A deep-learning-enabled diagnosis of ovarian cancer – Authors’ reply

Gao, Yue; Li, Huayi; Chen, Lingxi; Wu, Yuan; Ma, Ding; Gao, Qinglei

Published in:
The Lancet Digital Health

Published: 01/09/2022

Document Version:
Final Published version, also known as Publisher’s PDF, Publisher’s Final version or Version of Record

License:
CC BY-NC-ND

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1016/S2589-7500(22)00145-5

Publication details:

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher’s copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.
A deep-learning-enabled diagnosis of ovarian cancer

Authors’ reply

We appreciate Ben Van Calster and colleagues’ methodological concerns about our pelvic ultrasound-based deep convolutional neural network model for ovarian cancer diagnosis.1 Our study included data for healthy control patients in the model development, because familiarising the model with healthy ovaries with high anatomical and morphological variabilities (eg, follicle or corpus luteum) during the menstrual cycle prepares the deep convolutional neural network model to enable detection of abnormalities and differential diagnosis. Moreover, the discriminative performance of the model was validated in two external validation datasets that did not include data for healthy ovaries (an AUC of 0·870 for the dataset from Jingzhou First People’s Hospital1 and 0·831 for the dataset from Xiangyang Central Hospital [unpublished]). The ages of the 86 074 consecutively recruited healthy women included from eight Chinese hospitals in the study are provided in the appendix (p 1). We included women (aged ≥18 years) with ultrasound images showing no abnormalities in the adnexa. Adnexal images were confirmed at the local hospital. We excluded images less than 224 × 224 pixels. For eligible images (≥224 × 224 pixels), we resized to 256 × 256 pixels, cropped centrally to 224 pixels, and normalised by subtracting the mean and dividing by SD. Data augmentation embraced random resize, crop, horizontal flip, rotation, colour jittering, and normalisation. The Grad-CAM algorithm was used to identify regions that contributed the most to the model prediction.2 Representative images and corresponding saliency heatmaps are exhibited in the appendix (p 1). Three sonographers were invited to inspect randomly selected images (n=35) and their saliency heatmaps. The percentage that the sonographers agreed on the regions of malignancy that were captured by saliency heatmaps was 91·4–94·3%. Annotations within raw images contributed little to the model prediction (appendix p 1). We thought it debatable that the poor calibration might lead to undertreatment. First, good discrimination does not always ensure good calibration, and vice versa. Calibration performance is dispensable in binary diagnosis tasks3 and most suitable to evaluate prognostic models that predict the risk in the future.4 Second, the poor calibration was partly attributed to the threshold of malignant probability (0·14), which was determined when the deep convolutional neural network model achieved the best AUC during development. Third, the discriminative performance was validated in two independent external datasets.

We declare no competing interests.

Copyright © 2022 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license.

Yue Gao, Huayi Li, Lingxi Chen, Yuan Wu, Ding Ma, *Qinglei Gao
qingleigao@hotmail.com

National Clinical Research Centre for Obstetrics and Gynaecology, Cancer Biology Research Centre (Key Laboratory of the Ministry of Education) and Department of Gynaecology and Obstetrics, Tongji Hospital, Wuhan 430030, China (YG, HL, YW, DM, QG); Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China (YG, HL, YW, DM, QG); City University of Hong Kong Shenzhen Research Institute, Shenzhen, China (LC)


