# A Standardized Framework for Neuroimaging-Based Traumatic Brain Injury Outcome Prediction: Methodological Considerations and Best Practices

Ingrid Sørensen[1,*], Mikkel Rasmussen[1], Sofia Lindholm[1], Jakob Nygaard[1,2], Elena Petrova[1], Anders Holm[1], Anna Galli[1]

[1]NeuroTrauma Computational Imaging Group, Aalborg University, Denmark
[2]Department of Neurology, Aalborg University Hospital
[*]Corresponding author: ingrid.sorensenborg@gmail.com

## Abstract

Traumatic brain injury (TBI) affects millions worldwide, with outcomes ranging from complete recovery to severe disability or death. Accurate prediction of patient outcomes is crucial for clinical decision-making, resource allocation, and family counseling. While neuroimaging-based predictive models show promise, the field lacks standardization in data acquisition, preprocessing, feature extraction, and model validation. This paper presents a comprehensive methodological framework for developing and validating neuroimaging-based TBI outcome prediction models. We address critical considerations including multi-site data harmonization, quality control protocols, clinically relevant outcome definitions, and uncertainty quantification. Our framework emphasizes reproducibility, interpretability, and clinical translatability. We provide concrete recommendations for each stage of the predictive modeling pipeline and discuss common pitfalls and their solutions.

**Keywords:** traumatic brain injury, neuroimaging, outcome prediction, methodology, machine learning, clinical translation

## 1. Introduction

Traumatic brain injury represents one of the most significant public health challenges globally, with an estimated 69 million individuals sustaining TBI each year. The heterogeneity of injury mechanisms, pathophysiology, and patient characteristics makes outcome prediction particularly challenging. Traditional prognostic models relying on clinical variables alone achieve modest predictive accuracy, motivating the integration of advanced neuroimaging biomarkers.

Recent advances in magnetic resonance imaging (MRI) and computational methods have enabled extraction of quantitative biomarkers reflecting microstructural damage, connectivity disruption, and neuroinflammatory processes. However, the translation of these biomarkers into clinically useful predictive tools has been hampered by methodological inconsistencies across studies. Issues including small sample sizes, single-site validation, inconsistent outcome definitions, and lack of uncertainty quantification limit the generalizability and clinical adoption of proposed models.

This paper aims to address these limitations by presenting a standardized methodological framework for neuroimaging-based TBI outcome prediction. We draw on our group's experience developing and validating predictive models across multiple cohorts and synthesize best practices from the broader neuroimaging and machine learning literature.

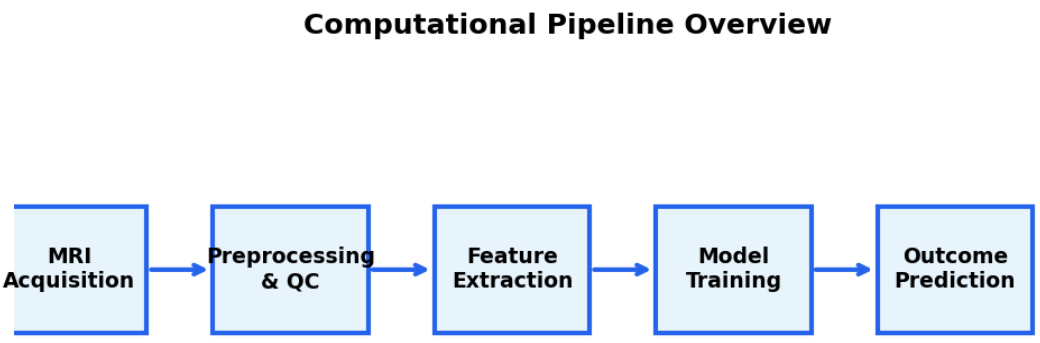


**Figure 1.** Overview of the proposed computational pipeline for TBI outcome prediction. Each stage incorporates specific quality control measures and standardization procedures.

## 2. Data Acquisition and Harmonization

Successful multi-site studies require careful attention to acquisition protocol harmonization. We recommend establishing minimum requirements for scanner field strength (≥1.5T, preferably 3T), spatial resolution (≤1mm isotropic for structural imaging), and sequence parameters. For diffusion tensor imaging, a minimum of 30 gradient directions with b-values of 1000-2000 s/mm² provides adequate angular resolution for robust tensor estimation.

Post-acquisition harmonization using tools such as ComBat or deep learning-based methods can reduce site-related variance while preserving biologically meaningful signal. We recommend including traveling phantoms or overlapping subjects across sites when feasible to empirically quantify and correct for scanner effects.

## 3. Preprocessing and Quality Control

Preprocessing pipelines should be fully automated and version-controlled to ensure reproducibility. For structural MRI, standard steps include bias field correction, skull stripping, tissue segmentation, and spatial normalization. For diffusion imaging, eddy current and motion correction, followed by tensor or higher-order model fitting, are essential. We strongly recommend visual quality control at multiple stages, supplemented by automated quality metrics.

Common quality control metrics include signal-to-noise ratio, motion parameters (framewise displacement), and registration accuracy. Establishing explicit exclusion criteria before analysis prevents bias from post-hoc data selection. We recommend excluding scans with framewise displacement exceeding 2mm or signal dropout affecting regions of interest.

