PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Development and validation of a prediction rule for patients suspected of acute coronary syndrome in primary care: a cross-sectional study
AUTHORS	Wouters, Loes; Zwart, D.L; Erkelens, Daphne; Adriaansen, Elisabeth; den Ruijter, Hester; De Groot, Esther; Damoiseaux, Roger; Hoes, Arno; van Smeden, Maarten; Rutten, Frans

VERSION 1 – REVIEW

REVIEWER	Steinar Hunskaar Universitetet i Bergen Det medisinsk-odontologiske fakultet, Departmnt of Global Public Health and Primary Care
REVIEW RETURNED	07-Jun-2022

GENERAL COMMENTS	The aim of this study was to develop and validate a symptom-based prediction rule for early recognition of acute coronary syndrome (ACS) in patients with acute chest discomfort who call out-of-hours services (OOH) for primary care. This is an important task, as correct classification of patients with low, medium, or high risk for ACS may be of vital importance. In many countries primary care has a gatekeeping role, usually by a two-step system: First the contact to a call center for a first triage, usually by a nurse or another qualified person. The contact is usually for degree of urgency (time categories), not diagnosis. Then, a doctor (GP or OOH-doctor) examines the patient and decides for referral or not (both degree of urgency and proposed diagnosis). Symptom-based diagnosis of ACS is difficult. Symptom-based prediction rules for diagnosing ACS in general practice and other prehospital settings are scarce. Therefore this study is very welcome. The study is impressive, with 2,192 patients calling for chest pain and/or discomfort. The design is advanced for this kind of study, and data were based on backed up recordings of telephone triage conversations. This is seldom, and secures a higher quality than using written recordings. The authors have performed advanced and adequate analyses in order to develop the prediction rule. I accept the choice of using sex as an interaction term, although many studies show different symptom presentation og ACS by gender. The analyses have resulted in a prediction model with good discrimination and calibration. The authors claim that the prediction rule may replace the existing telephone triage rules for patients with acute chest discomfort in general practice and OOH. Although the study has great value and is well performed, the written paper still has some possible improvements. I also have some comments and suggestions to be considered by the authors. 1. The new prediction rule is intended to eventually replace

	e.g. the Manchester Triage Standard (MTS, in the Netherlands used as the NTS). However, there is a lack of description, results presentation, and discussion about the principal differences between these two approaches (degree of urgency and diagnostic probability). 2. MTS/NTS is described as having poor sensitivity and specificity. The same specifications for the final model should be presented in a comparative way, and also discussed together with the 1% risk level. Over and under triage as concepts could be used more specifically. 3. Also, there is insufficient description on how the new instrument can be used in a triage setting. The MTS has simple questions. What about the new rule? How will it look like? A diagram, a mathematical equation, a program on the computer? 4. The results section shows the clinical presentation of the patients before the prediction rule is calculated. This is both useful and necessary. These results in itself a very important clinical epidemiological basis for risk and probabilities. I presume the data have not been presented before? (ref 17-18-19) 5. The tables need some adjustments and clarification a. ACS by gender and age groups should be presented (per 5 years?) b. There is a lot of double presentation of results between text and tables. This should be sorted out and avoided

REVIEWER	Pieter-Jan Vlaar Catharina Hospital
REVIEW RETURNED	22-Jun-2022

GENERAL COMMENTS	Interesting cross-sectional study on early recognition of acute coronary syndrome patients. My main comments are 1. Only ACS was included as endpoint. As not all patients were sent to an ED and not all patients had follow-up (31,6%), vital status / mortality should also be analysed / included (preferably as combined primary endpoint). 2. Definition of the primary endpoint, in particular NSTEMI should be given (was for example an type 2 infarction excluded, this because coronary spasm was classified as non cardiac LTE). 3. No external vallidation was performed. Minor - An optimal threshold for ACS was investigated, how many of the non cardiac LTE were missed when applied in the current population. - Several POC with high sensitive assays are already commercially avaliable (for example Siemens)

VERSION 1 – AUTHOR RESPONSE

Response to reviewers' comments - manuscript BMJ Open – 2022-064402

* Comments reviewer 1.1: The new prediction rule is intended to eventually replace e.g. the Manchester Triage Standard (MTS, in the Netherlands used as the NTS). However, there is a lack of description, results presentation, and discussion about the principal differences between these two approaches (degree of urgency and diagnostic probability).

