Radiel Health (Team 15)

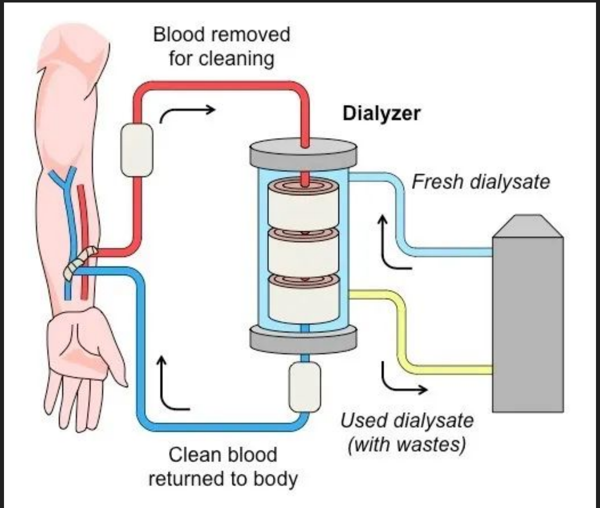
Rishabh Sharma, Nicholas Jiang, Sarvesh Sivakumar

Mission and Pipeline Recap

Who are we again?

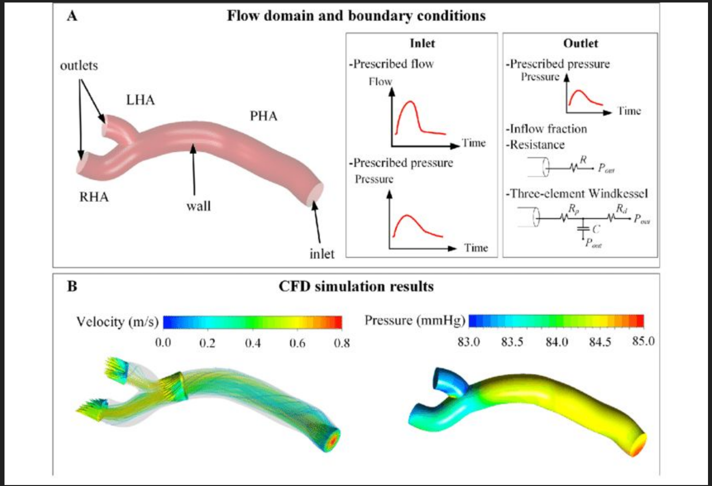
Why would Medicine need this?

- Medical diagnosis is very observational, but attempts to add more quantitative metrics have been underway.
- One of these initiatives is using Computational Fluid Dynamics (CFD) to predict key parameters such as Wall Shear Stress.
    - In a procedure where AV Fistulas are created for Kidney Dialysis, 70% of fistulas failed to mature after a year (thicken appropriately in relation to the rest of the surrounding arteries).
    - After the addition of CFD metrics, this failure rate dropped to 10-20%!



Limitations of CFD and Our Role

- **Specialized Expertise Required**
  - Setting up and interpreting simulations requires trained engineers and clinicians.
- **High Compute Time and Cost**
  - Patient-specific runs can take ~6 hours and ~$2 000 per case



- Our solution addresses limitations of traditional CFD.
- Our first focus is on improving patient outcomes in dialysis.
- Machine learning predicts key flow parameters rapidly.
- This technology lowers costs and resource burdens.

Project Abstract

Can a machine-learning surrogate, trained on high-fidelity CFD outputs from patient CT scans or ultrasounds, predict key flow parameters (pressure, velocity, shear stress) in seconds with accuracy comparable to traditional Ansys Fluent simulations?

Why This Matters:

- ***Enables Real-Time Clinical Decision Making***
- ***Improves Patient Outcomes***
- ***Lowers Cost & Resource Burden***

Our Solution

How do we plan on solving this problem?

Full User Pipeline

1. Clinician starts off with a series of patient ultrasounds or CT scans

2. Platform turns those images to a 3D Mesh

3. Model is run and prediction of hemodynamic values on new mesh is outputted.

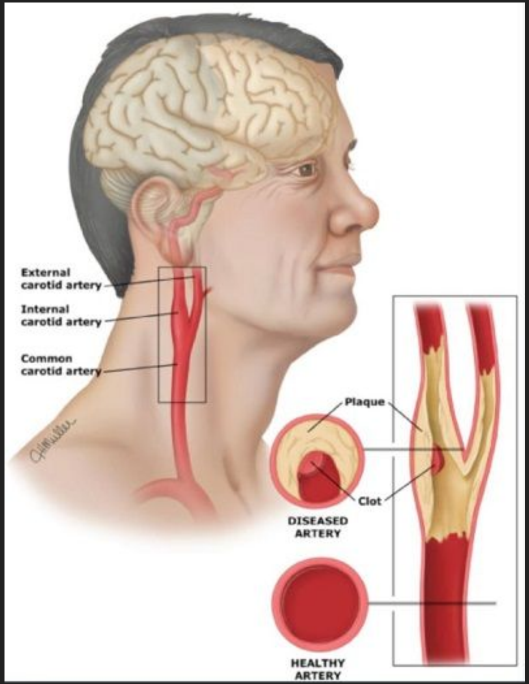
4. Those values are then used by clinicians to better inform their decisions

Mentor Update

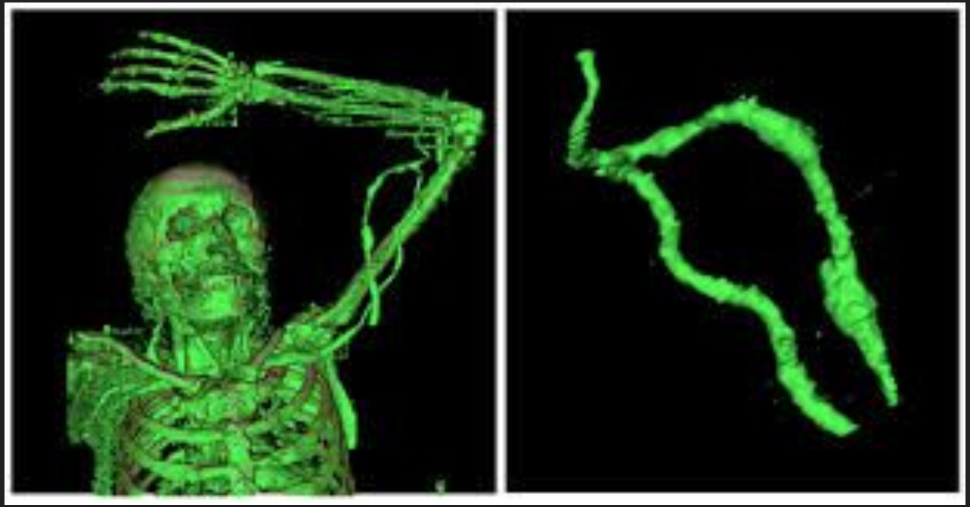
- Dr. Sean Peterson → MME Faculty at UWaterloo and running the UW Fluid Flow Physics group, mentoring us in the development of models.
- Dr. Aaron Fenster → Medical Physicist from UWO, specializes in Vascular Imaging and Image Guidance. He was key to the invention of 3D Ultrasound!

