A core outcome set for the evaluation of treatments for twin-twin transfusion syndrome

Helen Perry[1,2], James M N Duffy[3,4], Keith Reed[5], Ahmet Baschat[6], Jan Deprest[7,8], Kurt Hecher[9], Liesbeth Lewi[7], Enrico Lopriore[10], Dick Oepkes[11], Asma Khalil[1,2]
On behalf of the International Collaboration to Harmonise Outcomes for Twin-Twin Transfusion Syndrome (CHOOSE)

[1] Vascular Biology Research Centre, Molecular and Clinical Sciences Research Institute, St George's University of London, Cranmer Terrace, London SW17 0RE, United Kingdom.
[2] Fetal Medicine Unit, Department of Obstetrics and Gynaecology, St George's University Hospitals NHS Foundation Trust, Blackshaw Road, London SW17 0QT, United Kingdom.
[3] Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Rd, Oxford OX2 6GG, United Kingdom.
[4] Balliol College, University of Oxford, Oxford OX1 3BJ, United Kingdom.
[5] Twin and Multiple Births Association (TAMBA), The Manor House, Manor Park, Church Hill, Aldershot GU12 4JU, United Kingdom.
[6] The Johns Hopkins Center for Fetal Therapy, 600 North Wolfe, Nelson 228 Baltimore, MD 21287, Maryland, USA.
[7] Department of Obstetrics and Gynecology, University Hospitals of KU Leuven, Herestraat 49 3000 Leuven, Belgium.
[8] Institute for Women's Health, University College London, London, United Kingdom.
[9] Department of Obstetrics and Fetal Medicine, University Medical Center Hamburg-Eppendorf, Neues Klinikum, Gebäude O10 Martinistr. 52, 20246 Hamburg, Germany.
[10] Department of Pediatrics, Leiden University Medical Center, K-06-35 P.O. Box 9600, 2300RC Leiden, The Netherlands.
[11] Division of Fetal Medicine, Department of Obstetrics, Leiden University Medical Center, K-06-35 P.O. Box 9600, 2300RC Leiden, The Netherlands.

Correspondence to: Professor Asma Khalil MRCOG MD MSc (Epi) DFSRH Dip (GUM) Fetal Medicine Unit, Department of Obstetrics and Gynaecology, St. George's University Hospitals NHS Foundation Trust, Blackshaw Road, London SW17 0QT, United Kingdom. akhalil@sgul.ac.uk

Running Head: TTTS core outcome set
Key Words: Consensus study, core outcome set, Twin-Twin Transfusion Syndrome, modified Delphi method, and modified nominal group technique.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/uog.20183

ABSTRACT

Background: A minimum data set, known as a core outcome set, should standardise outcome selection, collection, and reporting across future twin-to-twin transfusion syndrome research.

Methods: An international steering group including healthcare professionals, researchers, and patients, guided the development of this core outcome set. Potential core outcomes, identified through a comprehensive literature review, were entered into a three round Delphi survey. Healthcare professionals, researchers, and patients were invited to participate. Consensus outcomes were entered into a modified nominal group technique to identify the final core outcome set.

Results: One hundred and three participants, from 29 countries, participated in a three-round Delphi survey. Of those 88 completed the three rounds. Twenty-two consensus outcomes were identified and entered into a modified nominal group technique. Eleven healthcare professionals, two researchers, and three patients prioritised 12 core outcomes. Fetal core outcomes included live birth, fetal loss (including miscarriage, stillbirth, and termination of pregnancy), recurrence of twin-twin transfusion syndrome, twin anaemia polycythaemia syndrome and amniotic band syndrome. Neonatal core outcomes included gestational age at delivery, birthweight, neonatal mortality, brain injury syndromes, and ischaemic limb injury. Maternal core outcomes included maternal mortality and admission to level two or three care.

Conclusions: Embedding the core outcome set within future twin-to-twin transfusion syndrome research could make a substantial contribution to advancing the usefulness of research. Standardised definitions and measurement instruments are now required for individual core outcomes.

INTRODUCTION

Management options for twin-to-twin transfusion syndrome (TTTS) have evolved rapidly over the last 40 years and include expectant management, termination of pregnancy, amnioreduction, septostomy and fetoscopic laser surgery. Although fetoscopic laser surgery has been shown to be associated with improved neurological outcomes[1], different techniques continue to be evaluated and non-invasive experimental therapies are being developed.[2,3] Potential interventions require robust evaluation.

Across women's health, complex issues including a failure to take into account the perspectives of key stakeholders when selecting outcomes, variations in outcome definitions and measurement instruments, and outcome reporting bias makes research evidence difficult to interpret, undermining the translation of research into clinical practice.[4–9] A recent systematic review of published TTTS research has demonstrated many of these issues.[10]

Standardisation of outcome selection, collection, and reporting could help reduce research waste and improve the quality of evidence by allowing comparison and combination of results from individual studies and reducing potentially harmful outcome-reporting bias and study publication bias.[11–15] A core outcome set is a collection of standardised outcomes and outcome measures for any given condition.[8] They can be collected in a standardised manner and reported consistently in publications.[16–18] Development of a core outcome set requires identifying potential outcomes for inclusion, determining core outcomes by taking into account the views of key stakeholders and finally, determining how the included outcomes should be measured.[18]

Given the development of new treatments for TTTS and the rarity of the condition, a core outcome set is required to ensure the results of all studies can be compared, contrasted, and combined to guide future clinical practice. Therefore, the objective of this study was to develop a core outcome set for TTTS.

