Technical Information Report

AAMI TIR24: 1999/(R)2025

Acquisition and use of physiologic waveform databases for testing of medical devices

AAMI
Advancing Safety in Health Technology

Acquisition and use of physiologic waveform databases for testing of medical devices

Approved 18 August 1999 and reaffirmed 19 September 2019 and 16 May 2025 by
AAMI

Abstract: This report defines the nomenclature, ingredients, and principles needed to develop, annotate, evaluate, and use physiologic waveform databases in developing and testing medical devices.

Keywords: waveform, physiologic, algorithms, arrhythmia

Published by

AAMI
901 N. Glebe Road, Suite 300
Arlington, VA 22203

© 2000 by the Association for the Advancement of Medical Instrumentation

All Rights Reserved

Publication, reproduction, photocopying, storage, or transmission, electronically or otherwise, of all or any part of this document without the prior written permission of the Association for the Advancement of Medical Instrumentation is strictly prohibited by law. It is illegal under federal law (17 U.S.C. § 101, *et seq.*) to make copies of all or any part of this document (whether internally or externally) without the prior written permission of the Association for the Advancement of Medical Instrumentation. Violators risk legal action, including civil and criminal penalties, and damages of $100,000 per offense. For permission regarding the use of all or any part of this document, contact the Copyright Clearance Center.

Printed in the United States of America

ISBN 1–57020–132–3

AAMI Technical Information Report

A technical information report (TIR) is a publication of the Association for the Advancement of Medical Instrumentation (AAMI) Standards Board that addresses a particular aspect of medical technology.

Although the material presented in a TIR may need further evaluation by experts, releasing the information is valuable because the industry and the professions have an immediate need for it.

A TIR differs markedly from a standard or recommended practice, and readers should understand the differences between these documents.

Standards and recommended practices are subject to a formal process of committee approval, public review, and resolution of all comments. This process of consensus is supervised by the AAMI Standards Board and, in the case of American National Standards, the American National Standards Institute.

A TIR is not subject to the same formal approval process as a standard. However, a technical committee and the AAMI Standards Board do approve a TIR for distribution.

Another difference is that although both standards and TIRs are periodically reviewed, a standard must be acted upon—either reaffirmed, revised, or withdrawn—and the action must be formally approved usually every 5 years but at least every 10 years. For a TIR, AAMI consults with a technical committee about every 5 years after the publication date (and periodically thereafter) for guidance on whether the document is still useful—that is, to check that the information is relevant or of historical value. If the information is not useful, the TIR is removed from circulation.

A TIR may be developed because it is more responsive to underlying safety or performance issues than a standard or recommended practice or because achieving consensus is extremely difficult or unlikely. Unlike a standard, a TIR permits the inclusion of differing viewpoints on technical issues.

CAUTION NOTICE: This AAMI technical information report may be revised or withdrawn at any time. Because it addresses a rapidly evolving field or technology, readers are cautioned to ensure that they have also considered information that may be more recent than this document.

Contents

Page

Committee representation vii

Acknowledgments vii

Foreword viii

Introduction ix

1 Scope 1

2 Normative reference 1

3 Definitions 1

4 Database requirements 2

4.1 Intended use of database 2
4.2 Clinical requirements 2
4.2.1 Population 2
4.2.2 Study design 2
4.3 Engineering requirements 2
4.4 Annotation requirements 3
4.5 Archive requirements 3

5 Waveform acquisition and synthesis 3

5.1 Overview 3
5.2 Signal issues and requirements 3
5.2.1 Distortion 3
5.2.2 Distortion classifications 4
5.2.3 Skew 6
5.2.4 Duration 6
5.2.5 Frequency translation 6
5.3 System design issues and requirements 6
5.3.1 Sampling theory considerations 7
5.3.2 Architectural issues 21
5.3.3 Channel acquisition guidelines 22
5.4 Storage 23
5.4.1 Raw signals—analog databases 24
5.4.2 Waveforms—digital databases 24
5.5 Archive 24
5.5.1 Environmental considerations 24
5.5.2 Format 25
5.5.3 Data integrity 25
5.5.4 Applicable standards 25
5.6 Database annotation 25
5.6.1 Annotation process 25
5.6.2 Annotation rules 26
5.6.3 Data presentation specifications 26
5.7 Maintenance and distribution 27

6 Application of waveform databases to testing 27

6.1 Evaluation of performance 27
6.2 Test objectives 27
6.3 Algorithm versus device testing system 28
6.4 Sufficiency and validity 28
6.4.1 Sufficiency 28
6.4.2 Validity—correct conclusions from results 29
6.4.3 Precautions, limitations, interpolations, and interpretations (limited scope of results) 29

Annexes

A CSE ECG Reference Library (Measurement Database) 31
B CSE ECG Reference Library (Diagnostic Database) 33
C MIT-BIH Arrhythmia Database 36
D Noise Stress Database 38
E European ST-T Database 39
F American Heart Association Database for Evaluation of Automated Ventricular Arrhythmia Detectors 41
G Massachusetts General Hospital/Marquette Foundation Waveform Database 42
H Creighton University Ventricular Tachyarrhythmia Database 43

