ABSTRACT
Objective: Cardiac marker sensitivity depends on chest pain duration at the time of sampling. Our objective was to estimate the sensitivity, specificity, and likelihood ratios of early CK–MB and myoglobin assays in patients presenting to the emergency department (ED) with nondiagnostic ECGs, stratified by the duration of ongoing chest pain at the time of ED assessment.
Methods: This was a prospective observational study carried out in 10 US and 2 Canadian EDs. Patients >25 years of age with ongoing chest pain and nondiagnostic ECGs were stratified by pain duration (0–4 h, 4–8 h, 8–12 h, >12 h). CK–MB and myoglobin assays were drawn at T = 0 (ED assessment) and T = 1 hr. Patients were followed for 7–14 days to identify all cases of acute myocardial infarction (AMI). ED test results were correlated with patient outcomes.
Results: Of 5005 eligible patients, 565 had AMI. Pain duration was 0–4 h in 3014 patients, 4–8 h in 961, 8–12 h in 487, and >12 h in 543. Marker sensitivity increased with pain duration, ranging from 28%–77% for CK–MB and 39%–73% for myoglobin. The maximal sensitivity achieved by a T = 0 assay was 73%, and this was in patients with 8–12 or >12 h of ongoing pain. No combination of tests achieved 90% sensitivity in any pain duration strata.
Conclusions: Regardless of chest pain duration, single assays and early serial markers (0+1 hr) do not rule out AMI; therefore, serial assays over longer observation periods are required. Likelihood ratios derived in this study will help physicians who use Bayesian analysis to determine post-test AMI likelihood in patients with chest pain.

RÉSUMÉ
Méthodes : Il s’agissait d’une étude prospective d’observation menée dans dix départements d’urgence américains et deux départements d’urgence canadiens. Les patients âgés de >25 ans accu-

*St. Paul’s Hospital and the University of British Columbia, Vancouver, BC; †Royal Columbian Hospital, New Westminster, BC; ‡Henry Ford Heart and Vascular Institute, Detroit, Mich.; §University of Cincinnati Medical Centre, Cincinnati, Ohio; ¶Ohio State University, Columbus, Ohio; ‡Wayne State University, Detroit, Mich.; **University of Washington, Seattle, Wash.; ††Willian Beaumont Hospital, Royal Oak, Mich.

Received: Apr. 1, 2002; final submission: June 15, 2002; accepted: July 7, 2002

This article has been peer reviewed.
Introduction

Resource limitations pressure physicians to admit fewer chest pain patients to acute care units, but medicolegal factors demand that they discharge fewer with unrecognized acute coronary syndromes. Consequently, more patients undergo emergency department (ED) diagnostic protocols, which are based largely on the use of cardiac marker assays to “rule out” myocardial infarction. One large study in 4 teaching hospitals concluded that emergency physicians rely heavily on the results of single marker assays, although these have been shown to have poor sensitivity. The National Academy of Clinical Biochemistry (NACB) recently recommended that physicians employ 2 cardiac markers to evaluate patients with chest pain — an early marker that is reliably increased within 6 hours and a definitive marker that is elevated within 6 to 9 h.

CK–MB and the troponins are considered definitive markers. They are highly specific for myocardial injury, but are released slowly during infarction. Sensitivity at the time of ED presentation ranges from 14% to 76% for CK–MB and 10% to 67% for troponins. With serial testing, sensitivity improves to 68%–100% for CK–MB and to 57%–100% for troponins. The troponins have additional value for risk stratification of patients with unstable angina, but different troponin assay techniques generate different quantitative results, and there is no widely accepted troponin threshold for myocardial infarction.

Myoglobin, a nonspecific marker of muscle injury, is released rapidly during acute myocardial infarction (AMI) and provides greater early sensitivity, from 26%–58% at presentation, to 79%–100% with serial assays. It has been proposed as the early marker for ED “rule-out” protocols, included in recently developed commercial marker panels, and advocated by the NACB panel as the most conveniently measured early marker. But myoglobin’s early sensitivity may not be adequate to rule out AMI and the clinical benefit of adding myoglobin to other, more specific, markers remains unclear.

