

Exercise Test Interpretation

Willi Kaiser¹, Martin Findeis¹, Rami Lehtinen^{2,3}, Terho Lehtimäki^{4,5}, Jari Viik⁶

¹GE Medical Systems, Freiburg, Germany, ²Tampere University of Applied Sciences, ³Tampere University Hospital, Finland, ⁴Department of Clinical Chemistry, Tampere University Hospital and University of Tampere, ⁵Medical School, Tampere, Finland, ⁶Department of Biomedical Engineering, Tampere University of Technology

Abstract

An exercise test delivers a large number of measurements that are valuable in risk prediction, in detecting coronary artery disease, and in describing the functional exercise response of a patient. However, it is difficult to have comprehensive knowledge of all measurements and their thresholds.

The Exercise Test Interpretation (XTI) program compares exercise measurements against established thresholds and provides statements and reasoning texts, as well as explanations of the statements when thresholds are exceeded. A concise, clear, and accurate overview of an exercise test is provided.

For validation, the FINCAVAS database (Tampere University, Finland) containing mortality, angiographic, and other clinical data has been used.

1. Introduction

Increased appreciation of the prognostic and diagnostic value of exercise testing resulted in the development of new exercise measurements, for instance, T-wave alternans (TWA) [1], ST/HR hysteresis [2], and heart rate recovery (HR recovery) [3]. Together with the exercise measurements used before, such as metabolic equivalent (METS) [4] and ST depression [4], an exercise test delivers a large number of measurements. The measurements have varying quality and usability and are valuable in risk prediction, in detecting coronary artery disease, and in describing a patient's functional exercise response. However, it is difficult to have comprehensive knowledge of all measurements and their thresholds. The Exercise Test Interpretation program is a solution for the problem.

The Exercise Test Interpretation program compares the exercise measurements against established thresholds and provides statements and reasoning texts, that explain the statements when thresholds are exceeded. A concise, clear, and accurate overview of an exercise test is

provided.

The Exercise Test Interpretation program creates statements on mortality prediction, on exercise induced arrhythmias, on functional response, and on ischemia / coronary artery disease. In addition, an overall statement is created, namely 'probably normal exercise test response', 'borderline exercise test response', or 'abnormal exercise test response'.

Among other databases, the FINCAVAS database from Tampere University, Finland, [5] has been used for validation. The database contains exercise ECGs, mortality data, angiographic results, and other clinical data.

2. Methods

2.1. Exercise measurements

The measurements collected during an exercise test (see table 1), together with other data (see table 2), are input for the Exercise Test Interpretation.

In general, the measurements can be classified as:

- Measurements for risk prediction
- Measurements for functional response assessment
- Measurements for ischemia and coronary artery disease (CAD) assessment

The classification is rather coarse and several measurements are meaningful not only in one group, but in two or even all three groups. For example, a measurement describing the functional response can also be useful for risk prediction.

Measurements are derived from the 12 standard leads. The limb leads are applied in Mason Likar position.

The measurements are T-Wave alternans (TWA) [1,8], ST/HR hysteresis [2,9], HR recovery [3,6,7], Heart rate reserve used [6,7,10], Frequent ventricular ectopics in recovery [11], Metabolic equivalent (METS) [4,7,8], Duke Treadmill Score [12], Blood pressure [4], Double product [4,13], ST/HR index [14], ST/HR slope [14], ST level [15], ST slope [15], ST index [15], ST integral [15],

QRS duration [4], and the arrhythmias atrial fibrillation (AFIB), exercise induced left bundle branch block (LBBB), exercise induced ventricular tachycardia (VT), and exercise induced supraventricular tachycardia (SVT).

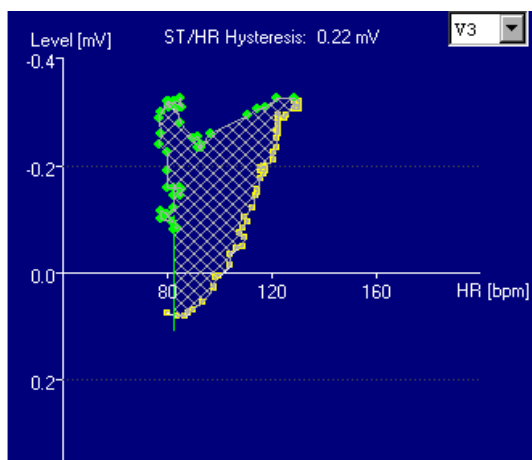


Figure 1. ST/HR hysteresis. The area between the lower curve (exercise phase) and the upper curve (recovery phase) divided by the difference of the peak exercise HR and the minimum HR during a 3-minute recovery period is the ST/HR hysteresis.

