Multimodality radiomics and deep learning for outcome modeling: application in head & neck cancer

Jan Seuntjens, Ph.D., FCCPM, FAAPM, FCOMP
Medical Physics Unit, McGill University
Montreal, Canada

Acknowledgements

• Post docs - RAs
  - Avishek Chatterjee
  - Martin Carrier-Vallières
  - Alexandre Berton
• Former grad students
  - André Diamant-Boustead
  - Marc-André Renaud

McGill clinician collaborators

• C. Reinhold
• H. Moffitt Cancer Center & Research Institute

International collaborators

• S. Rezvani
• I. Naqa
• H. Lee Moffitt Cancer Center & Research Institute
• O. Morin

Outline

• Radiomics for outcomes prediction
• Repeatability and reproducibility
• Radiomics and head & neck cancer – benchmark study
• Deep learning radiomics
• Extension to multiple modalities and clinical data
• Summary
Hypothesis 1: The genomic heterogeneity of aggressive tumors translates into heterogeneous characteristics at the anatomical scale.

Hypothesis 2: Intratumoral heterogeneity at the anatomical scale can be captured using quantitative image analysis.

Radiomics = use of “texture” information in images

- Example of GLCM textures

Texture analysis is concerned with the spatial distribution (patterns) of gray level variations within an image.

The radiomics world

Machine Learning Model

Outcome prediction

Adapted from (Lambin P et al., Eur J Cancer 48, 2012)
Soft tissue sarcoma – lung metastases prediction model

Repeatability and reproducibility

- **Repeatability** = measure of precision under identical or near-identical conditions and acquisition parameters
- evaluated by "test-retest" analysis
- 31 CT datasets Reference Image Database to Evaluate Therapy Response (RIDER)
- **Reproducibility** = better to assess overall robustness
  - Imaging system
  - Imaging parameters
  - Reconstruction
  - ROI delineation
  - Feature extraction and feature qualification

Measuring CT scanner variability of radiomics features

Reproducibility of radiomics for deciphering tumor phenotype with imaging

- Imaging parameters that affect edge sharpness significantly affect radiomic features
Learning from scanners: Bias reduction and feature correction in radiomics

- CT scanner variability is large compared to the interpatient variability in the NSCLC tumors for some features.
- 2/3 of the radiomic features depend on the exposure setting of the scanner. Models can correct for this in a large part. Scanner SNR correction will result in more reliable radiomics predictions in NSCLC.

Image biomarker standardization initiative (IBSI)
- Independent international collaboration working towards standardizing the extraction of biomarkers from imaging for the purpose of high-throughput quantitative image analysis.

- Open software packages and standardized implementations (e.g., IBSI) should be used to ensure reproducibility.
- Models and features should be tested to determine added prognostic and predictive accuracy compared to accepted clinical factors.
- Features should be tested for underlying dependencies using statistical analysis or by perturbing the data in controlled ways.
- Image quality (e.g., artifacts) should be assessed in a preprocessing step and contouring information included.
Head & neck radiomics
Benchmark study

Example of deep learning radiomics model

Deep learning radiomics in head & neck cancer outcome
CNN filter activation and texture features

Performance

Table 2. Validation set results compared to Valdivieso et al. setting set results on the same patient cohort. Balanced accuracy is defined as the average of the specificity and sensitivity. For each radiomic feature, the study calculated specificity and sensitivity based on thresholds optimized in the training set, while the benchmark study [25] performed meta-analysis adjustments during training and then used a single probability threshold of 0.5 in the testing phase. DM: Distant metastasis; LRF: Local-regional failure; OS: Overall survival.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Balanced Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM</td>
<td>0.69</td>
<td>0.73</td>
<td>0.71</td>
</tr>
<tr>
<td>OS</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Extension to multiple imaging modalities
Summary

- We are only in the early stages of outcome modeling using these newer techniques, and far away from clinical implementation – data federation
- We emphasized standardization in the radiomics steps with the goal of better reproducibility
- We may build successful models but we have to recognize that there is a large variability of factors influencing clinical outcomes. We have to be careful with early generalizations.
- Outcome modeling hinges on the quality of the data. Each patient experience must be carefully documented and stored to contribute to accurate models for future patients’ outcome.
MEDomics
medomics.ai
https://youtu.be/2030Pdgm3_4

Synergy between medical image analysis, machine learning, deep learning, natural language processing and distributed learning