Previous marker studies are limited by small sample size, inappropriate patient spectrum, poor follow-up of discharged patients, and lack of patient stratification by pain duration. Consequently, few data are available regarding the early diagnostic strength of marker assays in patients with differing pain duration. The current study was a sub-study of SMARTT (the serial markers, acute myocardial infarction and rapid treatment trial), a randomized trial assessing the clinical impact of early serial (0 + 1 h) CK–MB and myoglobin assays on the use of thrombolytic therapy. The objective of this sub-study was to estimate the sensitivity, specificity and likelihood ratios (LRs) for early CK–MB and myoglobin assays in ED patients with nondiagnostic ECGs, stratified by the duration of continuous chest pain at the time of ED assessment. Our hypotheses were that, in patients with ongoing pain, marker sensitivity would increase with pain duration, and that, in patients with 8 to 12 h of continuous pain, serial cardiac marker sampling over a 1-h interval would achieve high sensitivity for AMI.

Methods

Setting and patients

This was a prospective survey carried out at 12 university hospitals in Canada and the United States in 2000–2001.
and community hospital EDs (10 US, 2 Canadian). Con-
senting patients, 25 years and over, who had ongoing chest
pain consistent with possible acute coronary syndrome
were eligible. Patients were excluded if their pain was ob-
viously noncardiac (based on clinical presentation or chest
x-ray findings), if their pain had resolved prior to evalua-
tion, if they were suspected of drug or alcohol abuse, or if
their initial ECG was diagnostic of myocardial infarction.
All patients provided written informed consent, and the
study was approved by the investigational review boards at
all participating hospitals.

Clinical evaluation and stratification
Emergency physicians performed the clinical assess-
ment, determined study eligibility, and completed a stan-
dardized data form documenting patient demographics,
cardiac risk factors, duration of ongoing chest pain, pro-
visional (ED) diagnosis and patient disposition. ECGs
were performed on all patients and interpreted in blinded
fashion at the Ischemia Monitoring Core Laboratory,
Duke Clinical Research Institute, Durham, NC. Tracings
were considered diagnostic of AMI if they showed ST-
segment elevation >1 mV in 2 contiguous limb leads or
>2 mV in 2 contiguous anterior leads. Patients were
stratified into 4 groups, based on the duration of continu-
ous pain (0–4 h, 4–8 h, 8–12 h and >12 h) at the time of
ED assessment (T = 0).

Follow-up and outcome assessment
Hospitalized patients were followed for the duration of
their hospital stay. Within the protocol, AMI was de-
defined using WHO criteria, requiring evolution of ECG
changes or characteristic CK–MB rise and fall docu-
mented by serial assays. The study did not mandate spe-
cific inpatient testing regimes or change local practices.
In uncertain cases, if inadequate data had been gathered
to fulfill WHO criteria, we accepted the clinical diagno-
sis made by the treating cardiologist. Patients dis-
charged from the ED were followed at 7 to 14 days by
telephone or letter, to identify readmission, AMI or
death. Follow-up (after discharge) marker assays and
ECGs were not required.

Markers
Myoglobin and CK–MB assays were drawn at T = 0 and
T = 1 h and tested using Baxter Stratus II analyzers
(Dade International). CK–MB levels >6 ng/ml and myo-
globin levels >100 ng/ml were considered positive. Each
test result was correlated with the corresponding patient’s
outcome (AMI vs. no AMI) and determined to be true-
positive, false-positive, true-negative, or false-negative.
Sensitivity, specificity, predictive values, and LRs were
determined for T = 0 assays, T = 1-h assays and early ser-
ial (T = 0+1 h) assays.