Measurements	Measurement points/ranges	Leads
T-Wave alternans	4 ranges	12
ST/HR hysteresis	1	12
HR recovery	1	-
Heart rate reserve used	1	-
Frequent ventricular ectopics in recovery (FVE recovery)	1	-
Metabolic equivalent (METS)	1	-
Duke Treadmill Score	1	-
Blood pressure	4 ranges	-
Double (rate pressure) product	1	-
ST/HR index, ST/HR slope	1	12
ST level, ST slope, ST index, ST integral	4 points and 1 range	12
QRS duration	4 points and 1 range	-
Arrhythmias (AFIB, LBBB, VT, SVT)	4 ranges	-

Table 1. Measurements, number of measurement points, measurement ranges, and leads. Measurement points are: start exercise, peak exercise, the first minute and the third minute of recovery. Measurement ranges are pretest, exercise, recovery phase, and all phases together.

Some measurements are available from the following points: start exercise (end of the pretest phase), peak exercise, the first minute and the third minute in recovery phase. Other measurements are maximum values from

ranges of pretest, exercise, recovery phase, and from all phases. Measurements are available from every lead, or are lead independent. There are more than 350 measurements (see table 1).

- Age, gender, weight, height
- Angina score
- Medication (beta blocker)
- Lead descriptors
- Stress test device (treadmill or bicycle)

Table 2. Other data

2.2. Exercise Test Interpretation (XTI)

The Exercise Test Interpretation (XTI) is a program consisting of rules and a rule interpreter. The rule interpreter receives the input data, combines them with the rules, and creates statements and reasoning texts.

The intention of the statements and reasoning texts is to provide a concise, clear, and accurate overview of the results of an exercise test.

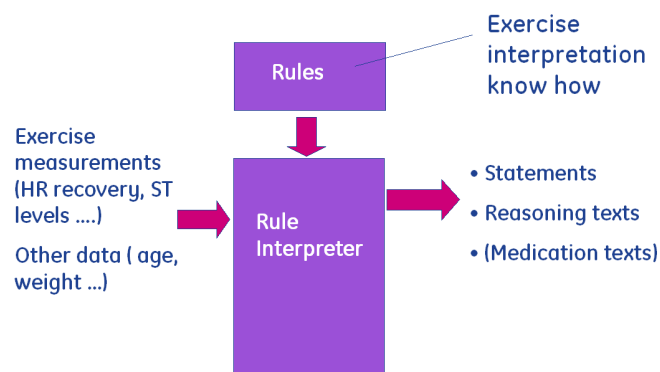


Figure 2. XTI program structure. The exercise measurements and other data are combined with the rules. Then statements and reasoning texts are created.

The input data are exercise measurements (see table 1) and other data (see table 2). Based on the input data and the rules, the rule interpreter creates

- statements on risk prediction,
- statements on functional response,
- statements on ischemia (coronary artery disease),
- technical statements.

Except for the overall statements, statements are accompanied by a reasoning text. In general, the reasoning texts conform to the rules.

Probably increased risk of malignant arrhythmia because T-wave alternans $\geq 65 \mu V$ in [V3]

Figure 3. Example of a statement with reasoning text. The first line is the statement and the second line is the reasoning text.

Medication texts are additional reasoning texts. They are only appended to medication sensitive statements when the corresponding medication name was entered.

Chronotropic incompetence
because HR reserve used < 80 %
Possible cause: Medication beta-blocker

Figure 4. Example with medication text. The first line is the statement, the second line is the reasoning text, and the third line is the medication text

In addition, one of the following overall statements is created:

- abnormal exercise test response, or
- borderline exercise test response, or
- probably normal exercise test response.

A statement does not appear when it is not significant. This means a statement is created only if the corresponding rule is true. The statement “Probably increased risk of malignant arrhythmias,” for example, appears only when the T-wave alternans value is greater than or equal to 65 μ V.

