

## ANEXO 1. Metodología Delphi

El cuestionario constaba de 66 aseveraciones (o ítems) divididos en 5 bloques (Tablas 1-5). Las preguntas estaban constituidas por aseveraciones (afirmativas o negativas). Cada ítem sintetizaba algún aspecto sobre el que la evidencia científica resultaba insuficiente

Los ítems se evaluaron por medio de una escala ordinal de nueve puntos (mínimo 1, pleno desacuerdo; y máximo 9, pleno acuerdo), según el formato recomendado por el método de la RAND/UCLA. Las categorías de respuesta se agrupan en tres regiones (1–3 = desacuerdo, 4–6 = ni acuerdo, ni en desacuerdo, 7–9 = acuerdo). El encuestado puede puntuar del 1 al 9 cada ítem pero también puede añadir observaciones o comentarios, expresar dudas en la redacción del ítem o sugerir nuevas propuestas de aseveraciones.

El cuestionario se cumplimentó de manera on-line en dos rondas de votación siguiendo las recomendaciones generales del método Delphi. Tras la primera ronda de votación, se realizó el análisis estadístico de los resultados para conocer si existía consenso en el acuerdo, en el desacuerdo o si no lo había en cada una de las aseveraciones. Aquellas aseveraciones en las que no se alcanzó un consenso fueron votadas de nuevo en una segunda ronda por todos los miembros del panel de expertos que participaron en la primera ronda de votación. Entre ambas rondas, los panelistas fueron informados de los resultados de la votación de la primera ronda. Aquellos ítems que generaron dudas por ser ambiguos o confusos fueron reformulados entre la primera y la segunda ronda por el comité científico. En el análisis de la segunda ronda, se aplicaron los mismos criterios estadísticos que en la primera. Para reducir posibles sesgos, los miembros del comité científico no participaron en la valoración del cuestionario en ninguna de las dos rondas.

### Análisis de resultados según la metodología Delphi

Para considerar una aseveración como consensuada o no consensuada se tuvo en cuenta en qué región de tres puntos se situaba la mediana de las puntuaciones. Se consideró consensuado un ítem en el acuerdo si la mediana estaba en la región 7–9 y se consideró consensuado en el desacuerdo si la mediana estaba incluida en la región 1–3. Si la mediana estaba en la región 4–6 el ítem se consideró no consensuado. Además para considerar un ítem como consensuado se tenían que cumplir otros dos criterios: el número de panelistas que votaba fuera de la región 1–3 o 7–9 tenía que ser menor de  $1/3$  del total y el intervalo intercuartílico (IIC) de la mediana de las respuestas tenía que ser  $\leq 4$ .