Our Data Update!

- As we're working with medical data, getting access to it is not trivial…
- However, Dr. Fenster's team at UWO has this dataset of 600 carotid arteries in 3d mesh form.
    - *This is perfect for us!* Having such a high amount of vascular meshes will be more than enough data for us to train a model.
    - *But wait? It's not the same artery from the AV Fistula procedure?* → Our main objective is to prove that's it's possible to predict CFD metrics in vascular structures. This model will allow us to generate this proof of concept and a working architecture!
- Currently, we're in the final stages of getting ethics approval. As this dataset has been in countless of other studies in Ontario, approval is almost guaranteed (🤞).



Impact of Progress on Model Development

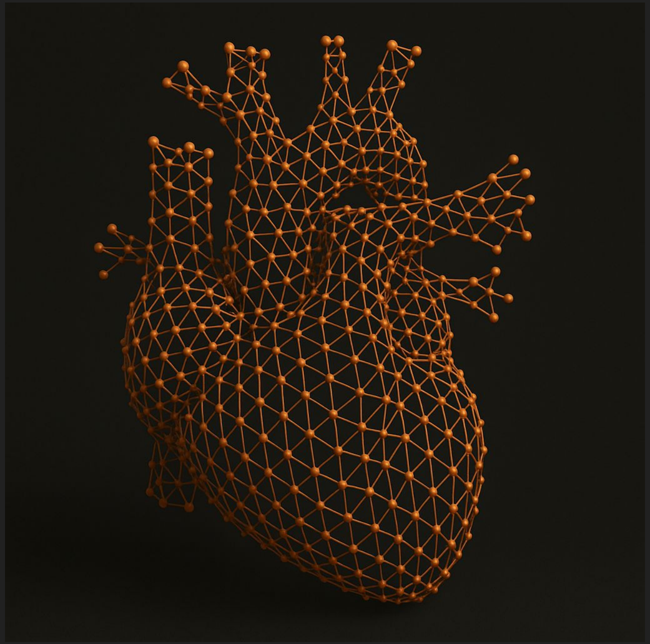


- In the interim, we’re using a public dataset of 6 AV Fistula meshes to test out different architectures.
- We’re using CFD ground truth for comparison.
- What is acceptable error?
  - Clinical input is needed, which Dr.Fenster provides!

So…What’s the Progress on our Goals?

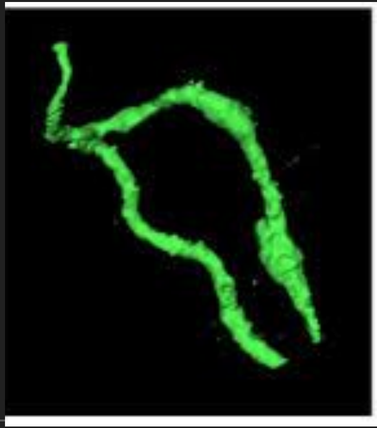
What We Defined as Success (Overview)

1. Creating an accurate model simulation of a vascular artery with a neural network (PINN/GNN)
    - Prediction accuracy will be measured against performance of existing systems
2. Prediction generalization on a new unseen dataset
3. Complete our user pipeline



Goal: Predict Accurate AV Fistula Metrics

- **Progress:** Not as far as we thought…
- **Why?**
  - We're only using a public dataset of 6 meshes (not a lot of data)
  - No state-of-the-art (SOTA) for our machine learning models to compare against.
  - We only have CFD ground truth for comparison.
  - What is acceptable error?
    - We didn't have an expert opinion
- **Resolution?**
  - Transition to carotid artery dataset from Prof. Aaron Fenster → generate proof of concept
  - Make use of Dr. Fenster's expertise for what is clinically acceptable error
  - Create AV Fistula data after success with proof of concept success under supervision of Dr.Fenster's lab.



Goal: Ensure the model generalizes to new, unseen datasets

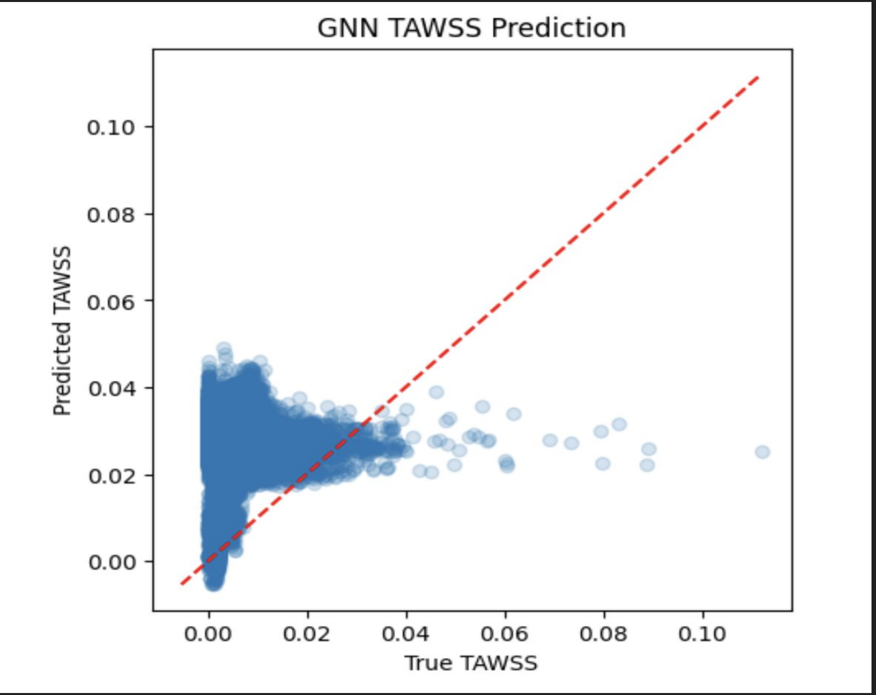
- Target: Prediction accuracy on unseen (within-distribution) data should be within 5% of training accuracy (if lower)