METHODS

The study was prospectively registered with The Core Outcome Measures in Effectiveness Trials (COMET) initiative (registration number: 921) and the detailed protocol for the development of this core outcome set has been previously published.[19] The protocol was informed by the COMET initiative handbook and other core outcome set development studies.[18,20–27] The National Research Ethics Service has advised that ethical approval is not required for this study.

Information Sources

A systematic review of the literature was performed by searching Cochrane Central Register of Controlled Trials (CENTRAL), Embase, and Medline from inception to August 2016.[10] The detailed methods are reported elsewhere, briefly, outcomes were extracted from randomised trials and observational studies evaluating treatments for TTTS in monochorionic-diamniotic twin pregnancies and monochorionic-triamniotic, and dichorionic-triamniotic triplet pregnancies.[28] Additionally, all steering group members were requested to submit any outcomes they felt should be included in the Delphi survey which had not been reported in previous published studies. Lay definitions were developed for individual outcomes informed by previous core outcome set development studies, published patient information, and expert opinion. The definitions were reviewed by public communication experts from the Twin and Multiple Birth Association prior to use.

Consensus-building process

We used a three-round Delphi survey designed using online software (DelphiManager, University of Liverpool). The DelphiManager is established software used to deliver online Delphi surveys.[29,30] The survey was piloted by representatives from each stakeholder group. The first round of the survey was open for four weeks with reminders sent to participants to complete data entry on days seven, 14, and 21. The second round was open for four weeks with personalised reminders sent for completion of data entry. The final round remained open for four weeks.

On registering to take part in the survey, participants were invited to complete a demographic questionnaire. Participants were asked to score individual outcomes on a nine-point Likert scale anchored between one (labelled 'of limited importance for making a decision') and nine (labelled 'critical for making a decision'). This scale was devised by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group to facilitate the ranking of outcomes according to their importance and has been adopted widely by core outcome set developers.[31] At the end of the first round, participants were invited to suggest additional outcomes for consideration in subsequent rounds of the survey. For the second and third round of the survey, the percentage of scores by different stakeholder groups were displayed in addition to the individual participants previous scores (Figure 1).

Consensus was defined a priori using the COMET initiative's 15%/70% definition:

[1] Consensus in (classify as a core outcome): Over 70% of participants in each stakeholder group score outcome 'critical for decision making' (score seven to nine) and less than 15% of participants in each stakeholder group score outcome 'of limited importance for decision making' (score one to three).

[2] Consensus out (do not classify as a core outcome): Over 70% of participants in each stakeholder group score outcome domain 'of limited importance for decision making' (score one to three) and less than 15% of participants in each stakeholder group score outcome domain 'critical for decision making' (score seven to nine); or

[3] No Consensus (do not classify as a core outcome): Anything else.[18]

Participants

Participation in the Delphi survey was voluntary. A steering group committee comprising specialists in the field of TTTS determined that three main stakeholder groups were required to inform the core outcome set; patients or relatives with experience of TTTS, healthcare professionals with expertise in TTTS and researchers with expertise in TTTS. Adverts were placed on social media platforms by steering group members and Twin and Multiple Birth Association prompting participation. Email invitations were sent to members of the International Society for Twin Studies, participants of a previous Delphi survey, and colleagues and contacts of the steering group.

Analysis of Results

The survey response results and demographic details were analysed using statistical software (SPSS 25.0; SPSS Inc., Chicago, Illinois, Untied States). The results after the third round were displayed as the percentage of stakeholders that scored the outcome as critical (7-9).

Modified nominal group technique

We elected to use the modified nominal group technique in a face-to-face consultation meeting for the development of this core outcome set. This technique allows all opinions to be considered at the start and, through discussion, presents information in a hierarchy of perceived importance.[32] Other core outcome sets have used this technique successfully.[33,34] Healthcare professionals, researchers, and patients, who lived in the United Kingdom and had completed all three rounds of the Delphi survey were invited to participate and participants from overseas were invited to join by teleconference.

