RECOMMENDATIONS FOR THE DIAGNOSIS AND MANAGEMENT OF CHURG-STRAUSS SYNDROME

METHODS AND SCORING

Structure of working party

The working party will be composed of three panels. The Core panel will have overall responsibility for the recommendations and developing the project. It will comprise the 3 chairs, secretary, Associated secretaries and international experts in CSS.

The Consultant panel will advise on specialist aspects of the guidelines including epidemiology, biopathology, internal medicine, rheumatology, respiratory medicine, cardiology, nephrology, neurology, allergology, immunology, dermatology, ophtalmology, gastroenterology, pathology, imaging, pharmacology, pediatrics.

The Review panel will review the documents and will comprise all members of the core and consultant panels plus international experts in CSS and presidents of patients’ associations.

Definition of processes

The process of recommendation development will be defined at a meeting of the core panel. It will include the following steps: 1) question formulation, 2) evidence collection and synthesis (core and consultant panels), 3) answer to questions and recommendations, with grading of recommendation strength (core and consultant panels), 4) circulation of documents, and first version of the recommendations (core and consultant panels), 5) first formal review with scoring of agreement and proposals for modifications (core, consultant and review panels), 6) integration of proposals (core panel), 7) second formal review with re-assessment of agreement (core,
consultant and review panels), and 8) final revision (core panel). The whole process will aim at obtaining an objective assessment of the following components 1) strength of evidence, 2) magnitude of benefit, 3) strength of recommendation, and 4) strength of expert consensus.

**Question formulation**

Members of all panels will be asked to identify questions which should be addressed by the Taskforce. All questions will be reviewed at a meeting of the core panel to achieve a precise formulation. Each question will be then attributed to a member of the core or the appropriate consultant panel for detailed work-up.

**Evidence collection and synthesis**

A Medline literature search will be performed for the period 1968-2009, including articles in all languages with available description of methodology. The search strategy terms will be defined. The consultant panels will be asked to perform additional searches as needed. Papers not recorded in the Medline database (published before 1968) will be added from the personal files of the participants.

For each specific question, relevant data will be synthesized in an evidence table, including reference, study design, population type, cohort size, inclusion and exclusion criteria, main results, and conclusions.

**Answer to questions, recommendations, and grading**

Based on available evidence, an answer will be formulated by the panelist in charge of the question. The strength of evidence will be graded as detailed here. This system is based on the following 2 components:

1) *Evidence quality*
Good evidence based on good randomised controlled trials or meta-analyses

Fair evidence based on other controlled trials, or randomised controlled trials with minor flaws

Low evidence based on nonrandomized, case-control, or other observational studies

Expert opinion evidence based on the consensus of a carefully selected panel of experts in the topic field.

2) *Net benefit*

- Substantial
- Intermediate
- Small/weak
- None
- Conflicting
- Negative

The net benefit will be based on evaluation of the risk / benefit ratio for the patient population (not for individual patients)

When appropriate, a recommendation for diagnostic and therapeutic intervention will be formulated for each specific question. The strength of the recommendation was expressed as follows:

A = strong recommendation
B = moderate recommendation
C = weak recommendation
D = negative recommendation
I = no recommendation possible (inconclusive)
E/A = strong recommendation based on expert opinion only
E/B = moderate recommendation based on expert opinion only
E/C = weak recommendation based on expert opinion only
E/D = negative recommendation based on expert opinion only

The strength of the recommendation will be determined according to evidence quality and net benefit, using the following predefined table (table a1.1) (McCrory et al, Chest 2004, 126:11S-13S):

<table>
<thead>
<tr>
<th>Quality of evidence</th>
<th>Net benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Substantial</td>
</tr>
<tr>
<td>Good</td>
<td>A</td>
</tr>
<tr>
<td>Fair</td>
<td>A</td>
</tr>
<tr>
<td>Low</td>
<td>B</td>
</tr>
<tr>
<td>Expert opinion</td>
<td>E/A</td>
</tr>
</tbody>
</table>

Table a1.1 Method for determining the strength of a recommendation according to evidence quality and net benefit.

