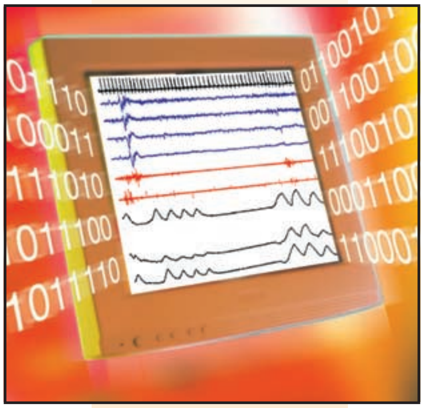

PhysioNet: A Web-Based Resource for the Study of Physiologic Signals

Free Access to a Signals Archive and a Signal Processing/ Analysis Software Library Fosters Online Collaboration

George B. Moody, Roger G. Mark, Ary L. Goldberger
Harvard-MIT Division of Health Sciences and Technology
and
Massachusetts Beth Israel Deaconess Medical Center

On August 1, 1999, researchers at Boston's Beth Israel Deaconess Medical Center, Boston University, McGill University, and MIT inaugurated a new resource for the biomedical research community, under the auspices of the National Center for Research Resources of the US National Institutes of Health. This resource, intended to stimulate current research and new investigations in the study of complex biomedical signals, has the following three closely interdependent components.

- PhysioNet is an online forum for dissemination and exchange of recorded biomedical signals and open-source software for analyzing them, by providing facilities for cooperative analysis of data and evaluation of proposed new algorithms. In addition to providing free electronic access to PhysioBank data and PhysioToolkit software via the World Wide Web (http://www.physionet.org/), PhysioNet offers service and training via on-line tutorials to assist users at entry and more advanced levels. PhysioBank is a large and growing archive of well-characterized digital recordings of physiologic signals and related data for use by the biomedical research community.
- PhysioBank currently includes databases of multiparameter cardiopulmonary, neural, and other biomedical signals from healthy subjects and patients with a variety of conditions with major public health implications, including sudden cardiac death, congestive heart failure, epilepsy, gait disorders, sleep apnea, and aging. These databases will grow in size and scope and will eventually include signals from selected in vitro and in vivo experiments, as developed and contributed by members of the research community.
- PhysioToolkit is a growing library of software for physiologic signal processing and analysis, detection of physiologically significant events using both classical techniques and novel methods based on statistical physics and nonlinear dynamics, interactive display and characterization of signals, creation of new databases, simulation of physiologic and other signals, quantitative evaluation and comparison of analysis methods, and analysis of nonequilibrium and nonstationary processes. A unifying theme of the research projects that contribute software to PhysioToolkit is the extraction of "hidden" information from biomedical signals, information that may have diagnostic or prognostic value in medicine, or explanatory or predictive power in basic research. All PhysioToolkit software is available in source form under the GNU General Public License (http://www.fsf.org/copyleft/gpl.html).

This article aims to introduce PhysioNet as a resource to the biomedical research community. After a capsule summary of its history and goals, we will discuss what PhysioNet offers to researchers, describe some of the technology needed to support these functions, and conclude with observations gleaned from PhysioNet's first year of service.

Origins and Objectives of PhysioNet

As noted elsewhere in this issue, the authors have shared long-term interests in understanding the dynamics of human

0739-5175/01/$10.00©2001IEEE May/June 2001

physiology and the implications of dynamical change in diagnosis and treatment of pathophysiology. Our research has often required physiologic signal databases, and we have collected many such databases, beginning in the 1970s. Recognizing the importance and value of these databases to others working in our field, we distributed copies of them, beginning with the MIT-BIH Arrhythmia Database in 1980 [1-3].

