Radiomic feature reproducibility for cervix MRI: impact of number of observers and patient datasets

Rhianna Brown[1,2,3], Amy Walker[3,2,4], Karen Lim[5,4], Shalini Vinod[2,5,4], Viet Do[5,4], Chelsie O'Connor[6], Jaqueline Veera[7], Nira Borok[8], Peter Metcalfe[1], Dean Cutajar[1], Lois Holloway[3,2,4]

[1]University of Wollongong, Wollongong, Australia. [2]Ingham Institute for Applied Medical Research, Liverpool, Australia. [3]Liverpool and Macarthur Cancer Therapy Centre, Department of Medical Physics, Liverpool, Australia. [4]University of New South Wales, South Western Sydney Clinical School, Liverpool, Australia. [5]Liverpool and Macarthur Cancer Therapy Centre, Department of Radiation Oncology, Liverpool, Australia. [6]Chris O'Brien Lifehouse, Department of Radiation Oncology, Camperdown, Australia. [7]Peter MacCallum Cancer Centre, Department of Radiation Oncology, Bendigo, Australia. [8]Liverpool and Macarthur Cancer Therapy Centre, Department of Radiology, Liverpool, Australia

Purpose

- Using a combination of clinical features and radiomic features (RF) have been shown to improve performance of prediction outcome models[1,2]
- The RF utilised in these models needs to be reproducible to interobserver contour variation (ICV)
- Current studies that have assessed this impact for cervix MRI RF have only used 2-3 observers[3,4,5,6]

Aim: to assess the reproducibility of RF from cervical cancer MRIs to ICV with up to 6 observers considering the trade-off between number of datasets and number of observers.

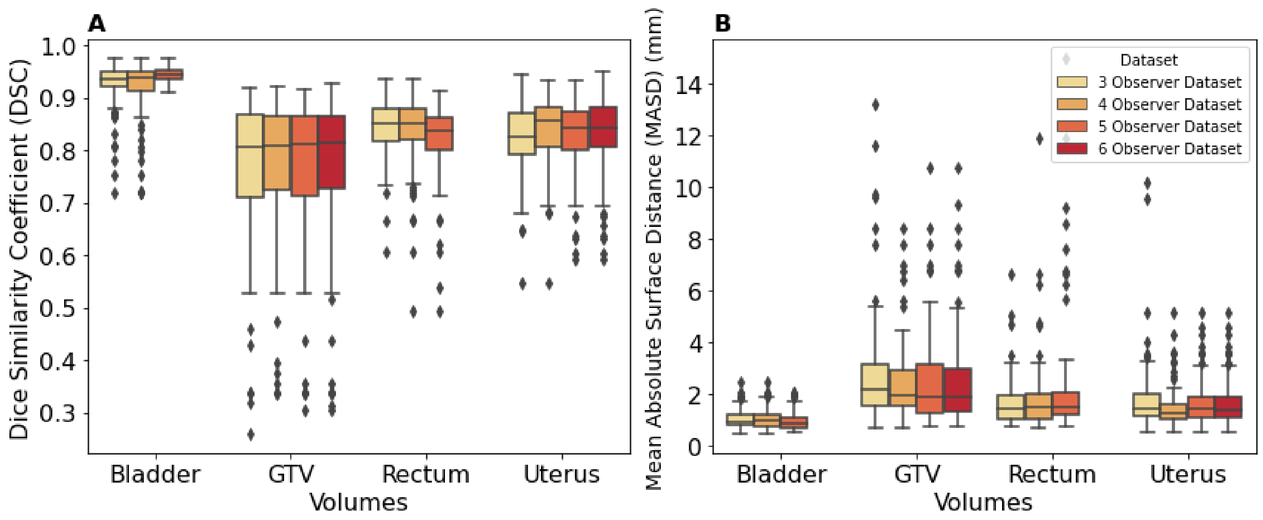

Figure 1: Interobserver contour variations of each of the volumes and different observer datasets. A) DSC of the volumes. B) MASD of the volumes

