Building consensus: thresholds for delivery in TRUFFLE-2 randomized intervention study

Fetal cerebral Doppler changes are associated with adverse neonatal outcome in fetal growth restriction (FGR)[1,2]. However, it remains unclear as to which Doppler thresholds should trigger delivery. As preparation for a randomized intervention trial in late-onset FGR, we describe how we identified Doppler thresholds for cerebral blood flow redistribution (the ratio of umbilical and fetal middle cerebral (MCA) artery PI) at which European experts in FGR were willing to randomize women to immediate delivery or expectant management.

We have identified no publications specifically outlining the most robust methods for reaching consensus regarding randomization thresholds for randomized controlled trials (RCT). Various methods exist[3], including but not limited to the Delphi technique[4], the nominal group technique (NGT)[5] and the consensus development conference[6]. In many instances, the technique is adapted to fit the needs or composition of the group[7]. The Delphi technique, originally developed in the 1950s, is a structured approach to reaching consensus within a group of experts, involving multiple rounds of questionnaires[8]. It can take many months to complete the process, which relies solely on opinion without cross-referencing evidence. The NGT is a structured meeting in which individuals independently list down their ideas and share these within a group, followed by in-depth discussion and finally a private vote[3].

None of these methodologies was suitable for reaching consensus amongst our group as we presented preliminary data from the Trial of Umbilical and Fetal Flow in Europe (TRUFFLE)-2 feasibility study[9] and, hence, suggestions were based on preliminary evidence rather than opinion, and neither technique was practical given the large number of TRUFFLE investigators. Our process of reaching consensus was divided into three stages, as detailed below.

Initial discussion of thresholds

Specific MCA and umbilical artery Doppler PI ranges were first discussed at a study meeting in September 2018 in Turin, Italy, at which 47 investigators were present. It became apparent that participants would be willing to randomize (and potentially deliver) women at an earlier gestational age only for more extreme Doppler values, hence gestational-age-related graduated thresholds were required.

Analysis of feasibility study

The data from a prospective feasibility study were presented at the Turin meeting[9]. A variety of thresholds derived from previous discussions were considered. For each, we compared the incidence of composite adverse outcome between those women who had an abnormal Doppler at least once and those who never had an abnormal Doppler. The Doppler parameter threshold that showed the highest relative risk (RR) for composite adverse outcome was suggested as the most appropriate candidate to be used in a future randomized study.

Consensus meeting

We carried out an exercise with participants from a group of 52 European investigators at a face-to-face meeting in March 2019 in Leuven, Belgium. After presentation of the analysis of the TRUFFLE-2 feasibility data, we used an iterative structured survey followed by voting to determine which umbilical artery and MCA Doppler PI thresholds were acceptable for use in a prospective intervention study. The participants voted on whether they were willing to randomize women at different Doppler PI values for different gestational-age epochs between 32 + 0 and 36 + 6 weeks.

Voting took place using a proprietary web-based application accessible to all attendees on smart devices (https://www.mentimeter.com/). Using a unique survey code, participants were able to view each question as it was projected on the screen, and anonymously and independently cast their vote. Doppler scenarios were presented as combinations of umbilical artery PI and MCA-PI. The corresponding derived ratio of umbilical artery PI to MCA-PI (i.e. the umbilicocerebral ratio (UCR)) was not displayed, thereby encouraging decisions to be based on Doppler parameters that were familiar to the investigators.

Fifty-two participants voted on each of the scenarios. For each question, voting was left open until more than 85% of participants had cast their vote (45 votes or more). The proportion of participants who voted on each scenario ranged from 88% (46/52) to 96% (50/52). Eleven scenarios with varying Doppler thresholds were presented for gestational ages ranging from 32 + 0 to 36 + 6 weeks. For all but one scenario, there was over 70% agreement on the willingness to randomize women (Figure 1).

The response to Question 2 demonstrates that experts were not willing to use a single UCR threshold for all gestational ages (Figure 2). The small majority of participants voted 'no' in Question 2, with a UCR of above 0.9, ruling out the use of 0.8 as an absolute threshold for delivery at all gestational ages. An adjustment of UCR was required at 34 weeks to ensure there was consensus regarding willingness to randomize patients to either immediate or delayed delivery. For all scenarios from 34 weeks of gestation onwards, there was consensus