During the past 25 years, as the technology supporting storage and transfer of large data collections has evolved, we shared our data using 8-inch diskettes (1977-1981), quarter-inch cartridge tapes (1978-1983), half-inch 9-track magnetic tapes (1980-1989), CD-ROMs (1989 to date), and CD-R (1994 to date). Our data collections rapidly outgrew the capacities of all of the physical media. As high-speed access to the Internet became more widely available, many researchers who wished to use our data were able to obtain samples of it, first using our anonymous FTP server (1991-1999), and then via the World Wide Web (1995 to date). In 1997, the capacities of affordable disk drives finally began once more to approach the size of our data collections, and we began to contemplate the possibility of creating a massive online resource of physiologic signals and software for studying them.

The National Institutes of Health's National Center for Research Resources (NIH NCRR) agreed to fund the creation and maintenance of PhysioNet as the major public outreach component of the Research Resource for Complex Physiologic Signals [4], a multicenter collaboration led by H. Eugene Stanley at Boston University, Leon Glass at McGill University, and the authors. The creation of much of the current contents of PhysioBank and PhysioNet has been stimulated by research within the Resource, which includes:

- assessment of signal quality in multiparameter data (identification of reliable and unreliable signals based on their interrelationships);
- multivariate trend analysis and forecasting, with applications in intensive care;
- quantification of transient and local properties of nonstationary physiologic signals;
- quantification of cross interactions among multiparameter signals;
- detection of changes that may precede the onset of catastrophic physiologic events, including epileptic seizures and sudden cardiac death;
- quantification of the dynamics of physiologic control;
- mathematical and physiological modeling of physiologic control mechanisms;
- identification of new measures related to nonlinear dynamics and fractal scaling that have diagnostic or prognostic use in life-threatening cardiopulmonary pathologies.

Rigorous review of the data and software available from PhysioNet is essential so that researchers can confidently make use of these materials. For this reason, we attempt to make explicit the extent to which all data and software have been reviewed, by assigning each item to one of the following three classes.

- Class 1 databases and software are fully supported. Class 1 databases have been carefully scrutinized and are accompanied by comprehensive sets of annotations that have been derived by at least two experts working independently, with all discrepancies resolved either by consensus of the original annotators, or by independent review of a third expert. Class 1 software has been extensively and rigorously tested. We will correct and document any remaining errors and encourage users to bring these to our attention.
- Class 2 databases and software are archival copies of materials that support published research, contributed by authors or journals. We will maintain copies of the original data and software together with corrections submitted by the authors. We encourage users to report errors directly to the authors.
- Class 3 databases include collections of data that may have been less thoroughly studied than those in class 1 but that may be of interest to the research community. These databases include works in progress, to which users are invited to contribute. In some cases, these databases may be archived on their creators' websites. Class 3 software includes code that may need further testing or development; again, users are invited to dig in and help their creators transform these works in progress into robust and useful tools for research.

PhysioBank

The PhysioBank archives currently contain approximately 40 GB of digitized physiologic signals, most with annotations, organized into 21 major collections (databases):

Multiparameter Databases

These databases include a variety of digitized physiologic signals in each recording.

[Class 3] MIT-BIH Polysomnographic Database [6]. This is a collection of recordings of multiple physiologic signals during sleep. Subjects were monitored in the Sleep Laboratory at Boston's Beth Israel Hospital (BIH; now the Beth Israel Deaconess Medical Center) for evaluation of chronic obstructive sleep apnea syndrome, and to test the effects of constant positive airway pressure (CPAP), a standard therapeutic intervention that usually prevents or substantially reduces airway obstruction in these subjects. The database contains over 80 hours' worth of four-, six-, and seven-channel polysomnographic recordings, each with an ECG signal annotated beat-by-beat, and EEG and respiration signals annotated with respect to sleep stages and apnea. Included is the [Class 2] Santa Fe Time Series Competition [5] Data Set B (data extracted from the MIT-BIH Polysomnographic Database).

