Surgical Intervention in Patients with Tubo-Ovarian Abscess: Clinical Predictors and a Simple Risk Score

Yuval Fouks, MD, Aviad Cohen, MD, Udi Shapira, MD, MPH, Neta Solomon, MD, Benny Almog, MD, and Ishai Levin, MD, BPharm

From the Lis Maternity Hospital, Department of Obstetrics and Gynecology, Sourasky Medical Center, Tel Aviv, Israel (Drs. Fouks, Cohen, Solomon, Almog, Levin), and Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel (Dr. Shapira).

ABSTRACT Study Objective: To identify the clinical characteristics associated with surgical intervention in patients with tubo-ovarian abscess (TOA).

Design: Retrospective cohort study (Canadian Task Force classification II-2).

Setting: Tertiary university-affiliated hospital.

Patients: Three hundred thirty-five patients were diagnosed with TOA based on sonographic and clinical criteria. Patients who underwent surgical intervention were compared with patients managed conservatively.

Intervention: Electronic medical records were used to identify patients who were diagnosed with TOA between 2007 and 2015. All patients received the same antibiotic regimen upon admission. The data extracted included microbial and pathologic reports. Clinical characteristics such as disease severity and outcomes were compared. The clinical predictors available on patient admission for surgical intervention were identified retrospectively. A logistic regression was used to determine the independent predictors of treatment failure. A risk score was created by giving a nominal weight to each predictor. The score was validated by a random bootstrap analysis. An additional validation cohort that consisted of patients diagnosed with TOA during the 2 years after the original study period was applied to the final score.

Measurements and Main Results: The following variables of patients who underwent surgical intervention in comparison with those successfully treated and were enrolled into the score analysis differed significantly: age at admission (odds ratio [OR], 2.1; 95% confidence interval [CI], 1.3–3.5), mean leukocytosis at admission (OR, 2.2; 95% CI, 1.3–3.6), ultrasonographic measurement of abscess diameter (OR, 3.6; 95% CI, 2.0–6.3), and the presence of bilateral abscess (OR, 2.2; 95% CI, 1.3–3.9). Risk groups A, B, C, and D were positively correlated with the need for surgical intervention. Those in the highest risk group D had an antibiotic failure rate of 92%, as compared with those with the lowest risk group, in which there was a 20% risk of antibiotic failure.


Keywords: Antibiotics; Culture; Risk score for tubo-ovarian abscess; Pelvic inflammatory disease; TOA

Tubo-ovarian abscess (TOA) is an infectious mass of the adnexa that forms as a late complication in as many as 30% of pelvic inflammatory disease (PID) cases [1,2]. It is associated with significant patient morbidities, such as pain, sepsis, and recurrent hospitalizations [3], and may lead to long-term complications, such as pelvic adhesions, infertility, ectopic pregnancies, and chronic pelvic pain [4,5]. Some of these complications mandate patient admission and intravenous antibiotic treatment upon diagnosis. Studies have shown that treatment with broad-spectrum antibiotics is effective in 34% to 87.5% of the women affected with TOA [6,7]. Surgical intervention is needed whenever antibiotic treatment is inadequate and in patients who present with suspected TOA rupture [6,8]. The interventional measures indicated in cases of antibiotic treatment failure are performed via either laparoscopy, laparotomy, or imaging-guided drainage [9,10].

Although TOA is not uncommon and carries potential long-term implications, risk assessment tools with which to estimate a patient’s relative risk for conservative treatment failure are not in common practical use. Studies have discussed the associations between antibiotic
failure in cases of TOA using several patient demographic as well as clinical characteristics [11–13]. Several studies have shown that TOA size was significantly associated with and inversely proportional to the need for surgical treatment, and patient age at admission was positively proportional to the need for surgical treatment in another study [14,15].

The attempt to construct a clinical risk score for TOA conservative treatment failure is based on the assumption that the prognosis of TOA may depend on the extent of the disease on diagnosis and on the clinical response to antimicrobial treatment during the first days after admission [9,10,16]. The early recognition of a patient’s risk for antibiotic treatment failure may assist clinicians in deciding to change treatment strategy and reduce patient morbidity and long-term complications in TOA. The aim of the present study was to determine the clinical characteristics associated with the need for surgical intervention by using a simple clinically based risk score system.

