TMI, GRACE and alternative risk scores in Acute Coronary Syndromes: A meta-analysis of 40 derivation studies on 216,552 patients and of 42 validation studies on 31,625 patients

Fabrizio D’Ascenzo a,⁎, Giuseppe Biondi-Zoccai b, Claudio Moretti a, Mario Bollati a, Pierluigi Omedè a, Filippo Sciuto a, Davide G. Presutti a, Maria Grazia Modena b, Mauro Gasparini c, Matthew J. Reed d, Imad Sheiban a, Fiorenzo Gaita a

a Division of Cardiology, University of Turin, Turin, Italy
b Division of Cardiology, University of Modena and Reggio Emilia, Modena, Italy
c Department of Mathematics, Politecnico di Torino, Italy
d Emergency Medicine Research Group Edinburgh (EMeRGE), United Kingdom

Abstract

Background: Acute coronary syndromes (ACS) represent a difficult challenge for physicians. Risk scores have become the cornerstone in clinical and interventional decision making.

Methods and results: PubMed was systematically searched for ACS risk score studies. They were divided into ACS studies (evaluating Unstable Angina; UA, Non ST Segment Elevation Myocardial Infarction; NSTEMI, and ST Segment Elevation Myocardial Infarction; STEMI), UA/NSTEMI studies or STEMI studies. The c-statistics of validation studies were pooled when appropriate with random-effect methods. 7 derivation studies with 25,525 ACS patients and 15 validation studies including 257,654 people were formally appraised. Pooled analysis of GRACE scores, both at short (0.82; 0.80–0.89 I.C 95%) and long term follow up (0.84; 0.82–0.87; I.C 95%) showed the best performance, with similar results to Simple Risk Index (SRI) derivation cohorts at short term. For NSTEMI/UA, 18 derivation studies with 56,560 patients and 18 validation cohorts with 56,673 patients were included. Pooled analysis of validations studies showed c-statistics of 0.54 (95% CI=0.52–0.57) and 0.67 (95% CI=0.62–0.71) for short and long term TIMI validation studies, and 0.83 (95% CI=0.79–0.87) and 0.80 (95% CI=0.74–0.89) for short and long term GRACE studies. For STEMI, 15 studies with 134,557 patients with derivation scores, and 17 validation studies with 187,619 patients showed a pooled c-statistic of 0.77 (95% CI=0.71–0.83) and 0.77 (95% CI=0.72–0.85) for TIMI at short and long term, and a pooled c-statistic of 0.82 (95% CI=0.81–0.83) and 0.81 (95% CI=0.80–0.82) for GRACE at short and long terms respectively.

Conclusions: TIMI and GRACE are the risk scores that up until now have been most extensively investigated, with GRACE performing better. There are other potentially useful ACS risk scores available however these have not undergone rigorous validation. This study suggests that these other scores may be potentially useful and should be further researched.

© 2012 Elsevier Inc. All rights reserved.

Keywords:
Acute coronary syndrome
Risk score
Meta-analysis
Prognosis
Percutaneous coronary revascularization
Surgical coronary revascularization
Systematic review

1. Introduction

Acute coronary syndromes (ACS) represent a wide clinical spectrum, ranging from unstable angina (UA) to ST Elevation Myocardial infarction (STEMI). There is heterogeneity of...
The use and development of dedicated scores to discriminate patients at high risk of serious adverse events from low risk ones has been suggested and encouraged by many cardiology expert groups in order to allow accurate therapeutic and diagnostic decision making. [3] Risk assessment remains crucial as the benefits of more aggressive and costly treatments are greatest in patients at higher risk of adverse clinical events [4–6]. Much effort has therefore been put into designing risk scores for ACS patients. The two most commonly used being the Global Registry in Acute Coronary Events (GRACE) [4] and the Thrombolysis in Myocardial Infarction (TIMI) [7] scores. Both are derived from landmark ACS studies and have undergone wide prospective evaluation. More recently other scores have been designed to focus on clinical risk assessment and to improve the selection of patients for clinical and interventional procedures.

Despite the presence of many validation studies confirming the validity of GRACE and TIMI in multiple clinical settings, to our knowledge there has been no meta-analysis to systematically compare their discriminatory performance. We therefore aimed to undertake a systematic review to assess ACS risk evaluation scores in order to determine the most accurately performing.

2. Methods

Current guidelines, including the recent Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) amendment to the Quality of Reporting of Meta-analyses (QUOROM) statement, as well as recommendations from The Cochrane Collaboration and Meta-analysis Of Observational Studies in Epidemiology (MOOSE) were followed during the course of this work [6–10].

2.1. Search strategy


Abstracts from scientific meetings and references of all included studies were also searched and appraised.