Progress: On Track

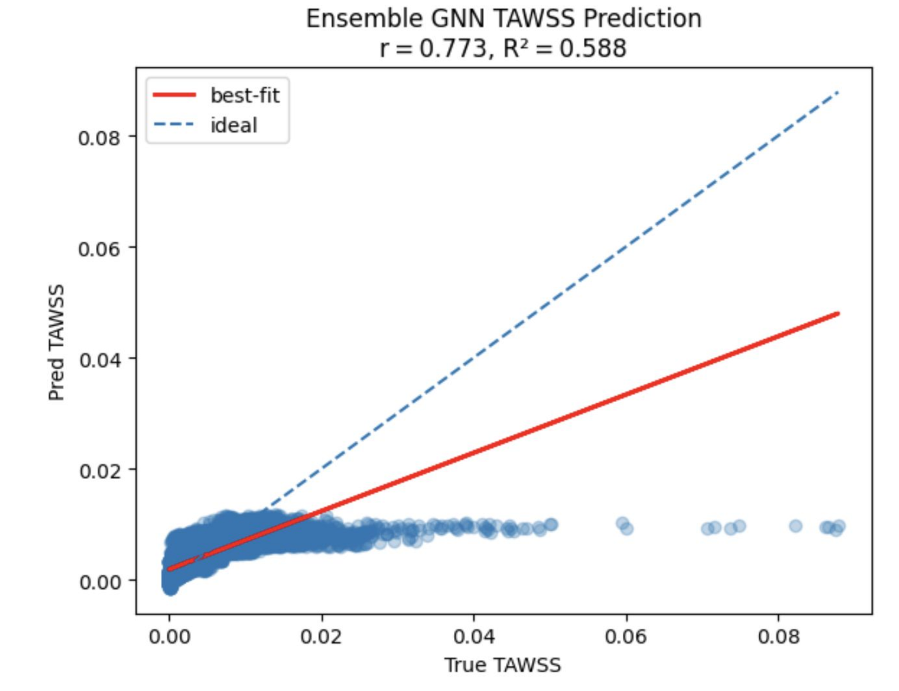
- Employed leave-one-dataset-out strategy: train on 5 datasets, test on 1 never-seen dataset.
- This setup simulates real-world deployment on unseen data.

Our Progress

GNN Only

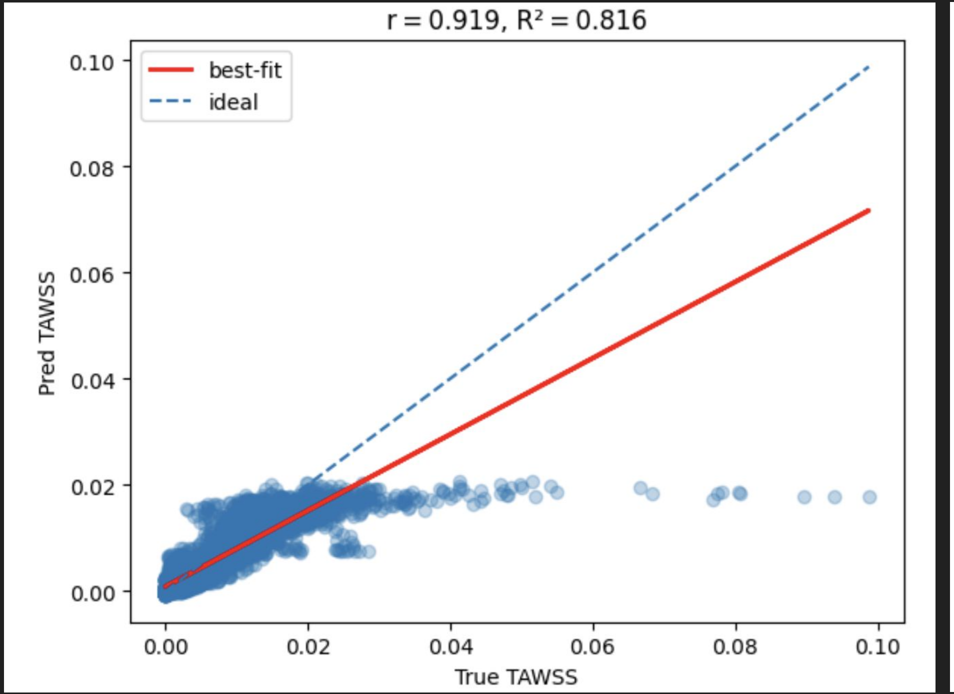


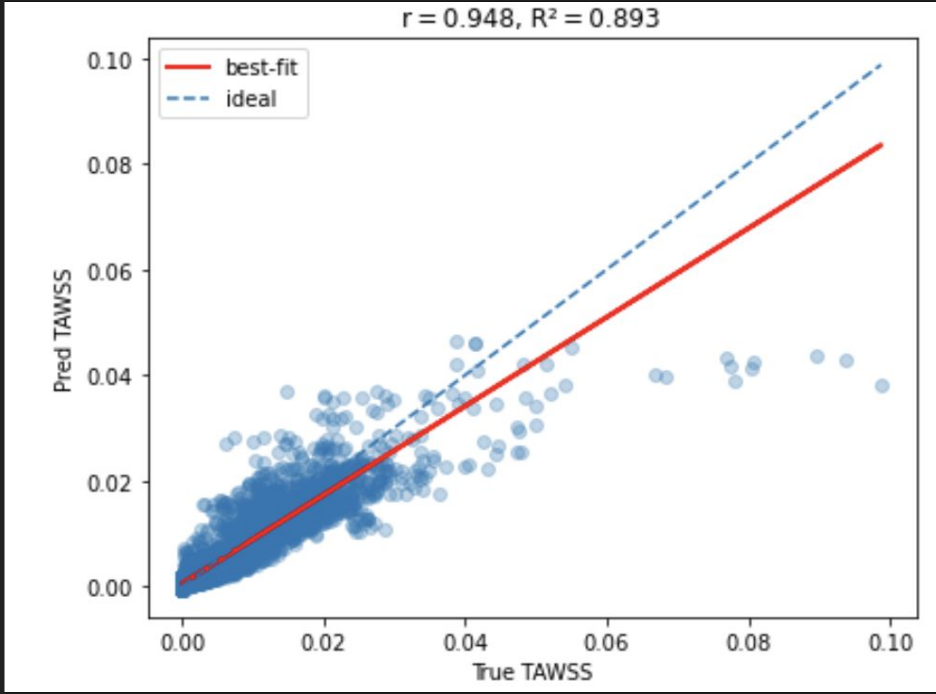
GNN + MPNN +Ensemble



GNN + MPNN (With more Layers)

