Development of a Long Term Database for Assessing the Performance of Transient Ischemia Detectors

F Jager[1], GB Moody[2], A Taddei[3], G Antolic[4], M Zabukovec[1], M Skrjanc[1], M Emdin[3], and RG Mark[2]

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

Abstract

We have begun to develop a new annotated long term ambulatory ST-T database. The aim of the database is to be a reference set containing a number of well documented ischemic ST episodes, axis-related non-ischemic ST episodes, episodes of slow ST level drift and mixed episodes to support development and evaluation of detectors capable of accurate differentiation of ischemic and non-ischemic ST events, as well as basic research into mechanisms and dynamics of ischemia. We discuss selection criteria, define the events of interest, and describe the annotation procedure.

1. Introduction

Ambulatory electrocardiographic (AECG) monitoring is widely used for analysis of transient ST-segment and T-wave changes compatible with ischemia. Most AECG instruments do not attempt to distinguish between ischemic and non-ischemic ST and T changes, however, because of a lack of standard definitions of transient ST-T events and knowledge about their meaning.

In order to study these events, and to evaluate and compare automated methods for their detection and interpretation, the ICP group in Pisa defined diagnostic criteria for transient ST and T changes, and a protocol for annotating them [1]. This group took the leading role in the development of the European Society of Cardiology ST-T Database (ESC DB) [2], which was the first generally available set of well-characterized, representative ECG recordings with documented ischemic and non-ischemic ST and T changes. The ESC DB has proven to be an invaluable tool for designers and evaluators of automated ischemia detectors. Its availability has stimulated extensive research and publication in this field during the past several years, including recognition algorithms based on time-domain analysis, the Karhunen-Loève Transform (KLT), neural networks, and fuzzy logic.

The ESC DB contains 90 two-hour, two-channel ambulatory records with 368 documented transient ischemic ST episodes, but only 11 non-ischemic ST episodes. Non-ischemic ST episodes, which are of no clinical interest per se, account for many of the false positives of automated ischemic ST detectors. Thus it is particularly important to understand these events and to define their distinctive characteristics in order to improve detector performance. The small number of non-ischemic episodes in the ESC DB does not permit exhaustive study of these differences, however.

Furthermore, our previous study on characterization of transient ST segment changes in the ESC DB[3], revealed two additional types of important ST events. We found three cases of "mixed episodes" (non-ischemic episodes containing ischemic episodes within), and 17 cases of significant ($>100\mu$V) slow drift of ST deviation level (15 of which also contain ischemic episodes). We also described striking and varied temporal patterns of transient ischemic ST changes. These observations provoke questions regarding the relationships between these patterns and the underlying mechanisms that are responsible for ischemia. We cannot answer these questions definitively, however, since we are not able to observe more than a handful of repetitions of each pattern in the two-hour segments of the ESC DB.

We have therefore begun to prepare a new, long-term, ambulatory ST database, in order to support the development and evaluation of detectors capable of more accurate differentiation of ischemic and non-ischemic ST changes, and to provide more examples of non-ischemic episodes, episodes of slow ST level drift and mixed episodes.

2. Methods

The Long Term ST Database (LTST DB) is being developed by the joint efforts of our research groups in Ljubljana, Pisa, and Cambridge. It is planned to contain up to 70 annotated 2-channel records, each 24 hours in duration, obtained from AECG recordings. These will include approximately 30 of the 90 24-hour recordings from which the two-hour excerpts in the ESC DB were obtained. The records have been selected to represent "real world" data as much as possible, while documenting significant numbers of ischemic and non-ischemic ST events. The annotation protocol is compatible with that developed for the MIT-BIH Arrhythmia and ESC Databases, but we have extended it to permit more detailed descriptions of non-ischemic ST events.

We obtain accurate human annotations of ST events using special-purpose interactive editing software developed by the FCIS group in Ljubljana, and using the general-purpose WAVE software system developed by the second author[4]. Each record also includes a compact clinical summary, with technical information about the recording as well as relevant clinical information (e.g., electrolytes, medications, and pathology). When complete, the database and associated utility software will be published on CD-ROMs in the standard MIT-BIH format, as also used for the ESC DB.

The recordings chosen for the original ESC DB were selected to include examples of baseline ST displacement resulting from conditions such as hypertension, ventricular dyskinesia, and effects of medication. From these recordings, we will include in the LTST DB those originally recorded by the ICP group.