The modified nominal group technique was delivered through a half-day consensus development meeting. Prior to the meeting, patient participants were offered the opportunity to discuss the study's background, the Delphi survey results, the scope of the meeting and the outcome terminology. The meeting was chaired by an experienced facilitator. Attendees provided demographic details, including age, gender, and ethnic group. At the start of the meeting, the study background and aims of the meeting were outlined and the results of the Delphi survey were presented. All potential core outcomes reaching the standardised consensus definition were entered into the process and participants were able to enter other potential core outcomes which had not reached the standardised consensus definition. Each participant was asked to explain which outcomes they felt most and least strongly about including in the final core outcome set. Following the initial discussion, outcomes were divided into three initial categories: (1) outcomes to be considered for inclusion in the final core outcome set; (2) outcomes where opinions on inclusion were divided; and (3) outcomes which should not be considered for inclusion in the final core outcome set. Participants were invited to discuss the ordering of the outcomes within each category and could move outcomes between the categories. During the discussion, participants were encouraged to consider reformulating outcomes that were similar or where it would improve clarity. They were asked to consider the relative importance of individual outcomes in relation to each other, the feasibility of collecting and reporting each outcome and the availability of suitable definitions and measurement instruments for each outcome. Following the discussion, the final core outcome set was agreed.

RESULTS

Modified Delphi method

One hundred and three participants, from 29 different countries, completed the round one survey (Table 1). Eighty-nine participants completed round two and 88 participants completed all three rounds.

Seventy-one outcomes, organised into six thematic domains, were entered into the first round of the Delphi survey. Twenty-one additional outcomes that were suggested by participants during round one were added to the second and third rounds.

Following completion of the three-round survey, 22 outcomes fulfilled the *a priori* consensus definition (Table 2). No outcomes met the exclusion criteria. The full scoring for round three is shown in Supplementary appendix 1. The flow of outcomes and participants are shown in Figure 2.

Nominal group technique

Sixteen people attended the stakeholder consensus meeting. Attendees included patients (n=3), healthcare professionals (n=11), and researchers (n=2) (table 3). Twenty-two consensus and four additional outcomes were discussed in the meeting. Eight outcomes were reformulated. The group prioritised 12 core outcomes and one aspirational outcome.

Core outcomes included:

- Live birth
- Loss during pregnancy or before hospital discharge including late miscarriage, termination of pregnancy, stillbirth, and neonatal mortality
- Subsequent death of a co-twin following single twin demise at the time of treatment
- Recurrence of TTTS
- Twin Anaemia Polycythaemia Syndrome;
- Amniotic band syndrome
- Gestational age at birth
- Birthweight
- Neonatal brain injury (arterial or venous infarction, cystic periventricular leukomalacia ≥ grade II, intraventricular haemorrhage ≥grade III, porencephalic cysts, severe ventricular dilatation ≥97th centile)
- Ischaemic limb injury
- Maternal mortality
- Requirement for admission to a level two (representing the need for single organ support, excluding mechanical ventilation) or three (representing the need for ≥ two organ support, or mechanical ventilation) care setting

Aspirational Outcome:

- Neurodevelopment at 18 to 24 months of age.

DISCUSSION

Summary of study findings

Using consensus science methods, 57 healthcare professionals, 18 researchers, and 28 people with lived experience of TTTS from 29 countries contributed to the development of a core outcome set for TTTS. Using a modified Delphi method, the long list of 92 potential core outcomes was reduced to 22 consensus outcomes. Using a modified nominal group technique, 22 consensus outcomes, and four additional outcomes suggested by participants were prioritised to 12 core outcomes and one aspirational outcome.

Most studies evaluating treatments for TTTS have not reported these clinically important outcomes consistently. For example, 17 (17%) reported recurrence of TTTS, 31 (31%) reported live birth, and two studies (2%) reported maternal mortality.[10] Implementation of this core outcome set for TTTS should ensure that future research reports the outcomes deemed important to all stakeholders, limit selective outcome reporting, and informs the development of clinical practice guidelines. The Core Outcomes in Women's and Newborn's Health (CROWN) initiative, supported by 80 speciality journals, including *Ultrasound in Obstetrics and Gynaecology*, is a resolve to implement core outcome sets and as such, researchers should be motivated to consider this TTTS core outcome set when designing future studies in order to comply with publication requirements.[35]

Strengths and Limitations

Our study has several strengths. Firstly, we used robust consensus-building methodology, informed by the COMET Initiative and other core outcome sets,[18,20–25] to converge a long list of potential core outcomes, derived from a systematic review and expert opinion, into a focus of twelve, clinically-important outcomes for TTTS. Secondly, 103 stakeholders from 29 countries of five continents took part in the Delphi survey, demonstrating global participation in the study. The COMET Initiative has highlighted that in the majority of core outcome sets, participants come only from Europe and/or North America (154 studies, 68%) with participants from the remaining continents participating in less than a third of studies (73 studies, 32%). [36,37] Finally, this study included people with lived experience of TTTS at every stage. Patients assisted with the design of the study by shaping which outcomes were included in the Delphi survey, participated in each round of the Delphi survey and participated in the modified nominal group technique during the consensus-development meeting. This should ensure the final core outcome set is relevant to all stakeholders of TTTS and subsequent research is designed to support the views of parents as well as healthcare professionals and researchers.