Methods

Study Design and Study Population

A retrospective study was performed in a single, tertiary, university-affiliated medical center. The study was approved by the local institutional review board.

The electronic medical records of patients admitted with complicated PID or suspected TOA between January 1, 2007 and December 31, 2015 were reviewed. A diagnosis of TOA was made in patients who fulfilled the PID criteria in accordance with Centers for Disease Control and Prevention criteria [16] and additionally exhibited ultrasonographic indications of TOA (sonographic evidence of TOA or tubo-ovarian complex). Patients were excluded due to uncertainty in the diagnosis of TOA: 136 patients were excluded who were diagnosed with TOA/suspected PID with no clear-cut diagnosis of TOA at discharge. Twenty-six patients underwent diagnostic laparoscopy with no evidence of TOA, 6 had a final diagnosis of corpus luteum, 7 had endometrioma, 3 had space-occupying lesions (in pathology reports), and 2 had primary appendicitis.

Clinical outcomes, such as disease characteristics, treatments, and the short-term outcomes of patients, were compared between patients who underwent surgery and patients who were successfully managed with antibiotics. Data of interest obtained included patient demographics and medical histories (including known risk factors for PID, such as marital status, recent medical or surgical procedures, such as ovum pickup during fertility treatments, dilatation and curettage, intrauterine device insertion), and past pelvic infection, patient vital signs, blood panel including inflammatory markers, and the results of blood, urine, cervicovaginal swabs and abscess cultures. The sonographic dimensions of the complex were measured and described in terms of the largest diameter of the tubo-ovarian mass.

Additional data included pathologic findings, antibiotic treatment adequacy, length of hospital stay, and need for surgical intervention.

Treatment

Antibiotic treatment was initiated on admission [17] after obtaining cervical, vaginal, intrauterine device, and blood cultures. The antibiotic treatment protocol in patients admitted for TOA during the entire study period was uniform and was followed in all cases. The antibiotic regimen included intravenous third-generation cephalosporin (ceftiraxone) at a dose of 1000 mg once a day, intravenous metronidazole at 500 mg 3 times a day, and oral doxycycline at 100 mg twice a day [18] (in cases of cephalosporin or metronidazole allergies, clindamycin at 900 mg was given 3 times daily). During study period there were no registered cases of allergy preventing the use of doxycycline. Adequate treatment was defined when susceptibility of the pathogenic isolates to at least 1 of the antibiotics administered was proven. Inadequacy was defined as resistance to all 3 drugs administered. In cases of isolates sensitive to the initial empirical treatment, a wide spectrum of coverage was still continued without omitting antibiotics. Antibiotic treatment failure was defined in cases when the patient did not improve clinically despite treatment within 72 hours [19] and/or there was an acute clinical deterioration with clinical evidence of sepsis or ruptured abscess and generalized peritonitis. In cases of inadequate treatment, the initial empiric treatment was upgraded to cover the specific resistant isolate.

During the study period our default surgical intervention for patients who failed antibiotic treatment was laparoscopic abscess drainage. A transvaginal ultrasound–guided drainage of abscesses was performed in selected cases of patients with a history of multiple abdominal interventions or extensive pelvic endometriosis. Treatment success was considered when patients had clinical resolution of symptoms, no clinical illness, apyretic state, and a sonographic evidence of disease regression.

Statistical Analysis

Descriptive statistics are expressed as number with percentages, mean and standard deviation (SD), or median and interquartile ranges. Student t tests were used to compare normally distributed parametric variables between the groups. χ² tests and Fisher’s exact test were used for categorical variables with small samples. All tests were 2-sided, and a p < .05 was considered significant.