[Class 3] MIMIC Database [7]. This database contains continuous recordings of multiple physiologic signals gathered from intensive care unit (ICU) monitors. Currently, the PhysioBank archives contain 72 complete records from this database. Each contains one to three ECG signals sampled at 500 Hz, together with other monitored signals typically including one or more blood pressure signals (e.g., arterial blood pressure, pulmonary arterial pressure, and central venous pressure), and respiration sampled at 125 Hz. The records also include measurements, computed by the ICU monitors at intervals of 1.024 s, of parameters such as heart rate, systolic and diastolic blood pressure, and respiration rate; these measurements are now available for all 121 records of the database, including multiple recordings of some of the 90 subjects. The lengths of these records vary but average nearly 40 hours each; the total length is approximately 4658 hours. The signals from the remaining 49 records will be added to the archives during 2001.

[Class 3] Apnea-ECG Database. This database has been assembled for Com-

PhysioNet is the major public outreach component of the multicenter Research Resource for Complex Physiologic Signals.

puters in Cardiology Challenge 2000 (see below). It consists of 70 ECG recordings, each typically eight hours long, with accompanying sleep apnea annotations obtained from study of simultaneously recorded respiration signals, which are included for eight of the recordings.

ECG Databases

[Class 1] MIT-BIH Arrhythmia Database [1-3]. This database consists of 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. Twenty-three recordings were chosen at random from a set of 4000 24-hour ambulatory ECG recordings collected from a mixed population of inpatients (about 60%) and outpatients (about 40%); the remaining 25 recordings were selected from the same set to include less common but clinically significant arrhythmias that would not be well represented in a small random sample. We have made ten of the first group of records and 15 of the second group (records 100-107, 118-119, and 201-219) available in their entirety. The complete reference annotation files and the first 10 min of the signal files for each of the remaining records are also available.

[Class 1] MIT-BIH Noise Stress Test Database [8]. This consists of 12 half-hour records, consisting of two records from the MIT-BIH Arrhythmia Database to which calibrated amounts of noise have been added, for the purpose of assessing the noise tolerance of automated analyzers.

[Class 1] European ST-T Database [9-10]. This standard database for evaluation of algorithms for detection of ST and T-wave changes consists of 90 two-hour recordings, with beat, rhythm, signal quality, and ST and T-wave change annotations. Forty-eight of the 90 recordings, and the complete set of 90 reference annotation files, have been contributed to PhysioBank by the European Society of Cardiology and the creators of the Database.

[Class 2] QT Database [11]. This database has over 100 15-minute two-lead ECG recordings (many excerpted from other databases), with onset, peak, and end markers for P, QRS, T, and (where present) U waves of 30-50 selected beats in each recording.

[Class 2] Post-Ictal Heart Rate Oscillations in Partial Epilepsy [12]. This consists of seven annotated single-lead ECG recordings, with times of seizures indicated.

[Class 2] Exaggerated Heart Rate Oscillations During Two Meditation Techniques [13], with additional data from spontaneously and metronomically breathing controls, and from highly trained athletes. These data include times of beat occurrence only (the original ECGs are not currently available).

[Class 2] BIDMC Congestive Heart Failure Database [14]. This contains 15 recordings, each about 20 hours long, from subjects with severe congestive heart failure (New York Heart Association class 3-4).

[Class 3] Fantasia Database [15]. This consists of heart-beat time series (no ECG signals) from five young and five elderly subjects, all healthy, in sinus rhythm during a resting state (two hours each).

[Class 3] MIT-BIH Malignant Ventricular Arrhythmia Database [16-17]. This consists of 22 excerpts of two-channel ECG recordings, each 35 min in length, containing 89 episodes of ventricular tachycardia, 60 episodes of ventricular flutter, and 42 episodes of ventricular fibrillation. This database contains rhythm and signal quality annotations only (no beat annotations).

[Class 3] Creighton University Ventricular Tachyarrhythmia Database. This database, collected by Floyd Nolle at the Creighton University Cardiac Center, contains 35 records, each of which shows the onset of ventricular fibrillation. Most of these records are original digital recordings from patient monitors; each is 8.5-min long and contains one ECG signal.

