Development of a scoring system using a statistical model to predict cure status in patients with cutaneous leishmaniasis

Mehri Khoshhali1,2, Sayed Mohsen Hosseini3, Mohammad Ali Niforoughzadeh1, Fariba Jaffary4, Azadeh Zolfaghari Baghbaderani4
1Skin and Stem Cell Research Center, Tehran University of Medical Sciences, Tehran; 2Department of Biostatistics and Epidemiology, School of Public Health, Isfahan University of Medical Sciences; 3Skin Diseases and Leishmaniasis Research Center, Department of Biostatistics and Epidemiology, Isfahan University of Medical Sciences; 4Skin Diseases and Leishmaniasis Research Center, Isfahan University of Medical Sciences, Isfahan, Iran

INTRODUCTION

Leishmaniasis is an infectious disease caused by the protozoa of Leishmania species, which is transmitted by a female sandfly bite.[2,3] Leishmaniasis is classified into three groups including cutaneous, mucocutaneous, and visceral. Cutaneous leishmaniasis (CL) is the most common form. About 1.5 million new cases of CL occur per year and more than 90% of them are observed in seven developing countries including Iran, Afghanistan, Syria, Saudi Arabia, Brazil, and Peru.[2-4] CL causes lesions on the exposed parts of the body. These lesions are usually painless but can become painful if they become secondarily infected. Most lesions develop during a few weeks of the sandfly bite, but they may also seem to several months later.[3] When the lesions cure, they may leave table and deep scars which can cause mental problems. Therefore, evaluating the severity of CL at each visit of patient is important to select suitable treatment that can reduce the size of the lesions with minimal scarring.

Methods of evaluating the severity of skin diseases are often subjective, which makes a difference in results. Therefore, to keep objectivity in observations, scores are applied to evaluate the severity of skin diseases.[7] This is particularly important for monitoring the response to therapy and for evaluating the efficacy of new drugs. In recent years, scoring systems have been developed for some skin diseases. Agarwal et al. suggested pemphigus area and activity score for the clinical assessment of severity and progression of pemphigus vulgaris.[8]

Kimbrogh-Green et al. developed melasma area severity index for the assessment of melasma. Ferriman and Gallwey suggested scoring system for hirsutism in women. Valencia et al. reported a score for prognosis of antimonials therapeutic failure in ulcerative CL patients treated with sodium stibogluconate (SSG) using the logistic regression.

To develop a clinically helpful scoring system, it has to keep several criteria: It should utilize readily available and confirmable clinical information, it should have been developed and validated in the population to whom it is to be used, and it should be free from confounding factors.

Since the use of clinical scores provides a valuable tool for clinical management and orients physicians to select the most suitable treatments, the purpose of the present longitudinal study was to develop a scoring system for predicting cure status in CL patients based on influential predictors using a statistical model. The generalized estimating equations (GEEs) approach was applied to this longitudinal data. Evaluation of model was performed using split-sample validation method.

**MATERIALS AND METHODS**

This study is an analysis of data collected from Skin Diseases and Leishmaniasis Research Center (Isfahan, Iran) in 2011–2012. Dataset includes 199 CL patients. Their information was involved gender, age, morphology of the lesion including flat and other types (papule, nodule, plaque, and others), number of lesions, size of lesions (length of lesion × width), lesions’ location including head and neck, body, hands, and legs, type of treatment including systemic, topical, oral and alone visit, visit times of patients during therapy, induration status (grouped in four levels), and cure status which had been defined as four ordered categories and considered as outcome variable. We regarded CL longitudinal data as three-level structure; the Level 1 units were the repeated occasions of measurement, the Level 2 units were lesions of CL patients, and the Level 3 units were CL patients. Hence, we performed a GEE ordinal logistic regression. The GEE approach denotes an extension of the generalized linear model to analyze correlated data. In this approach, the correlation between correlated measurements is modeled by assuming a working correlation matrix. The GEE models make estimates of model coefficients for predictors that are averaged over clusters whereas allowing residuals to correlate within clusters.

