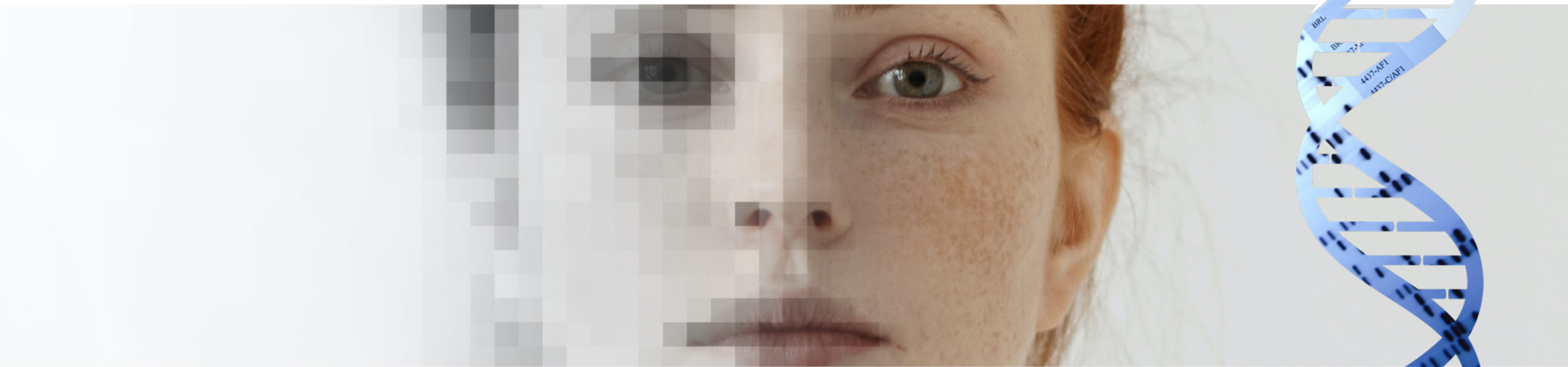
Automated Tumor Burden Assessment from CT-scans

Thomas Bengtsson
Skander Jemaa, Rick Carano, Jill Fredrickson, Shelby Wyatt,
Tina Nielsen, Alex de Crespigny
Genentech/Roche

BBSW, Nov 8, 2019

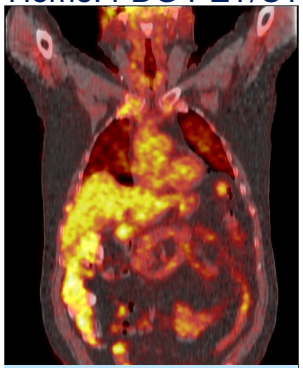


A Revolution at the Interface of Radiology and AI

What does all this mean for drug development?

Radiographic endpoints are a standard part of many of our trials, especially in oncology..

Heme: FDG PET/CT



Metabolic Response by **Cheson or Lugano** criteria

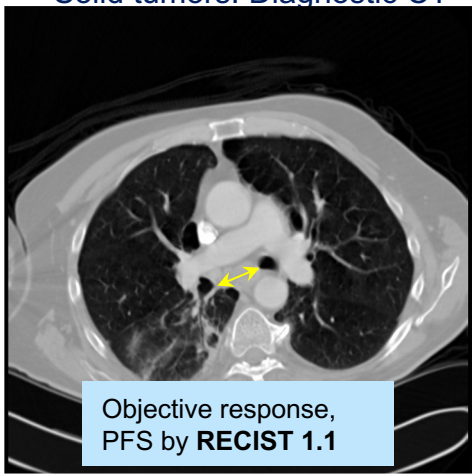
Current paradigm is central image review, with **adjudication rates 10--60%**

