Check for updates

Ultrasound Obstet Gynecol 2016; **47**: 652–656
Published online 8 April 2016 in Wiley Online Library (wileyonlinelibrary.com).

Letters to the Editor

Interobserver agreement in assessment of polycystic ovarian morphology using pattern recognition

In daily clinical practice, ultrasonographic diagnosis of polycystic ovarian (PCO) morphology is frequently made 'at first glance', based on pattern recognition. This is a subjective impression of the examiner[1], without precise quantification of total follicle number, follicle size, follicle distribution or ovarian volume. The classification of an ovary as being polycystic may have important clinical and therapeutic consequences. The aim of this study was to evaluate the interobserver agreement for the diagnosis of PCO morphology using only the subjective assessment of an examiner (or 'pattern recognition').

Offline analysis of ultrasound images was conducted using the web-based electronic data capture software framework Clinical Data Miner (CDM)[2]. CDM enables ultrasound images to be shown alongside questions and multiple-choice categories. For each patient, a single two-dimensional (2D) ultrasound image of an ovary was evaluated. In total, 40 different images from women of reproductive age (days 3–5 of a spontaneous cycle) were integrated into the CDM. Five gynecologists (N.R.F., D.D.N., T.B., T.V.B. and D.T.) with experience in gynecological ultrasound were asked to evaluate the images online. Images were shown in random order without clinical information and the examiners scored each ovary as 'typical' PCO morphology, 'possible' PCO morphology or 'not' PCO morphology. Subsequently the same 40 images were presented again in random order and the examiners had to specify follicle distribution as peripheral or random, follicle size as uniform or non-uniform and stromal appearance as central echogenic, diffuse echogenic or of normal appearance (Figures 1 and 2). Interobserver agreement was assessed using Fleiss' kappa coefficient[3]. Bias-corrected estimates and standard errors for the kappa coefficient were obtained by means of the jackknife subsampling technique. Fleiss' kappa values of < 0 indicate poor agreement, 0–0.2 slight, 0.2–0.4 fair, 0.4–0.6 moderate, 0.6–0.8 substantial and > 0.8 almost perfect agreement.

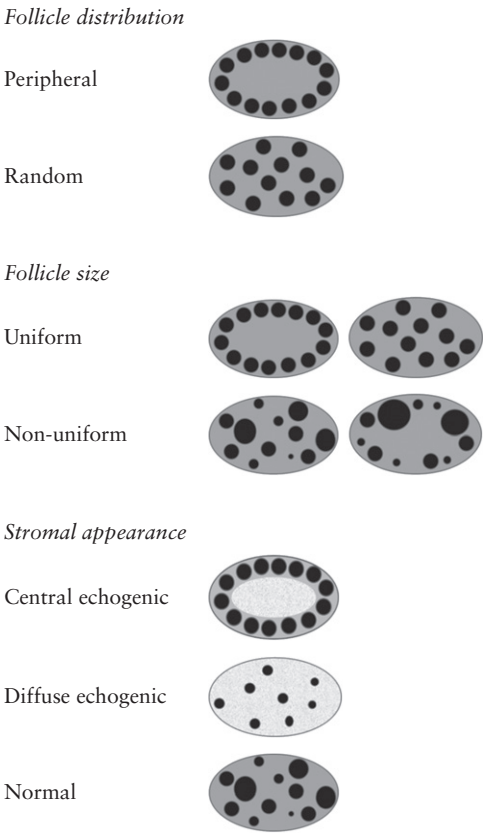

Figure 1 Pictograms illustrating follicle distribution, follicle size and stromal appearance, which were integrated into the Clinical Data Miner software framework[2] alongside multiple-choice questions and two-dimensional ultrasound images, to determine interobserver agreement for assessing polycystic ovarian morphology.