Technical statements are created to alert the user when no ECG is available, when the standard leads are incomplete, or when the exercise phase and/or recovery phase are too short.

2.2.1. Rules

The rules of the Exercise Test Interpretation determine when a certain interpretive statement is displayed. A rule consists of the result variable (for example, CAD, see below), which can be true or false, and terms, which are connected by logical ANDs. A term consists of an exercise measurement, a threshold, and a comparison operator. Different rules can have the same result variable. Those rules are connected by logical ORs. Example:

CAD = ST/HR hysteresis > 0.02 mV and HR reserve used < 70%

CAD = ST/HR hysteresis > 0.02 mV and METS < 2.0

Rules have been created for the following statements:

- Probably increased risk of morbidity / mortality
- Probably increased risk of malignant arrhythmias
- Exercise induced serious arrhythmia
- Significant chronotropic incompetence
- Chronotropic incompetence
- Insufficient exercise capacity
- Abnormal blood pressure response
- Insufficient rate pressure response
- Ischemia, probably caused by coronary artery disease (CAD level 1)

- ST-segment changes (CAD level 2)
- Unspecific ST-segment changes (CAD level 3)
- No ECG
- Standard leads incomplete
- Warning! Results are questionable

The rules and their terms are listed in [16].

2.2.2. Data and validation

Among other databases, the FINCAVAS database from Tampere University, Finland, [5] has been used for validation of the Exercise Test Interpretation. This database contains patient records with exercise ECGs, endpoints, angiographic references, arrhythmia references, and other clinical data. Endpoints are cardiovascular death and overall mortality according to the WHO ICD-10 classifications [17]. The angiographic reference for coronary artery disease is a stenosis of 50% or greater, in at least one of the main coronary arteries. The arrhythmia references have been created by ECG experts.

Validation was done retrospectively by feeding the exercise ECGs into the program. With every exercise ECG, a set of measurements was computed and classified by generating statements. The measurements and the statements were compared with the endpoints, with the angiographic references, and with the arrhythmia references. Statistical analysis was performed with the statistics program MedCalc 11.3.3.3.0.

3. Validation results

The clinical relevance and prediction and/or diagnosis capabilities of the measurements have been proved in the literature [1-15]. The implementation of all measurements and rules was validated and documented in-house. It would be too much to list all validation results here. As examples, the following unadjusted validation results have been achieved with statistical analysis:

- Overall mortality prediction: Hazard ratio 4, $p < 0.0001$, 2336 patients (first statement)
- Cardiovascular death prediction: Hazard ratio 6, $p < 0.0001$, 2336 patients (second statement)
- Ischemia (coronary artery disease) detection (CAD level 1): Sensitivity 58%, specificity 94%, 556 patients
- Exercise induced arrhythmias (LBBB, VT, SVT, AFIB) are rare in the FINCAVAS database. In 2336 patients, all of them have been detected correctly.

4. Discussion and conclusions

In recent years, international research on exercise testing has provided new and valuable exercise measurements. However, the motivation to implement

them in exercise test products was not very high. First, the user is overloaded with additional exercise measurements. Second, there was not a strong market demand, because many physicians simply rely on ST depression. We think that the Exercise Test Interpretation will help to follow research on exercise testing more closely, will effectively provide our customers with research results, and will therefore contribute to a higher quality in exercise test assessment.

As a first approach, the published thresholds were taken, but it became clear that with those thresholds too many exercise tests are positive. We changed some of the thresholds with the help of the corresponding ROC-curves. In addition, we added terms to rules to preselect patient groups. For instance, patients with a high METS value are excluded from mortality risk prediction [8].

Pharmacological tests or pharmacological tests combined with mechanical tests have a weak heart rate response, especially when beta blockers have been given in addition. Therefore, some rules have to be blocked. It is clear that further investigation is necessary to improve the interpretation quality for pharmacological and combined pharmacological tests.