There are some limitations to our study. Firstly, participation in the Delphi survey was dependent on both an understanding of the English language and computer and internet access and proficiency. This may have limited potential participants from taking part. Translating the survey into other languages and providing the survey in a paper or interview format are potential solutions to this limitation. However, when designing this study, we balanced the ideal design with the resources available and made a pragmatic decision to use the online format in English only. Secondly, although we achieved better global participation than many other core outcome sets,[36,37] the large majority (n= 70, 68%) of participants were from Europe with only six participants (5.8%) each from both South America and Asia. Similarly, it would be preferable to have more equal representation in the different stakeholder groups, and the imbalance between professionals and patients or

relatives could have created bias. However, as previously mentioned, we have included patients at each stage of this study to ensure their opinions were heard. Leaving the survey open for longer may have increased the number of patient participants, but we found little increase despite the reposting of social media adverts after the first time. Finally, there remains some uncertainty regarding aspects of consensus-building methodology, including the most appropriate definition of consensus.[18,38] Whilst the best definition is still unknown, we feel that by defining consensus *a priori*, and reporting our results in line with this agreed definition, we have assured validity in our methodology.

Clinical and research implications

The final stage of this project will be to determine definitions and measurement instruments for each outcome. This will be done in a systematic, objective manner following recommendations from the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative.[39] Potential definitions will be collated from formal definition development initiatives, national and international guidelines and systematic reviews and randomised controlled trials. We have previously reported variation in the definitions of outcomes across studies of treatments for TTTS and will consider each of these definitions in the process.[10] This inventory of potential definitions will be entered into a consensus-building workshop attended by healthcare professionals, researchers and patients with experience of TTTS. Potential measurement instruments will be identified from national and international guidelines and systematic reviews and randomised controlled trials and quality assessed using COSMIN framework.[39]

It should be acknowledged that this is a *core* outcome set and researchers are not limited to reporting only these outcomes depending on the scope of their study. In time, it may become clear that other outcomes should be routinely reported, for example if currently experimental therapies become mainstay treatments.[3] By formalising neurodevelopmental impairment as an aspirational outcome, we hope that researchers and funders will recognise the importance of designing and funding future studies to allow this to be reported.

Conclusion

This core outcome set for studies evaluating treatments for TTTS has been determined through a systematic review of existing literature and a consensus-building exercise. It should inform future research studies allowing easier comparison and combination of results. Clear definitions and tools for measurement are now required for each included outcome.

CHOOSE collaborators

Christina Coroyannakis, Fetal Medicine Unit, Department of Obstetrics and Gynaecology, St George's University Hospitals NHS Foundation Trust, Blackshaw Road, London SW17 0QT, United Kingdom.

Donovan Duffy, Neonatal Unit, St George's University Hospitals NHS Foundation Trust, Blackshaw Road, London SW17 0QT, United Kingdom.

Elisa Giallongo, Fetal Medicine Unit, Department of Obstetrics and Gynaecology, St George's University Hospitals NHS Foundation Trust, Blackshaw Road, London SW17 0QT, United Kingdom.

Helen Turier, Twin and Multiple Births Association (TAMBA), The Manor House, Manor Park, Church Hill, Aldershot GU12 4JU, United Kingdom.

Joseph Toms, Fetal Medicine Unit, Department of Obstetrics and Gynaecology, St George's University Hospitals NHS Foundation Trust, Blackshaw Road, London SW17 0QT, United Kingdom.

Karen Lundie, West Middlesex University Hospital and Chelsea Westminster Hospital NHS Foundation Trust, Twickenham Road, Isleworth, TW7 6AF

Kujan Paramanantham

Luke Burton

Sophie Windsor, Lewisham and Greenwich NHS Trust, Queen Elizabeth Hospital , Stadium Road, Woolwich, London, SE18 4QH

Sijo Francis, Neonatal Unit, St George's University Hospitals NHS Foundation Trust, Blackshaw Road, London SW17 0QT, United Kingdom.

ACKNOWLEDGEMENTS

We would like to thank all participants of the Delphi Survey for sharing their opinions and giving their time to this study. We would like to acknowledge the following people for giving their time and expertise in the consensus-building meeting: Christina Coroyannakis, Donovan Duffy, Elisa Giallongo, Helen Turier, Joseph Toms, Karen Lundie, Kujan Paramanantham, Luke Burton, Sophie Windsor and Sijo Francis.

FUNDING

This study was supported by TAMBA. HP is supported by a grant from the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South London at King's College Hospital NHS Foundation Trust. The views expressed

are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. JD is funded by the Great Ormond Street Hospital Charity fund, the Wellcome Trust (WT101957) and Engineering and Physical Sciences Research Council (ESPRC) (NS/A000027/1) via GIFT-surg. LL is a clinical researcher funded by the Fonds Wetenschappelijk Onderzoek Vlaanderen (FWO 1804718N).