Response from author 1.1. We agree with the reviewer this is an important aspect, but nevertheless, both are strongly linked. Degree of urgency is related to the likelihood of certain diseases causing symptoms. Patients with chest pain who are at high risk of underlying ACS receive a high urgency allocation not only because possibly circulatory problems may occur within a short time duration, but also the risk of life-threatening arrhythmias and the opportunity of percutaneous intervention with 'time is muscle' is considered. In other words, the 'urgency approach' is entangled with the 'diagnostic probability approach'.
We have accordingly added the following sentence to the discussion:"However, there is a principal difference between diagnostic probability and categories of urgencies. Risk prediction provides a continuous value of the probability of disease, while urgency level categorisation is based on the interpretation of how risk probability can be translated in urgency, and time within a patient should be seen and treatment delivered."

* Comment reviewer 1.2: MTS/NTS is described as having poor sensitivity and specificity. The same specifications for the final model should be presented in a comparative way, and also discussed together with the 1% risk level. Over and under triage as concepts could be used more specifically.
Response from author 1.2: Although the NTS was derived from the MTS, it is not similar. MTS was created for use at the ED and importantly includes physical items, which explicitly are not incorporated in the NTS which is created for telephone triage.
We agree with the reviewer that it is illustrative to compare the diagnostic accuracy of the NTS with the new prediction rule. We therefore performed decision curve analyses (p6 statistical analyses) to show the net benefit of the final model compared to the NTS (previously appendix figure 2). We also added this figure in the main text as Figure 4. In addition, we created an illustrative table of diagnostic test accuracy for various model-based risk thresholds of the final model (previously appendix-table 8), and added this table to the main text as Table 4. Finally, we incorporated in the main text the visual runway plot of diagnostic accuracy measures of the final model (previously appendix-figure 3, now Figure 5).
We are very hesitant about overtriage and undertriage and therefore try not to use it. It is either used very subjectively when the urgency allocation of triage nurses is compared to that experts, or 'incorrectly' by considering everybody referred who show eventually not to have an ACS or other LTE as overtriage, and the other way around those without an ACS/LTE who received a low urgency as undertriage. In other words, it is based on knowing the diagnosis and thus hindsight-bias. At the triage this information is of course not available. Thus, should a patient 65 years old with acute pressing heavy chest pain for 15 minutes, radiating to the shoulder blades, with sweating not have received a high urgency and direct referral to the ED with an ambulance if at the ED the troponin levels should be normal? Is this overtriage? And is this a correct triage if indeed the troponin levels are increased?
Because we plan to perform an implementation study for the new diagnostic model applying action research, we have added the following sentence:"The development of this diagnostic model is the necessary first step towards an implementation study in which this model is adapted to urgency levels that can be applied by triage nurses during telephone triage at the OHS-PC. "

* Comment reviewer 1.3: Also, there is insufficient description on how the new instrument can be used in a triage setting. The MTS has simple questions. What about the new rule? How will it look like? A diagram, a mathematical equation, a program on the computer?
Authors response 1.3: The model will be 'translated' into simple questions. These will substitute (some of the) key questions now available in the NTS. We agreed with the managing board of the NTS that we will do this first in a pilot using action research in a study called 'Safety First Action study'. We have added in the article: "The diagnostic model needs to be 'translated' in simple yes/no questions that can be incorporated in the existing NTS and a personalized risk prediction for age and gender is generated. Some older questions will then be substituted. We are aiming to do so in an implementation study applying action research."