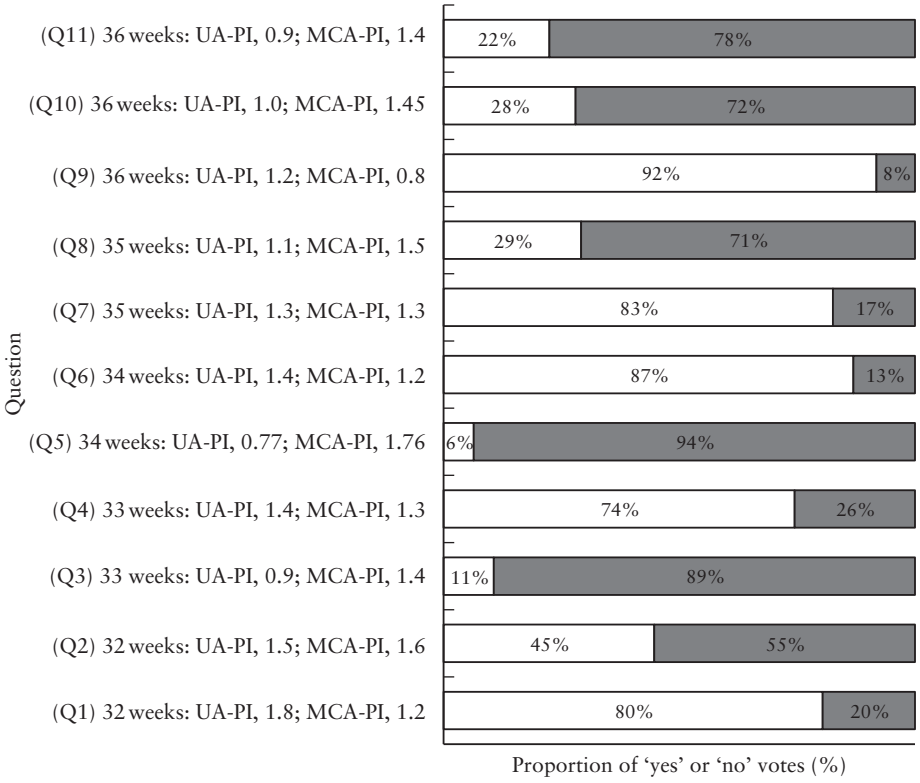

Figure 1 Results of vote on randomization thresholds of umbilical artery (UA) and fetal middle cerebral artery (MCA) pulsatility indices (PI) for willingness to randomize to immediate delivery in pregnancies with late-onset fetal growth restriction, according to gestational age, in TRUFFLE-2 consensus meeting in Leuven, Belgium in March 2019. □, 'yes' (willing to randomize); ■, 'no' (not willing to randomize).

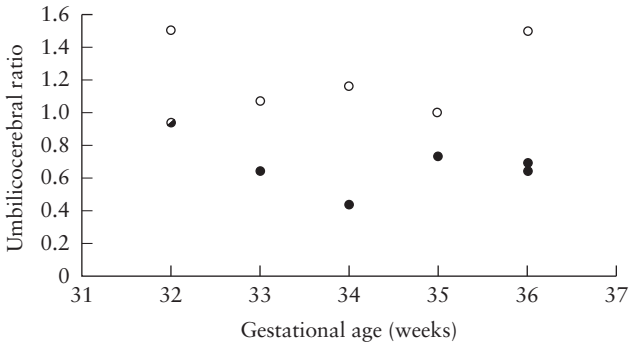

Figure 2 Results of vote on randomization thresholds of umbilical artery and fetal middle cerebral artery pulsatility indices, expressed as umbilicocerebral ratio, for willingness to randomize to immediate delivery in pregnancies with late-onset fetal growth restriction, according to gestational age. ○, majority (≥ 70%) voted 'yes' (willing to randomize); ●, majority voted 'no' (not willing to randomize); ◒, 'divided' vote (no majority agreement reached).

regarding the willingness to randomize patients with Doppler results corresponding to an UCR above 0.8. For scenarios below 34 weeks, there was willingness to randomize women only for a more extreme UCR of above 1.0.

We have shown how knowledge from an observational study can be used to inform a consensus process for determining thresholds for randomization in a RCT. The TRUFFLE-2 investigators demonstrated a spontaneous method of reaching consensus, which included aspects of the NGT and consensus development conference, in which participants were given suggestions from analysis of the feasibility study, an in-depth discussion was facilitated and an anonymous vote was held. We believe this to have been a practical and robust method for defining randomization thresholds. The strengths of this method are the multimodal approach to reaching consensus and derivation of Doppler thresholds based on evidence, not solely on expert opinion. Furthermore, by making these deliberations public, we aim to disseminate the results of this consensus-building process.

A potential weakness is that the consensus method was developed prospectively and, hence, was not defined clearly in advance[3]. Although not specified *a priori*, 70% agreement has generally been defined as achieving consensus in published literature[4,10,11]. Furthermore, the process was carried out by European experts in FGR from the TRUFFLE-2 Study Group. It might therefore be argued that the experts were 'self-defined'. There is discussion over who is regarded as an expert in the Delphi process. Fink suggested that 'An expert should be a representative of their professional group, with either sufficient expertise not to be disputed or the power required to instigate the findings'[12]. The participants who voted on Doppler thresholds comprised independent fetal medicine specialists with a particular interest in FGR and its management, attending from centers across 11 European countries.

