

Manual Segmentation of Uterine Niches in Transvaginal Ultrasound: Instruction Protocol and Inter- and Intra-Observer Variability Assessment

B.L. Post^{1,2*}, V. Jumutc^{3*}, B.G.S. Schor², D. Bliznuks³, F.A. Groenman², J.A.F. Huirne², R.A. de Leeuw²

* Shared first author

1 Department of Obstetrics and Gynecology, Amsterdam UMC, location AMC and VUMC, Amsterdam, the Netherlands

2 Amsterdam Reproduction and Development (AR&D), Research Institute, Amsterdam, The Netherlands

3 Institute of Applied Computer Systems, Riga Technical University, Riga, Latvia

Contact: b.post@amsterdamumc.nl

More info:



Objectives

No standardized method currently exists for precisely delineating uterine niche borders on transvaginal ultrasound (TVUS). Accurate and reproducible segmentation is essential for clinical practice, research, and the development of AI-based models. This study aims to **develop and validate a standardized protocol for manual uterine niche segmentation on transvaginal ultrasound and to quantify inter- and intra-observer variability.**

Methods

Three observers performed two rounds of manual uterine niche segmentation on TVUS images. The first round followed basic guidelines. Discrepancies were used to develop a standardized protocol, after which segmentation was repeated. Inter- and intra-observer variability were assessed using Dice and NSD metrics.

Results

Inter-observer agreement improved significantly after protocol implementation, with Dice scores increasing from 0.58-0.68 to 0.74-0.78 and NSD from 0.66-0.76 to 0.76-0.80. Intra-observer agreement also improved (Dice: 0.75 to 0.83; NSD: 0.84 to 0.88). Difficult cases reduced agreement, while their exclusion further improved segmentation consistency.

Conclusion

- Our standardized protocol enables consistent and reproducible segmentation for clinical, research, and AI applications.
- Protocol use improved inter- and intra-observer reliability.
- Common difficulties included oversimplified shapes, challenges with still images, exclusion of branches, and an unclear definition of the posterior border.

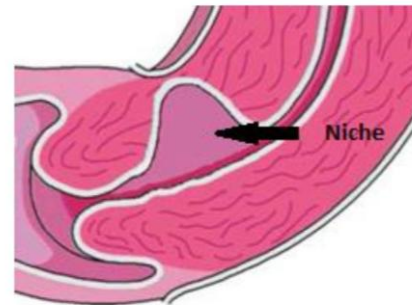


Figure 1: Uterine niche, a cesarean scar defect

Dice: Before vs After Protocol

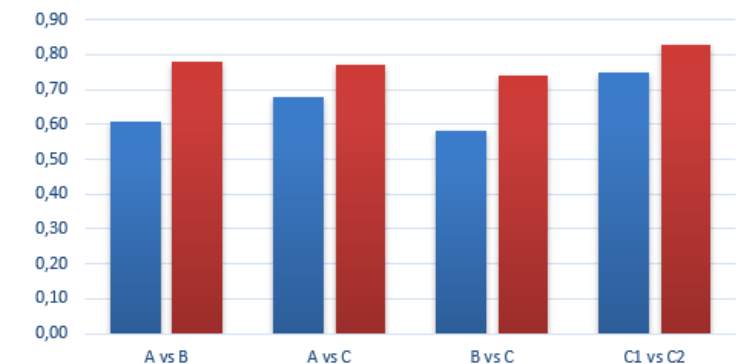


Figure 2: Dice score (overlap metric) before and after protocol

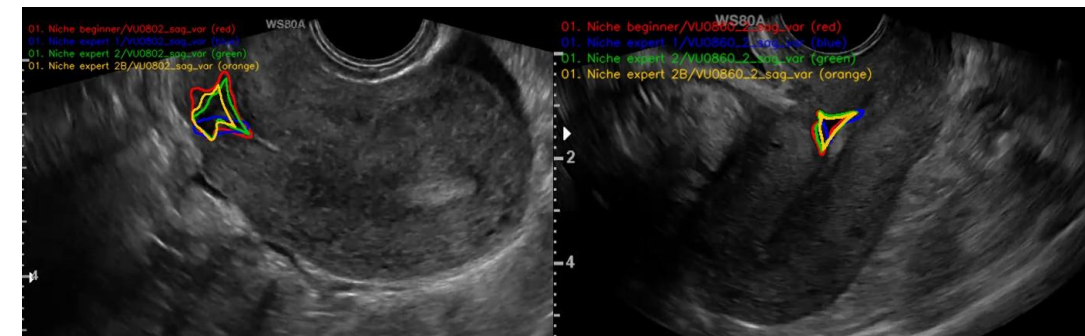


Figure 3: Examples of segmented niche and inter- and intra-observer variability