Prediction of Resistance to Intravenous Immunoglobulin in Children With Kawasaki Disease

Maskit Bar-Meir,1,2 Itai Kalisky,1 Andrei Schwartz,3 Eli Somekh,3,4 Diana Tasher4; Israeli Kawasaki Group

1Pediatrics and Infectious Diseases Division, Shaare-Zedek Medical Center, Jerusalem, Israel; 2Faculty of Medicine, Hebrew University, Jerusalem, Israel; 3Pediatrics and Pediatric Infectious Diseases, Wolfson Medical Center, Holon, Israel; 4Sackler School of Medicine, Tel-Aviv University, Israel

**Background.** Approximately 10%–20% of patients with Kawasaki disease (KD) are refractory to initial intravenous immunoglobulin (IVIg) therapy, and these “nonresponders” are at higher risk of coronary artery abnormalities. Early identification of these patients, who may benefit from additional therapy, is challenging. The aim of the present study is to identify predictors for IVIg resistance.

**Methods.** We reviewed clinical records of 312 consecutive KD patients from 9 medical centers in Israel (development dataset) and 186 patients from additional 5 centers (validation dataset). Using multivariate analysis, we identified predictors of IVIg resistance. A small prospective cohort of consecutive KD patients from a single medical center was used to test the accuracy of the predictors.

**Results.** Coronary artery abnormalities in the initial echocardiogram and presenting before day 5 of fever were independent predictors of IVIg nonresponse. Using either of these variables generated an area under the receiver-operating-characteristics curve of 0.7 (95% confidence interval [CI], 0.6–0.8). Sensitivity to predict nonresponse was 81% (95% CI, 67–90) and specificity was 50% (95% CI, 44–56). Similar results were found in the validation dataset and in the small prospective cohort.

**Conclusions.** Coronary artery abnormalities in the initial echocardiogram and presenting before day 5 of fever show high sensitivity in identifying IVIg nonresponders among our KD patients.

**Keywords.** coronary arteries; glucocorticoids; sensitivity; specificity.

The standard therapy for Kawasaki disease (KD) with intravenous immunoglobulin (IVIg) is effective in preventing coronary artery aneurysms (CAA), when given within 10 days of illness onset. However, 2%–5% of patients develop CAA despite prompt treatment [1]. Approximately 10%–20% of patients are refractory to initial IVIg therapy, and these “nonresponders” are at higher risk of CAA [2]. Corticosteroids, as well as infliximab, were used as adjunctive primary therapy for KD, in conjunction with IVIg and aspirin, to patients predicted to be nonresponders [3, 4]; however, selectively targeting patients who may benefit from adjunctive therapy is challenging [5]. Several Japanese scoring systems identify high-risk Japanese KD patients [6–8], but these were shown to be inaccurate in a North American cohort [9]. Moreover, a scoring system developed in San Diego in the setting of ethnically diverse population was also found to be relatively inaccurate, with sensitivity of 73% and specificity of 62% to predict IVIg resistance [10]. It was suggested that subtle genetic differences affect the performance of these prediction tools [5]. In a previous study, we have shown that the epidemiological and clinical characteristics of KD in Israel were similar to those reported for the white population in Europe and the United States [11]. The aim of the present study was to identify predictors for IVIg resistance in our Jewish and Arab patient population.

**METHODS**

**Study Population**

We retrospectively reviewed clinical records of consecutive KD patients treated from 1996 to 2009 in 9 medical centers in Israel. These data comprised the “development” dataset used to identify predictors for nonresponse. A separate retrospective review of patients treated for KD from 2001 to 2007 in 5 medical centers in Central Israel comprised the “validation” dataset, on which we tested the prediction accuracy. A third prospective cohort of consecutive patients treated for KD in a single medical center was used to test the accuracy of the predictors.

