manually marked intervals. The time series can be displayed at scales that permit as little as 1 minute or as much as 24 hours to be seen at once, while viewing between 1 second and 1 minute of the original or time-averaged ECG signals.

Annotating is done in three phases by cardiologists from the UMC in Ljubljana, from the BIDMC in Boston and from the ICP in Pisa by exchanging annotation files via the Internet in a "round robin" fashion. (Each site works on one third of records during each phase.) The ***first phase*** entails deriving correct ST level functions on the basis of manually determined positions of the isoelectric level and the J point. Using ST level functions and time series of KL coefficients derived during automated preprocessing, annotators first manually set "dummy" annotations. Time-averaged heart beats are then computed over 16-second windows surrounding each dummy annotation. Annotators manually adjust the positions of the isoelectric level and J point for these average beats. The positions of the isoelectric level and J point thus obtained estimate these two positions in the original ECG signal at the heart beat corresponding to the dummy annotation. Dummy annotations are set roughly prior to ST episodes, at the extrema of ST episodes, after the ST episodes, prior to axis shifts, after the axis shifts, otherwise roughly every 20 minutes. After storing dummy annotations, a postprocessing procedure automatically estimates the positions of the isoelectric level and J point for all heart beats that were marked as clean during preprocessing (by linearly interpolating the two positions over time), derives a time-averaged (16 seconds) heart beat for each individual clean heart beat, and computes ST amplitudes at J, J+20ms, J+40ms, J+60ms, J+80ms, J+100ms, and J+120ms point, on these averaged beats according to the estimated positions of the isoelectric level and the J point. The ST level function is then computed by taking the measurements obtained on time-averaged beats at J+80ms (or J+60ms if heart rate exceeds 120bpm), resampled (0.5 Hz), and smoothed (7 point moving average). These ST level functions replace those derived during preprocessing. The ***second phase*** consists of setting global-reference and many local-reference annotations in the ST level function for each ECG lead throughout the records. By inserting the local references, the annotators define the reference ST segment level function, used to derive ST deviation functions for automatic detection of ST episodes. In the ***third phase***, annotators review and correct the reference ST level function, and label ST episodes. Significant mixed episodes are annotated manually in the ST level function with regard to the preceding global-reference annotation, while other significant individual ST episodes and significant local episodes are annotated in the ST deviation function using automatic procedure. Finally, the annotators manually review and corrects these ST episodes.

3. Discussion and conclusions

Database development is complex, resource intensive, and time consuming. Among 68 collected, digitized and preprocessed ambulatory records so far, we have selected 42 (of these, QRS markers were manually reviewed/corrected during the tape scanning procedure for 10 records) of which 20 were annotated. When complete, the LTST DB will be published on CD-ROMs in the standard format used for the ESC DB. Each record includes a header file, a signal file, QRS and ST reference annotation files, a reference annotation file with measurements obtained on average heart beats and attached back to individual heart beats, and a clinical summary. Interactive graphic tools have allowed us to work nearly without paper, and facilitated international collaboration via the Internet. Reviewing and correcting the annotations after their automatic derivation instead of fully manual annotating proved to be much faster and more convenient for human experts. It is necessary to emphasize that the project benefits from the past expertise, resources and experiences of the three groups. The database is expected to be ready for release in 18 months.

Acknowledgment

This publication is based on work sponsored by Medtronic, Inc., Minneapolis, USA. Presentation of this