GNN + UNet + MPNN





```
Folds: 0it [00:00, ?it/s]

Folds: 0it [00:11, ?it/s]
Fold 1/3 Epochs:   0%|          | 0/2 [00:11<?, ?it/s]
Fold 1/3 Epochs:  50%|█████     | 1/2 [00:11<00:11, 11.04s/it]
Epoch 1: Train 0.1043 / Test 0.1432

Folds: 0it [00:21, ?it/s]

Folds: 0it [00:21, ?it/s]
Fold 1/3 Epochs:  50%|█████     | 1/2 [00:21<00:11, 11.04s/it]
Folds: 1it [00:21, 21.98s/it]
Epoch 2: Train 0.0947 / Test 0.1697
No improvement for 2 epochs, stopping early.
Fold 1 Best Loss: 0.1362


Folds: 1it [00:33, 21.98s/it]
Fold 2/3 Epochs:   0%|          | 0/2 [00:11<?, ?it/s]
Fold 2/3 Epochs:  50%|█████     | 1/2 [00:11<00:11, 11.08s/it]
Epoch 1: Train 0.1310 / Test 0.1465

Folds: 1it [00:44, 21.98s/it]

Folds: 1it [00:44, 21.98s/it]
Fold 2/3 Epochs:  50%|█████     | 1/2 [00:22<00:11, 11.08s/it]
Folds: 2it [00:44, 22.21s/it]
Epoch 2: Train 0.1311 / Test 0.1313
No improvement for 2 epochs, stopping early.
Fold 2 Best Loss: 0.1223


Folds: 2it [00:56, 22.21s/it]
Fold 3/3 Epochs:   0%|          | 0/2 [00:12<?, ?it/s]
Fold 3/3 Epochs:  50%|█████     | 1/2 [00:12<00:12, 12.15s/it]
Epoch 1: Train 0.0942 / Test 0.0795

Folds: 2it [01:08, 22.21s/it]

Folds: 2it [01:08, 22.21s/it]
Fold 3/3 Epochs:  50%|█████     | 1/2 [00:24<00:12, 12.15s/it]
Folds: 3it [01:08, 22.91s/it]
Epoch 2: Train 0.0937 / Test 0.0799
No improvement for 2 epochs, stopping early.
Fold 3 Best Loss: 0.0796
```

```
Loaded 3 folds → ensemble ready

Ensemble Test MSE: 0.089569
```

(0.1362+0.1223+0.0796)/3=0.1127

0.1127 Train vs 0.08957 test => test does better

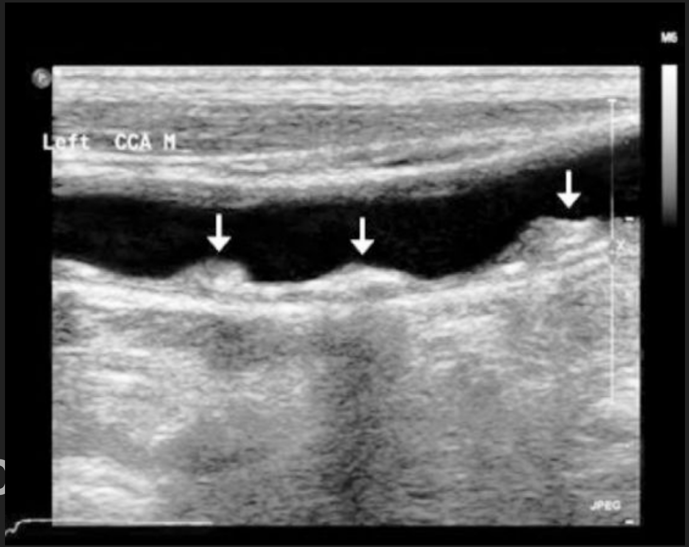
Goal: Complete User Pipeline

Progress: On Track (~50%)

Completed: We already have a working prototype for part 3 of our user pipeline.

To be done:
We are working with Prof. Aaron Fenster to acquire the tool for converting ultrasound to 3D mesh.

We need to plan out and execute how to integrate this tool into our pipeline



Full User Pipeline

1

Clinician starts off with a series of patient ultrasounds or CT scans

2

Platform turns those images to a 3D Mesh

3

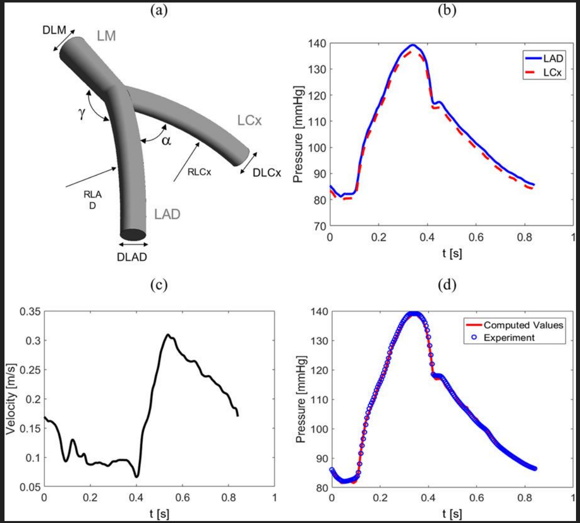
Model is run and prediction of hemodynamic values on new mesh is outputted.