[Class 3] MIT-BIH Supraventricular Arrhythmia Database [18]. This contains 78 excerpts of two-lead ECG recordings, each 30-min long, selected to supplement the examples of supraventricular arrhythmias in the MIT-BIH Arrhythmia Database.

[Class 3] MIT-BIH Atrial Fibrillation Database [19]. This database consists of 25 records, each ten-hours long, containing about 300 episodes of atrial fibrillation and 40 episodes of atrial flutter. The ECG signals are currently available for only one of the records; the others are represented by annotation files only, from which RR-interval time series can be derived.

[Class 3] MIT-BIH ST Change Database [20]. This contains 28 one- and two-lead ECG recordings, mostly from exercise stress tests.

[Class 3] MIT-BIH Normal Sinus Rhythm Database. This database is 18 two-lead ECG recordings, each between 20 and 24 hours in length, from subjects without diagnosed cardiac abnormalities, with beat annotations. Also available are two recordings that were excluded from this database because of the presence of occasional ectopic beats.

[Class 3] MIT-BIH ECG Compression Test Database [21]. This database contains a set of 168 short (20-s) records, exhibiting a wide variety of arrhythmias, conduction defects, and noise. Many records were chosen because characteristics of the signal or noise may be expected to pose a problem for an ECG compression method that is not exactly invertible.

[Class 3] MIT-BIH Long-Term Database. This database contains seven complete long-term two- and three-lead ECG recordings, ranging in length from 14 to 24 hours.

Gait Databases

These databases contain walking stride interval (gait cycle duration) time series.

[Class 2] Gait Maturation Database [22]. This contains stride interval time series from 50 healthy children, ages 3-14, walking at a self-determined normal pace for 8 min around a 400-m running track.

[Class 2] Gait in Aging and Disease Database [23-25]. This includes stride interval time series from 15 subjects: five healthy young adults (23-29 years old), five healthy old adults (71 -77 years old), and five older adults (60-77 years old) with Parkinson's disease. Data were collected from the healthy subjects as they walked in a roughly circular path for 15 min and from the subjects with Parkinson's disease as they walked for 6 min up and down a long hallway.

PhysioToolkit

The core of PhysioToolkit is the WFDB (WaveForm DataBase) software package, an open-source successor to earlier software written by the first author for use with the MIT-BIH Arrhythmia Database. This package includes command-line tools for a wide variety of classic signal-processing functions (e.g., filtering, sampling frequency conversion, signal averaging, QRS detection, power spectral density estimation), all written in C and usable on Linux or MS-Windows PCs and Unix workstations, among others. Reference implementations of novel algorithms for physiologic signal processing, simulation and modeling, and time-series analysis, such as applications for deriving a respiration signal from one or more ECG signals [26-27], and for detrended fluctuation analysis [28-29], are another major component of PhysioToolkit.

WAVE is an extensible interactive graphical environment for manipulating sets of digitized signals with optional annotations, usable under Linux, Solaris, and other environments supporting X11 and the open-source XView toolkit. WAVE supports graphical annotation editing (we and others have used it extensively for development of most of the currently available annotated databases of physiologic signals), high-quality printing, flexible control of external signal-processing analysis programs (any such application can be made into a plugin), and two-way communication with web browsers (annotations can be hyperlinks, and web browsers can use WAVE as a viewer for annotated signal files). A subset of the features of WAVE is available in WVIEW, an interactive viewer for MS-Windows.

