Comparison of Severity Scorings in Oral Food Challenges With Cow’s Milk

Nieminen O1, 2, Palosuo K1, 2, Kukkonen K1, 2, Mäkelä M1, 2
1Helsinki Skin and Allergy Hospital, University of Helsinki, Helsinki, Finland
2Helsinki University Hospital, Helsinki, Finland

doi: 10.18176/jiaci.0913

Key words: Oral food challenge. Cow’s milk allergy. Severity scoring.
Symptom. Allergic reaction.

To the Editor:

Cow’s milk allergy (CMA) is the most common food allergy in early childhood. Its diagnosis is based on clinical history, sensitization, and oral food challenge (OFC). Different scoring systems have been developed to objectively assess OFC symptoms and their severity. However, the large number of scoring systems reflects the difficulty of fully standardizing assessment of symptoms [1]. We compared 2 widely used methods for assessing symptoms, namely, that of Hourihane et al [2] and that of Sampson [3] and investigated their differences in open OFCs with cow’s milk. Furthermore, we compared these scores with the Food Allergy Severity Score (FASS) [4], the only validated scoring system developed to date.

We challenged 135 Finnish children (median [IQR] age, 1.8 [1.3-3.7] years) with cow’s milk at the Helsinki University Skin and Allergy Hospital, as previously described [5], and evaluated the reactions according to the PRACTALL consensus criteria [6]. Symptom severity was assessed retrospectively by author ON with the severity scoring system of Hourihane et al (HSS) [2], that of Sampson (SSS) [3], a modified HSS (mHSS) [7], and a modified SSS (mSSS) (sFile S1). Additionally, FASS was applied for all the reactions using R studio and the code provided by the developers of FASS [4]. All symptoms evaluated according to the PRACTALL criteria and their respective severity scorings appear in sTable S2.

We challenged 135 Finnish children (median [IQR] age, 1.8 [1.3-3.7] years) with cow’s milk at the Helsinki University Skin and Allergy Hospital, as previously described [5], and evaluated the reactions according to the PRACTALL consensus criteria [6]. Symptom severity was assessed retrospectively by author ON with the severity scoring system of Hourihane et al (HSS) [2], that of Sampson (SSS) [3], a modified HSS (mHSS) [7], and a modified SSS (mSSS) (sFile S1). Additionally, FASS was applied for all the reactions using R studio and the code provided by the developers of FASS [4].

Analysis because of inconclusive symptoms. The Figure shows the severity of the reactions according to the SSS, HSS, mSSS, mHSS, and FASS (ordinal format, 3 grades [oFASS-3]) [4].

Agreement on the severity of the reaction between SSS and HSS was weak (weighted Cohen κ, 0.496). Unlike SSS, HSS was affected by the cumulative reactive dose. Abdominal pain was recognized only by HSS, while inspiratory stridor was not identified by SSS. SSS identified many symptoms not recognized by HSS, namely, flushing, oral allergy syndrome (OAS), lip swelling, nausea, diarrhea, loss of bowel control, dysphagia, upper airway symptoms (sneezing, rhinorrhea, nasal congestion), cardiovascular symptoms (tachycardia, hypotension, severe bradycardia, dysrhythmia, cardiac arrest), dyspnea, cyanosis, feeling of impending doom, light headedness, and change in activity level. Symptoms not recognized by either scoring system included persistent cough, itching and rubbing of eyes and nose, and decrease in arterial blood saturation.

Some symptoms were scored differently. Vomiting was graded as moderate by HSS, but as mild (1 episode of vomiting) or moderate (>1 episode) by SSS. Angioedema was moderate with HSS but mild with SSS, whereas generalized

Analysis because of inconclusive symptoms. The Figure shows the severity of the reactions according to the SSS, HSS, mSSS, mHSS, and FASS (ordinal format, 3 grades [oFASS-3]) [4].

Agreement on the severity of the reaction between SSS and HSS was weak (weighted Cohen κ, 0.496). Unlike SSS, HSS was affected by the cumulative reactive dose. Abdominal pain was recognized only by HSS, while inspiratory stridor was not identified by SSS. SSS identified many symptoms not recognized by HSS, namely, flushing, oral allergy syndrome (OAS), lip swelling, nausea, diarrhea, loss of bowel control, dysphagia, upper airway symptoms (sneezing, rhinorrhea, nasal congestion), cardiovascular symptoms (tachycardia, hypotension, severe bradycardia, dysrhythmia, cardiac arrest), dyspnea, cyanosis, feeling of impending doom, light headedness, and change in activity level. Symptoms not recognized by either scoring system included persistent cough, itching and rubbing of eyes and nose, and decrease in arterial blood saturation.