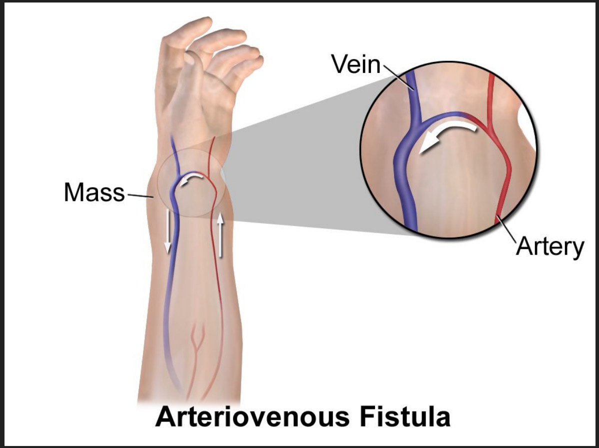
4

Those values are then used by clinicians to better inform their decisions

Making Use of Model Predictions

- Key Metrics
    - Oscillatory Shear Index (OSI)
        - Quantifies directional changes in wall shear; higher OSI indicates disturbed flow
        - Time-Averaged Wall Shear Stress (TAWSS)
    - Time-Averaged Wall Shear Stress (TAWSS)
        - Measures mean shear stress magnitude; lower TAWSS suggests low-flow regions
- Quantify flow disturbance and shear magnitude to map vessel “openness”





Making Use of Model Predictions

- **Arterial “Openness”**
  - Combined OSI & TAWSS map where the vessel is well-perfused vs. prone to flow disruption
- **Risk Stratification**
  - High-OSI/low-TAWSS zones ⟶ ↑ risk of thrombosis (clot formation)
  - Persistent low-flow areas ⟶ ↑ risk of stenosis (vessel narrowing or closure)
- **Predicting Maturation Probability**
  - Use pre-operative OSI/TAWSS profiles to estimate access maturation likelihood
  - Quantify probability of successful fistula maturation based on regional hemodynamics
- **Surgical Site Selection**
  - Evaluate potential access locations
  - Recommend sites with the most favorable OSI/TAWSS signature
  - Minimizes predicted thrombosis and stenosis rates

Challenges & Solutions (Overview)

- Data scarcity (6 AVF meshes)
  - Solution: Data Calibration
- Scattered Teamwork (What are each of us doing?)
  - Solution: More granular Project Management + Discord Forums
- Costly CFD data generation

Problem: Data scarcity / Solution: Data Calibration

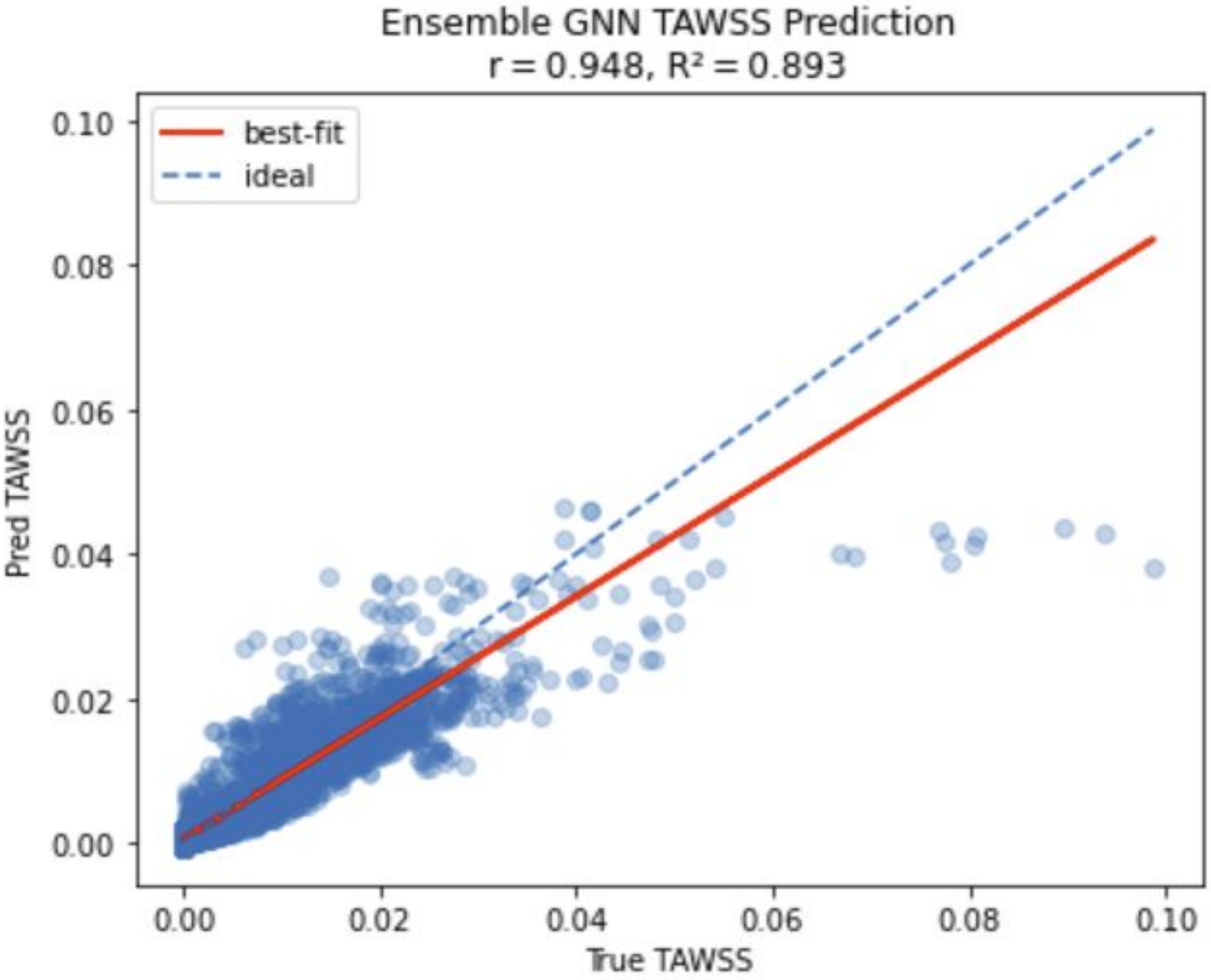
Description:

With only 6 meshes, model improvements were hard to find. Scaling up the network reduced training error but didn't improve generalization. We can't just sit around while data is scarce.

