1. **What if you do a standard intervention review not expecting any useful indirect evidence (or are doing an update) and then find that a NMA would help answer the question e.g. if you have some adverse events data. Would you always do the NMA in a separate review or would it be reasonable to add one post hoc?**
   There are two ways to incorporate an NMA: (1) do it now and be explicit about all post hoc changes and decisions; (2) keep it as a pairwise review, note in the discussion that a NMA might be more appropriate, and conduct NMA in the future update. Methodological and statistical considerations for network meta-analysis need to be given some thought and the author team needs to have the necessary experience.

2. **The rankings are based in statistical basis but not important clinical differences may not reflect a fare position in a ranking. How can we deal with this? Do you think that minimal important differences should be mandatory in the protocol?**
   Minimal important difference should not be mandatory – it depends on the clinical field, but reviewers should be aware their MCID is zero. If the results requires a certain MCID to be meaningful then this should be specified. An approach to produce and present graphically P-scores that accounts for the MCID has been recently developed by Mavridis et al (2019, Biometrical Journal).
   Note that MCID is not an appropriate term here. MCID refers, as we understand, to within-patient differences on an outcome, not on the outcome of group of patients with different drugs. The minimum important difference is what we are after I think.

3. **Why do we need to use the word 'transitivity' if the assumption is the same as the usual pairwise meta-analysis assumption on "sufficient homogeneity" across clinical characteristics as well as statistical? Why not just state that we ensured clinical homogeneity across studies?**
   The authors do not need to use the word transitivity, as long as they understand the issues regarding the assumption, and this is made clear in the relevant PICO section. Transitivity goes beyond homogeneity between two interventions.

4. **Decisions on lumping or splitting can be crucial for heterogeneity and should be guided by clinical reasoning and supported by statistical considerations (eg assessment of heterogeneity once data available). Can this be included as something that will be explored once the data are assembled (guided by clinical experience)?**
   Yes, this can be included as something that will be explored once the data has been extracted or it could be one form of sensitivity analysis. It is important for authors to give this some though during the protocol process and if it is decided that decisions will be made later the authors should include this detail.

5. **When there is evidence of inconsistency, is it enough to just state this? Should the whole review be reconsidered if evidence is not coherent?**
   As for tests for statistical heterogeneity in a standard pairwise meta-analysis, tests for detecting incoherence often lack power to detect incoherence when it is present. Empirical evidence also shows it is often assessed using inappropriate methods. Section 11.4.4.4 of the handbook discusses forming conclusions about incoherence. The whole review does not need to be reconsidered if the evidence is not coherent but possible explanations should be sought and it can also be taken account in assessing the certainty of the evidence.

6. **Have you considered the burden of pairwise assessments for GRADE in large networks? Are you making comprehensive (ie large) networks infeasible?**
   For large networks, it is recommended that CINeMA be used.

7. **For the funnel plot, is there a minimum number of comparisons that you need to have? Is there also a minimum number of studies in each comparison that you would expect?**
   For the number of comparisons, it makes sense to follow the standard guidance of '10', so 10 comparisons for a **comparison adjusted funnel plot**. As long as there are studies to undertake the network meta-analysis, then this is not an issue at this time.

8. **Is it expected that the searches are done within 12 months? It is clear that a NMA takes much longer than a standard review.**
   Yes this is true and I think we just need to assess on a case by case basis.

9. **Would it be better to start off the consideration of suitability at title registration stage?**
   Yes, it would be useful to start thinking about whether a network meta-analysis is appropriate at the title registration stage and whether the author team has the necessary expertise.

10. **How do we manage a review which the authors want to convert to a NMA. Should we start from scratch and get the team to complete a new protocol?**
    Ideally, yes the authors should start with a new protocol as there are many additional issues to consider when undertaking a network meta-analysis. The protocol template will be available soon which should help authors think about relevant issues in each section of the protocol.

11. **How can we identify the minimum heterogeneity, differing from general meta analysis? Should the dosage, frequency be totally the same in one intervention?**
    Heterogeneity should be identified using the standard approaches for pairwise meta-analysis as described in Section 10.10 of the handbook. For NMA, it is common to assume that the amount of heterogeneity is the same for every comparison (section 11.4.3.2) but the dosage and frequency doesn't need to be exactly the same – only similar. In other words, this is a similar consideration as a pairwise meta-analysis. Would you combine different doses/frequency in a pairwise meta-analysis?

12. **What if an NMA is possible for one outcome but not for others?**
    This is often the case. Outcomes that will be analysed using network-meta-analysis should be stated in the protocol. If this is not possible, for example due to lack of data or violation of assumptions, then this should be stated in the review.

13. **Is there a possibility to generate a ranking including multiple outcomes?**
    The approach in the paper mentioned above (Mavridis et al 2019) also produces a ranking that accounts for multiple outcomes (e.g. benefit-risk) by computing cumulative probabilities from a multivariate normal distribution. However, multivariate meta-analysis is not commonly used in practice and it poses problems with the estimation of between-study correlations.

    Other ways of presenting rankings for multiple outcomes are available in the literature. One is, for example, the clustered ranking plot for two outcomes proposed by Chaimani et al (2013, PLOS One) where a cluster analysis of the mean effects or SUCRA values for two different outcomes (e.g. effectiveness and acceptability) is used to "group" treatments according to their similarity with regard to the two outcomes. The treatments can then be plotted based of their SUCRA values for the two outcomes (x and y axis) by using different colours or symbols to represent a different cluster. A different approach was used by Hong and others (2013, Med Dec Making) where the rank probabilities (Bayesian posterior

probabilities of each treatment taking each possible rank) are presented as rankograms for three different outcomes.

14. **Will it be possible to integrate the interactive SoF in RevMan web?**
    This is currently being looked into for ReVMan Web.

15. **It is expected an extension of GRADE for assessing the certainty of evidence in NMA?**
    The certainty of the evidence can be assessed using CINeMA or GRADE. Please see previous webinars for details (https://training.cochrane.org/network-meta-analysis-learning-live-webinar-series).

16. **Is the NMA protocol template in RevMan Web?**
    The NMA protocol template is available as a RevMan 5 file and can be imported into RevMan Web via Archie.