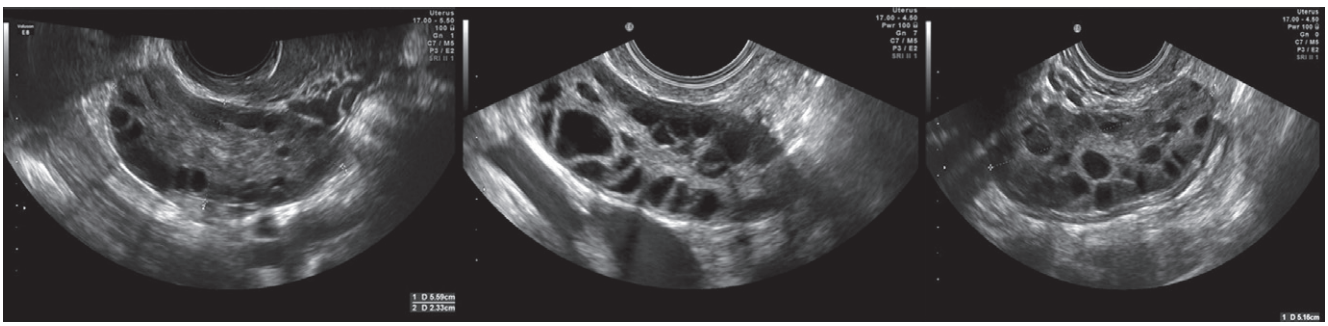

Figure 2 Ultrasound images of ovaries with > 12 follicles, showing differences in follicle distribution, follicle size and stromal appearance in each ovary.

Copyright © 2015 ISUOG. Published by John Wiley & Sons Ltd.

The Fleiss' kappa coefficient bias-corrected estimate (standard error) for the diagnosis of PCO morphology was 0.36 (0.08) and for defining follicle distribution, follicle size and stromal appearance was 0.80 (0.07), 0.34 (0.09) and 0.41 (0.07), respectively.

Experienced observers achieved a fair agreement when defining PCO morphology subjectively based on a single 2D ultrasound image. Knowledge of a patient's menstrual cycle, body mass index and/or androgenic symptoms may bias the interpretation of ultrasound images. This study was restricted to the evaluation of the interobserver agreement on a single image, being the most basic prerequisite of pattern recognition. We expected real-time scanning or three-dimensional volume evaluation[4] to worsen the interobserver agreement, as examiners have to evaluate images from different sections through the ovary. This is yet to be confirmed prospectively. The actual follicle count remains useful in the assessment of ovarian reserve.

The modest interobserver agreement suggests that published data on ultrasound features indicative of PCO morphology should be interpreted with caution. We acknowledge 'polycystic' is misleading and 'multifollicular' may be the more appropriate term[5], but this issue is beyond the scope of our study. Our data reinforce the importance of standardization of ultrasound features for defining PCO morphology in order to improve interobserver agreement and thus clinical relevance.

Acknowledgments

T.B. is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. D.T. is a Senior Clinical Investigator of the Research Foundation Flanders (FWO).

D. Van Schoubroeck*†, N. Raine-Fenning‡, A. J. F. Installé§¶, D. De Neubourg†, B. De Moor§¶, T. Bourne†**, T. Van den Bosch† and D. Timmerman†
†*Department of Obstetrics and Gynecology, University Hospitals KU Leuven, Herestraat 49, 3000 Leuven, Belgium;*
‡*Division of Child Health, Obstetrics & Gynaecology, University of Nottingham, Nottingham, UK;*
§*KU Leuven, Department of Electrical Engineering ESAT, STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Leuven, Belgium;*
¶*iMinds Medical IT, Leuven, Belgium;*
**Queen Charlotte's and Chelsea Hospital, Imperial College London, London, UK*
Correspondence. (e-mail: dominique.vanschoubroeck@uzleuven.be)

DOI: 10.1002/uog.15768

References

1. Ameye L, Timmerman D, Valentin L, Paladini D, Zhang J, Van Holsbeke C, Lissoni AA, Savelli L, Veldman J, Testa AC, Amant F, Van Huffel S, Bourne T. Clinically oriented three-step strategy for assessment of adnexal pathology. *Ultrasound Obstet Gynecol* 2012; **40**: 582–591.
2. Installé AJF, Van den Bosch T, De Moor B, Timmerman D. Clinical Data Miner: An Electronic Case Report Form System with Integrated Data Preprocessing and Machine-Learning Libraries Supporting Clinical Diagnostic Model Research. *JMIR Med Inform* 2014; **2**: e28.
3. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychological Bulletin* 1971; **76**: 378–382.
4. Lam PM, Raine-Fenning N. The role of three-dimensional ultrasonography in polycystic ovary syndrome. *Hum Reprod* 2006; **21**: 2209–2215.
5. Martins WP, Kollmann M, Raine-Fenning N. Counting ovarian follicles: updated threshold for diagnosis of hyperandrogenic anovulation. *Ultrasound Obstet Gynecol* 2014; **44**: 131–134.

Copyright © 2015 ISUOG. Published by John Wiley & Sons Ltd.

14690705, 2016, 5, Downloaded from https://obgyn.onlinelibrary.wiley.com/doi/10.1002/uog.15768 by Ku Leuven, Wiley Online Library on [27/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License