Some symptoms were scored differently. Vomiting was graded as moderate by HSS, but as mild (1 episode of vomiting) or moderate (>1 episode) by SSS. Angioedema was moderate with HSS but mild with SSS, whereas generalized

### Table

<table>
<thead>
<tr>
<th>Symptom Category</th>
<th>SSS</th>
<th>HSS</th>
<th>mSSS</th>
<th>mHSS</th>
<th>oFASS-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild (1-2)</td>
<td>58*</td>
<td>7</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate (3)</td>
<td>20</td>
<td>2*</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe (4-5)</td>
<td>0</td>
<td>0</td>
<td>9*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure. Bar chart representing the severity of oral food challenge reaction assessed based on the 5 different scoring systems (n=97). The severity of reactions graded similarly by both SSS and HSS is marked with an asterisk. The scores leading to the corresponding severity grades are shown in parenthesis. SSS indicates Sampson severity score [3]; HSS, Hourihane severity score [2]; mSSS, modified Sampson severity score (sTable S1); mHSS, modified Hourihane severity score [7]; oFASS-3, Food Allergy Severity Score, ordinal format, 3 grades [4].

© 2024 Esmon Publicidad

urticaria was either mild or moderate by HSS, depending on the cumulative reactive dose, but was graded as mild with SSS.

In our cohort, 20 of the 81 reactions scored as mild by SSS were scored as moderate by HSS. This was due to vomiting in 10 cases, generalized urticaria elicited by a small dose in 6 cases, and abdominal pain in 4 cases (instead graded as OAS in 2, nausea in 1, and nasal congestion in 1 with SSS). Seven cases of rhinorrhea were moderate according to SSS but mild with HSS, as local skin reactions were used for grading. One case of dyspnea scored as severe by SSS was moderate with HSS owing to abdominal pain. Eight cases of persistent cough were not recognized by either scoring system. Instead, they were graded based on vomiting in 2, local skin reactions in 3, and generalized urticaria in 1 case. The remaining 2 cases were scored according to rhinorrhea with SSS and local skin reactions with HSS.

The differences between SSS and HSS were especially pronounced in gastrointestinal symptoms. Vomiting was graded differently, and OAS was not recognized by HSS. In addition, HSS always graded abdominal pain as moderate, whereas SSS neglected it altogether. Relatively common symptoms such as persistent cough and ocular itching were neglected by both scorings. HSS, which was originally developed for peanut challenge, does not recognize upper airway symptoms, dysrhythmias, or changes in activity level. In addition, identification of severe respiratory symptoms and laryngeal symptoms depends largely on individual interpretation in HSS.

Our mSSS (sFile S1) added mild and severe abdominal pain as recognized symptoms (grades 2 and 3, respectively) and graded intermittent and constant sneezing and rhinorrhea as mild (grade 2) and moderate (grade 3) symptoms [8]. Ocular itching was added as a mild symptom [9]. Persistent cough and inspiratory stridor were added as severe grade 4 symptoms. Agreement between SSS and mSSS was moderate (weighted Cohen $\kappa$, 0.663). The results of the comparison between HSS and mHSS are shown in sFile S3. Their mutual agreement was weak (weighted Cohen $\kappa$, 0.570).

FASS recognizes symptoms according to the PRATCALL consensus criteria [6]. It uses the organs and systems affected rather than symptoms alone when grading the severity of a reaction. Grade 1 (OAS) is considered mild, grades 2 and 3 moderate (respectively, 1 and at least 2 of skin, nose/eye, digestive tract, and uterus affected), and grades 4 (larynx or bronchi) and 5 (cardiovascular or nervous system) as severe. Mild and moderate reactions assessed by SSS or HSS were moderate in FASS, apart from 12 cases considered severe. These included 1 case of weakness, 1 case of dyspnea (severe in SSS), and 10 cases of frequent coughing not recognized by either scoring system. All cases assessed as severe by either SSS or HSS were also severe in FASS.