Circulation of documents and first version of the recommendations

For each question, the work-up made by each core and consultant panelist will be circulated to the other panelists for extensive review. Issues raised by this review will be discussed at a meeting of the core and consultant panel and the first version of the guidelines will be then prepared for the first formal review.

First formal review with scoring of agreement

The first version of the guidelines will be sent to all panels for formal review. Panelists will be asked to make comments and suggestions as needed. When the available data are insufficient for evidence-based recommendations, the Delphi method will be used to develop recommendations based on expert opinions. The Delphi method is a formal
group consensus technique characterised by individual review of questions avoiding face-to-face interactions, expression of agreement on a visual analogic scale, statistical expression of results with feedback to panel members, and re-circulation of questions in several rounds to improve agreement.

For each question, panel members will be requested to express their agreement with each statement using the 9-point Likert scale, with 1 meaning extreme disagreement and 9 meaning extreme agreement. Panelists responses will be analysed and expressed in box-whiskers plots disclosing the median, interquartile range, and minimal and maximal values.

**Integration of proposals, second formal review and final revision**

Responses of the reviewers will be analysed at a meeting of the core panel. Based on statistics of agreement and proposals for changes, a second version will be produced by the core panel and will be re-sent to the reviewers for a second formal evaluation. This revised version will include statistics on agreement with the first version. Panelists responses to the second version will be again analysed and integrated in a final version made by the core panel.

In this final document, the **strength of the recommendations** based on expert opinion will express in the text by using predefined wording (“must”, “should”, “may”, “should not” and “must not”). This wording will be derived from the Likert scale statistics, using the following formal definitions as described in table a1.2.

<table>
<thead>
<tr>
<th>Management Options</th>
<th>Median  (Middle 50% Range)</th>
<th>Strength of Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferred management in</td>
<td>7–9(7–9)</td>
<td>&quot;Must&quot; if perfect consensus; &quot;should&quot;</td>
</tr>
<tr>
<td>most circumstances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acceptable management in many circumstances</td>
<td>7–9 (4–9)</td>
<td>&quot;Should&quot; if no preferred management exists; &quot;may&quot; if a preferred management exists</td>
</tr>
<tr>
<td>Acceptable management in certain circumstances</td>
<td>4–6 (4–9)</td>
<td>&quot;May&quot;</td>
</tr>
<tr>
<td>Acceptable management in rare circumstances</td>
<td>2 and 3 (1–≤ 4)</td>
<td>&quot;May&quot;</td>
</tr>
<tr>
<td>Inappropriate management</td>
<td>(1–3)</td>
<td>&quot;Must not&quot; if perfect consensus; &quot;should not&quot; otherwise</td>
</tr>
<tr>
<td>Indeterminate</td>
<td>All other median and range combinations including &quot;no consensus&quot;</td>
<td>No management recommendation</td>
</tr>
</tbody>
</table>

Table a1.2 Method for determining the description of the strength of a recommendation in the text.

In the final document, the **level of consensus** will be derived from Likert scale statistics, using the following formal definitions as described in table a1.3.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect consensus</td>
<td>All respondents agree on an answer</td>
</tr>
<tr>
<td>Very good consensus</td>
<td>Median and middle 50% (interquartile range) of</td>
</tr>
</tbody>
</table>
respondents are found at one integer (eg, median and interquartile range are both at 8) or 80% of respondents are within one integer of the median (eg, median is 8, 80% respondents are from 7 to 9)

Good consensus 50% of respondents are within one integer of the median (eg, median is 8, 50% of respondents are from 7 to 9) or 80% of the respondents are within two integers of the median (eg, median is 7, 80% of respondents are from 5 to 9).

Some consensus 50% or respondents are within two integers of the median (eg, median is 7, 50% of respondents are from 5 to 9) or 80% of respondents are within three integers of the median (eg, median is 6, 80% of respondents are from 3 to 9).

No consensus All other responses.

Table a1.3 Method for determining the level of consensus for recommendations.

References