Both retrospective chart reviews enrolled patients whose KD diagnosis complied with the 2004 American Heart Association (AHA) statement [2]. Criteria for a complete case definition included fever (temperature exceeding 38°C), accompanied by the presence of at least 4 of the following 5 findings: bilateral

**Predictors for IVIg Nonresponse in Kawasaki Disease**  JPIDS 2018:7 (March) 25
conjunctival injection, changes in the lips and oral cavity, nonpurulent cervical lymphadenopathy, polymorphous exan-
themata, and changes in the extremities. Incomplete KD cases
were diagnosed based on the AHA algorithm [2] (ie, fever and
2–3 clinical criteria, C-reactive protein [CRP] ≥3 mg/dL and/or
erthrocyte sedimentation rate [ESR] ≥40 mm/hour, and either
greater than or equal to 2 standard deviation (SD) units (z score)
from the mean, normal values were ≤2.5). This study was
reviewed and approved by each of the participating centers’ institutional review boards.

Data Analysis
Analysis was performed with SPSS software, version 22 (SPSS, Chicago, IL). Data are presented as mean ± SD for continuous
variables or as a percentage of patients with a given categorical variable. For all analyses, a 2-sided probability <.05 was
considered to be statistically significant. Univariable unpaired t test was used to determine whether white blood cell count,
percentage of neutrophils, absolute neutrophil count, hemoglo-
buin, platelet count, ESR, CRP, sodium, bilirubin, aspartate aminotransferase (AST), alanine aminotransferase, and albumin
deviation between the model and the validation dataset; however, only 10% of the validation dataset patients were Arab compared with 24% of the develop-
memt dataset.

RESULTS
Characteristics and Clinical Outcome of Patients
During the study period, 523 patients were discharged with a
diagnosis of KD in Israel. Eight patients were excluded
because they did not meet the diagnostic criteria, and 17
patients were not treated with IVIg, thus 498 patients were
included in this study (development dataset, N = 312; valida-
tion dataset, N = 186). Table 1 shows the characteristics and
clinical outcome of patients in both datasets. There were no
significant clinical differences between the development and
the validation dataset; however, only 10% of the validation dataset patients were Arab compared with 24% of the develop-
memt dataset. When the cohort was stratified by age (<1 vs ≥1 year old) or by gender, the sensitivity and specificity remained essentially the same. When the cohort was stratified by ethnicity, the sen-
sitivity and specificity in Arabs was somewhat lower (sensitivity
in Jews 83%, 95% CI, 66–92 vs in Arabs 70%, 95% CI, 39–89; specificity in Jews 52%, 95% CI, 45–58 vs in Arabs 42%, 95% CI,
30–55).

Next, the validation dataset was used to assess the accuracy of
the risk model in predicting nonresponse. The accuracy in
this dataset was similar, with area under the ROC curve of 0.69

Table 1. Baseline Characteristics and Clinical Outcome

<table>
<thead>
<tr>
<th>Variable</th>
<th>Development dataset N = 312</th>
<th>Validation dataset N = 186</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in months</td>
<td>31 ± 29.2</td>
<td>29 ± 25.1</td>
<td>.3</td>
</tr>
<tr>
<td>Range</td>
<td>1–182</td>
<td>1–147</td>
<td></td>
</tr>
<tr>
<td>Male gender, N (%)</td>
<td>202 (64)</td>
<td>111 (58)</td>
<td>.2</td>
</tr>
<tr>
<td>Ethnicity, %</td>
<td>Jews 76%</td>
<td>90%</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>Arabs 24%</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>Days of illness at presentation ≤5 days, N (%)</td>
<td>117 (38)</td>
<td>78 (42)</td>
<td>.3</td>
</tr>
<tr>
<td>Incomplete KD, N (%)</td>
<td>106 (34)</td>
<td>70 (37)</td>
<td>.4</td>
</tr>
<tr>
<td>IVIg nonresponders, N (%)</td>
<td>42 (13.4)</td>
<td>29 (15.5)</td>
<td>.5</td>
</tr>
<tr>
<td>Coronary artery abnormalities, N (%)</td>
<td>87 (28%)</td>
<td>65 (35%)</td>
<td>.1</td>
</tr>
</tbody>
</table>

Abbreviations: IVIg, intravenous immunoglobulin; KD, Kawasaki disease; SD, standard deviation.
Table 2. Comparison Between IVIg Responders and Nonresponders in the Development Dataset