Risk Score Construction

A receiver operating characteristic (ROC) curve was used to determine the cut-off, sensitivity, and specificity values for patient clinical characteristics with regard to the need for surgical treatment [20]. The discrimination of the
model, risk factors, and risk scores was assessed by means of the area under the ROC curve (AUC). The optimum sensitivity and specificity (Youden index [J]) of the variables were compared with the positive and negative likelihood ratios of several putative cut-off values. By using coordinates from the ROC, we choose putative diagnostic thresholds in addition to the maximum value point. Because all patients affected with TOA are hospitalized, we decided to maximize the test’s specificity at the expense of the test’s sensitivity. The cut-off value for age was rounded to the nearest whole integer of the maximum value point. A binary logistic regression analysis was performed for all significant variables in the initial univariate analysis. The Hosmer-Lemeshow goodness-of-fit statistic was used to assess the reliability of the models, where $p > .05$ indicated suitable calibration [21]. Each predictor in the final model was assigned a point value that corresponded to its odds ratio (OR) [22]. The points were computed as the nearest rounded whole integer of the selected predictors OR than divided by 2. The points were summed to create a risk score for each participant. The cohort was divided into 4 risk stratification groups, which were correlated with rates of failed antibiotic treatment.

**Validation of the Score**

The validation of the models was performed first with the derivation cohort, applying a bootstrap resampling with random computer-generated iterations of the study derivation cohort [23]. Additional validation was made by using a recent local dataset (validation cohort). The validation cohort consisted of patients admitted for TOA from January 1, 2016 to September 31, 2017. After applying the same scores to the validation cohort for each predictor, we examined the calibration of the risk categories by calculating the observed score versus the number of patients who experienced failed treatment in each risk class. We then compared the AUC of the model containing the risk score with the validation cohort. The predicted probability of the outcome for each member of the cohort was calculated on the basis of the final predictor model. All data analysis was performed using SPSS software Version 24 (IBM, Armonk, NY).

**Results**

**Incidence**

During the study period 515 patients were diagnosed with PID that met the hospitalization criteria [16] (Fig. 1). Of these, 335 patients (65%) received a final diagnosis of TOA. One hundred sixty-seven patients (49.8%) underwent surgical intervention subsequent to failed antibiotic treatment (surgical intervention group). The remaining 168 patients were treated with intravenous antibiotics only (conservative treatment group).

**Clinical Characteristics**

Table 1 presents the demographic and clinical characteristics of the surgical intervention group and conservative treatment group. Of note, mean white blood cell (WBC) count at admission and mean C-reactive protein (CRP) at admission were significantly higher in the surgical intervention group ($15.7 \times 1000/mm^3$ [± SD 6.0] vs $12.8 \times 1000/mm^3$ [± SD 6.2], $p = .0001$, and $111.2 \text{mg/L} [± \text{SD 82.3}]$ vs $85.2 \text{mg/L} [± \text{SD 90.2}]$, $p = .0009$, respectively).

**Microbial Characteristics and Adequacy of Treatment**

Seventy-one (21.1%) cervicovaginal and blood isolates were positive at time of admission. The common cultured isolates are shown in Supplementary Table 1. Empiric treatment administered at time of admission was adequate in 79.7% of cases ($n = 59$). Positive cultures of bacteria resistant to empiric treatment were significantly more common in the surgical intervention group as compared with the conservative treatment group ($34.2\%$ $[n = 12]$ vs $10\%$ $[n = 3]$, $p = .04$).

**Ultrasonographic Characteristics of TOA**

The rate of bilateral abscess was significantly higher in the surgical intervention group ($35.9\%$ $[n = 60]$ vs $18.5\%$ $[n = 31]$, $p = .0003$). In addition, the mean largest diameter of the abscess measured on admission was significantly higher in the intervention group as compared with the conservative treatment group ($66.3\text{ mm} [± \text{SD 21.1}]$ vs $47.4\text{ mm} [± \text{SD 20}]$, $p = .0001$).

**Surgical Treatment**

During the study period 167 patients underwent surgical interventions. Laparoscopy was performed in 89.8% ($n = 150$) of the surgical cases, of whom 60 underwent drainage only and 94 patients underwent salpingectomy. Thirteen cases of oophorectomies/partial oophorectomies were identified; 5 of them were the result of complicated operations/extensive TOA with no clear pelvic anatomic landmarks or technical inability to spare the ovaries. The median time elapsed from patient admission to surgical treatment was 3 days (interquartile range, 1–4). Nine patients (5.3%) underwent surgical intervention within the first 24 hours after admission, all of whom underwent laparotomy. Surgical indications for laparotomy were deteriorating clinical/hemodynamic state or the presence of extensive pelvic disease. Ultrasonographic or computed tomography-guided needle drains were performed in 4.1% of patients ($n = 7$); of these, 2 patients ultimately required laparoscopy. Surgical conversion from laparoscopy to laparotomy occurred in 4.1% of patients ($n = 7$). The hospitalization period was significantly longer in the surgical intervention group as compared with the conservative...
treatment group (7.2 days [± SD 4.0] vs 3.7 days [± SD 3.5], p = .0001).