2.2. Study selection

Retrieved citations were first screened independently by two unblinded reviewers (GBZ, FDA) at the title and/or abstract level, with disagreements resolved by consensus. The full text of all potentially relevant articles were then fully appraised using the following explicit selection criteria, which were pilot ed over the first 5 studies to ensure consistency and discrimination. Inclusion criteria were (all had to be met for inclusion): (i) Human studies, (ii) Studies investigating patients presenting to hospital with ACS (i.e. UA, NSTEMI and STEMI), (iii) Risk score derivation or validation studies (or both) and (iv) Studies appraising scores using multivariate analysis. Exclusion criteria were (any single one enough for exclusion): (i) Non-human setting, (ii) Duplicate reporting (in which case the manuscript reporting the largest sample of patients with ACS was selected, or if equal, the study with the largest number of overall patients), or (iii) studies reporting only multivariable predictors, without prediction score.

2.3. Data extraction

The following data were independently abstracted by two unblinded reviewers (GBZ, FDA) on pre-specified electronic forms, which were piloted over the first 5 studies to ensure consistency and discrimination, with disagreements resolved by consensus.

Studies were first divided according to ACS clinical presentation i.e. UA, NSTEMI and STEMI or UA/NSTEMI studies or STEMI studies. Information recorded included authors details, journal, year of publication, location of the study group, baseline, angiographic and procedural features, kind of revascularization (fibrinolysis, percutaneous coronary intervention, coronary artery bypass graft), short and long term rates of adverse events (including death, myocardial infarction, revascularization procedures and Major Adverse Cardiac Events; MACE), and risk score, with respective AUC (area under the curve), c-index or c-statistic with 95% confidence intervals. End points evaluated were c-index of the derivation risk scores and their performance when tested in validation cohorts.

2.4. Internal validity and quality appraisal

The quality of included studies was independently appraised by two unblinded reviewers (GBZ, FDA), on pre-specified electronic forms, which were piloted over the first 5 studies to ensure consistency and discrimination, with disagreements resolved by consensus.

Modifying the MOOSE item list in order to take into account the specific features of included studies [8], we separately abstracted and appraised study design, setting, data source and statistical methods for multivariable analysis, as well as, in keeping with The Cochrane Collaboration approach, the risk of analytical, selection, adjudication, detection and attrition bias (expressed as low, moderate, or high risk of bias, as well as incomplete reporting leading to inability to ascertain the underlying risk of bias).

2.5. Data analysis and synthesis

Continuous variables are reported as mean (standard deviation) or median (range). Categorical variables are expressed as n/N (%). Statistical pooling was performed according to a random-effect model with generic inverse-variance weighting and computing c-index of the validation scores with 95% confidence intervals using RevMan 5 (The Cochrane Collaboration, The Nordic Cochrane Centre, and Copenhagen, Denmark). Sensitivity analysis was performed to appraise small study bias by graphically inspecting funnel plots (Figure A, appendix,
three quarters of patients underwent a percutaneous revascularization, with rates of MACE ranging from 4.7% to 11% and of death from 4.2% to 11%. The short term TIMI AUC was 0.66 (95% CI = 0.64–0.68) and 0.73 (95% CI = 0.69–0.78) in derivation and validation cohorts respectively. GRACE short term derivation and validation AUC was 0.83 (95% CI = 0.82–0.84) and 0.82 (95% CI = 0.80–0.89). The Schiele et al. [16] score was the only one to perform similarly, with an AUC in the short term derivation cohort of 0.82 (95% CI = 0.80–0.82). The long term AUC of the GRACE score was 0.84, while for the Zhong et al. [15] score the AUC was 0.81 (95% CI = 0.71–0.86).

3.2. UA/NSTEMI

18 derivation studies [7,11,20,31–45] with 56,560 UA/NSTEMI patients and 18 validation cohorts [18,20,22,24,28,30,32–36,46–52] with 56,673 patients were included. As in ACS studies, validation cohorts included more NSTEMI patients than derivation ones (Table 1), with rates of PTCA ranging from 26 to 48%. Pooled analysis of TIMI validation studies showed an AUC of 0.54 (95% CI = 0.52–0.57) and 0.67 (95% CI = 0.62–0.71) at short and long term. AUC was 0.83 (95% CI = 0.79–0.87) and 0.80 (95% CI = 0.74–0.89) for GRACE validation studies (Figs. 3 and 4). The short term AUC for the Correia et al. [33] study was 0.82 (95% CI = 0.80–0.94) and for AMIS [20] was 0.87 (95% CI = 0.86–0.88). At long term no scores performed better than TIMI or GRACE, apart from several which had combined TIMI and GRACE with other variables.