In addition, we are selecting new AECG recordings from those obtained in routine clinical practice at Boston's Beth Israel Hospital (BIH) and at the ICP. Each selected recording must contain significant ($>100\mu$V) transient ST segment episodes corresponding to known or suspected ischemia, significant non-ischemic ST episodes, significant slow ST level drift, or mixed episodes. Recordings containing combinations of these events are preferred.

Both sets of analog recordings have been made using standard AECG recorders (the model of recorder used is documented in each case). The analog outputs of the playback units are passed through antialiasing filters and digitized. Since none of the AECG recorders preserves frequency content in the signals above about 45 Hz in the best cases, and closer to 30 Hz in typical cases, we digitize the records at 125 samples per second per channel. There is simply no additional information to be gained from using a higher sampling frequency for these recordings. The resolution is 12 bits, and the amplitude scale is 200 ADC units/mV for all signals.

As for the ESC DB, we defined *ST deviation* as a change in ST level relative to a reference level. Since some recordings exhibit fixed ST depression relative to the isoelectric level (due to prior infarcts, for example), it is not meaningful to define the significance of *transient* ischemic change in terms of *ST amplitude* (ST level relative to the isoelectric level) in these recordings. We identify the reference ST level by searching for a five-minute interval without significant variation in ST level as near as possible to the beginning of the record. Within this interval, a reference beat is selected and annotated for each ECG lead. The ST levels of these beats become the reference ST levels.

We define and annotate events independently on each channel, retaining the ESC DB's definition of *significant ST episodes*:

- 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.

Any significant event in the LTST DB must meet these criteria. The events of interest are ischemic episodes, non-ischemic episodes and episodes of slow ST level drift. *Ischemic episodes* typically exhibit a distinctive triangular pattern of ST deviation over time.

Based on our previous studies, we defined characteristics for *non-ischemic episodes* resulting from position-related (postural) changes in the cardiac electrical axis:

- A non-ischemic episode must exhibit a stable ST deviation level of less than 200 μV throughout.
- The episode must begin or end (or both) with a significant concurrent axis shift.

Axis shifts are best observed in time series of QRS morphology features.

Slow ST level drift is the most difficult event to recognize, especially if no other ischemic episodes are present. Drift may result from slow (non-postural) changes in the cardiac electrical axis, effects of medication on repolarization, or effects of changes in heart rate on repolarization. Since the cumulative effect of drift over periods ranging from 10 minutes to several hours may amount to a significant change of ST level (100μV or more), it cannot be ignored. Drift episodes are best identified from ST trend plots. Based on our previous studies, and the data at hand, we identify a *drift episode* as a significant ST episode that meets any of the following criteria:

- It contains one or more significant ischemic or non-ischemic episodes within.
- It is longer than the ischemic episodes in its neighborhood, and neither the temporal pattern of ST deviation nor the changes in ST morphology resemble those of the ischemic episodes.
- It appears due to rate-related ST-T changes.

In clinical practice, there is usually evidence independent of the ECG to support a diagnosis of ischemia. Hence it is likely that criteria such as those described above will miss *events of borderline significance* that would be considered ischemic in light of additional non-ECG

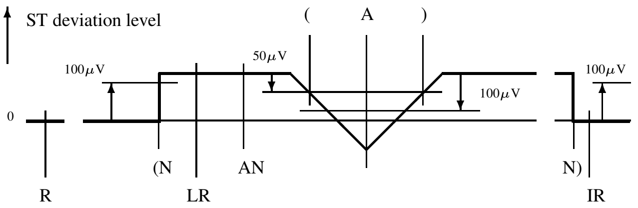

Figure 1. Schematic representation of a mixed episode (a non-ischemic episode containing an ischemic one). The shortened annotations shown here [R, LR, IR, (N, AN, (, A,), and N)] are defined in table 1.

Code	**M**eaning
([D|N] ST x $\pm$ 50, aaaa	*Beginning*
A [D|N] ST x $\pm$ dddd, aaaa [, s]	*Extremum*
[D|N] ST x $\pm$ 50, aaaa)	*End*
R ST x, aaaa	*Reference*
LR ST x, dddd, aaaa	*Local reference*
IR ST x, aaaa	*Initial reference*

Table 1. ST annotation codes used for the database. [D|N]: type of episode (D: drift, N: non-ischemic, none: ischemic); x: lead number (0 or 1); dddd, aaaa: ST deviation and amplitude in μV; s: subtype (+: borderline plus, $-$: borderline minus)

evidence. To account for these events, we also annotate episodes for which the maximum ST deviation is nearly 100μV, and which meet all of the other criteria for ischemic episodes. These are annotated as borderline plus or minus (respectively, with or without ST morphology change). Another category of events of borderline significance (borderline minus) is that satisfying the criteria described above but without ST morphology change.

