

## **Appendix 1 (as supplied by the authors): Methods (detailed version)**

### **Derivation Cohort**

We conducted a cross-sectional diagnostic study with a delayed-type reference standard<sup>1</sup> in a primary care setting with CHD as our reference condition which included stable CHD and Acute Coronary Syndrome.

We approached 209 Primary Care Providers in the State of Hesse and 74 (35%) agreed to participate in the study. Participating physicians were recruited from a network of research practices associated with the department of family medicine at Marburg University. Participating practices had to consecutively recruit every patient attending with chest pain either as presenting complaint or when asked by the physician. We emphasized the importance of recruiting every patient with chest pain irrespective of the presumed likelihood of CHD. The recruitment period lasted 12 weeks for each practice and was staggered in four waves between October 2005 and July 2006. Every patient above 35 years with pain localized on the anterior chest wall in the area between clavicles, lower costal margins and the posterior axillary lines was to be included. Patients whose chest pains had subsided for more than one month, whose chest pains had been investigated already and/or who came for follow-up of their chest pains were excluded.

This procedure, like the whole study protocol, was approved by the Ethics Committee of the Faculty of Medicine, University of Marburg. The study complies with the Declaration of Helsinki. Primary Care Providers took a standardized history and performed a physical examination according to a case report form that was piloted and modified accordingly. Index tests covered first impression of the patient, duration and temporal patterns of pain, character, localization and associated symptoms, known vascular diseases, risk factors, and relevant findings. Physicians also recorded their preliminary diagnoses, investigations, and management related to the patients' chest pain. Patients were contacted by phone six weeks and six months after the index consultation. Study assistants blinded to results of index tests asked about the course of the patient's chest pain and treatments including hospitalizations and drugs. Discharge letters from specialists and hospitals were requested from Primary Care Providers who were visited at four week intervals to check case report forms, recruitment logs, and compliance with study procedures. Random audits were performed by searching routine documentation of participating practices to identify cases of chest pain not included in the study.

Since the probability for CHD would be low in most patients, we considered an invasive reference standard, e.g. coronary angiography, as ethically not justified. Therefore a reference panel of one cardiologist, one PCP, and one research staff member of the Department of General Practice at the University of Marburg reviewed every patient's data after completed follow-up at 6 months. As a "delayed-type reference standard"<sup>1</sup> they decided on the presence/ absence of CHD at the time of patient recruitment (index consultation). The panel based its decision on all the results available after the follow up period (index questionnaire, the attending physician's provisional diagnosis, coronary angiography, if available, and non invasive tests like ECG, exercise test, or echocardiography). This design is based on the assumption that serious diseases like CHD would manifest themselves within the mentioned time period. CHD diagnosis was based on recommendations of the CHD guideline of the German program for disease management guidelines.<sup>2</sup> Disputes among participants of the reference panel were solved by consulting a second opinion.

The patient's medical history is part of the definition of acute and chronic CHD. However, providing the reference panel with clinical data recorded by Primary Care Providers

would have raised the possibility of incorporation bias. Therefore, the panel judged each patient first without the index test, i.e. blinded to the recorded clinical data including preliminary diagnoses, only using the information gathered at follow-up (blinded reference standard). In a second round, follow-up data were reviewed in a randomly changed order together with history and findings recorded by the attending physicians (unblinded reference standard). In patients who could not or only partly be reached for follow-up we contacted Primary Care Providers for relevant data. If sufficient data were available, the reference panel would still make a decision for those patients. Our analysis excluded trauma cases and patients with insufficient data for a reference decision with regard to CHD (see Figure 1 of the main article). For univariate analyses we calculated sensitivities, specificities, positive and negative predictive values, positive and negative likelihood ratios, and diagnostic odds ratios for all items covered by the case report form.