Solution:

Post-prediction stratified residual modelling and adjustment.

Figure 1: Scientist frustrated by lack of training data



Our best until now

Idea: Use the residuals from 70% of predictions to calibrate a model that adjusts the model solution and apply that to remaining 30%

What if we could fit a model that "nudges" the predictions in the right direction.

Fit linear correction

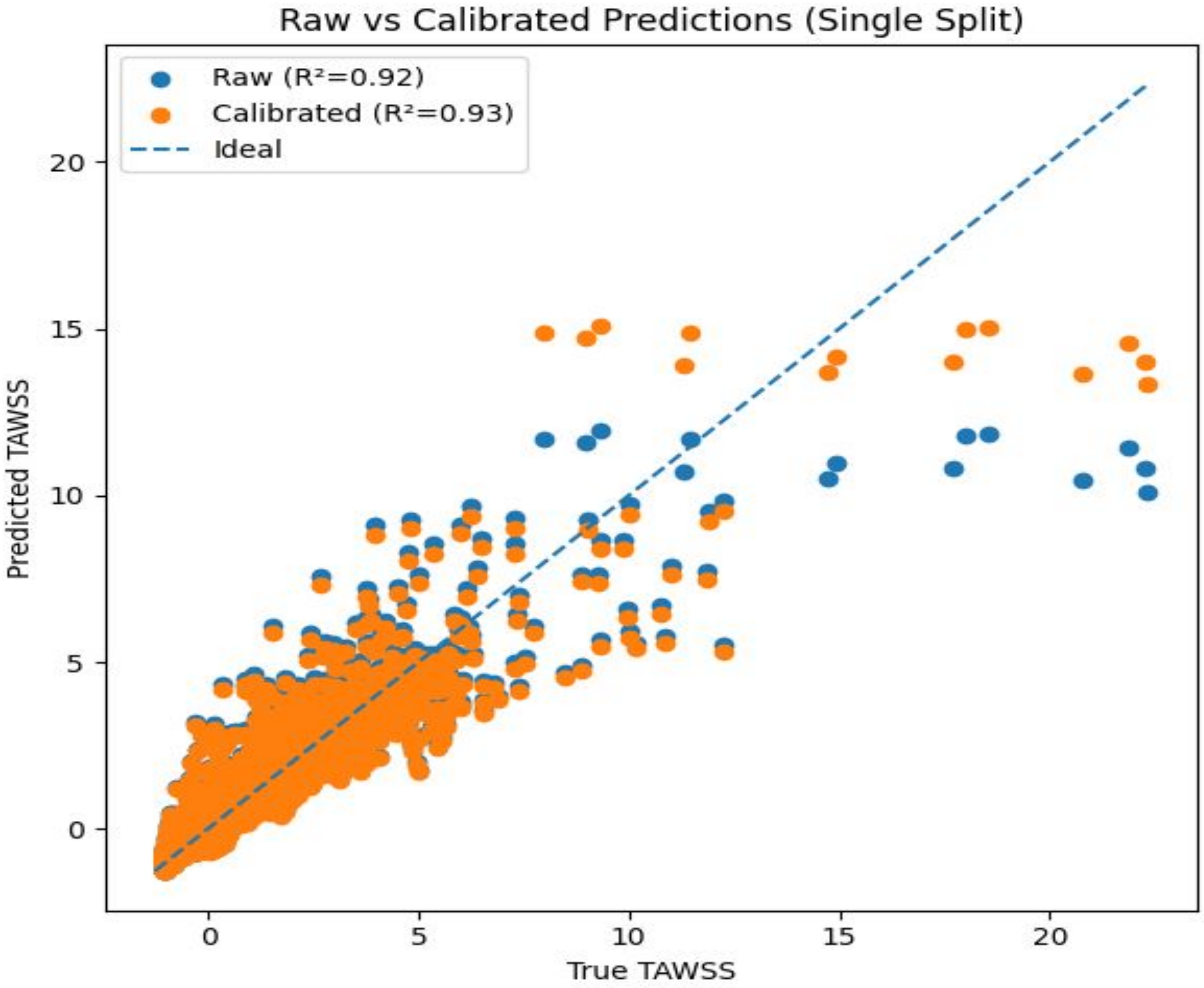
$$\min_{a,b} \sum (y_{\text{train}} - (a\, p_{\text{train}} + b))^2 \quad \Longrightarrow \quad \hat{y}_{\text{lin}} = a\, p + b.$$

Capture shape with residuals

$$r = y_{\text{train}} - \hat{y}_{\text{lin}}, \quad \delta = \text{Iso}(p_{\text{train}},\, r).$$

Apply to new data

$$y_{\text{cal}}(p) = a\, p + b + \delta(p).$$



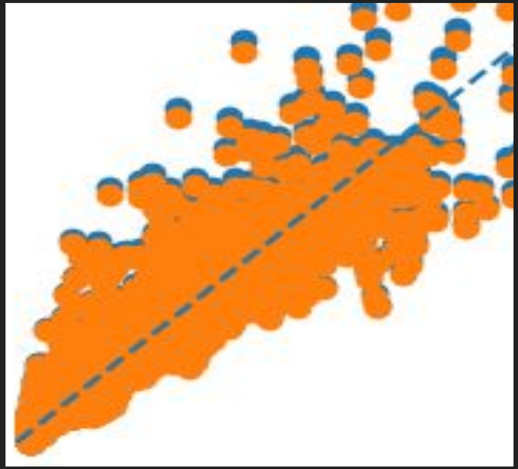
Results are better by ΔR^2 =0.004 compared to raw preds on 200 Monte Carlo runs

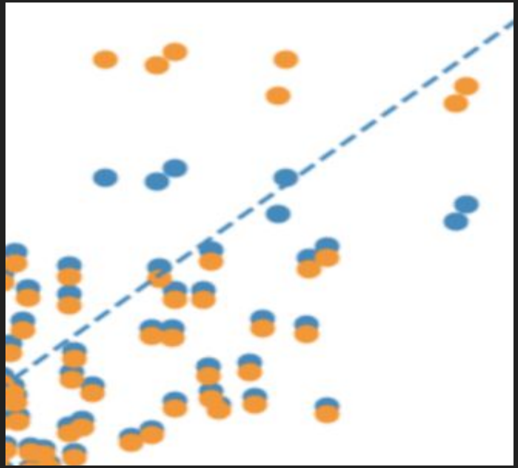
But notice that we're adjusting points in the wrong direction sometimes

Observation:

We actually don’t have much data in the mid region. See how “packed” the data is in the first part of the graph compared to the middle?