Table

1 Minimum f_s/BW ratios vs. A.A. filter order 23

Figures

1 Sampled signal spectra and aliasing 8
2 Ideal vs. practical anti-alias filter characteristics 9
3 Anti-alias filtering for signal oversampling 9
4 Anti-alias filtering for sigma-delta converters 10
5 Monochromatic, 20-Hz sine wave sampled at 500 samples per second 10
6 Amplitude spectrum for figure 5 11
7 Effect of irregular sampling of a monochromatic waveform 12
8 Amplitude spectrum for figure 7 12
9 Improper sample rate reduction 14
10 Improper sample rate reduction 15
11 Improper sample rate expansion of a 2-Hz sine wave sampled at 60 s/s expanded to 300 s/s and as if it were latched 16
12 Improper sample rate expansion 16
13 Proper sampled rate reduction 18
14 Proper sampled rate reduction 19

Committee representation

Association for the Advancement of Medical Instrumentation

Waveform Testing Committee

The AAMI Waveform Testing Committee developed this technical information report (TIR). Committee approval of the TIR does not necessarily imply that all committee members voted for its approval.

At the time this document was balloted, the **AAMI Waveform Testing Committee** had the following members:

Cochairs:	Carl A. Pantiskas
	Sandy Weininger
Members:	James J. Bailey, MD, National Institutes of Health
	Don Brodnick, Marquette Medical
	Richard Diefes, ECRI-MSLB
	Robert Donehoo, Critikon/Johnson & Johnson
	Stacy Gehman, Quinton Instrument
	George Moody, Massachusetts Institute of Technology
	William J. Murray, Siemens Medical
	Carl A. Pantiskas, Spacelabs Medical
	Cadathur Rajagopalan, PhD, Datascope Corp.
	William Saltzstein, Medtronic Physio-Control
	Kok-Swang Tan, PhD, Bureau of Radiation & Medical Devices Canada
	Lee Taylor, MD, Protocol Systems
	Sandy Weininger, U.S. Food and Drug Administration
Alternates:	David Browne, Critikon/Johnson & Johnson
	David Daly, U.S. Food and Drug Administration
	David Hengl, Siemens Medical
	William Midgette, U.S. Food and Drug Administration

NOTE—Participation by federal agency representatives in the development of this standard does not constitute endorsement by the federal government or any of its agencies.

Acknowledgments

The Waveform Testing Committee gratefully acknowledges the work of the Apnea Monitoring Committee in the areas of patient selection and database annotation as well as the contributions of Michael K. Leverich and Mustafa H. Sagiroglu in the area of signal processing. The committee would also like to acknowledge the works of the former cochairs: Don Brodnick of Marquette Medical, Stephen Daleo of Nellcor, and Richard Bowser, formerly of Creighton University.

Foreword

The objective of the *Acquisition and use of physiologic waveform databases for testing of medical devices* technical information report (TIR) is to provide an integrated overview of the methodology, technology, and potential pitfalls of acquiring, annotating, reconstructing, and using databases for the testing of medical devices and algorithms. It reflects the conscientious efforts of a group of concerned health care professionals, biomedical engineers, device manufacturers, and government representatives to develop a reference document for those involved in developing medical instrumentation.

Waveform databases either can be a collection of digital files stored on disk or other media or can be continuous analog data stored directly on tape media. Digital databases are often used to directly test the performance of software algorithms. Both reconstructed digital databases and analog databases are used to test subsystems or complete systems. Both database forms are valuable in the testing of medical instrumentation and algorithms and have unique advantages and disadvantages. Care must be taken to ensure that the database that is acquired or chosen to test the instrument or algorithm is suitable for its intended purpose.

Continuous patient signals can be complex in size, shape, and timing. Designing waveform sequences that totally represent all possible continuous patient signals is not feasible. Therefore, the use of waveform databases that are collected from the clinical environment is a great advantage in verifying the accuracy and performance of medical instrumentation.

This TIR is written primarily for those who collect, annotate, distribute, use, or evaluate the use of databases for developing and testing medical devices. This report is intended as a reference document for anyone involved with waveform database testing. It must be reviewed and updated periodically to integrate progressive technological developments.

 © 2000 Association for the Advancement of Medical Instrumentation ■ AAMI TIR No. 24:1999

Introduction

Physicians and associated health care professionals need to know how diagnostic and monitoring medical devices work for each subpopulation they encounter, including both the healthy and the diseased. Therefore, similar devices from different manufacturers must perform in a similar manner. If these health care professionals understand the common characteristics of different devices, they can predict how devices will work in particular cases and can rely on the validity of the results.

Medical device standards unify such needs and codify relevant device characteristics. A standards group evaluates a device in each aspect of its overall functionality. For standards purposes, devices are classified according to their specific medical purpose (e.g., the diagnostic electrocardiograph). Generally, such standards groups devise tests for multiple, diverse, and often quite subtle characteristics. Each individual test is then designed to quantify one or more particular aspects of the characteristics of that specific class of medical device.

