Critical Review Form
Diagnostic Test

A 2-hour thrombolysis in myocardial infarction score outperforms other risk stratification tools in patients presenting with possible acute coronary syndromes: comparison of chest pain risk stratification tools. Am Heart J. 2012 Oct;164(4):516-23.

Objective: **"to compare the 2-hour TIMI score with other ADPs [accelerated diagnostic protocols] in risk stratification of patients with symptoms suggestive of ACS [acute coronary syndrome] to identify patients suitable for early discharge to outpatient management." (p. 518)**

Methods: **This was a post-hoc analysis of data collected prospectively on patients presenting to the Emergency Department of Christchurch Hospital, a tertiary care facility in New Zealand, from November 2007 to April 2010. This was data from one of the centers contributing to the ASia Pacific Evaluation of Chest pain Trial (ASPECT). Patients were included if they presented with symptoms suggestive of ACS without an apparent noncardiac course. This was a convenience sample enrolled when a research nurse was available. Exclusion criteria included: age ≤ 18 years, inability to consent, refusal to participate, inability to contact for follow-up, and ST elevation on the initial ECG.**

All included patients underwent testing on arrival and at 2 hours post-presentation, including troponin, point of care CK-MB and myoglobin, and ECG. Based on this initial data, 10 different ADPs were calculated and patients categorized as either low risk or not low risk for each ADP. Standard of care included repeat troponin testing ≥ 8 hours post-presentation, and any additional testing was at the discretion of the treating physician, who was blinded to CK-MB and myoglobin results as well as any calculated ADPs.

Follow-up was conducted using the New Zealand death registry and a "national events search" (p. 518) that identifies any regional outpatient follow-up appointment or hospital admission within the nation. Patients also received follow-up telephone calls by research nurses at 30 to 45 days and 6 months. The primary endpoint was ACS diagnosed within 30 days. Diagnosis was made by adjudication of two research physicians blinded to CK-MB and myoglobin results, and calculated ADPs, and required either myocardial infarction (MI) or unstable angina. MI was defined as any of the following:

- **1 or more troponin ≥ 99th percentile with ≥ 20% rise or fall, in conjunction with symptoms of ACS.**
- **New ischemic ECG changes (ST deviation of ≥ 0.5 mm or T wave inversion ≥ 1 mm in 2 or more contiguous leads)**
- **New regional wall motion abnormality on ECHO.**

Unstable angina was defined as symptoms of cardiac ischemic with negative serial troponins and evidence of ischemia on the ECG or provocative testing (ST depression ≥ 2 mm, or ST depression ≥ 1 mm with ischemic symptoms).

A total of 1000 patients were included in the analysis, of whom 338 (33.8%) underwent stress testing and 303 (30.3%) underwent coronary angiography. A diagnosis of ACS was made in 362 (36.2%), including 242 (24.2%) with MI and 120 (12.0%) with unstable angina; 175 (17.5%) underwent coronary revascularization. The mean age was 65 and 59.6% were male.

Guide		Comments
I.	**Are the results valid?**	
A.	**Did clinicians face diagnostic uncertainty?**	Yes. This was a population of patients presenting to the ED with chest pain without ST elevation. There was a wide variety in terms of risk of acute coronary syndrome, and the clinicians faced significant uncertainty regarding the diagnosis.
B.	**Was there a blind comparison with an independent gold standard applied similarly to the treatment group and to the control group? (Confirmation Bias)**	N/A. There were no treatment or control groups, but rather this was a single cohort in whom the ADPs were tested. There is also no single recognized gold standard in the diagnosis of acute coronary syndrome.
C.	**Did the results of the test being evaluated influence the decision to perform the gold standard? (Ascertainment Bias)**	No. While there is no single accepted gold standard in the diagnosis of acute coronary syndrome, many would consider cardiac catheterization to be the gold standard. Clinicians in the study were not made aware of the results of the ADP calculations, and hence the test results did not influence the decision to perform further testing (including stress testing, thoracic imaging, or cardiac catheterization).
II.	**What are the results?**	
A.	**What likelihood ratios were associated with the range of possible test results?**	The test characteristics of the 3 ADPs with the best performance (with 95% CI) were as follows: **Christensen et al** • Sensitivity 99.4 (98.1-99.9) • NPV 93.8 (78.0-98.9) • Negative LR 0.12 (0.02-0.50) • Specificity 4.7 (3.9-5.0) • PPV 37.2 (36.7-37.4) • Positive LR 1.04 (1.02-1.05)

		Hess et al • Sensitivity 99.7 (98.3-100) • NPV 98.9 (99.3-99.9) • Negative LR 0.02 (0-0.13) • Specificity 14.1 (13.3-14.3) • PPV 39.7 (39.2-39.8) • Positive LR 1.16 (1.13-1.17) **2-Hour TIMI** • Sensitivity 99.2 (97.5-99.8) • NPV 98.1 (94.1-99.5) • Negaitve LR 0.04 (0.01-0.11) • Specificity 23.8 (22.9-24.2) • PPV 42.5 (41.8-42.7) • Positive LR1.30 (1.26-1.32) The ADP by Christensen et al had a high sensitivity, but a poor specificity. The ADP by Hess et al had a higher sensitivity and specificity, while the 2-hour TIMI had both a high sensitivity and a better specificity.
III.	**How can I apply the results to patient care?**	
A.	**Will the reproducibility of the test result and its interpretation be satisfactory in my clinical setting?**	Yes. All 3 ADPs with reasonable test characteristics rely on easily obtained clinical findings (history and physical exam), items in the past medical history, and test results (ECG findings, cardiac enzymes). All of these should be easily obtained and interpreted in our institutions, and in most institutions in the US.
B.	**Are the results applicable to the patients in my practice?**	Yes.
C.	**Will the results change my management strategy?**	Yes. Application of any of the 3 ADPs can predict a group of patients with very low risk of major adverse cardiac events. The extremely low specificity of the ADP by Christensen et al suggests that its application would not result in a significant reduction in provocative test ordering. Use of either the ADP by Hess et al or the 2-hour TIMI score would sufficiently rule out significant disease while at the same time reducing the number of provocative tests ordered.
D.	**Will patients be better off as a result of the test?**	Yes. By reducing the rate of provocative testing, we would reduce hospital/ED length of stay significantly, reduce cost, and reduce the rate of false positive testing in a very low-risk population. This would likely result in a reduction in rates of invasive testing (cardiac catheterization) as well as radiologic imaging (coronary CT angiography, nuclear stress testing, myocardial perfusion imaging).

Limitations:

1. **This was a retrospective application of several accelerated diagnostic protocols/clinical decision rules to data collected prospectively at a single institution of a multi-center trial. The chart review methods were not well-defined (Gilbert 1996 and Worster 2004).**

2. **It is unclear if application of the ADPs would impact practice, reduce test ordering, or affect patient outcomes. Further impact analyses would need to be performed.**

3. **There is no true "gold standard" in the assessment of outcomes in chest pain patients. Instead, a composite outcome (MACE) is used whose components are not necessarily equivalent in terms of patient importance. This practice has been called into question (Kip 2008).**

Bottom Line:

In this assessment of multiple ADPs on a population from a single site in Australia, 3 rules seemed to perform the best. The rule by Christensen et al had a high sensitivity, but a NPV of only 93.8% and a poor specificity of 4%. The rule by Hess et al and the 2-hour TIMI rule had high sensitivities and NPVs and more reasonable specificities. This was a retrospective validation of multiple ADPs on prospectively collected data in a single New Zealand institution. Further prospective validation of the 2-hour TIMI score and rule by Hess would need to be performed in various populations prior to widespread implementation.