Accelerating Development of Medical Imaging AI for BioPharma

Chaitanya Parmar
Senior Scientist, Computer Vision

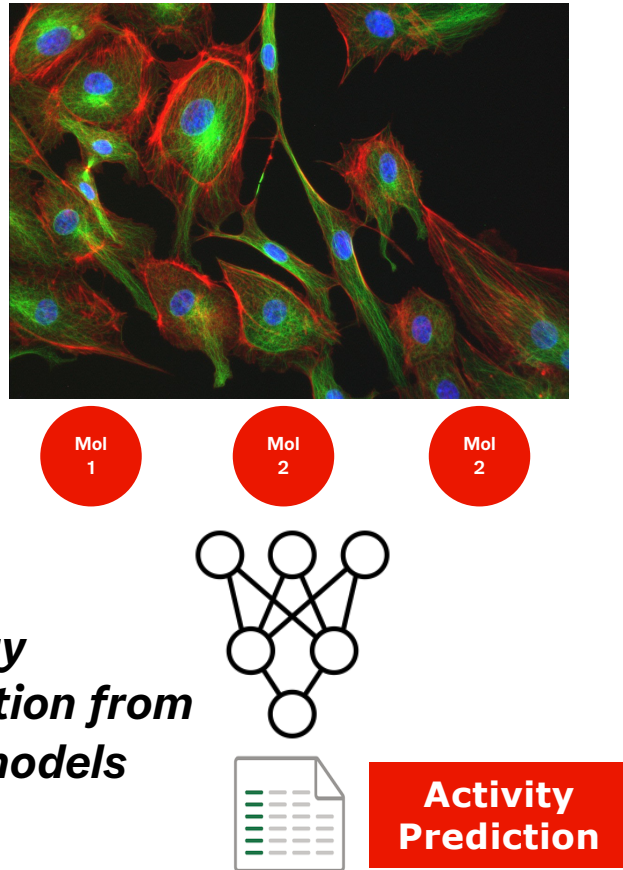**Johnson&Johnson**
Innovative Medicine

Data Science

Accelerating the BioPharma pipeline

There exist repositories of Imaging data from drug discovery to late-stage development that contain rich information and features that can accelerate all stages of the R&D pipeline

Genetic mutation
Molecule Activity
Drug Response
Tumor Subtypes
Disease Severity
*…*

Johnson&Johnson
Innovative Medicine
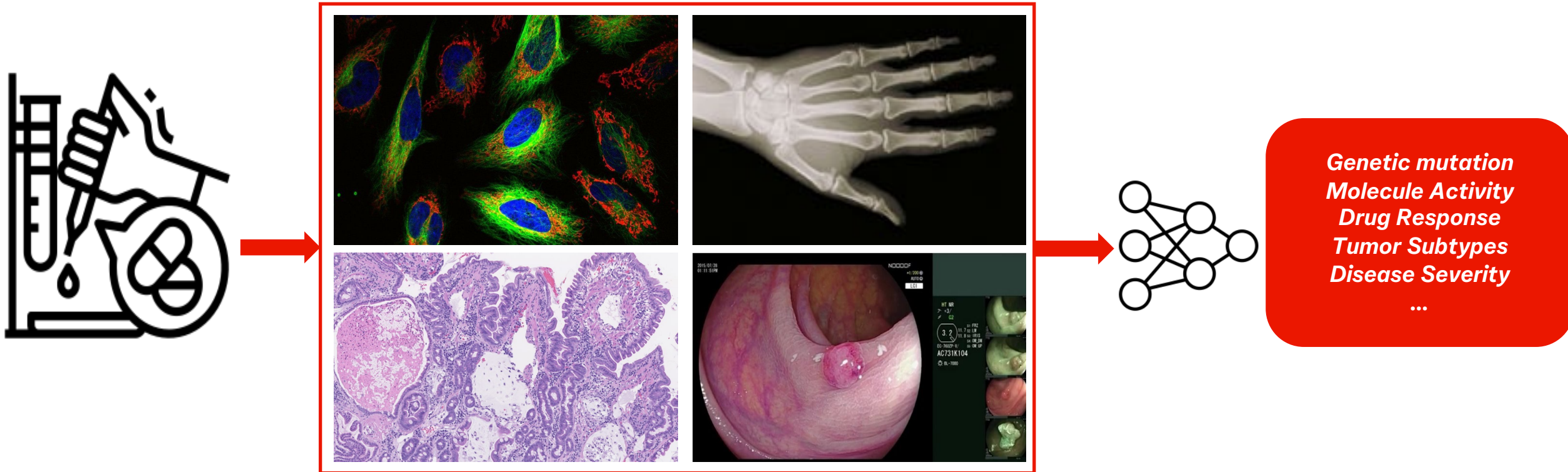
Large scope of applications of AI models