Automation: faster, cheaper, more reliable

Better endpoints, patient selection

New approvable endpoints, real-world impact

Solid tumors: Diagnostic CT



Objective response, PFS by **RECIST 1.1**

Auto CT Project Overview

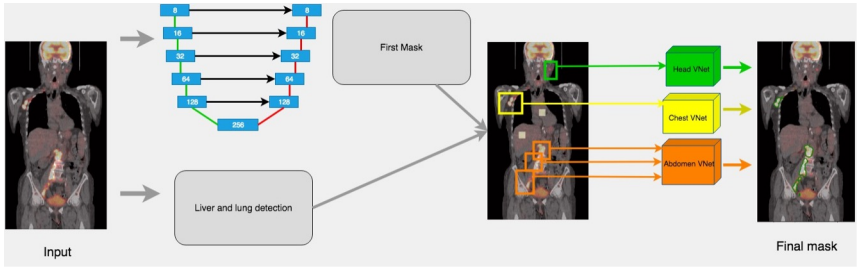
- **Background:**
  - Computed Tomography - Response Evaluation Criteria In Solid Tumors (CT-RECIST)
    - Radiological standard method to quantify treatment response in solid tumors
    - Tumor Quantification: **Identify up to 5 target lesions (most 2 per organ)**, Identify up to 10 non-target lesions, quantify longest diameter for target lesions
    - Clinical Assessment: Define response (Ex. CR, PR, SD, PD) based on **changes in target and non-target lesion btw time points**
  - CT-RECIST is an FDA accepted clinical trial endpoint for evaluating investigational therapies
- **Goal: Develop a fully-automated image analysis CT-RECIST algorithm**

What is possible?

- More informative and robust endpoints for trials
- Faster readouts → Replace IRFs,
- Tools to help clinicians to select optimal therapies

Automatic tumor burden assessment on FDG-PET

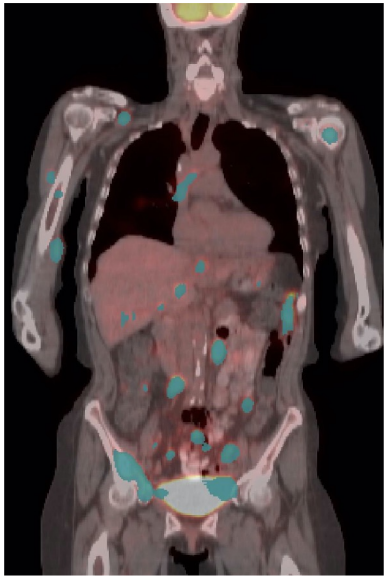
- Sparsity of disease (signal <1% body vol.) and lack of specificity of PET uptake relative disease requires multiscale residual learning DL architecture based on pyramidal dilations



- Spatial agreement* (trained on Goya, validated on Gallium)
  - DICE = 0.89
  - Corr. between reader- and predicted TMTV = 0.98
- Independent of IPI* (in Goya), a prognostic risk score based on {TMTV>330ml ∩ Bulky disease ∩ Nr. Lesions>12} was an independent predictor of PFS (HR=2.01; CI 1.42-2.86)

Ref: Jemaa et al. (SIIM 2019, ASH 2019)

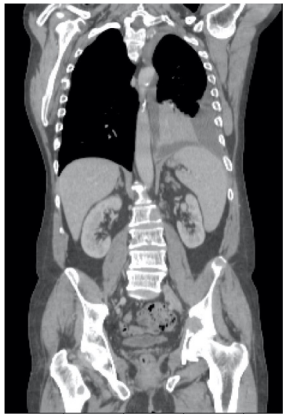
NSCLC pat. w/ ~90 lesions

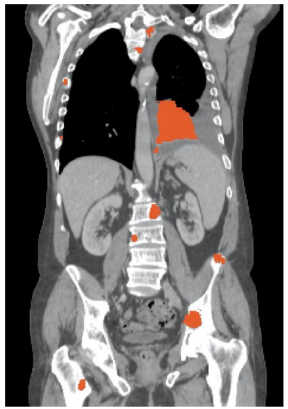


The model can process 1000 scans in < 1d
(*cf.* radiologist ~100 days)

Automatic tumor burden assessment on CT

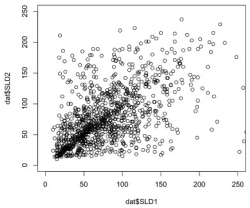
- Automatic lesion detection & segmentation on CT is a v. difficult problem, especially for advanced stage patients



