EP27.04 – Assessment of inter-operator variability of manual segmentation for analysis of the fetal face from 3D ultrasound scans

Clark A.1,2 Dall’Asta A.1,3 Biffi B.2 Wong T L.1 Fessey L.2 Paramasivam G.1 Dunaway D.4 Schievano S.4 Lees C.1,2
1. Queen Charlotte’s & Chelsea Hospital, Imperial College NHS Healthcare Trust 2. Imperial College, London 3. University Hospital of Parma, Italy 4. University College London GOS Institute of Child Health & Great Ormond Street Hospital for Children, London.

Introduction
Segmentation is a method used to divide an image/volume into multiple segments in order to facilitate analysis. We have previously demonstrated the feasibility of an automated segmentation (AS) technique for the quantitative assessment of fetal facial volumes. In order to improve the technique, additional manual segmentation (MS) is required. This study sought to assess the inter-operator variability of MS.

Methods
14 3D ultrasound (US) volumes of anonymised 3rd trimester fetal faces were acquired using GE Voluson E8 and E10 US equipment. An AS was created from the 3D volume, and refinement was performed manually by 3 different operators (op_1, op_2, op_3) using ITK-SNAP software.

A comparison of 14 MS was performed between op_1/op_2 and of 4 MS between op_1/op_3 and between op_2/op_3. The inter-operator variability was assessed using DICE scores to evaluate the percentage of overlap between 2 MS.

Paired samples Wilcoxon test was applied to the MS volumes to further assess the difference between operators.

Results
Median DICE scores between op_1/op_3 was 93% (IQR=2%, range 89%-97%), op_1 was found to slightly over-segment with respect to op_2 (p<0.01). Moreover, from the superimposed MS (Fig 1) the central facial region is the one with the best agreement. MS volumes and percentage of volume difference are plotted on Fig 1. Dice scores between the 3 operators were highly reproducible, median scores (±IQR) of 92% ±4 (op_1-op_2), 94% ±5 (op_1-op+3) and 94% ±5(op_2-op_3). Op_3 was found to slightly under-segment with respect to op_1 and over-segment with respect to op_2 although the difference was negligible (Fig. 2).

Conclusion
Good accuracy can be achieved by different operators performing manual adjustment of AS of ultrasound-derived 3D fetal faces.