Using the fitted model, probability or score of ordered categories for cure status was predicted as:

\[
P(Y \leq j | X) = \frac{1}{1 + e^{-\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \ldots - \beta_k X_k}}, j = 1, 2, \ldots, J - 1
\]

where \(J\) represents number of categories for outcome variable and \(P(Y \leq j | X)\) is cumulative probability from category 1 to category \(j\) and probability for category \(J\) is calculated as:

\[
\pi_j = 1 - (\pi_1 + \pi_2 + \ldots + \pi_{j-1})
\]

where \(L_j\) denotes the linear predictor of fitted model for category \(j\), \(L_j = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k\), which \(\beta_0\) denotes intercept for category \(j\) and \(\beta_1, \beta_2, \ldots, \beta_k\) denotes regression coefficients. \(X_1, X_2, \ldots, X_k\) represents predictive variables contained in the fitted model.

Thus, a CL patient belongs to the category that has highest probability among all categories. To evaluate the fitted model, split sample validation was used. In this method, dataset was split randomly into two parts; the training set includes a sample of 140 for estimation of the regression coefficients and the test set includes a sample of 59 for evaluating the performance of the score. The predicted categories for test set using regression coefficients resulting training set were compared with observed categories by physicians based on clinical information. High correct classification rate indicates good concordance of the score. Data analyses were performed using a statistical software package (IBM SPSS Statistics version 20, Tokyo, Japan).

**RESULTS**

Mean age of CL patients was 29.27 years with standard error 1.23% and 68.8% of patients were male. Outcome variable, cure status, was ordinal and includes four categories: No cure, initial cure, partial cure, and complete cure. Predictor variables were gender, age, morphology of the lesion, number of lesions, size of lesion, location, type of treatment, induration status, and times of visit. Table 1 shows results of GEE ordinal logistic regression.

The regression coefficients of the fitted model were used to calculate probability or score for cure status. Cumulative probability for category \(j\) can be calculated as:

\[
P(Y \leq j | Data) = \frac{1}{1 + e^{-L_j}}, j = 1, 2, 3
\]

where

\[
L_j = \beta_0 + 0.024 \text{gender} + 0.009 \text{age} - 4.705 \text{induration1} - 4.983 \text{induration2} - 10.744 \text{induration3} - 0.250 \text{location1} + 0.016 \text{location2} + 0.007 \text{location3} + 3.147 \text{morphology} - 1.771 \text{treatment1} - 0.778 \text{treatment2} - 0.865 \text{treatment3} + 0.004 \text{number} + 0.033 \text{size} + 0.039 \text{time}
\]

In this equation, \(\beta_0 = -11.121\), \(\beta_0 = -6.686\), and \(\beta_0 = 1.407\), gender = 1 for males and gender = 0 for females, induration1 = 1 for lesions that take induration at Level 1 and induration1 = 0 otherwise, induration2 = 1 for lesions that...
take induration at Level 2 and induration2 = 0 otherwise, induration3 = 1 for lesions that take induration at Level 3 and induration3 = 0 otherwise, location1 = 1 for lesions that are at head and neck and location1 = 0 otherwise, location2 = 1 for lesions that are at body and location2 = 0 otherwise, location3 = 1 for lesions that are at hands and location3 = 0 otherwise, morphology = 1 for lesions that are flat and morphology = 0 otherwise. Treatment1 = 1 for those who use treatment of systematic and treatment1 = 0 otherwise, treatment2 = 1 for those who use treatment of topical and treatment2 = 0 otherwise, treatment3 = 1 for those who use treatment of oral and treatment3 = 0 otherwise. Continues variables in equation take their real values.