All of this software uses a common library of functions (the WFDB library) for reading and writing annotation and signal files in a wide variety of supported formats. Among the supported signal file formats are 8-, 10-, 12-, and 16-bit signed and unsigned binary formats (including big- and little-endian formats and several bit-packed formats); signals in separate files, multiplexed files, or combinations of both; and signals sampled at the same or different sampling frequencies. Annotations are typically stored in the compact and extensible bit-packed format first defined for the MIT-BIH Arrhythmia Database and later used by most other physiologic signal databases, including the European ST-T Database and the Massachusetts General Hospital/Marquette Foundation Waveform Database, as well as in the format defined for the AHA Database for Evaluation of Ventricular Arrhythmia Detectors. (Translators are provided for converting binary signal and annotation files to text and vice versa, and for converting between binary formats.) It is not necessary for an application to be aware of the format of its input files, since all supported formats are transparently converted into a uniform representation by the library; and applications that create signal or annotation files can similarly generate files in any supported format simply by specifying the format at the time the files are created. User-written applications in C, C++, or Fortran can be linked to the WFDB library.

One of the most powerful features of the WFDB library is that it permits input not only from local files but also from remote web servers. The library includes HTTP and FTP client code, and when communicating with web servers such as Apache that support HTTP range requests, it is possible for a WFDB application such as WAVE to retrieve any desired portion of a given file without having to read all of it. The library finds its input files by searching a user-defined path, which may include any combination of local directories and web or FTP server directories; thus, it is easy to keep a set of working files locally and still have access to all of PhysioBank, simply by including "http://www.physionet.org/physiobank/database" in the search path following the local directories. The low-level HTTP and FTP client support is provided by the World Wide Web Consortium's free libwww library (http://www.w3.org/Library/), which is linked to the WFDB library if available (this feature can be disabled easily if desired).

The PhysioBank archives currently contain approximately 40 GB of digitized physiologic signals organized into 21 major collections.

Some PhysioToolkit software can be run by users on the PhysioNet web servers to provide analyses of their own locally stored data. These services permit use of the software without the need to install it first. For example, a user may submit an ECG recording to a QRS detector running on PhysioNet to obtain an RR-interval time series, a power spectrum of heart rate variability, a detrended fluctuation analysis, or a derived respiration signal.

PhysioNet on the World Wide Web

PhysioNet serves the research community from a master server located at MIT in Cambridge, Massachusetts, and a network of mirror servers in Italy, Israel, Spain, Slovenia, Taiwan, China, Brazil, and at two other US locations (see http://www.physionet.org/mirrors/ for a complete list). The mirrors provide improved access, especially where connections to the United States may be slow. Additional mirror sites are welcome.

PhysioNet servers are typically 200 MHz or faster x86-compatible PCs, running the Apache web server under Linux.

PhysioNet is fully indexed and searchable, using a slightly customized version of ht://Dig (http://www.htdig.org/). Mirroring is supported using rsync (http://rsync.samba.org/). Application services are provided using PhysioToolkit software and additional custom CGI scripts. All of the software used is freely available and open-source.

During PhysioNet's first year of service, the PhysioNet master server provided over 150 GB of data to visitors from over 15,000 locations in 90 countries.

The Future of PhysioNet

All data included in PhysioBank, and all software included in PhysioToolkit, are carefully reviewed. The classification of each item reflects the extent to which this review can be considered complete. Generally, class 3 materials have been reviewed only by their creators and require peer review, class 2 items may have received peer review but may not yet have stood the test of time, and class 1 data and software have been studied by many researchers over a period of years and may be expected to have few if any remaining flaws. As class 2 and 3 data and software receive review and corrections, they may be promoted to class 1 status following a public comment period. We invite the research community to participate in the ongoing review process. By sharing common data sets, and software in source form, the community benefits from access to materials that have been rigorously scrutinized by many investigators. We further invite researchers to contribute data and software for peer review and possible inclusion in PhysioBank and PhysioToolkit.

PhysioBank is beginning to function as a repository for selected physiologic signals and time-series data from published studies in peer-reviewed journals. As this repository grows, PhysioBank will increasingly serve as a "dynamical appendix" for published articles, giving readers access to the original data on which studies are based. A precedent for publication of such primary datasets has already been established by *Nature* and *Science* with respect to high-resolution bio-molecular structural data, which are now released at or before the time of publication of the articles describing these data.