Statistical analysis
Sensitivity (true-positive rate) and specificity (true-nega-
tive rate) were calculated using standard formulae.47 Posi-
tive likelihood ratios were determined using this formula:

\[
LR^+ = \frac{\text{sensitivity}}{1 - \text{specificity}}
\]

Negative likelihood ratios were determined using this for-
mula:

\[
LR^- = \frac{1 - \text{sensitivity}}{\text{specificity}}
\]

Intervals of 95% confidence were calculated around criti-
cal sensitivity, specificity, and LRs.

Results

Patients
Over a 22-month period, 8396 patients were enrolled in
the SMARTT pilot study and clinical trial. Of these,
400 (4.8%) had missing initial data that precluded
analysis, 355 (4.2%) had incomplete follow-up data or
were lost to follow-up. In total, 1804 (21%) were ineli-
gible for this sub-study because their pain was no
longer ongoing at the time of ED assessment, and 432
were ineligible because of diagnostic ST elevation on
their initial ECG. Of the remainder, 5005 had both T =
0 and T = 1-h samples drawn, therefore were included
in the analysis.

Table 1 summarizes patient characteristics. Pain duration
at the time of ED presentation was 0–4 h in 3014 patients,
4–8 h in 961, 8–12 h in 487, and >12 h in 543. Overall,
565 of 5005 patients had AMI, and the highest AMI rate
(16%) was in patients presenting at >12 h.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age, yr</td>
<td>60.0</td>
</tr>
<tr>
<td>Gender (% male)</td>
<td>57.6</td>
</tr>
<tr>
<td>Pain duration at ED presentation</td>
<td></td>
</tr>
<tr>
<td>0–4 h, no. (and %)</td>
<td>3014 (60)</td>
</tr>
<tr>
<td>4–8 h, no. (and %)</td>
<td>961 (19)</td>
</tr>
<tr>
<td>8–12 h, no. (and %)</td>
<td>487 (10)</td>
</tr>
<tr>
<td>&gt;12 h, no. (and %)</td>
<td>543 (11)</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of patients in the SMARTT trial
Sensitivity and specificity

Table 2 shows that at all time intervals CK–MB was more specific (96%–98%) than myoglobin (87%–93%). Sensitivities for both markers increased with pain duration, ranging from 28% to 77% for CK–MB and 39% to 73% for myoglobin. In patients with 0–4 h of pain, myoglobin was more sensitive (39% vs. 28%), in patients with 4–8 h of pain, sensitivities for the 2 markers were similar (61% vs. 56%), and in patients with over 8 h of pain, CK–MB was more sensitive (73% vs. 65%). The maximum sensitivity achieved by a single assay was 77%: this was the 1-h CK–MB draw in patients with 8–12 or >12 h of pain.*

Sensitivity was enhanced by combining markers and performing serial assays (Figs. 1 and 2). Table 2 shows that, in patients with >12 h of pain, 1 CK–MB sensitivity rose from 73% to 76% and myoglobin sensitivity rose from 65% to 70% if assays were repeated 1 h after presentation. In the same (>12 h) patient stratum, T = 0 sensitivity was 73% for CK–MB alone and 87% if both myoglobin and CK–MB were assayed. It is important to note, however, that no single assay achieved even 80% sensitivity and no combination of markers achieved 90% sensitivity in any (pain duration) strata.

Likelihood ratios

Negative LRs ranged from 0.24 to 0.74 for CK–MB and from 0.30 to 0.67 for myoglobin (Table 3). Both tests became stronger negative predictors as pain duration increased. Positive LRs ranged from 4.3–9.6 for myoglobin and from 9.3–37 for CK–MB.