**Predictors and Risk Score**

The diagnostic thresholds and the maximum value point of ROC, are shown in Table 2. The cut-off value for age was rounded to >35 years, which is the nearest whole integer of the maximum value point. To maximize the test’s specificity, the cut-off value for TOA largest diameter and mean WBC count at admission were ≥70 mm and ≥16 (x 1000/mm³), respectively.

A binary logistic regression analysis was performed using the variables found to be significant in the univariate analysis and their cut-off values based on ROC. As demonstrated in Table 3, 5 variables remained significantly associated with failed antibiotic treatment within the model: age at admission (OR; 95% confidence interval [CI], 2.1, 1.3–3.5), inadequate antibiotic treatment (OR, 4.0; 95% CI, 1.0–16.0), mean leukocytosis at admission (OR, 2.2; 95% CI, 1.3–3.6), the largest diameter of the abscess (OR, 3.6; 95% CI, 2.0–6.3), and the presence of bilateral abscess (OR, 2.2; 95% CI, 1.3–3.9). This model’s performance was statistically good (p = .9) based on the Hosmer-Lemeshow test of goodness-of-fit. For practical use we included only tests that were available to the clinician at the time of admission. Thus, only 4 variables in the regression were ultimately used as predictors. The included predictors were assigned scores in correlation with their relative weights and were calculated as the sum of the predictors’ weights as follows: risk score = [(1 x age (>36 years)) + (1 x mean WBC count (≥16 x 1000/mm³)) + (2 x abscess diameter (≥70 mm)) + (1 x bilateral abscess)]. The total scores for each individual ranged from 0 to 5. As illustrated in Table 4, scores of 1 to 2 and 3 to 4 were merged due to their relatively similar rates of antibiotic failure. The scores were used to define 4 risk groups: A = score of 0, B = score of 1 to 2, C = score of 3 to 4, and D = score of 5. The risk of failed antibiotic treatment was positively correlated with higher scores. A ROC curve was constructed to assess the risk scores’ correlation with final outcomes. The AUC of the risk score sum was .73 (95% CI, .77–.68).

**Score Validation**

The risk score was first validated via a random bootstrap resampling. One thousand randomly generated bootstrap
samples were drawn from the derivation cohort. With each bootstrap sample we modeled the prediction of the risk score given the selected predictors. The final risk score was validated by a recent cohort of all patients affected by TOA (the validation cohort) who met the same inclusion criteria and were hospitalized from January 1, 2016 to September 31, 2017 (Fig. 1). The validation cohort consisted of 92 patients with similar demographic and clinical characteristics to the study group. Patients in the validation cohort were managed with the same protocol and antibiotic regimen as the derivation cohort. The surgical intervention rate in the validation cohort group was 40.0% (n = 37) and did not statistically differ from the derivation cohort (p = .17). All 4 established predictors were extracted from the electronic medical files of the new validation cohort.