- Cost and time efficient clinical trials with AI based genetic biomarker predictors
- Faster identification of candidate molecules
- Prediction of drug response
- Developing new endpoints for clinical trials

Activity prediction from mice models

Activity Prediction

Johnson&Johnson
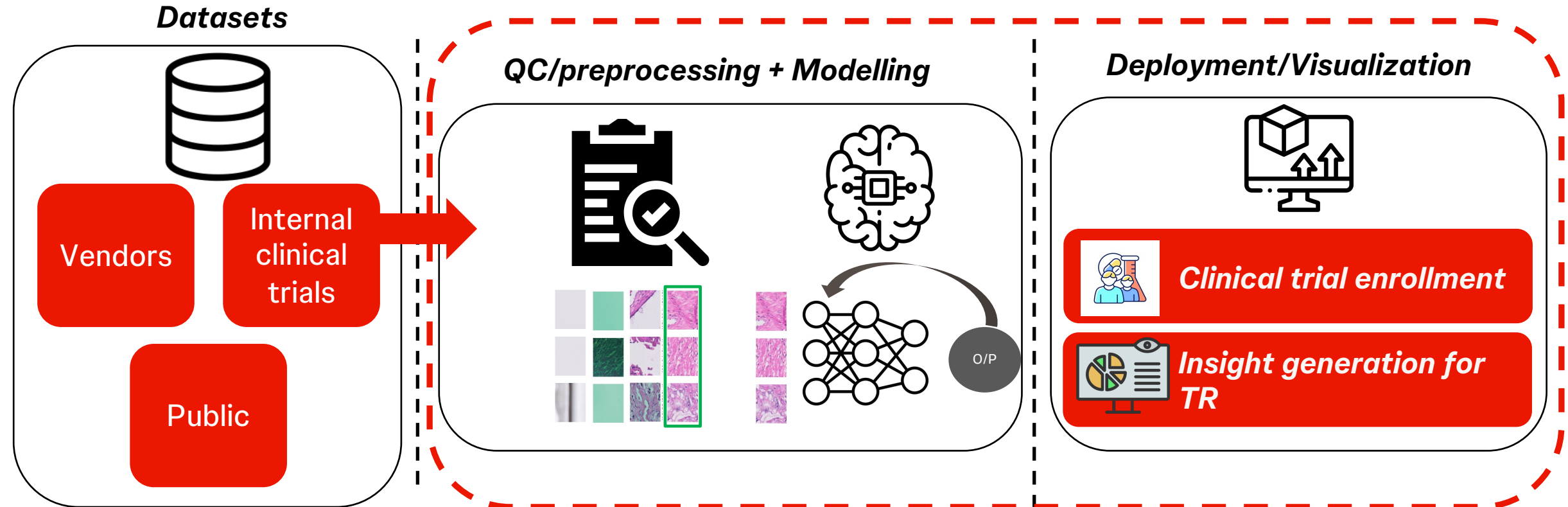Innovative Medicine
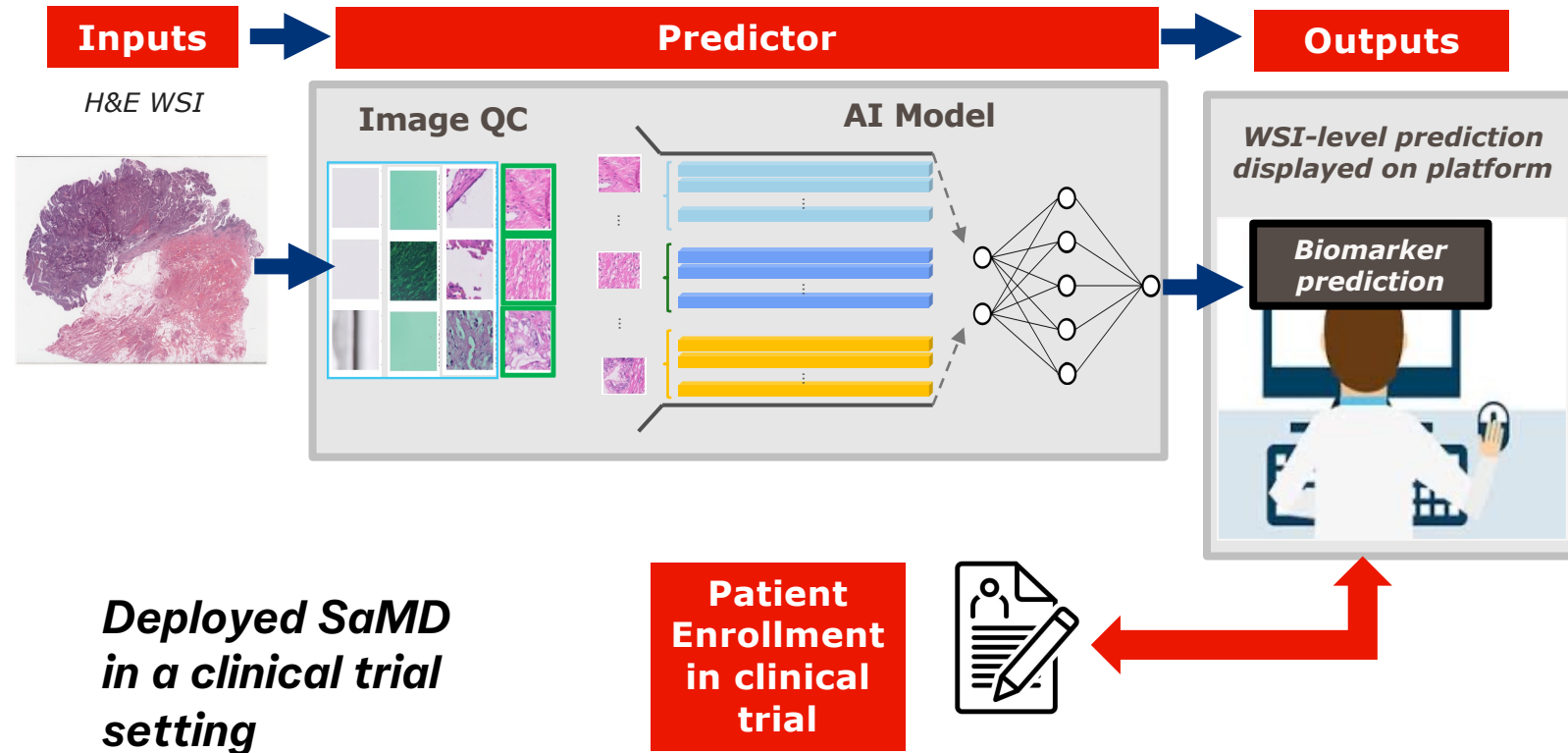
Deployed SaMD in a clinical trial setting