Equally important is the development of novel tools for data analysis, including techniques for signal processing, and algorithms for analyzing typically nonstationary and nonlinear data sets for clinical or research purposes. An archive of research software that is universally available, immediately accessible, and continuously refined through rigorous peer review is an ideal complement to the existing print media in which such algorithms are published. Consider, for example, the problem of errors in published algorithms for nonlinear analysis. Without a forum such as PhysioNet to alert users to such problems and correct them when they are discovered, progress in this field has been hampered by the difficulty of resolving differences in results that may result from errors in interpretation of algorithm descriptions, errors in algorithm implementation, or fundamental errors in algorithm design. Misleading and erroneous analyses follow. The open-source software development model in general, and PhysioNet in particular, offer an effective remedy for such problems and will also allow the developers and users of new or established software to let the community know promptly about updates or refinements.

Between February and September 2000, PhysioNet hosted the Computers in Cardiology Challenge 2000, an open competition focusing on the problem of detecting and quantifying sleep apnea using the ECG. We selected this topic for the first in a planned series of challenges because chronic sleep apnea is a common problem with major health implications, because it is typically diagnosed using techniques that are typically more cumbersome and expensive than ECG monitoring, and because studies during the past 15 years [26-27] have hinted at the possibility of detecting sleep apnea using features of the ECG.

The immediate goal of the challenge was to demonstrate that apnea can be detected reliably from the ECG in a large, representative, and well-characterized set of reference data. The use of a common data set (the Apnea-ECG Database, see above), consisting of ECG recordings that have been extracted from manually annotated polysomnograms and divided into equal training and test sets, permits reproducible evaluation and fair comparison of different analysis methods. Apnea annotations for the training set were made available for study; those for the test set were not released until the conclusion of the competition.

Entrants competed in two events (apnea screening, in which the object was to classify the test set recordings into apnea and nonapnea categories; and quantification, which required scoring each minute of each recording for the presence or absence of apnea). A majority of the entrants in the screening event correctly classified from 93% to 100% of cases; scores in the quantification event ranged from 75% to nearly 93%, a level comparable to inter-observer variability in apnea scoring based on the full set of signals normally available in polysomnograms, including nasal airflow, respiratory effort, and oxygen saturation signals. Details of the methods used were presented by 12 entrants at Computers in Cardiology 2000 and appear in the proceedings of the conference [30].

The more profound goal of this challenge is to leverage the opportunities presented by the PhysioNet resource to foster rapid progress by a widely distributed group of researchers on an important clinical problem. A major benefit of a research resource such as PhysioNet is to create a meeting place for new data and novel analytic methods. It is especially significant that several of the most successful entrants would not otherwise have had access to the data necessary for studying this topic. The spectacular success of this challenge is among the clearest demonstrations yet of how PhysioNet offers a new paradigm for catalyzing advances in subjects where access to data has traditionally been a barrier to entry.

Acknowledgment

This work was supported by a grant from the National Center for Research Resources of the National Institutes of Health (P41 RR13622).

George B. Moody is a research staff scientist in the Harvard-MIT Division of Health Science and Technology, and the designer and webmaster of PhysioNet. His research interests include automated analysis of physiologic signals, statistical pattern recognition, heart-rate variability, multivariate trend analysis and forecasting, and applications of artificial intelligence and advanced digital signal processing techniques to multiparameter physiologic monitoring.

Roger G. Mark, M.D., Ph.D., is distinguished professor of Health Sciences & Technology and professor of electrical engineering at MIT. He is a Co-PI of the Research Resource for Complex Physiologic Signals. Dr. Mark served as the co-director of the Harvard-MIT Division of Health Sciences & Technology from 1985 to 1996. Dr. Mark's research activities include physiological signal processing and database development, cardiovascular modeling, and intelligent patient monitoring. He led the group that developed the MIT-BIH Arrhythmia Database.