<table>
<thead>
<tr>
<th>Variable</th>
<th>Missing Data, N</th>
<th>IVIg Responders, N = 270</th>
<th>IVIg Nonresponders, N = 42</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in months</td>
<td>Mean ± SD</td>
<td>32 ± 29</td>
<td>29 ± 26</td>
<td>.5</td>
</tr>
<tr>
<td>Age &lt;1 year, N (%)</td>
<td>58 (21)</td>
<td>11 (26)</td>
<td>.5</td>
<td></td>
</tr>
<tr>
<td>Male gender, N (%)</td>
<td>176 (65)</td>
<td>26 (62)</td>
<td>.7</td>
<td></td>
</tr>
<tr>
<td>Days of illness at presentation &lt;5, N (%)</td>
<td>91 (34)</td>
<td>26 (62)</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Incomplete KD, N (%)</td>
<td>92 (34)</td>
<td>14 (33)</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Ethnicity (%)
- Jews: 7 / 75 / 75 .8
- Arabs: 24 / 25

Season, %
- Winter: 0 / 24 / 19 .1
- Spring: 27 / 26
- Summer: 24 / 26
- Fall: 25 / 26

Coronary artery abnormalities in the initial echo, N (%) | 4 / 67 (25) / 20 (47) .008

WBC, ×10^3/mm^3 | 16 / 15.4 / 15.4 .9
%Neutrophils | 46 / 62 ± 16 / 67 ± 17 .06
ANC | 46 / 8.7 ± 4.5 / 11 ± 6.3 .12
Hemoglobin | 48 / 11 / 10.8 .3
Platelet count, ×10^3/mm^3 | 16 / 42.7 ± 18 / 40.1 ± 15 .3
ESR, mm/hour | 96 / 71 ± 31 / 79 ± 34 .3
CRP, mg/L | 176 / 92 ± 83 / 94.5 ± 108 .9
Total bilirubin, mg/dL | 89 / 0.7 ± 1.4 / 1 ± 1.2 .2
AST, IU/L | 42 / 55 ± 76 / 75 ± 175 .2
ALT, IU/L | 65 / 60 ± 87 / 78 ± 94 .2
Sodium, mmol/L | 18 / 135 ± 36 / 134 ± 3 .03
Albumin, g/dL | 132 / 3.6 ± 0.5 / 3.3 ± 0.7 .01

Abbreviations: ALT, alanine aminotransferase; ANC, absolute neutrophil count; AST, alanine aminotransferase; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; IVIg, intravenous immunoglobulin; KD, Kawasaki disease; SD, standard deviation; WBC, white blood cells.

The main finding of this study is that IVIg nonresponders can be accurately identified in our population in 2 variables: (1) CA abnormalities (Z score ≥2.5) in the acute phase of illness and/or (2) presenting before day 5 of fever. It is widely accepted that the subset of IVIg-resistant patients is at highest risk for developing CAA and needs additional treatment to improve outcome [4]. Kobayashi et al [6] have developed a model to predict IVIg nonresponse, which includes 5 laboratory variables (sodium, AST, percentage of neutrophils, CRP, platelet count), age ≤12 months, and days of illness at initial treatment ≤4. This model identified IVIg nonresponders with a sensitivity of 86% and specificity of 68%. Area under the ROC curve was 0.85 [6]. Moreover, it was shown that patients predicted to have severe KD based on this risk score had better CA outcomes when prednisolone was added to the standard IVIg regimen. Of note, this study excluded patients with CA abnormalities at presentation. A recent study by Dominguez et al [12] found that CA abnormalities were present in most patients before the receipt of IVIg, thereby raising the question whether IVIg resistance is a risk factor for CA abnormalities or vice versa, or whether both are simply markers of severe disease. Either way, CA abnormalities may be the marker for high-risk patients. Excluding these patients from studies that aim at developing KD risk stratification may skew the prediction rules away from the high-risk patients.

Unfortunately, risk scoring systems from Japan, including the Kobayashi score, had a low sensitivity for predicting IVIg resistance in a multiethnic North American cohort [13]. A scoring system developed using retrospective data from the San Diego County KD patients had a sensitivity of 73% and specificity of 69% to predict nonresponse; however, this risk score was not validated on a different cohort [10].

In the present study, we aimed to identify predictors of IVIg nonresponse in our population, which is genetically unique and comprises Jewish (Ashkenazi and Sephardic descent) and Arab patients. We found that most variables that were used in the Japanese risk scores, including age, did not identify nonresponders in our cohort.