REFERENCES

1. Roberts D, Neilson JP, Kilby MD, Gates S. Interventions for the treatment of twin-twin transfusion syndrome. *Cochrane Database Syst Rev*. 2014;(1):CD002073. doi:10.1002/14651858.CD002073.pub3.
2. Slaghekke F, Lopriore E, Lewi L, Middeldorp JM, van Zwet EW, Weingertner A-S, Klumper FJ, DeKoninck P, Devlieger R, Kilby MD, Rustico MA, Deprest J, Favre R, Oepkes D. Fetoscopic laser coagulation of the vascular equator versus selective coagulation for twin-to-twin transfusion syndrome: an open-label randomised controlled trial. *Lancet (London, England)*. 2014;383(9935):2144-2151. doi:10.1016/S0140-6736(13)62419-8.
3. Shaw CJ, Civale J, Botting KJ, Niu Y, ter Haar G, Rivens I, Giussani DA, Lees CC. Noninvasive high-intensity focused ultrasound treatment of twin-twin transfusion syndrome: A preliminary in vivo study. *Sci Transl Med*. 2016;8(347):347ra95-347ra95. doi:10.1126/scitranslmed.aaf2135.
4. Duffy J, Hirsch M, Pealing L, Showell M, Khan KS, Ziebland S, McManus RJ, International Collaboration to Harmonise Outcomes in Pre-eclampsia (iHOPE). Inadequate safety reporting in pre-eclampsia trials: a systematic evaluation. *BJOG*. 2018;125(7):795-803. doi:10.1111/1471-0528.14969.
5. Duffy JMN, Hirsch M, Gale C, Pealing L, Kawsar A, Showell M, Williamson PR, Khan KS, Ziebland S, McManus RJ, International Collaboration to Harmonize Outcomes for Pre-eclampsia (iHOPE). A systematic review of primary outcomes and outcome measure reporting in randomized trials evaluating treatments for pre-eclampsia. *Int J Gynaecol Obstet*. 2017;139(3):262-267. doi:10.1002/ijgo.12298.
6. Duffy J, Hirsch M, Kawsar A, Gale C, Pealing L, Plana MN, Showell M, Williamson PR, Khan KS, Ziebland S, McManus RJ, iHOPE: International Collaboration to Harmonise Outcomes in Pre-Eclampsia. Outcome reporting across randomised controlled trials evaluating therapeutic interventions for pre-eclampsia. *BJOG*. 2017;124(12):1829-1839. doi:10.1111/1471-0528.14702.

7. Hirsch M, Duffy JMN, Kusznir JO, Davis CJ, Plana MN, Khan KS, International Collaboration to Harmonize Outcomes and Measures for Endometriosis JMN, Farquhar C, Hirsch M, Johnson N, Khan K. Variation in outcome reporting in endometriosis trials: a systematic review. *Am J Obstet Gynecol*. 2016;214(4):452-464. doi:10.1016/j.ajog.2015.12.039.

8. Duffy J, Rolph R, Gale C, Hirsch M, Khan K, Ziebland S, McManus R, International Collaboration to Harmonise Outcomes in Pre-eclampsia (iHOPE). Core outcome sets in women's and newborn health: a systematic review. *BJOG An Int J Obstet Gynaecol*. 2017;124(10):1481-1489. doi:10.1111/1471-0528.14694.

9. Duffy JMN, Ziebland S, von Dadelszen P, McManus RJ. Tackling poorly selected, collected, and reported outcomes in obstetrics and gynecology research. *Am J Obstet Gynecol*. September 2018. doi:10.1016/j.ajog.2018.09.023.

10. Perry H, Duffy JMN, Umadia O, Khalil A, International Collaboration to Harmonise Outcomes for Twin-Twin Transfusion Syndrome (CHOOSE). Outcome reporting across randomised trials and observational studies evaluating treatments for Twin-Twin Transfusion Syndrome: a systematic review. *Ultrasound Obstet Gynecol*. April 2018. doi:10.1002/uog.19068.

11. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan A-W, Cronin E, Decullier E, Easterbrook PJ, Von Elm E, Gamble C, Ghersi D, Ioannidis JPA, Simes J, Williamson PR. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias. Siegfried N, ed. *PLoS One*. 2008;3(8):e3081. doi:10.1371/journal.pone.0003081.

12. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet (London, England)*. 2009;374(9683):86-89. doi:10.1016/S0140-6736(09)60329-9.

13. Duffy J, Bhattacharya S, Herman M, Mol B, Vail A, Wilkinson J, Farquhar C, Cochrane Gynaecology and Fertility Group. Reducing research waste in benign gynaecology and fertility research. *BJOG An Int J Obstet Gynaecol*. 2017;124(3):366-369. doi:10.1111/1471-0528.14438.

14. Wilkinson J, Bhattacharya S, Duffy JMN, Kamath MS, Marjoribanks J, Repping S, Vail A, van Wely M, Farquhar CM. Reproductive medicine: Still more ART than science? *BJOG An Int J Obstet Gynaecol*. July 2018. doi:10.1111/1471-0528.15409.

15. Sielo F, Duffy JM, Townsend R, Khalil A. Addressing the variation in outcome reporting in high risk twin studies: The key to reducing research waste and improving clinical care. *Ultrasound Obstet Gynaecol*. 2018;(in press).

16. Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials*. 2007;8:39. doi:10.1186/1745-6215-8-39.

17. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Tugwell P. Developing core outcome sets for clinical trials: issues to consider. *Trials*. 2012;13(1):132. doi:10.1186/1745-6215-13-132.

18. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, Clarke M, Gargon E, Gorst S, Harman N, Kirkham JJ, McNair A, Prinsen CAC, Schmitt J, Terwee CB, Young B. The COMET Handbook: version 1.0. *Trials*. 2017;18(Suppl 3):280. doi:10.1186/s13063-017-1978-4.

19. Khalil A, Perry H, Duffy J, Reed K, Baschat A, Deprest J, Hecher K, Lewi L, Lopriore E, Oepkes D, International Collaboration to Harmonise Outcomes for Twin–Twin Transfusion Syndrome (CHOOSE). Twin-Twin Transfusion Syndrome: study protocol for developing, disseminating, and implementing a core outcome set. *Trials*. 2017;18(1):325. doi:10.1186/s13063-017-2042-0.

20. van 't Hooft J, Duffy JMN, Daly M, Williamson PR, Meher S, Thom E, Saade GR, Alfirevic Z, Mol BWJ, Khan KS, Global Obstetrics Network (GONet). A Core Outcome Set for Evaluation of Interventions to Prevent Preterm Birth. *Obstet Gynecol*. 2016;127(1):49-58. doi:10.1097/AOG.0000000000001195.

21. Duffy JMN, van 't Hooft J, Gale C, Brown M, Grobman W, Fitzpatrick R, Karumanchi SA, Lucas N, Magee L, Mol B, Stark M, Thangaratinam S, Wilson M, von Dadelszen P, Williamson P, Khan KS, Ziebland S, McManus RJ, International Collaboration to Harmonise Outcomes for Pre-eclampsia (iHOPE). A protocol for developing, disseminating, and implementing a core outcome set for pre-eclampsia. *Pregnancy Hypertens*. 2016;6(4):274-278. doi:10.1016/j.preghy.2016.04.008.

22. Hirsch M, Duffy JMN, Barker C, Hummelshoj L, Johnson NP, Mol B, Khan KS, Farquhar C, International Collaboration to Harmonize Outcomes and Measures for Endometriosis (iHOME). Protocol for developing, disseminating and implementing a core outcome set for endometriosis. *BMJ Open*. 2016;6(12):e013998. doi:10.1136/bmjopen-2016-013998.

23. Whitehouse KC, Kim CR, Ganatra B, Duffy JMN, Blum J, Brahmi D, Creinin MD, DePiñeres T, Gemzell-Danielsson K, Grossman D, Winikoff B, Gülmezoglu AM. Standardizing abortion research outcomes (STAR): a protocol for developing, disseminating and implementing a core outcome set for medical and surgical abortion. *Contraception*. 2017;95(5):437-441. doi:10.1016/j.contraception.2016.12.009.

24. Webbe J, Brunton G, Ali S, Duffy JM, Modi N, Gale C. Developing, implementing and disseminating a core outcome set for neonatal medicine. *BMJ Paediatr Open*. 2017;1(1):e000048. doi:10.1136/bmjpo-2017-000048.

25. Duffy JMN, Bhattacharya S, Curtis C, Evers JLH, Farquharson RG, Franik S, Khalaf Y, Legro RS, Lensen S, Mol BW, Niederberger C, Ng EHY, Repping S, Strandell A, Torrance HL, Vail A, van Wely M, Vuong NL, Wang AY, Wang R, Wilkinson J, Youssef MA, Farquhar CM. A protocol developing, disseminating and implementing a core outcome set for infertility. *Hum Reprod Open*. 2018;2018(3). doi:10.1093/hropen/hoy007.

26. Pergialiotis V, Durnea C, Elfituri A, Duffy J, Doumouchtsis SK, International Collaboration for Harmonising Outcomes, Research, and Standards in Urogynaecology and Women's Health (CHORUS). Do we need a core outcome set for childbirth perineal trauma research? A systematic review of outcome reporting in randomised trials evaluating the management of childbirth trauma. *BJOG*. 2018;125(12):1522-1531. doi:10.1111/1471-0528.15408.

27. Durnea CM, Pergialiotis V, Duffy JMN, Bergstrom L, Elfituri A, Doumouchtsis SK, CHORUS, an International Collaboration for Harmonising Outcomes, Research and Standards in Urogynaecology and Women's Health. A systematic review of outcome and outcome-measure reporting in randomised trials evaluating surgical interventions for anterior-compartment vaginal prolapse: a call to action to develop a core outcome set. *Int Urogynecol J*. October 2018. doi:10.1007/s00192-018-3781-5.

28. Duffy J, McManus RJ. Influence of methodology upon the identification of potential core outcomes: recommendations for core outcome set developers are needed. *BJOG*. 2016;123(10):1599. doi:10.1111/1471-0528.14219.

29. Sinha IP, Smyth RL, Williamson PR. Using the Delphi Technique to Determine Which Outcomes to Measure in Clinical Trials: Recommendations for the Future Based on a Systematic Review of Existing Studies. *PLoS Med*. 2011;8(1):e1000393. doi:10.1371/journal.pmed.1000393.