Discussion

This study confirms that marker sensitivity increases with pain duration and that, in patients with 8 to 12 h of ongoing pain, serial cardiac marker sampling over a 1-h interval achieves rel-

---

Table 2. Sensitivity, specificity and predictive values for cardiac markers stratified by chest pain duration

<table>
<thead>
<tr>
<th>Pain duration</th>
<th>CK–MB</th>
<th></th>
<th>MYO</th>
<th></th>
<th>CK–MB or MYO</th>
</tr>
</thead>
<tbody>
<tr>
<td>@ T = 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–4 h</td>
<td>0.28</td>
<td>0.39</td>
<td>0.45</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>4–8 h</td>
<td>0.56</td>
<td>0.61</td>
<td>0.52</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>8–12 h</td>
<td>0.73</td>
<td>0.63</td>
<td>0.95</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>&gt;12 h</td>
<td>0.73</td>
<td>0.65</td>
<td>0.96</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td><strong>0.28</strong></td>
<td><strong>0.39</strong></td>
<td><strong>0.45</strong></td>
<td><strong>0.45</strong></td>
<td><strong>0.28</strong></td>
</tr>
<tr>
<td><strong>Specificity</strong></td>
<td><strong>0.97</strong></td>
<td><strong>0.93</strong></td>
<td><strong>0.91</strong></td>
<td><strong>0.91</strong></td>
<td><strong>0.97</strong></td>
</tr>
<tr>
<td><strong>PPV</strong></td>
<td><strong>0.59</strong></td>
<td><strong>0.70</strong></td>
<td><strong>0.59</strong></td>
<td><strong>0.59</strong></td>
<td><strong>0.59</strong></td>
</tr>
<tr>
<td><strong>NPV</strong></td>
<td><strong>0.91</strong></td>
<td><strong>0.95</strong></td>
<td><strong>0.95</strong></td>
<td><strong>0.95</strong></td>
<td><strong>0.95</strong></td>
</tr>
</tbody>
</table>

* The >12-h pain duration subset was chosen for illustrative purposes because tests performed best in this group.

AMI = acute myocardial infarction; PPV = positive predictive value; NPV = negative predictive value; MYO = myoglobin

---

* The >12-h pain duration subset was chosen for illustrative purposes because tests performed best in this group.
atively high sensitivity for AMI. However, even in patients with prolonged pain, test sensitivity did not approach 100%.

Many authors suggest that myoglobin and CK–MB assays achieve excellent sensitivity in patients with 3–6 or 6–8 h of symptoms, respectively. A recent NACB position paper proposes the use of “an early marker that is reliably increased within 6 h and a definitive marker that is elevated within 6–9 h.” Based on this recommendation and previous studies, clinicians may feel that a single test can rule out AMI in patients with adequate symptom duration. Our data demonstrate, however, that neither the “early” nor the “definitive” markers are reliably elevated within the time frames suggested. The data also show that, regardless of pain duration, one marker assay does not rule out myocardial infarction, and that, if there is significant likelihood of AMI, serial sampling over longer time periods is necessary.

Myoglobin utility
Because myoglobin is released within 3–4 h of symptom onset, it has been advocated as an “early” marker of myocardial injury. The current study confirms that, in patients with 0–4 h of pain, myoglobin is more sensitive than CK–MB; however, in this time range, myoglobin sensitivity was insufficient to rule out AMI even if serial (0+1 h) assays were performed. By the time its sensitivity approached adequate levels (in patients with >8 h of pain), CK–MB was more sensitive. This suggests that myoglobin’s early sensitivity advantage may not be clinically important. Further, because myoglobin lacks specificity, it cannot be used to guide specific AMI therapy, and false positive myoglobin assays could inappropriately increase downstream investigation costs and monitored admissions.

