Abstract. Aim: The European Society of Cardiology guidelines for pulmonary embolism (PE) published in 2008 and updated in 2014 recommend a risk stratification including risk scores like Wells and the Geneva score. The utility and practicability of these scores are controversially discussed. Recently, in a trauma cohort and in spinal surgery patients, no correlation between Wells Score and PE diagnosis was found. The aim of the study was the evaluation of Wells and Geneva scores in patients presenting with chest pain, dyspnoea or syncope in an emergency department. Patients and Methods: We retrospectively examined 326 patients suspected of PE, including assessment, according to Wells and Geneva scores. Results: PE was detected in 13.5 %. The average Wells score was 1.0, the average Geneva score 3.9. The receiver operating characteristic (ROC) curve analyses showed for both scores a high significant area under the curve (Wells score 0.68; Geneva score 0.64). The association between the scores and the diagnosis of PE was calculated with logistic regression analysis and showed high significant odds ratios (OR) for both scores (Wells score 1.38; Geneva score 1.24). There was no significant difference between the area under the curve (AUC) of Wells score and Geneva score. Conclusion: The utility of Wells and Geneva scores for the evaluation of patients suspected of PE in an emergency patient cohort.

The problem of reliable detection of pulmonary embolism (PE) in emergency facilities has been the subject of scientific discussion since a long time (1-4). Rapid imaging diagnostics by computed tomography (CT) of the pulmonary arteries seem to be essential (5, 6). The paramount goal of treatment is early anti-coagulation immediately after PE detection (7). This is the background against which the European Society of Cardiology (ESC) guidelines of PE diagnosis and management were implemented in 2008 and updated recently (8). The pivotal element of these recommendations is low-threshold indication for CT, to be made using clinical and laboratory criteria. Clinical risk evaluation is based on scores, such as the Wells and Geneva scores (9, 10). Despite official recommendation, however, these scores are not used universally in emergency medicine. The clinical relevance, in particular of the Wells score, is a matter of discussion. Recently, in a trauma cohort, no correlation between Wells Score and PE diagnosis was found (11). Individual criteria of this score are imprecisely defined and considered to be of little practical use. Moreover, neither score includes important risk factors for PE, such as taking of contraceptives and thrombophilia (12-15).

The objective of the present investigation is to ascertain the significance of the Wells and Geneva scores for the clinical assessment of patients with suspected PE. The work focuses specifically on the question of correlation of these scores with the probability of the diagnosis of PE in patients presenting with chest pain, dyspnoea or syncope in an emergency Department.
Materials and Methods

Study population. In our study, the medical records of 326 patients suspected of PE and positive D-dimer who had presented in the emergency department of the University Hospital of Mannheim in the period from April 2010 to July 2011 were retrospectively analysed. The study included hemodynamically stable patients with symptoms of chest pain, dyspnoea or syncope (16-18). The standard comprises special anamnesis of risk factors and all clinical parameters of Wells and Geneva scores. D-dimer, as a parameter of acute coagulation activation, was routinely determined. Pulmonary artery CT was performed in all cases of any of the mentioned symptoms and positive D-dimer test. CT was considered contra-indicated in case of pregnancy, contrast medium allergy, higher-grade renal insufficiency (creatinine >1.5 mg/dl), manifest hyperthyreosis or current metformin therapy. The determination of D-dimer (Tina-Quant D-dimer assay by Roche Diagnostics, Mannheim, Germany) was based on a reference range of 0-0.5 mg/l. D-dimer levels >0.5 mg/l were considered as suspicious with regard to pulmonary embolism.

Statistical methods. The suitability of Wells and Geneva scores for diagnosis of PE was evaluated by receiver operating characteristic curve (ROC) analyses. To measure which of the scores is the best to diagnose PE, the areas under the curves (AUC) of these scores were compared by the method of Hanley et al. (19). In addition, to evaluate the association of both scores with PE, logistic regression analysis was performed using PE as the dependent and the score as independent variable. The calculations were performed with the IBM SPSS 22 statistical software (IBM SPSS Inc., Chicago, IL, USA). Continuous variables are expressed as mean±SD (standard deviation) and range (min-max). For non-continuous variables numbers, percentages (%) were used.