Ary L. Goldberger, M.D., is the director of the Margret and H.A. Rey Laboratory for Nonlinear Dynamics in Medicine, associate professor of Medicine at the Harvard Medical School, and the program director of the Research Resource for Complex Physiologic Signals. Dr. Goldberger has a long-standing interest in electrocardiography and noninvasive electrophysiology and is the author of two standard textbooks on ECG analysis. His research interests include fractal and nonlinear dynamical mechanisms in physiology and medicine.

Address for Correspondence: George B. Moody, MIT, Room E25-505A, Cambridge, MA 02139 USA. E-mail: george@mit.edu.

References

[1] R.G. Mark, P.S. Schluter, G.B. Moody et al., "An annotated ECG database for evaluating arrhythmia detectors" in "Frontiers of Engineering in Health Care–1982," in *Proc. 4th Annu. Conf. IEEE EMBS*, Long Beach, CA: IEEE Comput. Soc. Press, 1982, pp. 205-210.

[2] G.B. Moody and R.G. Mark, "The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it," in *Computers in Cardiology 1990*. Los Alamitos, CA: IEEE Comput. Soc. Press, 1991, pp. 185-188.

[3] G.B. Moody, R.G. Mark, "The Impact of the MIT-BIH Arrhythmia Database," *IEEE Eng. Med. Biol. Mag*., vol 20, no. 3, pp. 45-50, 2001.

[4] A.L. Goldberger, L.A.N. Amaral, L. Glass et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, p. e215, 2000. Available: http://circ.ahajournals.org/cgi/content/full/101/23/e215

[5] A.S. Weigend and N.A. Gershenfeld, Eds., *Time Series Prediction: Forecasting the Future and Understanding the Past.* Reading, MA: Addison-Wesley, 1994. Available: http://www.stern.nyu.edu/~aweigend/Time-Series/SantaFe.html

[6] Y. Ichimaru, G.B. Moody, "Development of the polysomnographic database on CD-ROM," *Psychiatry and Clin. Neurosci.*, vol. 53, no. 2, pp. 175-177, Apr. 1999.

[7] G.B. Moody and R.G. Mark, "A database to support development and evaluation of intelligent intensive care monitoring," in *Computers in Cardiology 1996*. Piscataway, NJ: IEEE Press, 1996, pp. 657-660. Available: http://www.physionet.org/physiobank/database/mimicdb/mimic-cic96/

[8] G.B. Moody, W.K. Muldrow, and R.G. Mark "A noise stress test for arrhythmia detectors," in *Computers in Cardiology 1984*. Long Beach, CA: IEEE Comput. Soc. Press, 1985, pp. 381-384.

[9] A. Taddei, A. Benassi, A. Biagini et al., "ST-T change analysis in ECG ambulatory monitoring: a European standard for performance evaluation," in *Computers in Cardiology 1987*. Long Beach, CA: IEEE Comput. Soc. Press, 1988, pp. 63-68.

[10] A. Taddei, G. Distante, M. Emdin et al., "The European ST-T Database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography," *Eur. Heart J.*, vol. 13, pp. 1164-1172, 1992.

[11] P. Laguna, R.G. Mark, A.L. Goldberger, and G.B. Moody, "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG," in *Proc. Computers in Cardiology 1997*. Piscataway, NJ: IEEE Press, 1997, pp. 673-676. Available: http://www.physionet.org/physiobank/database/qtdb/doc/.

[12] I.C. Al-Aweel, K.B. Krishnamurthy, J.M. Hausdorff et al., "Post-ictal heart rate oscillations in partial epilepsy," *Neurology*, vol. 53, no. 7, pp. 1590-1592, Oct. 1999.

[13] C.K. Peng, J.E. Mietus, Y. Liu et al., "Exaggerated heart rate oscillations during two meditation techniques," *Int. J. Cardiology*, vol. 70, pp. 101-107, 1999.