<table>
<thead>
<tr>
<th>Pain duration @ T = 0</th>
<th>T = 0 h assay</th>
<th>T = 1 h assay</th>
<th>0 or 1 h assay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0–4 h</td>
<td>4–8 h</td>
<td>8–12 h</td>
</tr>
<tr>
<td><strong>CK–MB</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR+</td>
<td>9.3</td>
<td>19</td>
<td>24</td>
</tr>
<tr>
<td>LR–</td>
<td>0.74</td>
<td>0.45</td>
<td>0.28</td>
</tr>
<tr>
<td><strong>MYO</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR+</td>
<td>4.3</td>
<td>8.7</td>
<td>7.0</td>
</tr>
<tr>
<td>LR–</td>
<td>0.67</td>
<td>0.42</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>CK–MB or MYO</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR+</td>
<td>4.5</td>
<td>9.1</td>
<td>7.1</td>
</tr>
<tr>
<td>LR–</td>
<td>0.61</td>
<td>0.29</td>
<td>0.25</td>
</tr>
</tbody>
</table>

MYO = myoglobin; LR+ = positive likelihood ratio; LR– = negative likelihood ratio
Bayesian analysis and likelihood ratios

Although this study shows that single and early serial assays do not reliably rule out AMI, it does not prove that all patients with chest pain require a uniform approach, with multiple marker assays over prolonged time periods. Bayesian logic tells us that different patients require different testing strategies based on their pretest clinical likelihood of disease. For example, in patients with high pretest likelihood, AMI can only be "ruled out" by a powerful negative test such as serial examination, serial ECGs and serial markers over 12–24 h, followed by other noninvasive or invasive modalities. In patients with low to moderate pretest likelihood, a weaker test may suffice — for example, serial ECGs and markers over a 6-h period. In patients with extremely low pretest likelihood (e.g., <1% chance of AMI), no marker testing may be necessary. Pretest likelihood, therefore, determines what type of diagnostic testing is necessary to carry the clinician to a positive or negative decision threshold.

Diagnostic "strength" is best expressed by a test’s LRs. Negative LR (LR–) reflect the test’s power to rule out disease, while positive LR (LR+) reflect its power to confirm disease. Armed with an estimate of pretest likelihood, clinicians can use LRs to determine post-test likelihood.†

The negative LRs determined in this study are modest, suggesting that these tests, used as described, are weak negative predictors. To illustrate, the strongest LR– seen in this study (the 1-h CK–MB assay patients with 8–12 h of pain) was 0.24. In a patient with 10 h of ongoing pain whose pretest clinical likelihood is 50%, post-test likelihood, after a negative CK–MB, would only fall to 20%. To reduce post-test likelihood to a more acceptable discharge threshold level of 2%, a much stronger test with an LR– of 0.02 would be required. No combination of tests in the current study approached this level of diagnostic strength.

In an ideal patient, with more than 8–12 h of ongoing pain, combining 0- and 1-h serial CK–MB and myoglobin assays provides an LR– of 0.15. Accepting a rule-out threshold of 2%, this combination of assays is strong enough to rule out AMI only if pretest likelihood is <10%. If the acceptable risk threshold (post-test likelihood) is lowered to 1%, then combined serial testing is only capable of ruling out patients who have a pretest likelihood of less than 5%. Therefore, our data suggest that early serial assays are only "sufficient" to rule out AMI in a small subset of patients who have more than 8–12 h of ongoing pain and who have very low pretest clinical likelihood of AMI.

Marker insensitivity

It is difficult to postulate why marker sensitivities failed to approach 100% in patients with 8 or more h of pain. Despite the history of continuous pain, some of these patients may have had unstable angina without infarction at the time of their ED visit, and evolved to AMI during the follow-up period. It is also possible that the history of chest pain duration is unreliable, even when collected prospectively, and we know of no study examining the interobserver reliability of chest pain duration.

Previous studies

Previous authors have reached more optimistic conclusions regarding cardiac marker utility. For several reasons, these conclusions should be examined critically. Some studies included patients with diagnostic ECGs. These patients tend to have more prolonged symptoms, more severe clinical illness and more marker leakage than those with nondiagnostic tracings. Assays will appear more sensitive if patients with diagnostic ECGs are included.