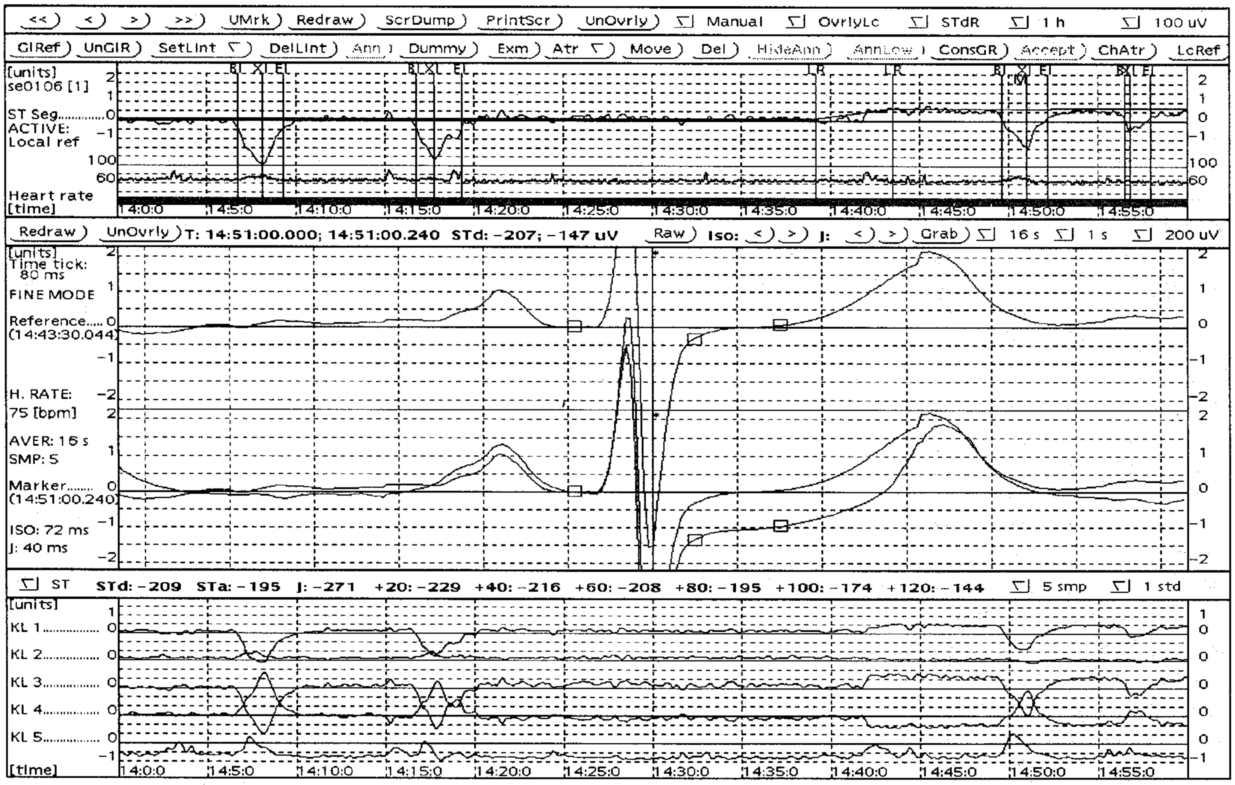

Figure 3. Appearance of the graphical editor SEMIA while annotating transient ST episodes in the ECG. *Upper:* ST level function (resolution: 100 μV/unit), piecewise linearly interpolated reference ST level function (above), heart rate (below), local reference annotations (LR) defining the knot points of the reference ST level function, and ST-episode annotations (BI, XI, EI) indicating four ischemic ST episodes; *Middle bar:* Values of ST deviation (STd) and ST level (in μV) at currently marked (M) time index (T); *Middle:* Time-averaged (over 16 seconds) ECG signals (resolution: 200 μV/unit) centered on the currently marked extremum (M) of the third episode (below, time: 14:51:0.24) and on the currently active local-reference annotation (above - also overlaid below, time: 14:43:30.044); *Lower bar:* Values of ST deviation (STd) between the two time-averaged heart beats at J+80ms (rightmost boxes in the middle panel), ST amplitude (STa) of the currently marked (M) average heart beat at J+80ms, and its ST amplitudes at J, J+20ms, J+40ms, J+60ms, J+80ms, J+100ms and J+120ms (in μV); *Lower:* Time series of ST segment KL coefficients (resolution: 1 std/unit).

work has been made possible by the Dean's Funds of the University of Ljubljana, Slovenia. We are also grateful for the support of the BIDMC Arrhythmia and Bioengineering Laboratories. We especially wish to thank Diane Perry, Nilo Roggero, Ing. Santa Ombrato and Ing. Gabriele Cecchetti from the BIDMC and from the ICP for their help in obtaining recordings, data acquisition and technical assistance.

References

[1] Taddei A, Distante G, Emdin M, Pisani P, Moody GB, Zeelenberg C, Marchesi C. The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. Eur. Heart J. 1992;13:1164–72.

[2] Jager F, Moody GB, Taddei A, Antolič G, Zabukovec M, Škrjanc M, Emdin M, Mark RG. Development of a long term database for assessing the performance of transient ischemia detectors. In: Computers in Cardiology 1996. Piscataway: IEEE Society Press, 1996;481–4.

Address for correspondence:
Franc Jager
Faculty of Computer and Information Science,
Tržaška 25, 1001 Ljubljana, Slovenia
Internet: `franc@manca.fri.uni-lj.si`