Patient Enrollment in clinical trial

- https://ascopubs.org/doi/abs/10.1200/JCO.2023.41.16_suppl.e15057
- https://aacrjournals.org/cancerres/article/83/7_Supplement/6554/721997/Abstract-6554-Prediction-of-MET-amplification-from

Data Science

3

Typical steps in a model lifecycle

Datasets

QC/preprocessing + Modelling

Deployment/Visualization

Clinical trial enrollment

Insight generation for TR

- *Data is ingested*
- *Data goes through rigorous QC pipelines*
- *Data is used to train models*
- *Models get deployed on internal/external platforms*

Johnson&Johnson
Innovative Medicine

Data Science

4

Modality Agnostic modelling

- *Given the wide variety of imaging modalities, each project can be modelled using its own codebase*
- *We find that the same modelling methods can be applied to solving different tasks in drastically different imaging modalities*

Weakly Supervised Learning can be used to predict biomarkers from both, H&E images as well as endoscopy videos
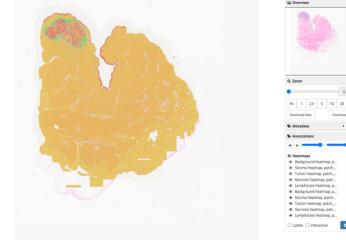
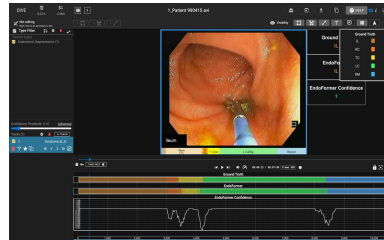Data Science
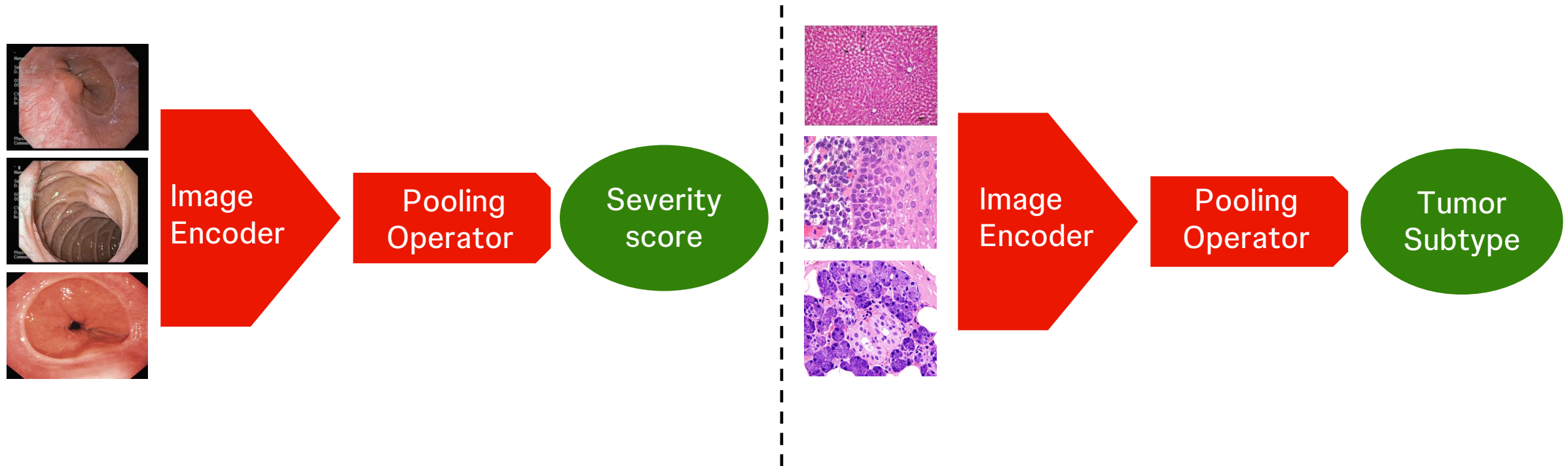
Internal suite of tools for all models