Methodology

- Data: 24 pretreatment cervical cancer T2W-MRIs
- Contours: 6 observers (5 radiation oncologist and 1 radiologist) delineated the bladder, gross tumour volume (GTV), rectum and uterus.
- Observer datasets: 9 MRIs were delineated by 6 observers, 18 by 4 observers and 24 by 3 observers, resulting in 3-, 4-, 5- and 6-observer datasets with varying number of images. The 6th observer delineated only the GTV and uterus.
- ICV: assessed with Dice Similarity Coefficient and Mean Absolute Surface Distance
- Radiomics: extracted from all observer volumes using *PyRadiomics*[7], an open-source python library.
- Reproducibility: assessed by calculating intraclass correlation coefficient estimates (ICC(2,1)) for each of the radiomics and each of the observer datasets with *Pinguoin*[8], open-source python library.
- RFs were categorised as excellent ICC≥0.90, good 0.75≥ICC>0.90, moderate 0.50≥ICC>0.75 and poor ICC<0.5.

Results

- Bladder has the least ICV, followed by the uterus and rectum, and lastly the GTV. (ICV shown in figure 1)
- Figure 2 shows the reproducibility of the 107 RF across all observer datasets and volumes.
- Only 23 and 55 RF had excellent or at least good reproducibility, respectively, across all volumes and observer datasets.
- The number of RFs with good or excellent reproducibility for the GTV decreased between the 3 and 5/6 observer datasets from 98 to 68
- Volumes with less ICV had more RFs with excellent reproducibility. For the bladder, 104/107 of RF had good or excellent reproducibility across all observer datasets. The uterus and rectum had 96 and 85 features, respectively.