Table 1

<table>
<thead>
<tr>
<th>Diagnosis at admission</th>
<th>Total</th>
<th>Conservative treatment only</th>
<th>Underwent surgery</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;34.2*</td>
<td>233</td>
<td>54.5 (127)</td>
<td>75 (69.3–82.6)</td>
<td></td>
</tr>
<tr>
<td>&gt;35</td>
<td>220</td>
<td>55.9 (123)</td>
<td>68.8 (60.6–75.2)</td>
<td></td>
</tr>
<tr>
<td>&gt;40</td>
<td>155</td>
<td>64 (96)</td>
<td>55.7 (47.8–63.3)</td>
<td></td>
</tr>
<tr>
<td>&gt;45</td>
<td>95</td>
<td>64.2 (61)</td>
<td>34.7 (27.9–42.1)</td>
<td></td>
</tr>
<tr>
<td>Largest TOA diameter, mm</td>
<td>57.5*</td>
<td>210</td>
<td>61.9 (130)</td>
<td>75.1 (69.7–83.2)</td>
</tr>
<tr>
<td>Mean WBC count, × 1000/mm³</td>
<td>13.25*</td>
<td>210</td>
<td>61.9 (130)</td>
<td>75.1 (69.7–83.2)</td>
</tr>
</tbody>
</table>
| *Inflexion point of the ROC (maximum = Sensitivity + Specificity – 1).
shown in Fig. 2. The scores for the validation cohort showed similar rates of antibiotic failure at each point as compared to the derivation cohort. Comparing the derivation and validation cohorts did not reveal significantly different rates of surgical intervention. A ROC curve was constructed to assess the risk scores' correlation with the final outcomes in the validation group. The AUC for the validation cohort was .75 (95% CI, .66–.85) (Fig. 3). The 4 risk groups, A, B, C, and D, were positively correlated with the need for surgical intervention in the derivation and validation cohorts (Fig. 4).

Discussion

In this retrospective cohort study, data from 425 women affected with TOA were analyzed. We created a simple model based on 4 clinical and sonographic features that were available at time of patient admission: age, WBC count level, largest diameter of TOA, and the presence of bilateral abscess on admission. Using these variables, clinicians may more adequately predict an eventual antibiotic failure and provide patients with prognostic estimation and recommendations regarding the need for surgical intervention in the days to come.

In previous studies, patient clinical characteristics and pelvic ultrasonographic findings have been identified as associated with or having the potential to predict failed antibiotic treatment in patients affected by TOA [10–15]. In our study abscess diameter was the strongest predictor of failed antibiotic treatment. This notion is supported by previous studies reporting that patients with larger TOA may have higher rates of surgical management [24,25]. The mean age of patients affected with TOA who underwent surgery was significantly higher in the intervention group than in the conservative treatment group. Increased patient age was previously reported to be associated with an increased rate of surgery [12]. The higher rate of reported surgical intervention in older patients may be partially explained by the potential clinical comorbidities in older age and in a more liberal attitude toward potential need for surgery that may lead to oophorectomy.

We found that bilateral TOA was an independent risk factor for antibiotic treatment failure. This sonographic feature was examined in previous smaller studies and was not found to be significantly associated with the need for surgical intervention [11].

Inflammatory markers, such as WBC counts and CRP levels at admission, were reported to be higher in patients that eventually required surgery [15]. In the present study, mean WBC counts and CRP levels at admission were significantly higher among patients in the intervention group. Mean CRP levels did not add any additional predictive value beyond that of mean WBC count in the regression analysis and were thus excluded from our final score.

The mean body temperature on admission and incidence of fever (<38°C) did not differ significantly between the groups in our study. These data are supported by the similar results seen in one recent study [26], whereas in another study patients with fever at admission ultimately underwent more surgical treatments [25]. In our study the presence of resistant

<table>
<thead>
<tr>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariable predictors and predictors weights</td>
</tr>
<tr>
<td>Predictor</td>
</tr>
<tr>
<td>Age &gt; 35</td>
</tr>
<tr>
<td>Largest diameter (≥70 mm)</td>
</tr>
<tr>
<td>Bilateral abscess</td>
</tr>
<tr>
<td>Mean WBC count &gt;16, × 1000/mm³</td>
</tr>
<tr>
<td>Mean CRP, mg/L</td>
</tr>
<tr>
<td>Inadequate antibiotic treatment</td>
</tr>
</tbody>
</table>

aCI = adjusted confidence interval, n/r = not relevant.

<table>
<thead>
<tr>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score points and risk groups</td>
</tr>
<tr>
<td>Score points</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>
bacteria was significantly and positively associated with the need for surgical intervention. The observation that a single pathogen linked to worse prognosis was not easily explained in light of the polymicrobial characteristics of TOAs and the high rates of pathogens considered as contaminant [27,28].

Our goal was to create a simple, easy-to-use scoring tool using predictors that are widely accepted. To our knowledge, no prior attempt has been made to create a risk index for predicting antibiotic treatment failure in cases of TOA. We chose a binary logistic regression model to simulate the first initial clinical decision based on the information obtained at the gynecological emergency department. When we reran the analyses using a decision tree algorithm, the same predictors were identified.