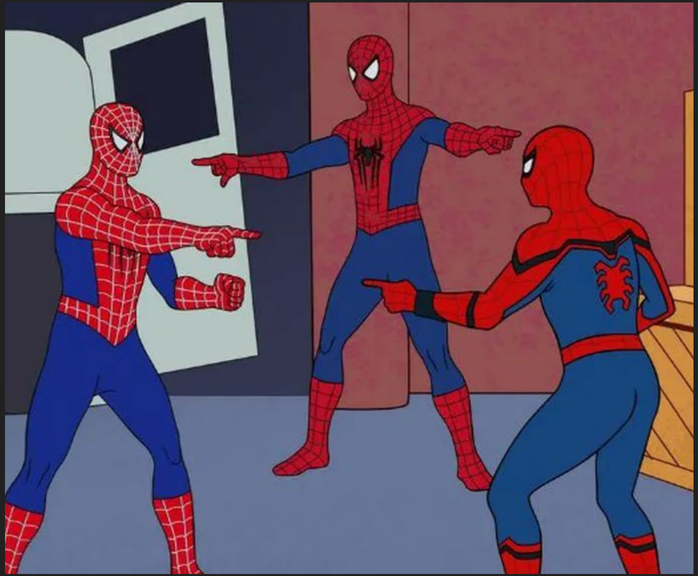
Can we say that the model is just as certain about the results of the first third of graph compared to middle?



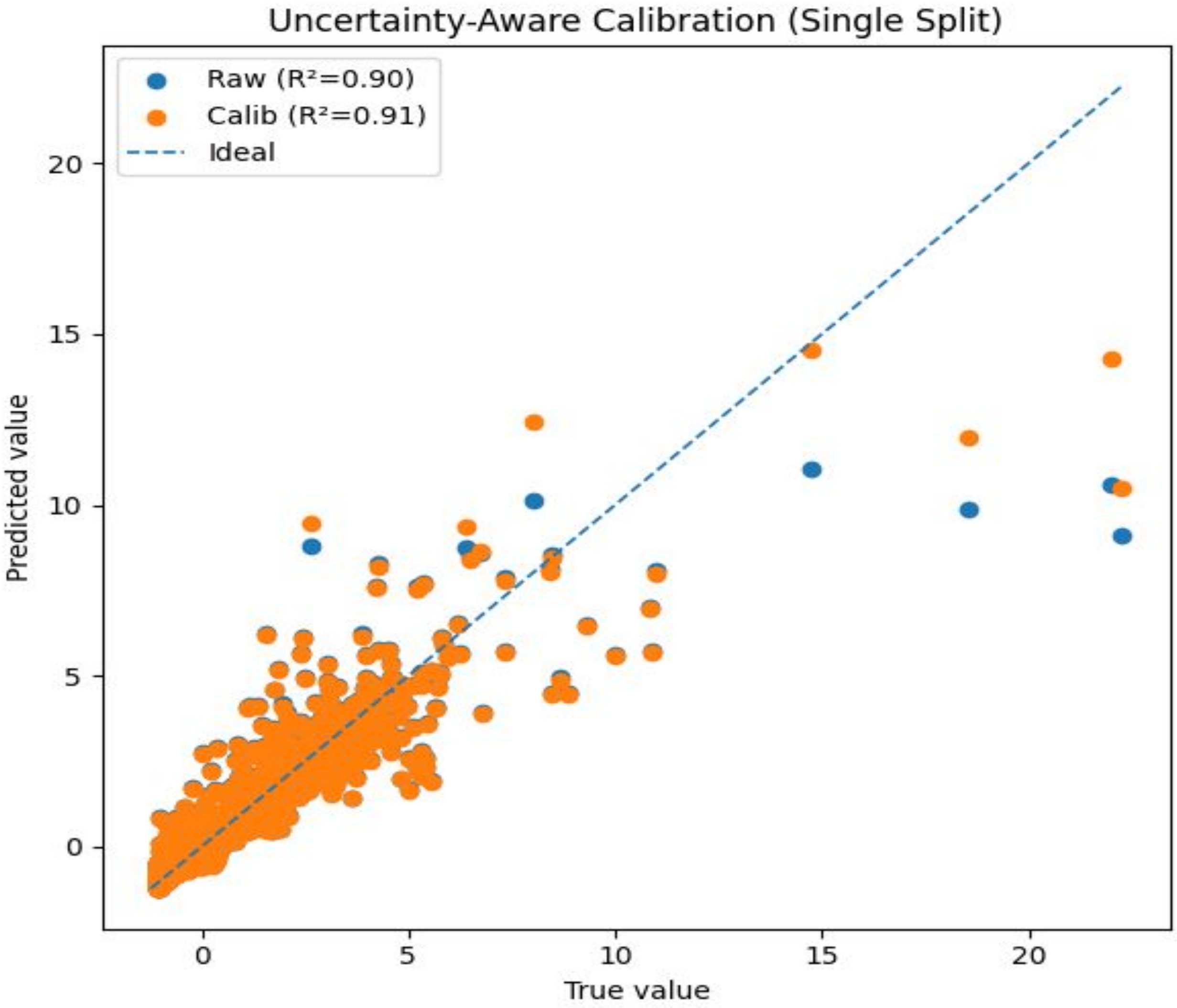


Quantifying Uncertainty

- How certain is the model of a prediction relative to others?
- Exist Bayesian methods but applying them wasn't feasible for us (needed to learn a lot of theory, etc.)
- We opted for a simple proxy that turned out to be effective
- Standard deviation of k fold model averaging (disagreement of model predictions)



$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$



Applies the adjustment less to data points of high disagreement (i.e. proxy for confidence)