Using oFASS-3 [4], no reaction was mild, since there were no patients with only oral or pharyngeal itching. Of these reactions, 76 (78%) were moderate and 21 (22%) severe (Figure). Using oFASS-5 [4], 54 of the moderate reactions (71%) were grade 2 and 22 (29%) were grade 3. Of the severe reactions, 17 (81%) were grade 4 and 4 (19%) were grade 5. Unlike SSS and HSS, FASS recognized both itching and rubbing of the eyes and nose and frequent coughing. Compared to SSS, FASS recognized abdominal pain and inspiratory stridor. It did not recognize nasal congestion, dyspnea, respiratory arrest, dysphagia, loss of bowel control, or feeling of impending doom. In addition, dysrhythmia, severe bradycardia, and cyanosis were not specified. The weighted Cohen $\kappa$ was 0.155. Unlike HSS, FASS recognized flushing, lip edema, OAS, nausea, upper respiratory symptoms (rhinorrhea, sneezing), changes in activity level, weakness, tachycardia, hypotension, and diarrhea. It also better specified laryngeal symptoms, collapse, and severe respiratory symptoms. The weighted Cohen $\kappa$ was 0.155. These $\kappa$ values are probably explained by the absence of patients with OAS as their only symptom, recognition of frequent coughing as a symptom, and the use of organs and systems to grade severity instead of symptoms alone. When oFASS-5 was compared with SSS grades 1 to 5, the $\kappa$ value increased to 0.261, thus further supporting the role of different views on OAS and coughing as explanations for the small $\kappa$ values.

Our results for the severity of reactions in 97 positive oral cow’s milk challenges show that different severity scoring systems generate different grades. Clinicians worldwide use varied scoring systems, further increasing heterogeneity. Even identical scoring systems show substantial interobserver variability for symptom assessment, with the $\kappa$ for mutual agreement being 0.31 to 0.46 [10]. Attempts to predict the severity of reactions using different markers, such as specific IgE, would benefit from a unified severity scoring system. Currently, only 1 severity scoring system has been properly validated [4]. As such, there is much anticipation for the DEFASE project undertaken by the World Allergy Organization to further our understanding of the severity of food allergies [1].

**Acknowledgments**

We wish to thank Janne Burman, MD for contributing to the conceptualization of this study. We also wish to thank the personnel of Helsinki Skin and Allergy Hospital for carrying out the challenges and the families for their participation in the study.

**Funding**

This work was supported by the Helsinki University Hospital Research Funds, the Allergy Research Foundation, the Pediatric Research Foundation, the University of Helsinki, and the Sigrid Jusélius Foundation.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**References**

Letters to the Editor

Manuscript received December 22, 2022; accepted for publication May 10, 2023.

Otso Nieminen

Skin and Allergy Hospital, University of Helsinki
and Helsinki University Hospital
PO Box 160, FIN-00029 HUS, Finland
E-mail: otso.nieminen@helsinki.fi

Impact of the ERS/ATS 2022 Guidelines for Interpretation of Lung Function Test Results When Assessing the Response to Biologics in Asthma

Pérez de Llano L1, Muñiz Fernández MC1, Dávila I2,3

1Pulmonology Department, Lucus Augusti University Hospital, Lugo, Spain
2Department of Biomedical and Diagnostic Sciences, Faculty of Medicine, University of Salamanca, Salamanca, Spain
3Allergy Department, University Hospital of Salamanca, Salamanca, Spain

doi: 10.18176/jiaci.0943

Key words: Biologicals. Asthma. Lung function. Evaluation.


To the Editor:

In 2022, the American Thoracic Society (ATS) and the European Respiratory Society (ERS) updated their standards for the interpretation of pulmonary function tests [1]. The main changes regarding spirometry are set out below.

1. The use of 80% predicted to define normal was no longer recommended. Instead, the general use of the lower limit of normal (LLN) or 5th percentile and the upper limit of normal (ULN) or 95th percentile was advocated (ie, Z-scores or percentiles). These guidelines now recommend the equations developed by the Global Lung Initiative (GLI) for referencing normal spirometry, diffusion capacity, and lung volumes. According to the GLI, bronchial obstruction should be diagnosed when the ratio of forced expiratory volume in 1 s (FEV1) to forced vital capacity (FVC) is not >5th percentile and FVC is >5th percentile. This recommendation is the result of an expert consensus aimed at identifying, in a standardized and unbiased way, values that fall outside the range of those expected in the general population. It also implies accepting that the change will result in 5% of healthy individuals being incorrectly classified as having an abnormal result. However, on the other hand, it overcomes the disadvantage of classifying a significant percentage of the elderly population as having obstructive disease.

The newly formulated concept of “clinical remission” [2] and several of the tools developed to quantify the response to biologics in asthma [3], eg, the FEOS score [4], incorporate lung function as one of the domains to be improved by treatment. In all cases, FEV1 was chosen as the parameter for estimating bronchial obstruction. However, its interpretation is based on outdated recommendations. Considering that scores to measure response should be simple, that it will be mandatory to assume some limitation (spirometry is not a simple technique to perform and interpret), that most published studies on biological response use FEV1, and that this parameter has also traditionally been used in the estimation of lung function trajectories in asthma patients, we propose, at least, to replace the 80% predicted cut-off point by the Z-score value (−1.65).

References


Manuscript received December 22, 2022; accepted for publication May 10, 2023.