Various significant *mixed* (compound) episodes require special treatment. This category includes non-ischemic episodes containing ischemic "sub-episodes," and drift episodes containing ischemic or non-ischemic sub-episodes. Figure 1 schematically shows such a mixed episode. The general trend of ST deviation during a mixed episode is more or less stable (typically between 100 and 200μV). While the boundaries of the mixed episode itself are determined by ST deviations relative to the reference ST level from the beginning of the record, those of the sub-episodes are defined within the context of the mixed episode by local ST deviations relative to a local reference beat. This local reference is selected and annotated immediately after the beginning of each mixed episode. Each ST annotation contains both the ST deviation relative to the initial (or local) reference, and the ST amplitude (deviation relative to the isoelectric level). Except for reference beat annotations, each ST annotation also contains information about the type of episode to which it belongs.

ST annotations are made manually with reference to the ECGs and to trend plots of heart rate and QRS and ST morphologic features. The trend plots[3] are produced using ARISTOTLE[5] for QRS complex detection and classification, followed by removal of baseline wander using a cubic spline approximation and subtraction technique, low-pass filtering by a 6-pole Butterworth filter (with a cut-off frequency of 55 Hz), and extraction of ST and QRS morphology features with heart rate. Next, the Ljubljana group's software for interactive ST analysis is used for rejecting abnormal beats and their neighbors, filtering of the feature time series (see figure 2, center trace), resampling, and smoothing (figure 2, top trace). This approach makes use of the representation power of the KLT series, while compact trend plots assure accurate detection of important as well as subtle events in the series [3]. Events are visually detected after the final preprocessing step. Annotations are made directly on the trend display (at user-selected scales from 2 minutes to 24 hours) after noise detection in the KLT space, with reference to the original ECG signals in the region of interest displayed using the WAVE system.

In the final phase of annotating, the exact locations of the ST annotations are determined by visual comparison of each clean beat of the original ECG signals in the region of an ST event with the reference beat at high resolution (see figure 2 at bottom). This requires manual determination of the isoelectric level and J point for each beat under consideration. As for the ESC DB, the ST segment amplitude and ST deviation level (according to the reference beat) are measured 80ms after the J point (or 60ms after the J point if the heart rate exceeds 120 bpm). The database will also include semi-automated measurements of ST amplitude at both J+60ms and J+80ms for each beat.

As of August, 1996, we had collected and digitized 50 24-hour records: 30 at BIH, 20 at the ICP. These 50 records have been preprocessed and 15 of them annotated by the FCIS group, and subsequently verified and corrected by a cardiologist. These 15 records contain 179 ischemic, 40 non-ischemic, 7 drift, and 10 mixed episodes.

The Ljubljana group's interactive ST analysis software is being developed in parallel with the database. Although the current version has been used in the annotation of 15 records, planned improvements should permit greater efficiency in the remaining work.

3. Discussion and conclusions

Analysis of transient ST events is considerably more complex than was believed before the development of