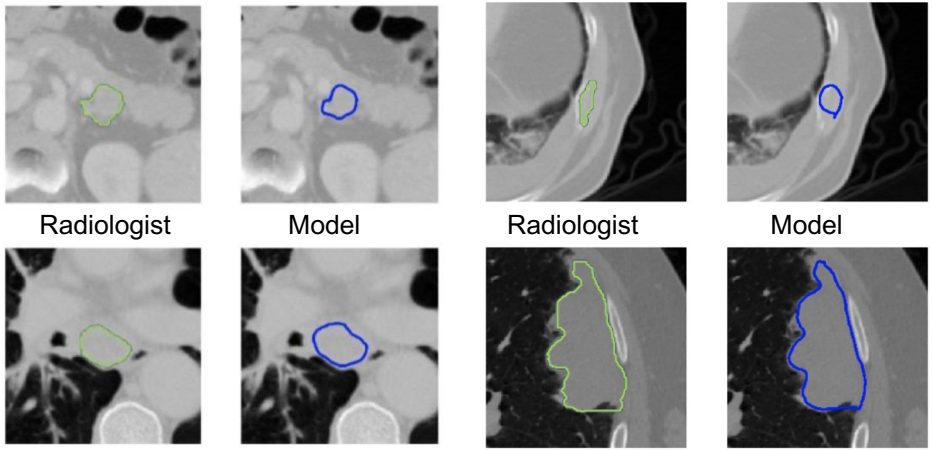


  - High number of lesions: up to 250
  - Substantial fraction of small lesions

Tumor selection problem leads to low inter-reader agreement



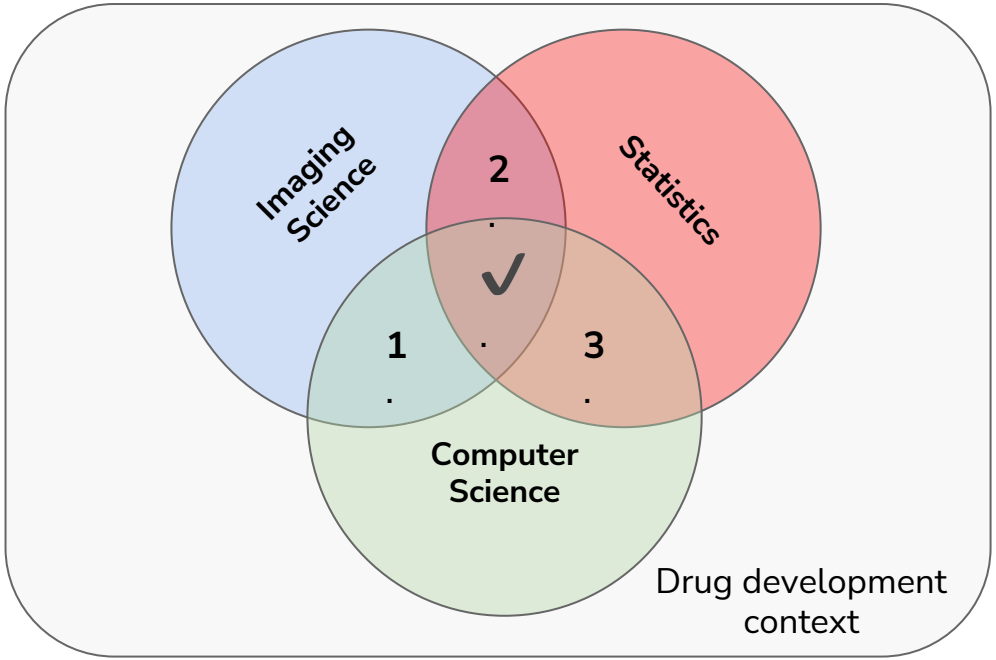
- A cascaded & memory efficient DL model has been designed to delineate tumors in Whole Body CT scans.



Results:

- For CT-R identified lesions (by radiologist), our **full-body model** has a voxel-wise sensitivity of 0.89 and a lesion level sensitivity of 0.94
- The segmentation achieves a 0.82 Dice Similarity Coefficient.

Cross-functional collaboration on AI and medical imaging



1. Lots of loops…
2. Traditional analysis of tabular data…
3. Analysis of imaging artifacts…

✓ required blend of imaging, CS, & Stats