Table 3. Multivariable Logistic Regression Analyses for Prediction of IVIg Nonresponse in the Development Dataset

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Logistic Coefficient (βi)</th>
<th>SE</th>
<th>Odds Ratio</th>
<th>95% CI Lower to Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days of illness at presentation ≤4</td>
<td>1.09</td>
<td>0.348</td>
<td>2.973</td>
<td>1.503–5.880</td>
</tr>
<tr>
<td>Coronary artery abnormalities</td>
<td>0.96</td>
<td>0.347</td>
<td>2.613</td>
<td>1.323–5.162</td>
</tr>
<tr>
<td>Intercept</td>
<td>−2.701</td>
<td>0.304</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; IVIg, intravenous immunoglobulin; SE, standard error.

(95% CI, 0.59–0.8); the sensitivity was 86% and the specificity was 41%. In this cohort, there was only 1 Arab patient who was an IVIg nonresponder and had 1 of the risk factors.

The correspondence of IVIg nonresponse with these 2 risk factors is shown in Figure 2. If either of these 2 risk factors were used prospectively to identify IVIg nonresponders, who may be candidates for additional therapy, 43% of the patients would receive unnecessary additional therapy (false positive); however, only 3% of the patients would be nonresponders, but they would not have received additional therapy (false negative).

We then examined the performance of these risk factors prospectively among 32 consecutive KD patients. The rate of nonresponders in this prospective cohort was 22% (7 of 32). Using either of these 2 predictors correctly identified 86% (95% CI, 49–97) of the nonresponders, whereas the specificity was 61% (95% CI, 43–78). Area under the ROC curve was 74% (95% CI, 60–87).

DISCUSSION

The main finding of the present study is that IVIg nonresponders can be accurately identified in our patient population using 2 variables: (1) CA abnormalities (Z score ≥2.5) in the acute phase of illness and/or (2) presenting before day 5 of fever. It is widely accepted that the subset of IVIg-resistant patients is at highest risk for developing CAA and needs additional treatment to improve outcome [4]. Kobayashi et al [6] have developed a model to predict IVIg nonresponse, which includes 5 laboratory variables (sodium, AST, percentage of neutrophils, CRP, platelet count), age ≤12 months, and days of illness at initial treatment ≤4. This model identified IVIg nonresponders with a sensitivity of 86% and specificity of 68%. Area under the ROC curve was 0.85 [6]. Moreover, it was shown that patients predicted to have severe KD based on this risk score had better CA outcomes when prednisolone was added to the standard IVIg regimen. Of note, this study excluded patients with CA abnormalities at presentation. A recent study by Dominguez et al [12] found that CA abnormalities were present in most patients before the receipt of IVIg, thereby raising the question whether IVIg resistance is a risk factor for CA abnormalities or vice versa, or whether both are simply markers of severe disease. Either way, CA abnormalities may be the marker for high-risk patients. Excluding these patients from studies that aim at developing KD risk stratification may skew the prediction rules away from the high-risk patients.

Unfortunately, risk scoring systems from Japan, including the Kobayashi score, had a low sensitivity for predicting IVIg resistance in a multiethnic North American cohort [13]. A scoring system developed using retrospective data from the San Diego County KD patients had a sensitivity of 73% and specificity of 69% to predict nonresponse; however, this risk score was not validated on a different cohort [10].

In the present study, we aimed to identify predictors of IVIg nonresponse in our population, which is genetically unique and comprises Jewish (Ashkenazi and Sephardic descent) and Arab patients. We found that most variables that were used in the Japanese risk scores, including age, did not identify nonresponders in our cohort.
Two variables differentiated responders and nonresponders: day of illness <5 at presentation and CA abnormalities (Z score ≥2.5) in the acute phase. When using either one of the 2 predictors—day of illness <5 at presentation or CA abnormalities in the acute phase—over 80% of IVIg nonresponders can be correctly identified; however, the specificity of this model is modest (≈50%). Practically, using these 2 variables in our small prospective cohort, we were able to identify 86% (6 of 7) of patients who would likely benefit from additional treatment, at the cost of unnecessarily treating 10 of 26 (38%) patients who would eventually respond to IVIg alone.