Brain
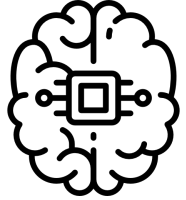
- Codebase to train/validate/test models
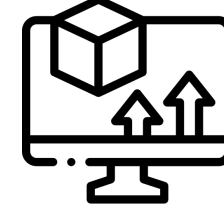- Version Release process with CI/CD pipelines and feature updates

View

- Platform for data and model visualization for multiple modalities including histopath, CTs, MRIs, X-Rays

Ship

- Codebase for dockerizing and deploying models on internal and external platforms

Johnson&Johnson
Innovative Medicine

Data Science

Histopathology in focus:

- *Histopathology Whole Slide Images(WSIs) and especially H&E-stained images are crucial in the R&D pipeline from discovery to development*
- *Predictive models solve many use cases:*

Predicting Genetic Mutations for accelerated Trial enrollment(Example: FGFR, cMET)

Tumor Subtyping for Translational Research efficiency

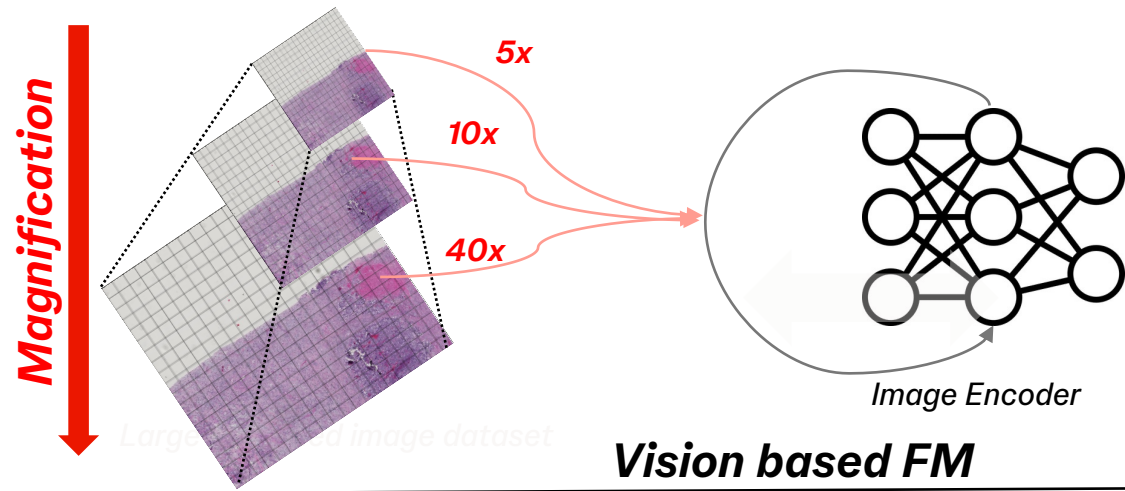**Tissue classification for Translational Research efficiency**
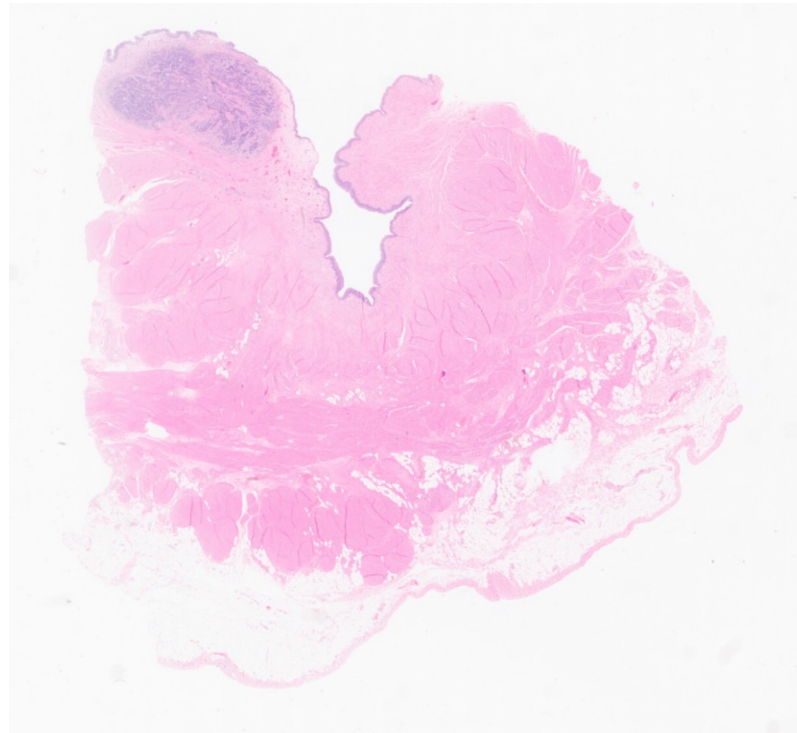
Predicting drug response from H&E Image