For example, for a female CL patient at the age of 20 years with 12 lesions, for a lesion in face with induration at Level 3, morphology of others, size of 4 cm², and used treatment of topical, on the 7th day, her probability for category 1, no cure, is

\[ P(Y = 1) = \frac{1}{(1 + e^{-0.018})} = 0.5045. \]

Cumulative probability for category 2 can be calculated as:

\[ P(Y \leq 2) = P(Y = 1) + P(Y = 2) = \frac{1}{(1 + e^{-4.485})} = 0.9895, \]

then probability for category 2, initial cure, becomes

\[ P(Y = 2) = 0.9895 - 0.5045 = 0.4850. \]

Cumulative probability for category 3 can be calculated as:

\[ P(Y \leq 3) = P(Y = 1) + P(Y = 2) + P(Y = 3) = \frac{1}{(1 + e^{-12.346})} = 0.9999, \]

then probability for category 3, partial cure, is

\[ P(Y = 3) = 0.9999 - 0.9895 = 0.0104 \]

and probability for category 4, complete cure, is

\[ P(Y = 4) = 1 - 0.9999 = 0.0001. \]

Based on calculated probabilities for each category of cure, this CL patient belongs to the category 1, no cure, because it has a higher probability.

Table 2 shows results of classification from split sample validation for the fitted model. The overall correct classification rate for GEE ordinal logistic regression was 0.82. Since values of observed and predicted categories were ordinal, the Spearman correlation coefficient was calculated to determine association between observed and predicted values which was 0.876 (P < 0.001). It shows strong association between the values of predicted and observed categories.

**DISCUSSION**

In the present study, we developed a scoring system to predict cure status in CL patients. Ordinal logistic GEE regression was applied to this longitudinal dataset. The significant predictors in ordinal logistic GEE regression were induration status, morphology of the lesion, type of treatment, size of lesion, age, and times of visit. Although gender and location were not significant in the fitted model, they are influential in cure of CL.[17] To adjust effect of variables on cure status, we applied all of mentioned predictor variables in model for calculating score. Based on split-sample validation method, there was good concordance between observed and predicted categories.

**Table 1: Results of ordinal logistic generalized estimating equation for longitudinal cutaneous leishmaniasis data**

<table>
<thead>
<tr>
<th>Outcome (complete cure as reference)</th>
<th>B</th>
<th>SE</th>
<th>Wald Chi-square</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cure</td>
<td>-11.121</td>
<td>0.7064</td>
<td>247.849</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Initial cure</td>
<td>-6.686</td>
<td>0.6579</td>
<td>103.257</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Partial cure</td>
<td>1.407</td>
<td>0.4009</td>
<td>12.315</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Gender (female as reference)**

<table>
<thead>
<tr>
<th>Induration (level 4 as reference)</th>
<th>B</th>
<th>SE</th>
<th>Wald Chi-square</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Induration1: Level 1</td>
<td>-4.705</td>
<td>0.6501</td>
<td>52.385</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Induration2: Level 2</td>
<td>-4.983</td>
<td>0.6697</td>
<td>55.360</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Induration3: Level 3</td>
<td>-10.744</td>
<td>0.7132</td>
<td>226.924</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Location (legs as reference)**

| Location1: Head and neck          | -0.250  | 0.2144 | 1.355           | 0.244   |
| Location2: Body                   | 0.016   | 0.2584 | 0.004           | 0.951   |
| Location3: Hands                  | 0.007   | 0.1463 | 0.002           | 0.962   |

**Morphology of the lesion**

<table>
<thead>
<tr>
<th>(others [no flat] as reference)</th>
<th>B</th>
<th>SE</th>
<th>Wald Chi-square</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>3.147</td>
<td>1.0677</td>
<td>8.690</td>
<td>0.003</td>
</tr>
</tbody>
</table>

**Type of treatment (visit as reference)**

| Treatment1: Systematic            | -1.771  | 0.3284 | 29.077          | <0.001  |
| Treatment2: Topical               | -0.778  | 0.3160 | 6.061           | 0.014   |
| Treatment3: Oral                  | -0.865  | 0.4630 | 3.493           | 0.062   |

**Size of lesions**

| Number                            | 0.033   | 0.0117 | 7.870           | 0.005   |

**Number of lesions**

| Age                               | 0.004   | 0.0023 | 3.699           | 0.054   |
| Time                              | 0.009   | 0.0043 | 4.570           | 0.033   |

| Time                              | 0.039   | 0.0034 | 129.357         | <0.001  |

SE = Standard error

\[ P(Y \leq 2) = P(Y = 1) + P(Y = 2) = \frac{1}{(1 + e^{-4.485})} = 0.9895, \]

then probability for category 2, initial cure, becomes

\[ P(Y = 2) = 0.9895 - 0.5045 = 0.4850. \]