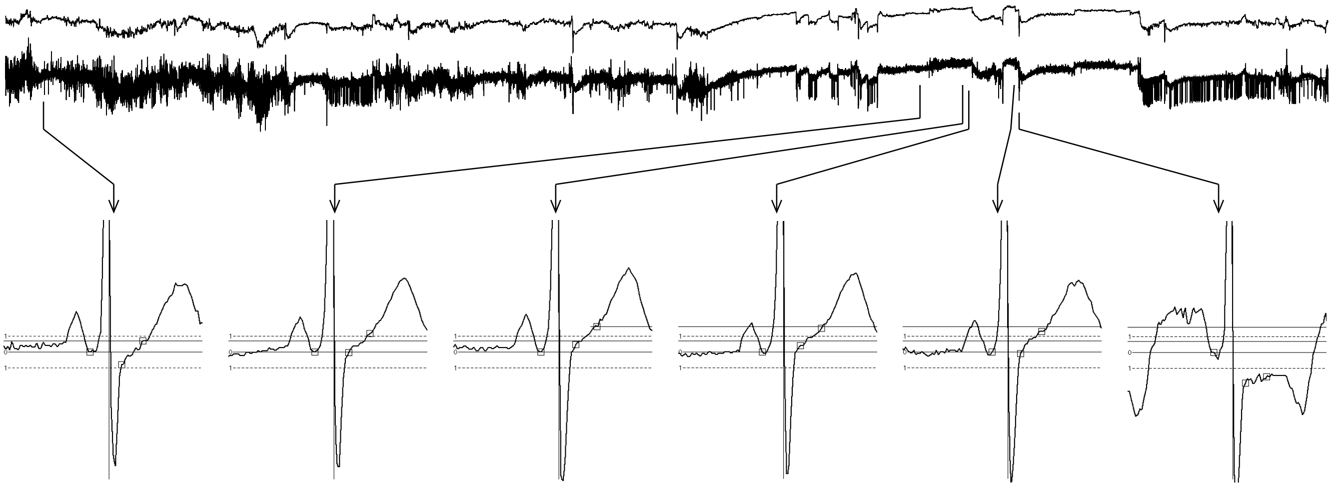

Figure 2. Example of a mixed episode (record 20612). The center and top traces show ST deviation over 24 hours, before and after smoothing. The sequence of beats below, shown as displayed by the FCIS annotation editing software, illustrates the evolution of the ST segment. From left to right: the reference beat, the beginning of the drift episode, the local reference beat, the extremum of the drift episode, the beginning of the ischemic episode, and the extremum of the ischemic episode. Dotted lines mark $\pm 100\mu$V.

the ESC DB. The publication of the ESC DB has given researchers a tantalizing view of temporal patterns in ST change that are as yet poorly understood but are likely to be of future clinical interest. At the same time, the ESC DB shows us examples of ST changes that confound most automated analysis techniques. These events appear to have the principal characteristics of ischemic ST episodes, yet on closer examination are clearly non-ischemic in nature.

We are developing a new long-term ST database as a complement to the ESC DB. It is important to observe that the LTST DB is not intended as a replacement for the ESC DB; its goals are different, and (because of its far greater size) it is not practical to annotate the LTST DB beat-by-beat as was done with the ESC DB. What we hope to accomplish is to better represent the wide variety of "real-world" data, including many more examples of mixed and non-ischemic episodes, and to permit researchers to study lengthy examples of quasi-periodic and other temporal patterns in ST change [3]. The LTST DB is intended to support the development of improved algorithms to differentiate ischemic from non-ischemic ST events, and (by its size) to permit more reliable prediction of clinical performance from first-order performance statistics.

Acknowledgment

This publication is based on work sponsored by the U.S.-Slovenian Science and Technology Joint Fund in cooperation with the Department of Health and Human Services, U.S.A. and the Ministry for Science and Technology of the Republic of Slovenia under Project number 95-158. Beside, it was sponsored by the Dean's Funds of the University of Ljubljana. We are also grateful for the support of the BIH Arrhythmia and Bioengineering Laboratories, and we especially wish to thank Diane Perry for her help in obtaining recordings.

References

[1] Taddei A, Benassi A, Biagini A, Bongiorni MG, Contini C, Distante G, Landucci L, Mazzei MG, Pisani P, Roggero N, Varanini M, Marchesi C. ST-T change analysis in ECG ambulatory monitoring: a European standard for performance evaluation. In Computers in Cardiology 1987. Washington: IEEE Computer Society Press, 1988; 63–68.

[2] 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; 13:1164–1172.

[3] Jager F, Moody GB, Mark RG. Characterization of transient ischemic and non-ischemic ST segment changes. In Computers in Cardiology 1995. Piscataway: IEEE Society Press, 1995; 721–724.

[4] Moody GB, Mark RG. The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it. In Computers in Cardiology 1990. Los Alamitos: IEEE Computer Society Press, 1990; 185–188.

[5] Moody GB, Mark RG. Development and evaluation of a 2-lead ECG analysis program. In Computers in Cardiology 1982. Los Angeles: IEEE Computer Society Press, 1992; 39–44.

Address for correspondence:

Franc Jager
Faculty of Computer and Information Science
Tržaška 25 / 1001 Ljubljana / Slovenia

Tel./Fax: +386-61-1768-362/1264-647
franc@manca.fri.uni-lj.si