

Sharing Data and Software on PhysioNet

George B. Moody <george@mit.edu>, Roger G. Mark <rgmark@mit.edu>
Massachusetts Institute of Technology, Cambridge, MA

Since its establishment in 1999, **PhysioNet** (<http://physionet.org/>) has provided unrestricted web access to a large variety of databases of time-varying physiologic signals and a wide range of relevant, rigorously tested, well-documented, open-source software that can run on any platform. PhysioNet's worldwide community of over 40,000 researchers, clinicians, educators, students, and medical instrument and software developers retrieve over 250 gigabytes of data per day. By providing free access to its unique and wide-ranging data and software collections, PhysioNet enables studies that lead to an average of over 100 scholarly publications per month (well over 7000 studies since 1999) by academic, clinical, and industry-affiliated researchers worldwide.

In early 2011, as the number of curated databases neared 50, we inaugurated **PhysioNetWorks**, a self-service virtual laboratory for development of data and software resources that will eventually become freely available via PhysioNet. It freely provides large, secure, and private workspaces with redundant backup to active researchers. The workspaces can easily be shared with colleagues anywhere. PhysioNetWorks encourages investigators to create well-organized and -documented, ready-to-share data and software repositories during the conduct of their research. When the research is complete and the major results have been published, or at any time the researcher wishes, the repository can be shared with the research community at large via PhysioNet.

In its first 18 months, PhysioNetWorks has attracted about 1400 members and has hosted 41 shared projects, with about 30 more in preparation. Data from six of the shared projects have been made freely available via PhysioNet; two of these six are continuing projects, and four are completed. The 37 currently active projects include data archives of multi-center studies as well as those of individual researchers; some are collecting data, others are engaged in annotating existing collections, and still others focus on development and evaluation of software for analysis and modeling of physiologic signals and time series.

Among the PhysioNetWorks projects most directly relevant to the interests of the MSM community are:

- *Physiological Origins of Baroreflex Dysfunction*: Data collected in a recent study that developed a computational model of baroreflex dysfunction [Bugenhagen SM, Cowley AW Jr, Beard DA. Identifying physiological origins of baroreflex dysfunction in salt-sensitive hypertension in the Dahl SS rat. *Physiological Genomics* 2010;42:23-41.] The data include non-invasive blood pressure recordings for 9 Dahl SS rats and 6 Dahl SS.13BN rats, under high and low salt conditions.
- *Abnormal ECG model and denoising tools*: This project aims to make a realistic model of ischemia and the MHD effect during MRI analysis. It is also developing tools to remove the artifact.
- *Continuous LVEF Monitoring by Aortic Pressure Waveform Analysis*: Data that were analyzed in: Swamy G, Kuiper J, Gudur MSR, Olivier NB, and R Mukkamala. Continuous Left Ventricular Ejection Fraction Monitoring by Aortic Pressure Waveform Analysis. *Annals of Biomedical Engineering* 37(6):1055-1068, June 2009.
- *Freezing of Gait in Parkinson's Disease*: Data showing episodes of freezing of gait (FOG), collected from patients with Parkinson's Disease (PD) using 3 accelerometers and 3 gyroscopes on the lower back. The purpose of this project is to test algorithms for detecting FOGs.
- *MIMIC II Clinical Database*: Comprehensive clinical data from tens of thousands of Intensive Care Unit (ICU) patients, including demographics, surrogate dates of hospital admission, discharge, and death (in or out of hospital); ICD-9, CPT, DRG, LOINC codes, hourly vital signs, fluid I/O, SAPS, SOFA, ventilator settings, IV meds; chemistry, hematology, ABGs, microbiology, cardiac catheterization, ECG, radiology, and echo reports; nursing progress notes and discharge summaries. This project has attracted a community of over 400 researchers engaged in a wide variety of studies.
- *Model of Heart Injury Physiology*: Models, experimental data, simulations, and publications related to the physiology of heart injury.
- *Noninvasive Intracranial Pressure Monitoring*: Data from patients in neuro-critical care to support development and validation of novel methods for noninvasive estimation of intracranial pressure.
- *Washout of Diffusible Indicators*: Data from Bassingthwaite JB, Strondell T, Donald DE. Estimation of Coronary Blood Flow by Washout of Diffusible Indicators. *Circulation Research*. 1968;23:259-278.

PhysioNetWorks offers modelers a secure repository in which models and data can be stored together, shared with collaborators, and eventually with the research community in fulfillment of NIH data-sharing goals. We welcome and encourage contributions from the MSM community.