The results section shows the clinical presentation of the patients before the prediction rule is calculated. This is both useful and necessary. These results in itself a very important clinical epidemiological basis for risk and probabilities. I presume the data have not been presented before? (ref 17-18-19) We thank the reviewer for giving us the opportunity to clarify this aspect. Reference 18 concerns the study design article of the 'Safety First study', describing the applied a multi-methods approach of multiple quantitative and qualitative studies all with the overarching aim of describing, understanding and improving telephone triage of patients suspected of acute cardiovascular diseases. Reference 17 and 19 concern two studies within the same 'Safety First' project, in which a smaller sample size was used for univariable analyses. This is the first study aimed at building a prediction rule with state-of-the-art multivariable analyses.

* Comment author 1.4: The tables need some adjustments and clarification
a. ACS by gender and age groups should be presented (per 5 years?)
b. There is a lot of double presentation of results between text and tables. This should be sorted out and avoided
c. Because of missing data it is very difficult to understand some of the distributions in the tables. I found an explanation in Appendix table 4, but there is a need of a better connection and explanation so that the confusing percentages are understood
d. Page 11 under Medical history: should the table reference be to Table 1 and not Table 2?
Authors response 1.4: We thank the reviewer for these suggestions.
a. We kindly refer to figure 2. In this figure we express the risk of ACS and relation to age and gender as a continuous value. We deliberately choose for risk prediction in a linear way because categorization would lead to loss of information and this we want to avoid. We have added a sentence to highlight this figure "The basic model shows that the risk of ACS increases with age for both sexes, with a notable peak risk for men at an age near 60 years and a more gradual increase in risk of ACS for women (Figure 2)"
b. We have decided to describe the key information from the tables in the text, using the style recommended by BMJ Open. If nevertheless the editor wants to do this otherwise, then we are willing to adjust it accordingly.
c. We agree with the reviewer our way of presentation percentages without absolute numbers has complicated 'checking' the correctness of the data. We therefore have now added the total numbers of the complete data per variable in table 1, appendix-table 4 and appendix-table 5 for clarification.
d. Thank you. We have corrected it to table 1 instead.

*Comment reviewer 1.5: In table 2 a series of p-values are presented. For ACS and LTEs there are many nonsignificant p-values. However, for such important events, the clinical significance based on the absolute results is of outmost interest and should be discussed. I would also think that a 2 x 4 chi squared analysis of ACS would show a significant distribution by gender?
Authors response 1.5: We agree with the reviewer that it is a 'balancing act' to discuss clinical relevance of non-significant differences, notably if the number of outcomes are rather small. But indeed, the differences between males and females for ACS (15.3% vs. 8.3%, $p<0.001$) and other LTEs (3.8% vs. 2.3%, $p=0.04$) is not only significant but seems also clinically relevant, even for the smaller difference selectively in other LTE. We have added the following in the results section: "In total 251 patients were diagnosed with ACS and 65 with other LTEs, and of clinical relevance is that both critical events occurred significantly more in males than females (15.3% vs.8.3%, $p<0.001$ for ACS, and 3.8% vs. 2.3%, $p=0.04$ for other LTEs, respectively). " We decide not to comment on the clinical relevance of the other p-values, because these are such small numbers that significance is more likely to be based on chance than real significance. Therefore a 2x2 table for gender would have in our view no added value.

*Reviewers comment 1.6: The base model with sex and age had an apparent c-statistic of 0.72. The large difference found in the gender-age distribution may have been commented more on, and also

discussed in association with the choice made about sex as interaction term. Did the actual result change the thinking? We agree with the reviewer that it is worthwhile to pay attention to this finding in the Discussion.
Authors response 1.6: We predefined our base model with age by a cubic spline function and sex as an interaction term. The results presented did not change our thinking, because it is nowadays well accepted that methodologically properly prediction models are not built data driven. (ref 27,29). We have added a sentence to the discussion section: “This may largely be explained by the addition of age, the strongest predictor of ACS. This is in line with the notion that the prevalence of ACS increases with age. 7 9 31 Importantly, in our study among people aged below 40, only one (0.4%) male patient had an ACS (UAP). For males to the age of 55 we found a peak risk of ACS of around 20% and remaining at this level with further age increase onwards. For females we found a gradual increase of risk with age with a maximum ACS risk of around 18% for those aged over 80 years.”