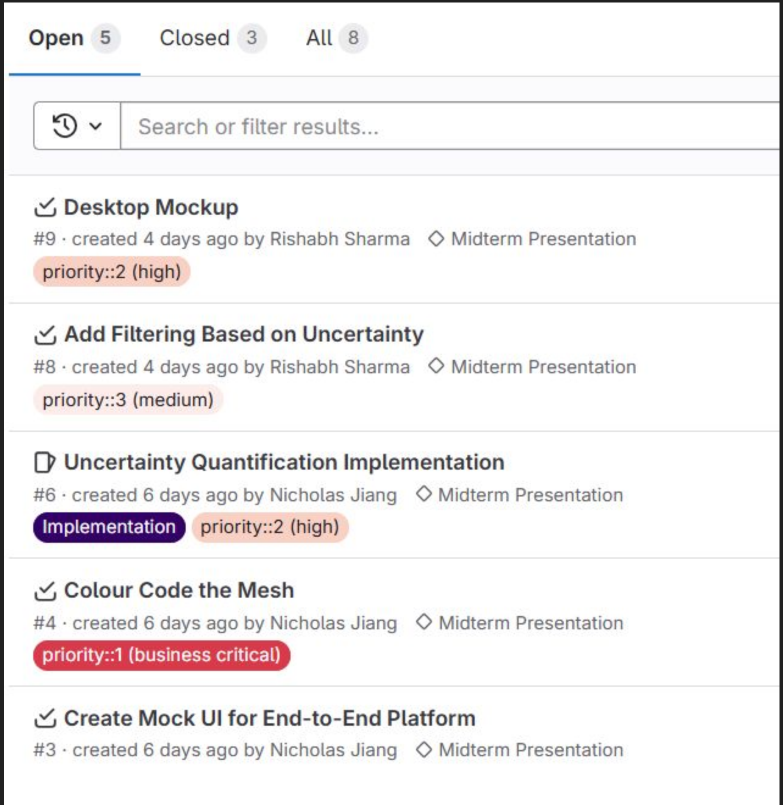
It can be seen that we adjust in the wrong direction less often (esp. in middle section)

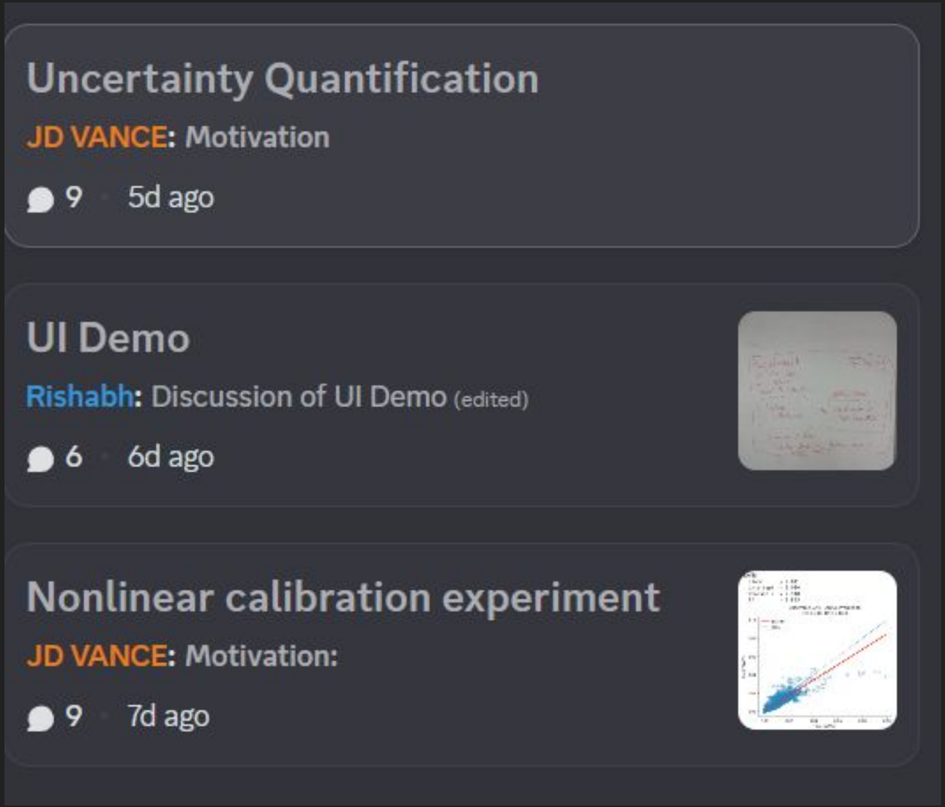
Improvement by $\Delta R^2 = 0.008$ compared to raw (doubled calibration performance)

Problem: Scattered Teamwork

- Originally, we all worked on our respective models locally + reconvened once a week to talk about results.
- We also didn't have a mentor to guide our work throughout the week and give expert advice on what we can do next.
- **Now, we:**
  - **Use Gitlab + Discord Forums to track tickets and progress**
  - **Meet with professor mentors on Wednesdays @ 9am**
  - **Meet with team Wednesdays @ 2pm and Saturdays @ 3pm**

Project Management Screenshots





Problem: High Cost for CFD data generation

- First, our team used to generate the 3D Mesh data manually, then run it through a CFD solver for us to get the data we need.
- However, this was costly, took a lot of time and we couldn't ensure medical accuracy to the meshes we made.

Problem: High Cost for CFD data generation

- Now, we have 600 vascular structures from ***real*** patient data available for us, ensuring high quality data.
- After training our preliminary data, we'll be working with Dr.Fenster's lab to create AV Fistula data!
    - Creating medically accurate "test dummies" and getting real patient data from the UWO's associated hospitals.

Lessons Learned

- **Start small and focused.**
  - One of the most powerful lessons from our mentor Prof. Sean Peterson was learning to narrow our scope**.**
  - We wanted to build out the entire pipeline on our own and this just wasn't feasible for the time we had. Instead, we focused on model predictions and interpretation and we are now looking to acquire the ultrasound to mesh part from Prof. Aaron Fenster
  - It's better to have a small project that is completed than a big project that isn't
- **Big ideas are cool but implementable ideas get done**
  - How to solve the model scarcity problem? Well we could have used GANs to try to sample from a distribution of synthetic data to train our models but this is very complex for mesh data and real valued predictions. Also very expensive computationally
    - Instead we opted for a simpler but achievable technique of model uncertainty quantification and validation.
  - In every project there's going to be a tradeoff between ambition and achievability

Remaining Work

- Group nodes with similar characteristics together to allow for easier human interpretation
- Integrate Prof. Fenster's ultrasound to mesh into our pipeline
- Train/Evaluate our model on the 600 carotid artery meshes