Tissue/Cell Segmentation and Quantification

Text Prompting for highlighting salient regions in H&E images

Johnson&Johnson
Innovative Medicine

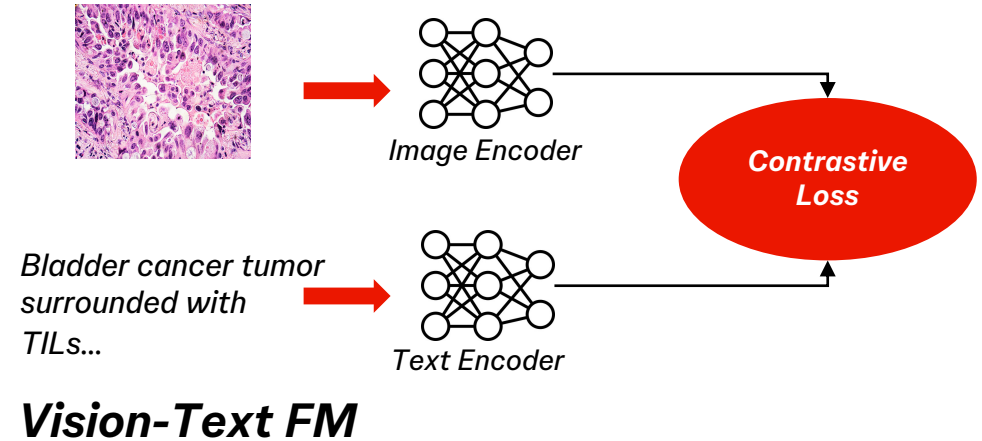Data Science

Training Foundational Models

Self Supervised objectives has shown a marked improvement in terms of performance and generalizability over ImageNet pretrained and end to end models trained from scratch

Magnification

5x
10x
40x

Large curated image dataset

Image Encoder

Vision based FM

Image Encoder

Bladder cancer tumor surrounded with TILs...

Text Encoder

Contrastive Loss

Vision-Text FM

- Specifications:

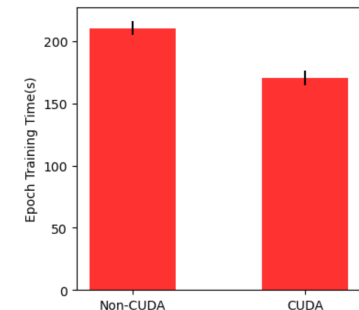| Modality	# of images	Model/# of params	Method(s)
H&E	75k WSIs, 100M+ patches	Vision Encoder: ViT-L/300M, Text Encoder: 80M parameter	DINOv2, CLIP