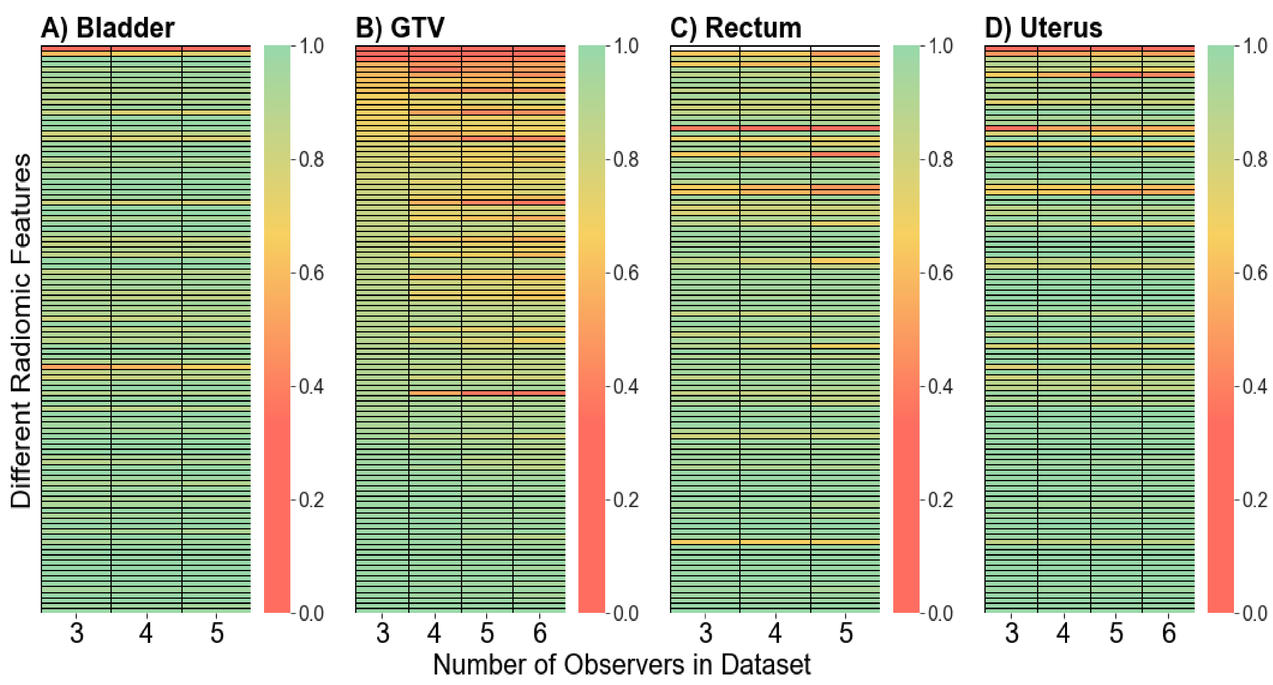

Figure 2. Reproducibility of the RFs from all volumes and observer datasets, represented by ICC values varying from 0 to 1

Conclusion

- For RF extracted from cervical cancer GTVs, the number of observers in the datasets impacted the number of RF with good or excellent reproducibility.
- As more observer contours are included in the analysis, an improved representation of the potential and reasonable contours is achieved. This also results in greater variation in contours, specifically for the GTV, therefore resulting in a less consistent overlap which affects the RF extracted from these varying contours.
- RFs utilised in modelling should be reproducible to ICV, therefore, avoiding the 52 RF which showed poor or moderate reproducibility overall to ICV is recommended

UNIVERSITY OF WOLLONGONG AUSTRALIA

CENTRE FOR MEDICAL RADIATION PHYSICS

NSW GOVERNMENT South Western Sydney Local Health District

[1]Y. Zhou, H.-L. Gu, X.-L. Zhang, Z.-F. Tian, X.-Q. Xu, and W.-W. Tang, 'Multiparametric magnetic resonance imaging-derived radiomics for the prediction of disease-free survival in early-stage squamous cervical cancer', Eur. Radiol., vol. 32, no. 4, pp. 2540–2551, Apr. 2022, doi: 10.1007/s00330-021-08326-6. [2]R. Zheng et al., 'An MRI-based radiomics signature and clinical characteristics for survival prediction in early-stage cervical cancer', Br. J. Radiol., vol. 95, no. 1129, p. 20210838, Jan. 2022, doi: 10.1259/bjr.20210838. [3]S. Fiset et al., 'Repeatability and reproducibility of MRI-based radiomic features in cervical cancer', Radiother. Oncol., vol. 135, pp. 107–114, Jun. 2019, doi: 10.1016/j.radonc.2019.03.001. [4]A. Traverso et al., 'Sensitivity of radiomic features to inter-observer variability and image pre-processing in Apparent Diffusion Coefficient (ADC) maps of cervix cancer patients', Radiother. Oncol., vol. 143, pp. 88–94, Feb. 2020, doi: 10.1016/j.radonc.2019.08.008. [5]Y. Liu et al., 'Radiomics analysis of apparent diffusion coefficient in cervical cancer: A preliminary study on histological grade evaluation', J. Magn. Reson. Imaging, vol. 49, no. 1, pp. 280–290, 2019, doi: 10.1002/jmri.26192. [6]H. Chen et al., 'Reproducibility of radiomics features derived from intravoxel incoherent motion diffusion-weighted MRI of cervical cancer', Acta Radiol., vol. 62, no. 5, pp. 679–686, May 2021, doi: 10.1177/0284185120934471. [7]J. J. van Griethuysen et al., 'Computational Radiomics System to Decode the Radiographic Phenotype', Cancer Res., vol. 77, no. 21, pp. e104–e107, Nov. 2017, doi: 10.1158/0008-5472.CAN-17-0339. [8]R. Vallat, 'Pingouin: statistics in Python', J. Open Source Softw., vol. 3, no. 31, p. 1026, Nov. 2018, doi: 10.21105/joss.01026.

rb672@uowmail.edu.au

/rhianna-brown3