All standards testing is necessarily parametric because each test quantifies some aspect of a device's characteristics. Even verification of a required label is dichotomized as "acceptable" or "not acceptable" if the device is physically inspected for conformity.

Traditional testing of medical devices has focused on certain device characteristics such as input impedance, frequency response, dynamic range, and common mode rejection. Within the confines of traditional engineering, these characteristics are easy to measure. Relevant medical device standards have codified acceptable ranges or limits for such characteristics. Assessing conformance within these limits is relatively straightforward.

Once the standard engineering parameters are measured, a device can be subjected to signals and loads that simulate actual applications to real patients, and the results can be evaluated. This process is the basis of operational testing. Even though such tests are parametric, they use simulated reality to assess actual clinical utility. Common operational tests include tests for accuracy, sensitivity and specificity of event detection, noise stress testing, and defibrillation protection and recovery.

Computerized medical devices that run diagnostic and classification algorithms on sampled analog signals are becoming increasingly prevalent. Tests for such devices demand careful extensions to the already established principles of traditional test design. For example, devices that actively sense and interpret a patient's condition must be tested for their ability to recognize certain clinical patterns (e.g., a cardiac monitor must sense the onset of ventricular fibrillation (VF) and respond with its highest priority alarm). Synthesizing waveform sequences that totally represent all possible continuous patient signals is not feasible. Hence, the need for properly documented databases of representative waveforms for testing medical devices becomes evident. For example, electrocardiogram (ECG) rhythm annotations can be determined by consensus review (e.g., MIT-BIH databases), whereas left ventricular hypertrophy requires non-ECG clinical evaluation (e.g., CSE database).

Waveforms that are used for assessment of such responses must be documented clinically or validated by expert medical opinion, according to established medical practice. Specific procedures used in validating such waveforms should be based on standards of medical practice for that purpose. Thus, if consensus review of computer-interpreted patient waveforms were medically accepted, it would be inappropriate to demand independent blind validation by isolated physicians who would be deprived of the preliminary computerized scan and interpretation. The same would be true for the CSE diagnostic database that was documented by non-ECG means, namely angiography, echocardiography, or cardiac enzyme studies.

All tests have characteristic associated costs and information yields, which are determined by the nature of each device and its intended use. Good medical device standards specify tests that appropriately evaluate the characteristics under scrutiny, yield repeatable and clinically relevant results, and impose reasonable costs on those who perform the tests.

Acquisition and use of physiologic waveform databases for testing of medical devices

1 Scope

This document is intended to define the nomenclature, ingredients, and principles to develop, annotate, evaluate, and use physiologic waveform databases for developing and testing medical devices. The TIR identifies issues that should be addressed in the design and development of a physiologic database. It discusses many major pitfalls that must be avoided. Annexes are included that describe several databases in detail. The database profiles that are presented here are intended to serve as a guide in the design, development, acquisition, and documentation of future waveform databases that may be used in the development and evaluation of medical devices and algorithms.

This report considers continuous electrophysiologic signals such as the electrocardiogram and electroencephalogram, as well as nonelectrophysiologic signals such as invasive blood pressure and respiratory tachograms. Medical devices that deal with intermittent data such as thermodilution cardiac output or oscillometric noninvasive blood pressure are not covered in this initial report, but may be included in a later report.

2 Normative reference

2.1 ASSOCIATION FOR THE ADVANCEMENT OF MEDICAL INSTRUMENTATION. *Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms.* AAMI EC57. Arlington (Vir.): AAMI, 1998. American National Standard.

3 Definitions

For purposes of this AAMI Technical Information Report, the following definitions apply:

3.1 ADC: Analog-to-digital converter.

3.2 aliasing: The source of distortion that may result when a signal is sampled at less than the Nyquist rate.

NOTE—Energy from higher-frequency signal components—higher than one-half the Nyquist rate—are folded into frequency components of less than one-half the Nyquist rate.

3.3 CSE: Common standards for quantitative electrocardiology.

3.4 DNRi: See instantaneous dynamic range.

3.5 FFT: Fast Fourier transform.

3.6 impedance: A measurement (in ohms) of electrical opposition to the flow of current in a given circuit.

3.7 instantaneous dynamic range (DNRi): The ratio of the largest signal a system will see to the smallest signal, usually the noise floor.

3.8 LSB: Least significant bit.

3.9 loop area: The actual surface area of a planar loop encircled by the signal pickup leads and the signal source.

NOTE—The amount of noise electromagnetically induced is directly proportional to the area of this loop. Therefore, its minimization is critical. This phenomenon applies to magnetic field coupling only.

3.10 morphology: The appearance of a waveform when viewed in the time domain or, more simply put, the waveform's shape.

3.11 Nyquist rate: The sample rate required to preserve signal content without distortion; at least two times the highest frequency component of the signal.

© 2000 Association for the Advancement of Medical Instrumentation ■ AAMI TIR No. 24:1999