*Reviewers comment 1.7: On page 16 there is an important discussion about the consequences of different risk level acceptance. I would suggest that the authors try to conclude with a specific recommendation for the use of their own prediction rule in the setting of telephone triage and also potentially for the GP/OOH doctor.
Author response 1.7: We agree with the reviewer this is the ultimate goal. But we also want to kindly refer to our response of comments 2 and 3.“The development of this diagnostic model is the first necessary step towards an implementation study in which this model is adapted to urgency levels that can be applied by triage nurses during telephone triage at the OHS-PC. “ And; “The diagnostic model needs to be ‘translated’ in simple yes/no questions that can be incorporated in the existing NTS. Some older questions will then be substituted. We are aiming to do so in an implementation study applying action research.”

*Reviewers comment 1.8: The reference list is adequate and contains recent studies. But the list is not edited in accordance with the journal’s formula. The whole list must be scrutinized. As for now, it gives an impression of an unfinished paper.
Author response 1.8: We thank the reviewer and have now all referred studies adjusted to the reference style of BMJ.

* Reviewers comment 2.1: Only ACS was included as endpoint. As not all patients were sent to an ED and not all patients had follow-up (31,6%), vital status / mortality should also be analysed / included (preferably as combined primary endpoint).
Authors response 2.1: We thank the reviewer to give us the opportunity to clarify on this item. We agree there is a potential risk of ‘misclassification’ in patients not referred to the ED. We tried to reduce this problem that is common with prediction research with routine care data by collecting medical information of the patient up to 30-days following the contact with the OHS-PC. Although thus, myocardial infarctions or unstable angina pectoris and other LTE may have been missed, these are likely small numbers and with clinically less critical outcomes. None of the patients died within 30-days of the contact with the OHS-PC. We described in the method section ‘Outcome’ ‘The final diagnoses were retrieved from the patient’s GP, and based on the GP’s electronic medical files which include ED and cardiologist discharge letters and notes from the OHS-PC contact. We used medical information up to 30-days following the contact with the OHS-PC to allow us to include diagnoses of ACS that were initially missed because the patient was not referred to the cardiologist the same day of the OHS-PC contact.’

* Reviewers comment 2.2: Definition of the primary endpoint, in particular NSTEMI should be given (was for example an type 2 infarction excluded, this because coronary spasm was classified as non cardiac LTE).
Authors response 2.2: In table 2 we provide data of the subtypes of diagnosis ACS. We in general do not have information on the distribution of types of myocardial infarction within the NSTEMI category,

e.g. based on epicardial obstruction or MINOCA based on either microvascular disfunction or coronary spasm. Importantly, however, from the triage nurse/GP perspective it is critical to urgently refer as optimal those with an ACS or other LTE. That means irrespective of whether a patient with ACS has a NSTEMI based on an epicardial coronary obstruction or MINOCA based on coronary spasm. In both cases they should be referred urgently to the ED/cardiologist.

*Reviewers comment 2.3: No external validation was performed.
Authors response 2.3: We agree with the reviewer this is a limitation of our study, as described in our limitations. However, we could perform extensive internal-external validation because we had data from nine sites. Because of the differences in prevalence rates of ACS the 'case-mix' differed substantially, but nevertheless our model remained to produce robust outcomes.
Moreover, at the point in time of execution of this study, no other primary care research data set was available to perform external validation. We have added a sentence to our final conclusion; "Although, we performed extensive internal-external validation making use of the datasets of nine sites with substantial differences in case mix, we will strive for formal external validation before it can be widely applied in everyday primary care practice."

VERSION 2 – REVIEW

REVIEWER	Steinar Hunskaar Universitetet i Bergen Det medisinsk-odontologiske fakultet, Departmnt of Global Public Health and Primary Care
REVIEW RETURNED	14-Aug-2022

GENERAL COMMENTS	The authors have adequately revised the manuscript according to my suggestions and the additions in the manuscript are relevant. The authors claim that the references now are adjusted to the BMJ Open style. However, I find errors in most references. The journal names are not correct (program error?) and there are also a number of other errors (capital letters in titles, book references not correct). The editor must decide if this has to be taken in the publishing phase as the authors do not seem to have the skills.