RAPIDS

Reduction in storage overhead by *~50%* and enabling multimagnification training
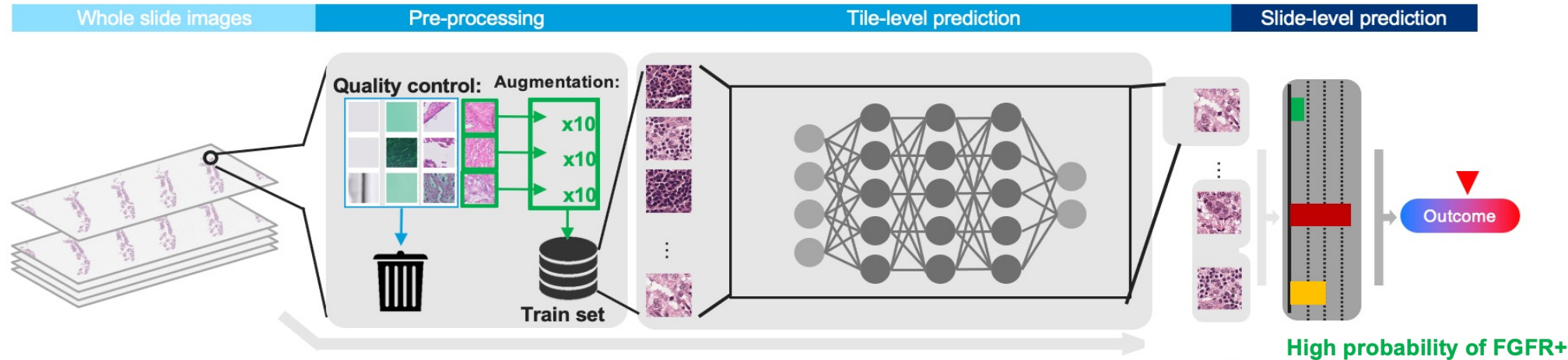
MONAI⁺
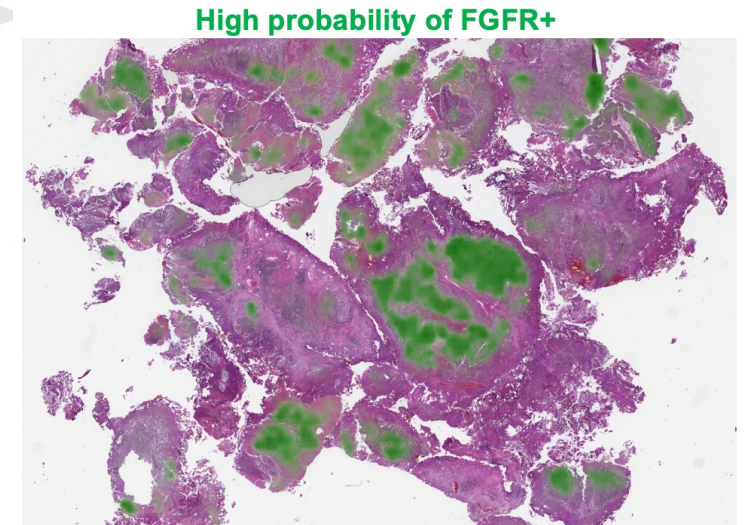
Faster training times *~20%* with CUDA based transforms

Data Science

Use Case: Predicting FGFR status from Bladder H&E Images

Model	Training Data(AUC)	MIBC Unseen Data(AUC)	NMIBC Unseen Data(AUC)
FM based Model	*0.84*	*0.82*	*0.83*
Non-FM based Model	*0.80*	*0.80*	*0.76*