References

- [1] Nieminen T, Lehtimäki T, Viik J, Lehtinen R, Nikus K, Kööbi T, Niemelä K, Turjanmaa V, Kaiser W, Huhtala H, Verrier RL, Huikuri H, Kähönen M. T-wave alternans predicts mortality in a population undergoing clinically indicated exercise test. *Eur Heart J*, 2007;28:2332-2337
- [2] Lehtinen R, Sievänen H, Viik J, Turjanmaa V, Niemelä K, Malmivuo J. Accurate Detection of Coronary Artery Disease by Integrated Analysis of the ST-Segment Depression/Heart Rate Patterns During Exercise and Recovery Phases of the Exercise Electrocardiography Test. *Am J Cardiol* 1996; 78:1002-1006.
- [3] Cole CR, Blackstone EH, Pashkow FJ, Snader CE, Lauer MS, Heart-Rate Recovery Immediately After Exercise As a Predictor of Mortality. *N Engl J Med* 1999;341:1351-7
- [4] Gibbons RJ, Balady GJ, Bricker JT, Chaitman BR, Fletcher GF, Froelicher VF, Mark DB, McCallister BD, Mooss AN, O'Reilly MG, Winters WL. ACC/AHA 2002 guideline update for exercise testing: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Exercise Testing). *Circulation* 2002;106:1883
- [5] Nieminen T, Lehtinen R, Viik J, Lehtimäki T, Niemelä K, Nikus K, Niemi M, Kallio J, Kööbi T, Turjanmaa V, Kähönen M. The Finnish Cardio-vascular Study (FINCAVAS): characterizing patients with high risk of cardiovascular morbidity and mortality. *Cardio-vascular Disorders* 2006, 6:9
- [6] Lauer MS. Exercise Testing, Part 1: Looking Beyond the ST Segment. *Cardiology Rounds*, Brigham and Women's Hospital, Boston, Massachusetts. May 2002 Volume 6, Issue 5.
- [7] Lauer MS, Exercise Testing, Part 2: The Value of Heart Rate Recovery. *Cardiology Rounds*, Brigham and Women's Hospital, Boston, Massachusetts. May 2002 Volume 6, Issue 6.
- [8] Minkkinen M, Nieminen T, Verrier RL, Leino J, Lehtimäki T, Viik J, Lehtinen R, Nikus K, Kööbi T, Turjanmaa V, Kähönen, M. Impaired exercise capacity predicts sudden cardiac death in a low-risk population: enhanced specificity with heightened T-wave alternans. *Annals of Medicine*, 2009;41:380-389
- [9] Lehtinen R, Viik J, Kaiser W, Niemi M, Kööbi T, Niemelä K, Turjanmaa V. Accurate Detection of Coronary Artery Disease by ST-segment/Heart Rate Hysteresis of the Exercise Electrocardiography Test. *American College of Cardiology*, vol 45, issue 3, suppl 1, page A158.
- [10] Lauer MS, Okin PM, Larson MG, Evans JC, Levy D. Impaired heart rate response to graded exercise. Prognostic implications of chronotropic incompetence in the Framingham Heart Study. *Circulation* 1996;93:1520-1526.
- [11] Frolkis JP, Pothier CE, Blackstone EH, Lauer MS. Frequent Ventricular Ectopy after Exercise as a Predictor of Death. *N Engl J Med* 2003;348:781-790.
- [12] Mark DB, Shaw L, Harrell FE, Hlatky MA, Lee KL, Bengton JR, McCants CB, Califf RM, Pryor DB. Prognostic value of a treadmill exercise score in outpatients with suspected coronary artery disease. *N Engl J Med* 1991;325:849-853
- [13] Driggers DA, Marchant D. Maximizing the exercise stress test, Critical factors that enhance its validity. *Postgraduate Medicine*, 1999;105
- [14] Kligfield P, Lauer MS. Exercise Electrocardiogram Testing: Beyond the ST Segment. *Circulation* 2006;114:2070-2082.
- [15] Froelicher VF, Myers J, Follansbee WP, Labovitz AJ. Exercise and the Heart. *Mosby-Yearbook*. Ed3,1993
- [16] GE Healthcare. Physician's Guide to GE Stress Systems. 2009353-001, Revision D.
- [17] WHO ICD-10. International Statistical Classification of Diseases and Related Health Problems. 10th Revision, Version for 2007.
<http://www.who.int/classifications/icd/en/>

Address for correspondence.

Willi Kaiser
 GE Medical Systems Information Technologies GmbH
 Munzinger Straße 5
 79111 Freiburg
 Germany
 Willi.kaiser@med.ge.com