[14] D.S. Baim, W.S. Colucci, E.S. Monrad et al., "Survival of patients with severe congestive heart failure treated with oral milrinone," *J. Amer. College of Cardiology*, vol. 7, no. 3, pp. 661-670, Mar. 1986.

[15] N. Iyengar, C-K. Peng, R. Morin et al., "Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics," *Amer. J. Physiology*, vol. 271, pp. 1078-1084, 1996.

[16] S.D. Greenwald, P. Albrecht, G.B. Moody, and R.G. Mark, "Estimating confidence limits for arrhythmia detector performance," in *Computers in Cardiology 1985*. Washington, DC: IEEE Comput. Soc. Press, 1986, pp. 383-386.

[17] S.D. Greenwald, "Development and analysis of a ventricular fibrillation detector," M.S. thesis, MIT Dept. of Electrical Eng. and Comput. Sci., 1986.

[18] S.D. Greenwald, "Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information," Ph.D. thesis, Harvard-MIT Div. of Health Sciences and Technol., 1990.

[19] G. B. Moody and R.G. Mark, "A new method for detecting atrial fibrillation using RR intervals," in *Computers in Cardiology 1983*. Long Beach, CA: IEEE Comput. Soc. Press, 1984, pp. 227-230.

[20] P. Albrecht, "S-T segment characterization for long-term automated ECG analysis," M.S. thesis, MIT Dept. of Electrical Eng. and Comput. Sci., 1983.

[21] G.B. Moody, R.G. Mark, and A.L. Goldberger, "Evaluation of the 'TRIM' ECG data compressor," in *Computers in Cardiology 1988*. Los Alamitos, CA: IEEE Comput. Soc. Press, 1989, pp. 167-170.

[22] J.M. Hausdorff, L. Zemany, C.K. Peng, and A.L. Goldberger, "Maturation of gait dynamics: stride-to-stride variability and its temporal organization in children," *J. Appl. Physiology*, vol. 86, pp. 1040-1047, 1999. Available: http://www.physionet. org/physiobank/database/gait-maturation-db/ jap/

[23] J.M. Hausdorff, P.L. Purdon, C.K. Peng et al. "Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations," *J. Appl. Physiology*, vol. 80, pp. 1448-1457, 1996.

[24] J.M. Hausdorff, S.L. Mitchell, R. Firtion et al., "Altered fractal dynamics of gait: reduced stride interval correlations with aging and Huntington's disease," *J. Appl. Physiology*, vol. 82, pp. 262-269, 1997.

[25] J.M. Hausdorff, M.E. Cudkowicz, R. Firtion et al., "Gait variability and basal ganglia disorders: stride-to-stride variations of gait cycle timing in Parkinson's and Huntington's disease," *Movement Disorders*, vol. 13, pp. 428-437, 1998.

[26] G.B. Moody, R.G. Mark, A. Zoccola, and S. Mantero, "Derivation of respiratory signals from multi-lead ECGs," in *Computers in Cardiology 1985*. Washington, DC: IEEE Comput. Soc. Press, 1986, pp. 113-116. Available: http://www.physionet.org/physiotools/edr/cic85/

[27] G.B. Moody, R.G. Mark, M.A. Bump et al., "Clinical validation of the ECG-derived respiration (EDR) technique," in *Computers in Cardiology 1986*. Washington, DC: IEEE Comput. Soc. Press, 1987, pp. 507-510. Available: http://www.physionet.org/physiotools/edr/cic86/

[28] C-K Peng, S.V. Buldyrev, S. Havlin et al., "On the mosaic organization of DNA sequences," *Phys. Rev. E*, vol. 49, pp.1685-1689, 1994.

[29] C-K. Peng, S. Havlin, H.E. Stanley, and A.L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series," *Chaos*, vol. 5, pp. 82-87, 1995.

[30] A. Murry, Ed., *Computers in Cardiology 2000*. Piscataway, NJ: IEEE Press, 2000.