Results

The baseline characteristics (age, gender, blood pressure, heart rate and oxygen saturation) are listed in Table I. PE was diagnosed in 44 of 326 cases (13.5 %). D-dimer testing showed a mean value of 2.98 mg/l ±4.76 (range=0.19-34.37). The average Wells score was 1.0±1.9 (range=0.0-11.0), the average Geneva score 3.9±2.7 (range=0.0-16.0). Table II shows the definition of the clinical probability of PE (high-intermediate-low) in Wells and Geneva scores. The usefulness of Wells score to differentiate patients with PE from those without was determined with ROC curve analyses (Figure 1). The AUC was 0.68 (95 % confidence interval (CI)=0.58-0.77; p<0.001). The association between Wells score and the diagnosis of PE was calculated with logistic regression analysis (Odds ratio (OR)=1.38; 95% CI=1.2-1.6; p<0.001). The usefulness of Geneva score for PE diagnostics is described by ROC curve analyses in Figure 2. The AUC was 0.64 (95% CI=0.54-0.73; p=0.004). The logistic regression analysis for calculating the association between Geneva score and PE resulted an OR of 1.24 (95% CI=1.1-1.4; p<0.001). There was no significant difference between the AUC of Wells score and Geneva scores (p>0.05).

Discussion

Under the ESC guidelines, for the first time internationally, valid standards for reliable detection of PE have been created (8).The Wells and Geneva scores form an essential part of the clinical assessment in the context of these recommendations (9, 10). However, the significance of these scores for the risk evaluation of patients with suspected PE is being questioned. The objective of the present study was to investigate the Wells and Geneva scores in a cohort of emergency patients with suspected PE in more detail with regard to their correlation with the probability of the diagnosis of PE.
Comparison of the two scores in detail shows, as already mentioned, that in both cases the parameters used include some very precisely defined ones (heart rate, risk factors) and also some whose definition is very vague. This concerns in particular the criterion of “alternative diagnosis less likely than PE” from the Wells score, which leaves ample scope for the user, thereby limiting the practicability of the score. Comparison of the scores using the data now collected reveals that patients with suspected PE on average receive higher ratings in the Geneva than in the Wells score. In particular, the group at intermediate probability is much larger according to the Geneva score – at the expense of the low-probability group – than according to the Wells score. This is due, on the one hand, to the fact that the criterion of age >65 years exists only in the Geneva score and is particularly often met here. On the other hand, the heart rate ranges in the Geneva score are lower than those of the Wells score and, thus, become relevant more often. In general, the specific assessment of the heart rate factor matches the data in literature that describe tachycardia in terms of hemodynamic compensation as a common symptom, occurring in about 25 % of all PE cases (20). However, our own data show that primary tachycardia, such as atrial fibrillation, is common in patients suspected of PE as well occurring at a rate of 10 %; thus, the value of the criterion of tachycardia for PE diagnosis is not entirely clear (21). The number of high-probability patients, which is, however, conspicuously low in both scores in comparison with the other risk groups, with values between 2 and 3 %, is comparable among the scores.

Considering the Wells score, in terms of its suitability for prediction of PE diagnosis, our data show a highly significant AUC of 0.68. Our logistic regression analysis also found a highly significant association between the Wells score and the diagnosis of PE (OR=1.38). From the statistical data, it may, therefore, be concluded that the Wells score correlates very strongly with the likelihood of PE diagnosis. The same applies to the Geneva score. Herein, too, a highly significant AUC (0.64), as well as a highly significant logistic regression analysis (OR=1.24), were found. Thus, the data show good correlation with the likelihood of PE diagnosis for the Geneva score as well. Hence, despite all the discussions regarding their suitability, both the Wells and the Geneva score generally appear useful for PE diagnosis in a classic emergency cohort. The two scores also show – despite all differences in detail – comparable AUC and are, therefore, to be regarded as fundamentally equivalent.

As mentioned above, the assessment of risk scores in the context of PE diagnosis is controversial. This is particularly true for the Wells score. For example, the previously cited article by Young et al. found no correlation between Wells score and PE detection by CT in a trauma cohort (11). In this study, the rate of detected PE was 16 % and the average Wells score between 3.2 and 3.5. These values were, thus, higher than those in our collective. A further study by Wang et al. in spinal surgery patients found no correlation between the Wells score and PE diagnosis either (22). In this population, the rate of detected PE was 28.8 % and the average score between 4.9
and 5.3. These figures were, thus, likewise significantly higher than those in our collective. By contrast, another study by Yap et al. concludes that the Wells score plays an important role for risk stratification of patients with suspected PE (23). As opposed to the surgical patients above, the cohort of this study was a classic emergency collective. However, here, PE was demonstrated not by CT but by scintigraphy, which is considered as a less reliable method (8). The rate of PE detected was lower in this study (9 %) than in our collective.

The data suggest that the patient collectives of the studies differed considerably in terms of their characteristics and, obviously, also in their amenability to assessment using the Wells score. In this context, particularly, surgical collectives seem to tend to higher scores. This may be interpreted as an indication that this group of patients are more prone to non-specific symptoms that may nonetheless be score-relevant. Such a dependence of the significance of the risk scores named on the respective patients cohort should be taken into account, in particular in the application of the Wells score. Therefore, despite of our own positive evaluation, uncritical application of Wells and Geneva scores for the PE diagnostics is certainly not expedient.

References