Kobayashi et al [6] noted that each of the variables included in their large well designed scoring model as a predictor of IVIg nonresponse was also reported to be a risk factor for CAA. In addition, CAA during the acute phase developed in their patients significantly more often in IVIg nonresponders than in IVIg responders. Therefore, unlike various laboratory and demographic variables that seem to be inconsistent as predictors of IVIg nonresponse (depending on the patient population examined), coronary abnormalities in the acute phase stands out repeatedly as strongly associated with nonresponse in all patient populations, including our own cohort. Because the Kobayashi et al study [6] excluded patients with CAA, this significant predictor is not taken into account when using this risk score system for identifying high-risk KD patients. Moreover, the study by Dominguez et al [12] demonstrated the early occurrence of CA abnormalities in a subset of KD patients who also had a higher rate of IVIg resistance. Our study lends further support to Dominguez et al [12]. Taken together, the findings from both studies suggest that patients with early CA abnormalities may benefit from a more aggressive initial therapy.

Early diagnosis (presenting before day 4–5 of illness) is strongly associated with IVIg nonresponse in multiple patient populations as well [6, 10]. It was suggested that these patients have a more severe disease, as demonstrated by higher scores in the disease severity index.

Because the Japanese scoring systems cannot successfully identify US children at higher risk for treatment failure, a nonselective intensification of primary therapy by the addition of either a single dose of intravenous methylprednisolone [14] or a single dose of infliximab [4] was assessed. The methylprednisolone trial did not show a significant difference in CA Z score due to the overall low rate of CAA, and the infliximab study could not show a significant difference in the rate of IVIg nonresponse, and it demonstrated a slightly greater reduction in Z score of the left anterior descending CA in the treatment group at week 2 but not at week 5. Therefore, because the overall rates of adverse outcome in treated KD patients are low, it seems that in a nonselected population of KD patients it will be difficult to demonstrate the benefit of intensified primary therapy, even when large numbers of patients are enrolled.

Figure 1. Receiver operator curve analysis shows an area under the curve of 0.7 for both datasets (95% confidence interval of 0.6–0.75 for the development dataset and 0.6–0.8 for the validation dataset).

Figure 2. The occurrence of intravenous immunoglobulin (IVIG) nonresponse by number of risk factors.
The major limitation of our study was its retrospective design, which is more likely to suffer from missing data. Furthermore, we relied on history obtained from the parents to determine the first day of fever. Because day of illness remains a predictor of nonresponse in our small prospective cohort, a systematic recall bias is less likely.

Another limitation of our study is the fact that echocardiograms were performed and interpreted by many different cardiologists. However, all of the pediatric cardiologists in Israel followed the AHA guidelines for the diagnosis of KD, and they use body surface area-adjusted coronary dimensions [2]; therefore, differences in expertise are unlikely to be systematically different in responders compared with nonresponders, and these are unlikely to cause a systematic bias of our results. Finally, the generalizability of our results to a more heterogeneous European or North American populations is unknown.

The 2 predictors of IVIg nonresponse performed better in Jews compared with Arab KD patients. This is not surprising because it was shown in the past that the accuracy of the risk scoring systems tends to vary among ethnically different populations 10.

The strength of our study is the validation of our findings on a different retrospective cohort, as well as on a small prospective cohort. Still, it only meets the requirements for a “level 3” clinical decision rule as defined by McGinn et al [15] and requires further validation in a different population of KD patients, as well as an impact analysis study. Another caveat for the implementation of our findings in clinical practice is the heterogeneity in glucocorticoid treatment protocols used in KD patients. A prospective multicenter study that will examine a more aggressive initial therapy for this subset of high-risk KD patients is indicated.

CONCLUSIONS

In summary, CA abnormalities in the initial echo and/or presenting before day 5 of illness accurately identify IVIg nonresponse in initial treatment of patients with KD. If validated in different cohorts, these 2 variables may be used for risk stratification of KD patients and for the selection of patients at higher risk of nonresponse who may benefit from adjunctive primary therapy.

Notes

Acknowledgments. The Israeli Kawasaki Group members are as follows: Shai Ashkenazi, Diana Averbuch, Dan Miron, Alex Guri, Michal Stein, David Greenberg, Daniel Glikman, Moshe Efrat, Barak Chaim, Ilan Youngster, Asher Barzilai, Itzhak Srugo, and Yael Shachor-Meyouhas.

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References