Once validated with other populations, our risk score model may have other applications, in addition to its prognostic value, such as indicating risk for use in cost-to-benefit estimations and the detection of populations prone to future infertility. As mentioned above, medical treatment alone was shown to be effective in 34% to 87.5% of women with TOA. It is accepted that 70% of TOA cases are resolved with the first antibiotic trial. Our intervention rate during the study period was apparently higher than that in previously reported studies. We believe this difference is due in part to prudent exclusion criteria regarding patients who were managed expectantly. Cases without clear-cut ultrasound diagnoses of TOA were excluded. As a result, the intervention group was relatively larger. Moreover, many patients in the study group had undergone prior inpatient medical or surgical treatments elsewhere and were later referred or readmitted to our tertiary medical center. Therefore, we believe that the percentage of more complicated/unresolved TOA cases in our cohort was relatively

\[\text{Fig. 2}\]
Antibiotic failure, derivation vs. validation cohorts.

\[\text{Fig. 3}\]
AUC of the derivation and the validation cohorts. (A) AUC = .73, 95% CI, .68-.77. (B) AUC = .75, 95% CI, .66-.85.
higher than in previous studies. The strengths of our study include its relatively large number of outcome events and participants recorded over more than 10 years. One additional strength is the fact that the antibiotic protocol was identical for all patients during the study period.

However, a number of methodologic limitations must be considered: First, its retrospective nature and that fact that it is a single-center study may limit the generalization of our results. Second, our laboratory does not routinely perform polymerase chain reaction for gonorrhea or chlamydia in patients admitted with a diagnosis of PID. However, these pathogens are targeted by the routine empirical antibiotic treatment administered, which includes third-generation cephalosporin and doxycycline. Furthermore, it is accepted that none of these pathogens is typically isolated from TOA cases. Third, surgical approach in women with TOA can range from interventional radiology guided drainage, laparoscopy, and laparotomy. Guided drainage is considered as the preferred approach by many gynecologists. Because the latter is not the method of choice in our department, this study lacks data regarding the role of guided needle drainage in cases of either antibiotic failure and or first-line treatment with antibiotics. However, laparoscopic approach was shown to be safe with an improved outcomes as compared to laparotomy [29]. Currently, data comparing interventional radiology guided drainage and laparoscopy in patients with TOA subsequently to antibiotic failure are lacking and the optimal approach for management of TOA is still debatable [30]. We believe that the uniform surgical treatment protocol over 10 consecutive years represents a substantial methodologic strength in our study. Nevertheless, the clinical question we addressed is whether antibiotic failure can be predicted. As such, reaching a decision for a given intervention by any means would have the same impact on clinical decision-making. Fourth, we could not externally validate the prediction score for TOA because of the lack of a previous cohort with similar predictors and cut-off values. Thus, we decided to validate our prediction score by using local recent retrospective data and then additionally validate the model via bootstrapping. To make our prediction model applicable to TOA-affected populations, we call on other gynecologic units to implement this score to further validate it in other populations.

In summary, we created a risk score for use to predict the probability of surgical intervention in cases of TOA treated with intravenous antibiotics. This model has potential clinical utility and should be validated in other populations.

**Supplementary materials**

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jmig.2018.06.013.

### Supplementary Table

<table>
<thead>
<tr>
<th>Score</th>
<th>Derivation cohort</th>
<th>Intervention rate (%)</th>
<th>Validation cohort</th>
<th>Intervention rate (%)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>61</td>
<td>20</td>
<td>21</td>
<td>17</td>
<td>.5</td>
</tr>
<tr>
<td>B</td>
<td>182</td>
<td>45</td>
<td>50</td>
<td>40</td>
<td>.3</td>
</tr>
<tr>
<td>C</td>
<td>78</td>
<td>77</td>
<td>18</td>
<td>73</td>
<td>.3</td>
</tr>
<tr>
<td>D</td>
<td>13</td>
<td>92</td>
<td>3</td>
<td>100</td>
<td>.6</td>
</tr>
</tbody>
</table>

**Fig. 4**

Prediction model of performance in score categories in the validation cohort.
References