Cumulative probability for category 3 can be calculated as:

\[ P(Y \leq 3) = P(Y = 1) + P(Y = 2) + P(Y = 3) = \frac{1}{(1 + e^{-12.346})} = 0.9999, \]

then probability for category 3, partial cure, is

\[ P(Y = 3) = 0.9999 - 0.9895 = 0.0104 \]

and probability for category 4, complete cure, is

\[ P(Y = 4) = 1 - 0.9999 = 0.0001. \]

Based on calculated probabilities for each category of cure, this CL patient belongs to the category 1, no cure, because it has a higher probability.

Table 2 shows results of classification from split sample validation for the fitted model. The overall correct classification rate for GEE ordinal logistic regression was 0.82. Since values of observed and predicted categories were ordinal, the Spearman correlation coefficient was calculated to determine association between observed and predicted values which was 0.876 (P < 0.001). It shows strong association between the values of predicted and observed categories.

**DISCUSSION**

In the present study, we developed a scoring system to predict cure status in CL patients. Ordinal logistic GEE regression was applied to this longitudinal dataset. The significant predictors in ordinal logistic GEE regression were induration status, morphology of the lesion, type of treatment, size of lesion, age, and times of visit. Although gender and location were not significant in the fitted model, they are influential in cure of CL.[17] To adjust effect of variables on cure status, we applied all of mentioned predictor variables in model for calculating score. Based on split-sample validation method, there was good concordance between observed and predicted categories.

**Table 2: A cross-tabulation of the predicted vs true values, from a generalized estimating equation ordinal logistic regression**

<table>
<thead>
<tr>
<th>Predicted categories</th>
<th>No</th>
<th>Initial</th>
<th>Partial</th>
<th>Complete</th>
<th>Correct classification rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>True categories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>85</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0.82</td>
</tr>
<tr>
<td>Initial</td>
<td>19</td>
<td>49</td>
<td>5</td>
<td>0</td>
<td>0.63</td>
</tr>
<tr>
<td>Partial</td>
<td>2</td>
<td>74</td>
<td>1</td>
<td>0</td>
<td>0.92</td>
</tr>
<tr>
<td>Complete</td>
<td>0</td>
<td>0</td>
<td>35</td>
<td>0</td>
<td>0.97</td>
</tr>
<tr>
<td>Overall correct classification rate</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Complete cure category showed the highest rate of correct classification (82%) and the lowest rate was for initial cure category (63%).

Valencia et al. reported a score for prognosis of antimonial therapeutic failure in ulcerative CL patients treated with SSG. Outcome variable in their cross-sectional study was binary and they calculated probability of treatment failure using a logistic regression. The present study was regarded as longitudinal and outcome variable, cure status in CL patients, was ordinal with four groups. Thus, ordinal logistic GEE regression was applied.

Maxwell et al. developed and validated a scoring system for patients undergoing hip fracture surgery using a logistic regression. They assessed goodness-of-fit of their score to the data using the Hosmer–Lemeshow statistic. They also performed sensitivity analysis using standard receiver operating characteristic (ROC) curve. Although the ROC curve is more informative than the classification table, it is complicated when outcome variable has more than two categories.

Bastuji-Garin et al. suggested a specific severity-of-illness score using logistic regression for cases of toxic epidermal necrolysis. They compared their score with the simplified acute physiology score and a burn scoring system.

A limitation of the present study was that there was no other scoring system for CL with multi-category outcome to compare our score with it. In addition, we obtained a scoring system in a small sample of CL patients from a single center which may influence the results. However, a useful scoring system should be applicable to different centers with similar populations.

CONCLUSION

This study suggested a scoring system predict cure status in CL patients. This predictive score presents useful benefits such as it relies on clinical characteristics and it is easily obtained by physicians or health workers.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

AUTHORS’ CONTRIBUTION

MN, AZ, and MH contributed in the conception and design of the work, performed data collection. MKh and MH contributed in the analysis, drafted the manuscript, performed significant revisions, approved the final version of the manuscript. All authors agreed to all aspects of the work.

REFERENCES