- *Published in ASCO 2023*

https://ascopubs.org/doi/abs/10.1200/JCO.2023.41.16_suppl.e15057

Johnson&Johnson
Innovative Medicine

Data Science

Use Case: Predicting cMET status from Lung H&E Images
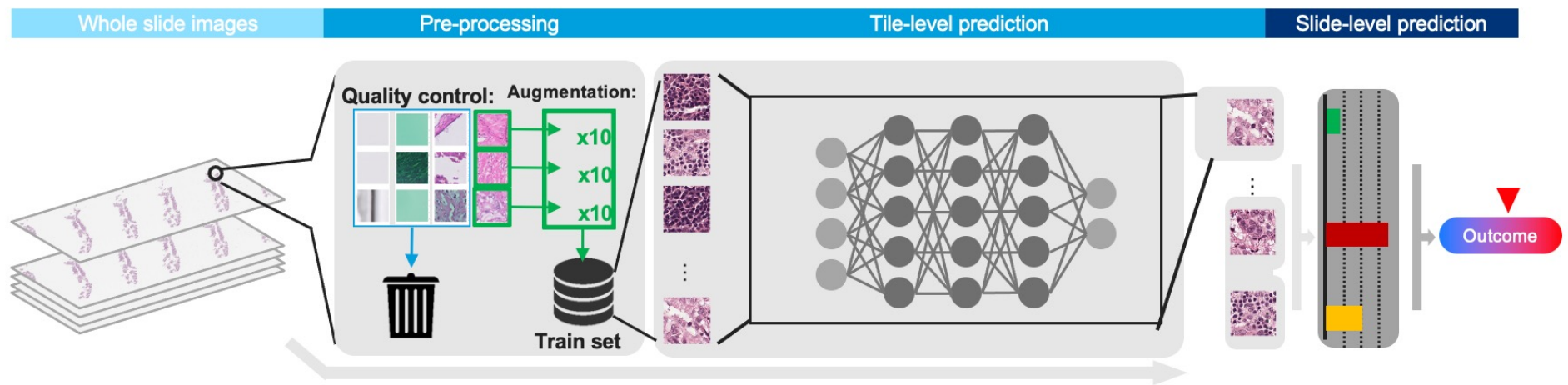
Model	Training Data(NGS)(AUC)	NGS Test Data(AUC)	FISH Test Data(AUC)
FM-based Model	*0.78*	*0.82*	*0.74*
Non-FM based Model	*0.78*	*0.78*	*0.66*

- *Published at AACR 2023*
- *Less discordance in performance on NGS and FISH data*
- *FM model generalizes better on FISH data despite model trained only on NGS data*

High Probability of cMET Amplification

Johnson&Johnson
Innovative Medicine

Data Science
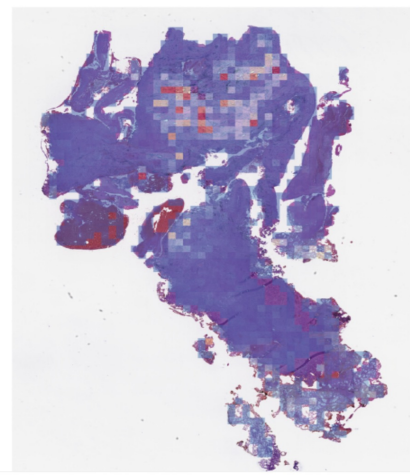
Beyond Histopathology and Conclusions

- *J&J Innovative Medicine has end to end integration of Vison based AI models that accelerate the pipeline from drug discovery to development*
- *From discovering molecules to accelerating trials, AI is integrated with the core programs to help patients get the best therapy and increase value for the organization*
- *Beyond histopathology, there are models integrated for Radiomics(X-Rays, CTs, MRIs), Skin Images, Endoscopy Videos and many more imaging modalities.*

Acknowledgments

JRD DSDH – Computer Vision:
Kristopher Standish
Albert Juan Ramon
Stephen Yip
Pablo Damasceno
Brendon Lutnick
Brandon Ginley
Krishna Chaitanya
Pooya Mobadersany
Erik Burlingame
Darshana Govind
Zijun Gao
Tharindu De Silva
Shreyas Patil
Amit Sharif Kamran
Fatemeh Koochakighermezcheshme

JRD DSDH – Platforms:
Io Flament
Carlos Csiszer
Eric Ho
Joseph Rivas
Pratik Patel

JRD DSDH – SLT:
Najat Khan
Tommaso Mansi

Data Science