Appendix S4 GRADE summary for efficacy of probiotics on mental health in patients with multiple sclerosis

(a) EDSS scores

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>ILLUSTRATIVE COMPARATIVE RISK</th>
<th>RELATIVE EFFECT</th>
<th>NO OF PARTICIPANTS</th>
<th>QUALITY OF THE EVIDENCE</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDSS scores</td>
<td>Assumed risk</td>
<td>Placebo</td>
<td>Probiotic</td>
<td>(95% CI)</td>
<td>Assumed risk</td>
</tr>
<tr>
<td>Follow-up, 3 months</td>
<td>The mean EDSS score in the treatment group was lower (2.4 to 0.03 points)</td>
<td>173</td>
<td>(3 studies)</td>
<td>very low</td>
<td>1.25</td>
</tr>
</tbody>
</table>

*The basis for the assumed risk (e.g. median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.
1. serious inconsistency due to high heterogeneity with 75% to 12.
2. serious indirectness due to different strains and intervention time in all studies.
3. serious imprecision due to the small sample size (<400 individuals) and wide confidence interval.

(b) BDI scores

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>ILLUSTRATIVE COMPARATIVE RISK</th>
<th>RELATIVE EFFECT</th>
<th>NO OF PARTICIPANTS</th>
<th>QUALITY OF THE EVIDENCE</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI score</td>
<td>Assumed risk</td>
<td>Placebo</td>
<td>Probiotic</td>
<td>(95% CI)</td>
<td>Assumed risk</td>
</tr>
<tr>
<td>Follow-up, 3 months</td>
<td>The mean BDI score in the treatment group was lower (19.5 to 12.0 points)</td>
<td>173</td>
<td>(3 studies)</td>
<td>very low</td>
<td>1.22</td>
</tr>
</tbody>
</table>

*The basis for the assumed risk (e.g. median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.
1. serious inconsistency due to high heterogeneity with 75% to 12.
2. serious indirectness due to different strains and intervention time in all studies.
3. serious imprecision due to the small sample size (<400 individuals) and wide confidence interval.

(c) GHQ scores

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>ILLUSTRATIVE COMPARATIVE RISK</th>
<th>RELATIVE EFFECT</th>
<th>NO OF PARTICIPANTS</th>
<th>QUALITY OF THE EVIDENCE</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>GHQ scores</td>
<td>Assumed risk</td>
<td>Placebo</td>
<td>Probiotic</td>
<td>(95% CI)</td>
<td>Assumed risk</td>
</tr>
<tr>
<td>Follow-up, 3 months</td>
<td>The mean GHQ score in the treatment group was lower (1.92 to 0.4 points)</td>
<td>173</td>
<td>(3 studies)</td>
<td>low</td>
<td>1.2</td>
</tr>
</tbody>
</table>

*The basis for the assumed risk (e.g. median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.
1. serious indirectness due to different strains and intervention time in all studies.
2. serious imprecision due to the small sample size (<400 individuals) and wide confidence interval.

(d) DASS scores
### Probiotics compared to Placebo for Multiple Sclerosis

**Patient or population:** Patients with multiple sclerosis

**Interventions:** Probiotic

**Comparator:** Placebo

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>No. of Participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of life</td>
<td>Assessed risk: Probiotic vs Placebo</td>
<td>0.41 (0.21 to 0.81)</td>
<td>168</td>
<td>very low</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Corresponding risk: Probiotic vs Placebo</td>
<td>0.72 (0.52 to 1.01)</td>
<td>(2 studies)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up: 3 months</td>
<td>The mean change in the intervention group was 6.72 standard deviations lower (1.13 to 2.32 lower)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The basis for the assumed risk (e.g., the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

**CIs:** Confidence intervals

**GRADE Working Group grades of evidence**

*High quality: Further research is very unlikely to change our confidence in the estimate of effect.*

*Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.*

*Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.*

*Very low quality: We are very uncertain about the estimate.*

1. serious imprecision due to different doses and intervention times in the two studies.
2. very serious imprecision due to the small sample size (< 450 individuals) and wide confidence interval.