30. Hsu C-C, Sandford BA. The Delphi Technique: Making Sense of Consensus - Practical Assessment, Research & Evaluation. *Pract Assessment, Res Eval*. 2007;12(10). http://pareonline.net/getvn.asp?v=12&n=10. Accessed June 18, 2018.

31. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schünemann HJ. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011;64(4):395-400. doi:10.1016/j.jclinepi.2010.09.012.

32. Harvey N, Holmes CA. Nominal group technique: An effective method for obtaining group consensus. *Int J Nurs Pract*. 2012;18(2):188-194. doi:10.1111/j.1440-172X.2012.02017.x.

33. Haywood KL, Griffin XL, Achten J, Costa ML. Developing a core outcome set for hip fracture trials. *Bone Joint J*. 2014;96-B(8):1016-1023. doi:10.1302/0301-620X.96B8.33766.

34. Ruperto N, Ravelli A, Murray KJ, Lovell DJ, Andersson-Gare B, Feldman BM, Garay S, Kuis W, Machado C, Pachman L, Prieur A-M, Rider LG, Silverman E, Tsitsami E, Woo P, Giannini EH, Martini A, Paediatric Rheumatology International Trials Organization (PRINTO), Pediatric Rheumatology Collaborative Study Group (PRCSG). Preliminary core sets of measures for disease activity and damage assessment in juvenile systemic lupus erythematosus and juvenile dermatomyositis. *Rheumatology*. 2003;42(12):1452-1459. doi:10.1093/rheumatology/keg403.

35. Khan K. The CROWN Initiative: journal editors invite researchers to develop core outcomes in women's health. *BJOG An Int J Obstet Gynaecol*. 2014;121(10):1181-1182. doi:10.1111/1471-0528.12929.

36. Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M, Williamson PR. Choosing Important Health Outcomes for Comparative Effectiveness Research: A Systematic Review. Scherer RW, ed. *PLoS One*. 2014;9(6):e99111. doi:10.1371/journal.pone.0099111.

37. Gorst SL, Gargon E, Clarke M, Blazeby JM, Altman DG, Williamson PR. Choosing Important Health Outcomes for Comparative Effectiveness Research: An Updated Review and User Survey. Garattini S, ed. *PLoS One*. 2016;11(1):e0146444. doi:10.1371/journal.pone.0146444.

38. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, Wales PW. Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol*. 2014;67(4):401-409. doi:10.1016/j.jclinepi.2013.12.002.

39. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, Williamson PR, Terwee CB. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” – a practical guideline. *Trials*. 2016;17(1):449. doi:10.1186/s13063-016-1555-2.

Figure legends:

Figure 1: Illustrative example of an outcome presented in round two.

Figure 2: The flow of participants and outcomes

Figure 1: Illustrative example of an outcome presented in round two.

Outcome 52. Preterm Birth

A baby born before 37 weeks of pregnancy.

		Not important (%)				Important but not critical (%)				Critical (%)			
Stakeholder	**Number**	**1**	**2**	**3**		**4**	**5**	**6**		**7**	**8**	**9**	
Healthcare professionals	57	0	0	4		13	0	56		21	3	3	
Researchers	18	0	0	0		0	8	67		18	2	5	
Parent or Carer	28	0	0	0		4	14	48		30	4	0	
→ **Please rescore**		□	□	□		□	□	□		□	□	□	

The percentage of participants scoring the outcome from every possible response from one to nine was presented. The orange column highlights the participant's score from the previous round.