3.3. STEMI

15 derivation studies [53–67] of 134,557 patients and 17 validation studies [18,20,22,28,30,33,46,57,60,61,68–72] with 187,619 patients were included. Between 71 and 100% of patients underwent PCI, and PTCA (Table 3) was more frequently performed in validation cohorts than in derivation cohorts. AUC (Fig. 6) was 0.77 (95% CI = 0.71–0.83) and 0.77 (95% CI = 0.72–0.85) for TIMI at short and long term, and 0.82 (95% CI = 0.81–0.83) and 0.81 (95% CI = 0.80–0.82) for GRACE at short and long term. At short term, the CADILLAC score [59]
had a comparable AUC as did studies by Chang et al. [54] (0.83; 95% CI = 0.82–0.84), Lee et al. [62] (0.86; 95% CI = 0.84–0.86) and Peterson et al. [65] (0.90; 95% CI = 0.89–0.91). Long term AUC values for APEX AMI [66], PAMI [53], Khan et al. [61], Damman et al. [55] and Urbonaviciene et al. [67] studies were also good. These studies, are all showing comparable performance to TIMI and GRACE scores, are all derivation studies which have not yet been externally validated (Figs. 1–5). Results of studies reporting AUC values inferior to GRACE or TIMI are reported in Table D, Appendix.

4. Discussion

The most important findings of our meta-analysis are: a) There is a striking difference in the rates of patients undergoing invasive revascularization between derivation and validation studies, b) TIMI and GRACE risk scores are the only scores which have been validated in all types of ACS, with the GRACE score performing better, c) many other risk scores, which show good performance in a derivation cohort, have not yet been evaluated in validation cohorts.

Derivation and validation studies evaluated in our review are quite heterogeneous from a methodological point of view. While about half of derivation studies consist of data derived from randomized clinical trials, almost all validation study data came from observational registries, most of them located in Europe and in North America. The application of data from highly selected cohorts to everyday life may undermine the reproducibility of these scores.

Patients in all studies are similar in baseline characteristics, with some differences in the rate of revascularization procedures. This was especially marked in studies evaluating UA/NSTEMI or STEMI alone with higher revascularization rates in validation cohorts. It is worth noticing that, apart from STEMI patients, and despite recent guidelines no more than half of overall patients have undergone percutaneous or surgical revascularization, with an important burden of unfavorable effects, both for short and long term outcomes, as recently demonstrated [73,74]. The effect of these management strategies in derivation cohorts could also affect variables resulting in independent predictors of adverse events, thus underlying the need of new scores using more contemporary databases.

Our work confirms that TIMI and GRACE risk scores are the only ones validated in multiple clinical setting, with GRACE showing a better performance with an AUC around 0.85. In all the studies, the highest AUC values are around 0.85. This is a satisfactory performance when compared to clinical scores for other medical conditions. [75–77].

Table 2

<table>
<thead>
<tr>
<th>Baseline characteristics of UA/NSTEMI studies.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Derivation studies</strong></td>
</tr>
<tr>
<td>[7,11,20,31–45] (18 studies, 56,560 patients)</td>
</tr>
<tr>
<td>Patients</td>
</tr>
<tr>
<td>Male gender</td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>Pts presenting with UA</td>
</tr>
<tr>
<td>Pts presenting with NSTEMI</td>
</tr>
<tr>
<td>Pts undergoing PTCA</td>
</tr>
<tr>
<td>Pts undergoing CABG</td>
</tr>
<tr>
<td>Follow up (days)</td>
</tr>
<tr>
<td>Mace</td>
</tr>
<tr>
<td>Death</td>
</tr>
<tr>
<td>AMI</td>
</tr>
</tbody>
</table>

Fig. 2. Short and long term Area Under the Curve for derivation and validation scores of ACS. For scores other than TIMI and GRACE, only those with a performance better than the latter were reported. (*in hospital and 30 days follow up; ** 185 days (158–234); *** 180 days (180–360)). Heterogeneity for pooled results: Tau² = 0.00; Chi² = 79.53, df = 6 (P<0.00001); I² = 92%.
Baseline characteristics of STEMI studies.

ACS and UA/NSTEMI studies, GRACE AUC is the highest in validation cohorts, both for evaluating short term outcomes and especially long term outcome which has been shown recently to be a challenge. [74] The only notable exception is the performance of TIMI with the addition of proBNP for long term outcome prediction in UA/STEMI [36], which performed better than GRACE. Another interesting finding is that some risk scores (both new scores and scores derived from TIMI with additional variables) show a comparable AUC to GRACE in their derivation cohorts, but without external validation their relevance is limited. For example, for studies evaluating ACS, the Schiele et al. [16] study and the Zhong et al. [15] study offer similar AUC to the AUC of Correia et al. [33] and AMIS [20] for UA/NSTEMI at short term. Interestingly only scores adding gdf15 or cystatin to GRACE [34] or pro-BNP to TIMI [36] perform similarly for patients with UA/NSTEMI. While the first two predictors may be difficult to exploit in everyday clinical practice, the latter could be very useful to guide management of these patients.