## 4. Feature Extraction

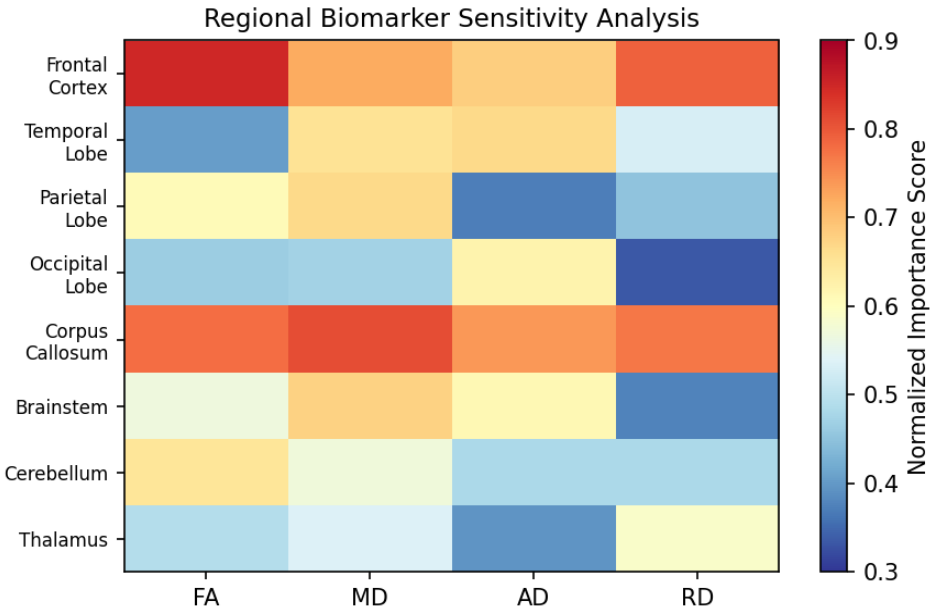


**Figure 2.** Regional biomarker sensitivity analysis showing the relative importance of different brain regions and diffusion metrics for outcome prediction. FA = fractional anisotropy, MD = mean diffusivity, AD = axial diffusivity, RD = radial diffusivity.

Feature extraction should be guided by both data-driven and hypothesis-driven approaches. Regions frequently implicated in TBI include the corpus callosum, brainstem, and frontal white matter. Standard atlases (e.g., JHU white matter atlas, AAL) enable consistent region definition across studies. We recommend extracting features at multiple spatial scales: voxel-wise, regional averages, and network-level measures.

## 5. Outcome Definition and Timing

The Glasgow Outcome Scale-Extended (GOS-E) remains the most widely used outcome measure in TBI research. We recommend dichotomizing GOS-E at the favorable/unfavorable boundary (GOS-E ≥5 vs <5) for primary analyses, with secondary analyses examining the full ordinal scale. Outcome assessment timing significantly impacts predictive accuracy; 6-month outcomes balance stability with clinical relevance for acute decision-making.

Mortality prediction requires separate consideration, as imaging biomarkers may have different relationships with survival versus functional outcomes. We recommend treating mortality as a distinct endpoint rather than combining with functional outcomes.

## 6. Model Development and Validation

Model selection should consider interpretability requirements alongside predictive performance. For clinical applications, models that provide feature importance and uncertainty estimates are strongly preferred. Ensemble methods (random forests, gradient boosting) offer good performance with inherent feature importance, while linear models maximize interpretability.

Validation strategy critically determines generalizability estimates. We recommend nested cross-validation for internal validation, with held-out external cohorts for final testing. Bootstrap confidence intervals should accompany all performance metrics. Area under the ROC curve (AUC) provides discrimination assessment, while calibration plots reveal systematic prediction biases.

| Validation Strategy | Use Case | Limitations |
|---|---|---|
| K-fold CV | Limited data, internal validation | Optimistic estimates |
| Nested CV | Hyperparameter tuning | Computationally intensive |
| LOOCV | Very small samples | High variance |
| External validation | Generalizability testing | Requires additional data |
| Temporal validation | Deployment simulation | May confound with drift |

**Table 1.** Comparison of validation strategies for predictive model development. CV = cross-validation, LOOCV = leave-one-out cross-validation.

## 7. Uncertainty Quantification

Clinical deployment requires honest uncertainty estimates. Point predictions alone are insufficient for high-stakes medical decisions. We recommend conformal prediction or Bayesian approaches to generate prediction intervals. Models should be calibrated such that stated confidence levels match empirical coverage rates.

Epistemic uncertainty (model uncertainty) and aleatoric uncertainty (inherent outcome variability) should be distinguished when possible. High epistemic uncertainty may indicate out-of-distribution inputs requiring clinical review, while high aleatoric uncertainty reflects fundamental unpredictability.

## 8. Clinical Translation Considerations

The gap between research models and clinical tools remains substantial. Successful translation requires integration with existing clinical workflows, interpretable outputs for non-expert users, and regulatory compliance. We recommend early engagement with clinical stakeholders to define acceptable performance thresholds and output formats.

Prospective validation studies with pre-registered analysis plans provide the strongest evidence for clinical utility. Decision curve analysis can quantify net benefit across threshold probabilities, informing the clinical contexts where model deployment is justified.

## 9. Conclusions

Neuroimaging-based TBI outcome prediction holds significant promise for improving clinical care. Realizing this potential requires rigorous methodological standards across the development pipeline. This framework provides concrete guidance for researchers developing predictive models, with emphasis on reproducibility, interpretability, and clinical relevance. Future work should focus on multi-site validation studies, integration of multimodal data, and prospective clinical trials of model-guided decision-making.

## References


1. Maas AIR, et al. Traumatic brain injury: integrated approaches to improve prevention, clinical care, and research. Lancet Neurol. 2017;16(12):987-1048.

2. Dewan MC, et al. Estimating the global incidence of traumatic brain injury. J Neurosurg. 2018;1-18.

3. Steyerberg EW, et al. Predicting outcome after traumatic brain injury: development and international validation of prognostic scores. PLoS Med. 2008;5(8):e165.

4. Fortin JP, et al. Harmonization of multi-site diffusion tensor imaging data. Neuroimage. 2017;161:149-170.

5. Johnson WE, et al. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118-127.