To arrive at a smaller subset of criteria we selected those index test items that had a  $p$  value  $< 0.05$  and likelihood ratios indicating at least moderate diagnostic accuracy, i.e.  $LR_{+/-} > 2$  or  $< 0.5$ . Those were included as independent variables in multivariable logistic regression analysis. The dependent variable was CHD. Known CHD, cerebrovascular diseases, and peripheral arterial disease were grouped into a combined variable “clinical vascular disease” which was classified “positive” if at least one of the single variables was positive. Variable selection was then conducted using a backward stepwise procedure ( $p < 0.05$ ). Odds of having CHD were compared between patients with and without each risk factor by calculating odds ratio with 95%-confidence intervals. Reliability (goodness of fit) of the model was estimated using the Hosmer and Lemeshow test. Rounding the coefficients of the logistic regression to the nearest unit, an initial score was defined and the area under the Receiver Operating Characteristic (ROC) curve (AUC) calculated. In a sensitivity analysis we simplified the score by stepwise reduction of contributing variables. The guiding principle of this process was to prevent significant changes in the AUC and in the misclassification rate of low risk patients. The final model consisted of 5 different variables (including one compound variable).

The prognostic ability of our score to discriminate between patients with and without disease was assessed by analyzing the AUC. We calculated estimates of diagnostic accuracy (sensitivity, specificity, likelihood ratios, and predictive values) for different cut points.

In order to correct for overfitting, we assessed internal validity by bootstrapping, replicating the whole analysis on 200 different bootstrap samples drawn from the original sample.

All analyses were based on the unblinded reference standard with missing data not replaced.

The robustness of the results was tested in sensitivity analyses with both blinded reference standard and missing data replaced by multiple imputations.<sup>3,4</sup> As one of the predictors, pain reproducible by palpation, was missing in 28% of the patients, a sensitivity analysis was performed by replacing the missing values by means of multiple imputation. Thus, 5 complete datasets were built, and the variation of the 5 imputations reflects the uncertainty with which the missing values can be predicted based on the observed values. The AUCs resulting in each of these 5 were averaged and compared to the real, incomplete dataset. All analyses were performed with SPSS software version 15.1 except the multiple imputation procedure which was performed with NORM Version 2.03 for Windows 95/98/NT.

## Validation cohort

External validation was performed on data of the TOPIC study, a multicenter clinical cohort of primary care patients with chest pain.<sup>5</sup> The primary goal was to study the frequency of occurrence, mode of presentation, etiology, and clinical characteristics, correlates, management, and outcomes of these patients. Patients were recruited from March to June 2001 from a research network of general practitioners in 58 independent urban and rural offices, as well as one primary care outpatient clinic from Western Switzerland. Physicians consecutively enrolled all patients over sixteen years of age who reported chest pain during their visits. The presence of chest pain was ascertained according to the usual practice of each Primary Care Provider. The case report form included 70 questions on history and clinical signs of chest pain. Additional follow-up information was obtained after three and twelve months. All final one-year diagnoses were retained and grouped in six categories including CHD. A group of clinicians discussed the cases in detail where the diagnosis “chest pain” was inconsistent or uncertain. Additionally ten percent of all case report forms were revised by the investigators in order to evaluate the consistency of the final diagnosis.

Four out of the five score variables could be directly derived from the validation data base. For the fifth variable (‘patient assumes cardiac origin of pain’) we used induced anxiety (defined as a positive answer to the question “Are you feeling very worried about your chest pain?”) as a proxy variable. Validation was performed with Stata 7.0 and SPSS 15.1.

## References

1. Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross-sectional study. *J Clin Epidemiol* 2003;56:1118-28.
2. Ollenschlager G, Lelgemann M, Kopp I. The German Program for Disease Management Guidelines: CHD guideline 2006. Short review. *Med Klin (Munich)* 2006;101:993-8.
3. Donders AR, van der Heijden GJ, Stijnen T, et al. A gentle introduction to imputation of missing values [review]. *J Clin Epidemiol* 2006;59:1087-91.
4. Streiner DL. The case of the missing data: methods of dealing with dropouts and other research vagaries. *Can J Psychiatry* 2002;47:68-75.
5. Verdon F, Herzig L, Burnand B, et al. Chest pain in daily practice: occurrence, causes and management. *Swiss Medical Weekly* 2008;138:340-7.