Many studies enroll only patients who are admitted to cardiac care unit settings and do not follow patients discharged from the ED. The result is a sampling bias, because inpatients differ systematically from unselected ED patients. Those admitted to hospital cardiac units tend to be higher risk, with more severe clinical presentation and a higher diagnostic ECG rate. They are further along on the time continuum; hence marker assays are more sensitive. For all these reasons, data derived from inpatient studies should not be generalized to the ED setting.

When discussing the time-dependent utility of cardiac marker assays, many physicians cite kinetic studies, which generally report excellent early sensitivity. But kinetic studies enroll patients with obvious myocardial infarction and diagnostic ECGs. This, too, is a different spectrum of patients from the ones who pose a diagnostic dilemma in the ED.

Other limitations of previous studies include failure to report symptom duration at the time of marker sampling, and small sample size, which reduces the precision of test accuracy estimates. Only 5 ED-based studies have enrolled more than 50 AMI patients, only 2 of these limited enrollment to patients with nondiagnostic ECGs, and only one followed up patients...
who were discharged from the ED.\textsuperscript{22}

In order to gather meaningful data with respect to the ED diagnostic utility of cardiac markers, the current study enrolled ED patients with nondiagnostic ECGs, prospectively determined symptom duration at the time of marker sampling, studied an adequate sample of AMIs, and followed outcomes in patients discharged from the ED.

Limitations

This study suggests that single markers and early (0+1 h) serial markers lack sensitivity, are relatively weak diagnostic predictors and have limited clinical utility. These conclusions cannot, however, be generalized to diagnostic protocols involving serial marker draws over longer time periods. In other words, one marker assay after 12 h of pain is less sensitive than 2 markers drawn at 6 and 12 h of pain.

In patients discharged from the ED, we conducted health records and telephone follow-up but did not require follow-up marker assays and ECGs; therefore, it is possible that some patients had unrecognized ischemic events. A more intensive detection process would probably have led to slightly lower marker sensitivity estimates than those reported. We did not study troponin assays, and cannot make conclusions about the diagnostic utility of early troponin testing; however, because troponins have similar release kinetics to CK–MB,\textsuperscript{2,29,35,49–54} it is likely that troponin assays would perform similarly at these early time intervals. Finally, the use of a core lab to perform assays may have introduced freeze/thaw artifact, which can potentially reduce detectable levels of biomarkers — particularly CK–MB.

Because our main objective was to characterize changes in test performance that occur related to pain duration, we excluded patients with intermittent or resolved pain. In these patients, when there is no distinct time of onset or relief, it is difficult to reliably define pain duration. Including patients with uncertain chest pain duration would have “contaminated” our primary results. As a result, the study findings can only be generalized to patients with ongoing pain. Patients with intermittent or resolved pain are less likely to have occlusive coronary thrombosis and myocardial necrosis; therefore, marker assays would probably have had even poorer diagnostic sensitivity in excluded patients.

Studies of diagnostic tests are most useful if they demonstrate the impact of diagnostic testing on clinical outcomes. If clinical sensitivity (96%–100%) is better than marker sensitivity (25%–89% in this study), then increasing the emphasis on early marker assays has the potential to decrease diagnostic sensitivity and influence physicians to incorrectly discharge patients with unstable acute coronary syndromes. Because this study was descriptive in nature, we cannot suggest that the use of ED markers had any impact, beneficial or detrimental, on patient outcomes. Future randomized trials that expose patient groups to different diagnostic strategies will provide better information about test utility.

Conclusions

Regardless of chest pain duration, single assays and early serial markers (0+1 h) do not rule out AMI; therefore, serial assays over longer observation periods are required. Myoglobin assays may have limited additional clinical utility relative to definitive markers like CK–MB and troponin. Likelihood ratios derived in this study will help physicians who use Bayesian analysis to determine post-test AMI likelihood in patients with chest pain.

Competing interests: None declared.

Acknowledgements: The SMARTT study was supported in part by an unrestricted grant from Genentech, Inc., South San Francisco, Calif., and Dade Stratus Inc., Miami, Fla.

References