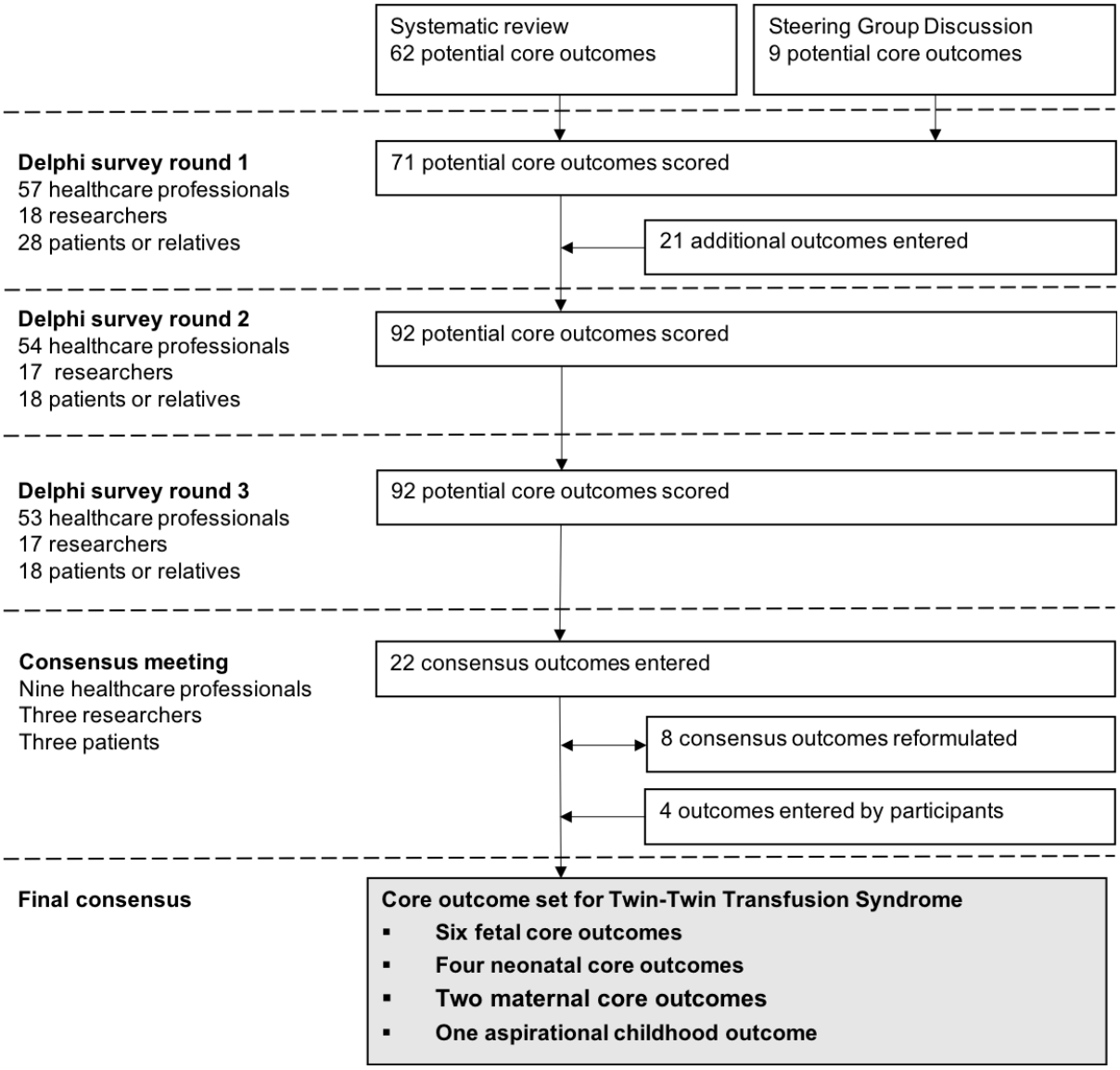

Figure 2: The flow of participants and outcomes

Table 1. Delphi survey participant characteristics

	Round 1 n=103	Round 2 n=89	Round 3 n=88	Withdrawals n=15
Stakeholder group, n				
Patients	18	10	10	8
Relatives of someone with TTTS	10	8	8	2
Healthcare professionals	57	54	53	4
Midwives	3	3	3	0
Neonatologists or Paediatricians	3	3	3	0
Obstetricians	16	15	15	1
Sonographers	2	2	2	0
Fetal Medicine Specialists	33	31	30	3
Researchers	18	17	17	1
Gender, n				
Male	43	39	38	5
Female	60	50	50	10
Age (years), n				
18 to 25	3	3	3	0
26 to 35	27	20	20	7
36 to 45	31	26	26	5
46 to 55	21	20	20	1
Over 56	21	20	19	2
Ethnic group, n				
Asian	10	8	8	2
Black	3	3	3	0
White	84	72	71	13
Any other ethnic group	3	3	3	0
Prefer not to say	3	3	3	0
Geographical location, n				
Africa; 0 countries	0	0	0	0
Asia; 5 countries	6	4	4	2
Australia; 2 countries	14	11	11	3
Europe; 15 countries	70	63	62	8
North America; 2 countries	7	6	6	1
South America; 5 countries	6	5	5	1
Resource setting, n				
High-income; 24 countries	96	83	82	14
Low- and middle-income; 5 countries	7	6	6	1

Table 2. The outcomes that fulfilled the *a priori* consensus definition

DOMAIN
Fetal outcomes
Recurrence of twin-twin transfusion syndrome
Death of a surviving cotwin after first twin death
Twin anaemia polycythaemia syndrome
Cerebral lesions
Brain injury of a surviving cotwin after first twin death
Antenatal Brain Injury
Offspring survival
Live birth
Miscarriage
Intrauterine death/Stillbirth
Neonatal mortality (death)
Perinatal mortality (death)
Infant mortality (death)
Mortality (death) of both twins
Neonatal outcomes
Gestational age at delivery
Stroke
Intraventricular haemorrhage
Periventricular leukomalacia
Early childhood outcomes
Cerebral palsy
Neurodevelopmental Impairment
Maternal outcomes
Maternal Mortality
Abruption
Amniotic Fluid Embolism

Table 3. Consensus development meeting participant characteristics

Stakeholder group, n	
Patients	3
Healthcare professionals	11
Midwives	3
Neonatologists or Paediatricians	3
Obstetricians	3
Fetal Medicine Specialists	2
Researchers	2
Gender, n	
Male	7
Female	9
Age (years), n	
16 to 25	0
26 to 35	6
36 to 45	7
46 to 55	3
Over 55	0
Ethnic group, n	
Asian	2
White	13
Mixed ethnic background	1