On the contrary, for STEMI patients many scores performed with similar AUC for both short term [54,59,62,65] and long term [53,55,61,67,66] outcomes. If this is confirmed in validation studies, they could provide physicians a powerful tool to discriminate high risk patients. This is particularly true for more recently derived scores which include patients treated with the most modern medical and interventional strategies. There were several studies that performed poorly. The reason for this can only be guessed however maybe related to poor selection of patients or to statistical methodology.

The present work has several limitations. We considered only studies that had at least one analysis performed to assess incremental predictive ability. Many other articles reporting only risk factors without a clear evaluation of prediction were excluded, and it is important to remember that empirical evidence in other fields, for example cancer, suggest that new predictors are almost always significant [78]. Moreover, as suggested from visual inspection by funnel plot, no publication bias was reported. (Figure A, appendix, web only figure). Most of the included studies reported a low or moderate risk of selection and attrition bias, while attrition and adjudication were mostly appraised as moderate. (Table A, B and C, appendix, web only tables). Heterogeneity ranged from low to high, thus we performed our analysis with random effect methods; however we also used fixed models, with no effect on AUC.

### Table 3
Baseline characteristics of STEMI studies.

<table>
<thead>
<tr>
<th></th>
<th>Derivation studies</th>
<th>TIMI validation studies</th>
<th>GRACE validation studies</th>
<th>Cadillac validation studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[53–67] (15 studies, 134,457 patients)</td>
<td>[18,20,33,46,57,60,61,68–72] (12 studies, 164,835 patients)</td>
<td>[22,28,30,46,50,51,68,69] (8 studies, 12,204 patients)</td>
<td>[68,69] (2 studies, 1360 patients)</td>
</tr>
<tr>
<td>Patients</td>
<td>2485 (1412–9690)</td>
<td>885 (553–7520)</td>
<td>602 (456–1495)</td>
<td>1033 (900–1770)</td>
</tr>
<tr>
<td>Male gender</td>
<td>74 (67–78)</td>
<td>62 (59–67)</td>
<td>67.5 (62–71)</td>
<td>65.5 (61.5–69)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>62.5 (62–66)</td>
<td>64 (61–67)</td>
<td>65 (61–69.5)</td>
<td>67.5 (62–69)</td>
</tr>
<tr>
<td>Pts undergoing fibrinolysis</td>
<td>100 (0–100)</td>
<td>15.5 (0–64)</td>
<td>74.5 (42.5–100)</td>
<td>36 (29–39)</td>
</tr>
<tr>
<td>Pts undergoing PCI</td>
<td>100 (12–64)</td>
<td>71 (38–100)</td>
<td>71 (58.6–71)</td>
<td>6 (3–9.7)</td>
</tr>
<tr>
<td>Pts undergoing PTCA</td>
<td>31 (25–64)</td>
<td>60.5 (30–72)</td>
<td>9.1 (8.1–10)</td>
<td>36 (29–39)</td>
</tr>
<tr>
<td>Pts undergoing CABG</td>
<td>21</td>
<td>6 (3–9.7)</td>
<td>9.1 (8.1–10)</td>
<td>6 (3–9.7)</td>
</tr>
<tr>
<td>Follow up (days)</td>
<td>30 (30–290)</td>
<td>270 (143–11)</td>
<td>180 (180–315)</td>
<td>270 (143–11)</td>
</tr>
<tr>
<td>MACE</td>
<td>5 (5–10)</td>
<td>5 (3.3–4.5)</td>
<td>3.7 (3.3–5.5)</td>
<td>5 (3.5–4.5)</td>
</tr>
<tr>
<td>Death</td>
<td>6 (4–7)</td>
<td>7.5 (7–12)</td>
<td>6.7 (6.3–14)</td>
<td>7.5 (6.3–14)</td>
</tr>
<tr>
<td>AMI</td>
<td>3 (2–4)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Revascularization</td>
<td>13 (12–20)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>


5. Conclusions

TIMI and GRACE are the risk scores that up until now have been most extensively investigated, with GRACE performing better. There are other potentially useful ACS risk scores available however these have not undergone rigorous validation. This study suggests that these other scores may be potentially useful and should be further researched.

Disclosures

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.cct.2012.01.001.