Final randomization thresholds chosen for the TRUFFLE-2 RCT are UCR above 1.0 from 32 + 0 to 33 + 6 weeks and UCR above 0.8 from 34 + 0 to 36 + 6 weeks. At the end of the consensus-building process, it was confirmed that all participants were satisfied with the result, which not only confirms the consensus, but also strengthens the legitimacy of the protocol for the planned multicenter interventional trial. By involving expert investigators in the process of defining randomization thresholds, hence giving all ownership of the protocol, we hope to improve recruitment and adherence to the protocol in the future randomized study. Simultaneously, it reassures participants and those external to the study that the choices of Doppler thresholds are based both on the best available evidence to date and on those that expert clinicians would be willing to use in their routine clinical practice.

B. Mylrea-Foley[1], A. Bhide[2], E. Mullins[1], J. Thornton[3], N. Marlow[4], T. Stampalija[5,6], R. Napolitano[4,7] and C. C. Lees[1,8]*

[1]*Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK;*
[2]*Fetal Medicine Unit, St George's University Hospitals NHS Foundation Trust and Molecular & Clinical Sciences Research Institute, St George's University of London, London, UK;*
[3]*School of Clinical Sciences, University of Nottingham, Division of Obstetrics and Gynaecology, Maternity Department, City Hospital, Nottingham, UK;*
[4]*UCL Elizabeth Garrett Anderson Institute for Women's Health, University College London, London, UK;*
[5]*Unit of Fetal Medicine and Prenatal Diagnosis, Institute for Maternal and Child Health, IRCCS Burlo Garofolo, Trieste, Italy;*
[6]*Department of Medicine, Surgery and Health Sciences, University of Trieste, Trieste, Italy;*
[7]*Fetal Medicine Unit, University College London Hospitals NHS Foundation Trust, London, UK;*
[8]*Centre for Fetal Care, Queen Charlotte's and Chelsea Hospital, Imperial College NHS Trust, London, UK*
*Correspondence.
(e-mail: c.lees@imperial.ac.uk)

References

1. Meher S, Hernandez-Andrade E, Basheer SN, Lees C. Impact of cerebral redistribution on neurodevelopmental outcome in small-for-gestational-age or growth-restricted babies: a systematic review. *Ultrasound Obstet Gynecol* 2015; **46**: 398–404.
2. Stampalija T, Arabin B, Wolf H, Bilardo CM, Lees C, Brezinka C, Derks JB, Diemert A, Duvekot JJ, Ferrazzi E, Frusca T. Is middle cerebral artery Doppler related to neonatal and 2-year infant outcome in early fetal growth restriction? *Am J Obstet Gynecol* 2017; **216**: 521.e1–13.
3. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: Characteristics and guidelines for use. *Am J Public Health* 1984; **74**: 979–983.
4. Hsu CC, Sandford BA. The Delphi technique: Making sense of consensus. *Pract Assess Res Eval* 2007; **12**: 1–8.
5. Harvey N, Holmes CA. Nominal group technique: An effective method for obtaining group consensus. *Int J Nurs Pract* 2012; **18**: 188–194.
6. Black N, Murphy M, Lamping D, McKee M, Sanderson C, Askham J, Marteau T. Consensus development methods: A review of best practice in creating clinical guidelines. *J Heal Serv Res Policy* 1999; **4**: 236–248.
7. Schneider P, Evaniew N, Rendon JS, McKay P, Randall RL, Turcotte R, Vélez R, Bhandari M, Ghert M. Moving forward through consensus: Protocol for a modified Delphi approach to determine the top research priorities in the field of orthopaedic oncology. *BMJ Open* 2016; **6**: e011780.
8. Delbecq AL, Van de Ven AH, Gustafson DH. Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes. *Group & Organization Studies* 1975; **1**: 180.
9. Stampalija T, Thornton J, Marlow N, Napolitano R, Bhide A, Pickles T, Bilardo CM, Gordijn SJ, Gyselaers W, Valensise H, Hecher K, Sande RK, Lindgren P, Bergman E, Arabin B, Breeze AC, *et al.*, on behalf of the TRUFFLE-2 Group. Fetal cerebral Doppler changes and outcome in late preterm fetal growth restriction: prospective observational cohort study. *Ultrasound Obstet Gynecol* 2020; **56**: 173–181.
10. Slade SC, Dionne CE, Underwood M, Buchbinder R. Standardised method for reporting exercise programmes: Protocol for a modified Delphi study. *BMJ Open* 2014; **4**: e006682.
11. Kleynen M, Braun SM, Bleijlevens MH, Lexis MA, Rasquin SM, Halfens J, Wilson MR, Beurskens AJ, Masters RSW. Using a Delphi Technique to Seek Consensus Regarding Definitions, Descriptions and Classification of Terms Related to Implicit and Explicit Forms of Motor Learning. *PLoS One* 2014; **9**: e100227.
12. Fink A. *The Survey Handbook - The Survey Kit* (2nd edn), Laughton D, Novak V, Selhorst J, Foster D (eds). SAGE: Thousand Oaks, CA, 2003; 184.