

## Assess the quality of obstetrical ultrasound of prenatally misdiagnosed congenital heart malformations: the utility of AI models at a glance

Enrico Corno<sup>1</sup>, Camille Dupont<sup>2</sup>, Ivan Voznyuk<sup>2</sup>, Andrea Dall'asta<sup>1</sup>, Tullio Ghi<sup>1</sup>, Edwin Quarello<sup>3</sup>

<sup>1</sup>Department of Medicine and Surgery, Obstetrics and Gynaecology Unit, University of Parma, Parma, Italy.

<sup>2</sup>Diagnoly s.r.l., France.

<sup>3</sup>Service de Gynécologie-Obstétrique-AMP, Hôpital Saint-Joseph-de-Marseille, Marseille, France.

DOI: 10.36129/jog.2024.S167

**Objective.** To assess the performance of the Deep Learning (DL) models in evaluating the completeness and quality of 2nd trimester ultrasound scans reports in cases of prenatal misdiagnosed congenital heart defect (CHD).

**Materials and Methods.** A retrospective analysis was performed on selected cases of prenatally misdiagnosed isolated CHD from the Marseille area, France, between 2020 and 2021. Those cases were selected based on postnatal CHD's diagnoses that are typically detectable through prenatal screening and require interventional cardiac catheterization and/or surgery within the first year of life. All foetal heart still pictures or video clips were collected, including low-quality images. Six views of interest were evaluated as being present or absent and their 52 related anatomical quality criteria as well.

**Results.** Overall, 20 cases of misdiagnosed CHD were analysed which correspond to 168 still pictures and 3 video

clips. 16 images were classified by the AI system as "unclassifiable" (100% of concordance with human). DL system incorrectly categorized the scans in only 2 cases. Global DL concordance with ground truth was 98.8% (169/171). The specific concordance related to quality criteria was 78.3% (1,479/1,889). The concordance rate between AI and clinician was systematically compared across the five key-views. DL system had a higher concordance rate in 4CV respect to LVOT-view, RVOT-view and 3V-view (Proportion-Z:4.3439, P-value = 0.00001; Proportion-Z:6.1558, P-value = 0.00001; Proportion-Z:6.5561, P-value = 0.00001, respectively).

**Conclusions.** The capacity of our DL algorithms to assess foetal heart views of interest and their anatomical quality criteria is interesting. The next step consists of assessing their usability and performance in daily practice.