RESEARCH QUESTIONS: SCIENTIFIC JUSTIFICATION

Proposed Internship Research - LIH (January–March 2026)

Prepared by: Moad Hani, MSc
PhD Candidate, University of Mons (UMONS), Belgium
Expected Defense: September 2026

RESEARCH QUESTION 1: GENETIC STRATIFICATION OF DISEASE SUBTYPES

Question: Do genetic variants (GBA, LRRK2, idiopathic) map to specific SuStaIn subtypes?

Hypothesis: GBA mutation carriers exhibit accelerated progression and preferentially cluster within the rapid-progression SuStaIn subtype.

Rationale

Parkinson's disease manifests across multiple genetic etiologies, with GBA and LRRK2 mutations representing the most prevalent risk factors. This research question investigates whether genetically distinct patient populations exhibit clinically distinguishable progression trajectories as captured by SuStaIn event-based modeling. Establishing genotype-phenotype correlations would validate the hypothesis that molecular genetic dysregulation manifests as observable clinical heterogeneity.

Methodological Assets

My prior validation of the SuStaIn algorithm on the PPMI cohort (N=1,358 patients) yielded two statistically distinct subtypes with 87% stage prediction accuracy. The LuxPark cohort provides direct access to patient-level GBA and LRRK2 mutation status (approximately 50 carriers per variant), enabling genetic stratification unavailable in public datasets. Longitudinal follow-up exceeding four years per patient ensures robust subtype assignment across repeated clinical assessments.

Translational Significance for LIH

This analysis validates Dr. Hemedan's Boolean modeling hypothesis linking genetic mutations to specific pathway dysregulation (FOXO1 in GBA-PD, mitochondrial dysfunction in both variants). Demonstrating that genetic variants correlate with clinically defined SuStaIn subtypes establishes the critical bridge between molecular genotype and observable clinical phenotype, confirming that

computational pathway predictions translate to measurable disease progression patterns.

Extension of Doctoral Research

My doctoral work employed only idiopathic PD patients without genetic stratification. RQ1 extends this methodology to incorporate genetic heterogeneity, demonstrating that SuStaIn identifies clinically meaningful patient subgroups even within genetically homogeneous subcategories - thereby enriching the precision medicine applications of this staging framework.

RESEARCH QUESTION 2: CLINICAL-MOLECULAR STAGING INTEGRATION

Question: Does SuStaIn stage correlate with Boolean pathway activity (FOXO1, mitophagy, glycolysis)?

Hypothesis: Advanced SuStaIn stages exhibit elevated FOXO1 activity and impaired mitophagy, consistent with Boolean-predicted signatures of insulin resistance-mediated neurodegeneration.

Rationale

Dr. Hemedan's Boolean networks predict temporal evolution of pathway dysregulation across disease stages based on transcriptomic and miRNA profiles. However, these molecular predictions require validation against longitudinal clinical outcomes. This research question tests whether computationally inferred pathway states correlate with empirically defined disease stages, thereby confirming that Boolean models capture biologically relevant progression dynamics rather than statistical artifacts.

Methodological Assets

The LuxPark cohort provides 810 SuStaIn-staged patients, each assigned to clinical stages 0–5 based on longitudinal biomarker progression. Access to 830 blood miRNA profiles enables pathway activity inference via established miRNA-pathway regulatory signatures. Multiple assessment timepoints per patient permit tracking of pathway evolution across stage transitions. Fairness-aware imputation methodology ensures missing miRNA values are handled consistently across demographic and genetic strata, preventing systematic bias.

Translational Significance for LIH

This represents the critical validation step for Dr. Hemedan's Boolean framework. Demonstrating that FOXO1 activity increases systematically with SuStaIn stage progression confirms that molecular models accurately reflect real

disease biology. Such validation transforms computational predictions into clinically credible biomarkers suitable for precision medicine applications.

Extension of Doctoral Research

My doctoral work focused exclusively on clinical staging without molecular integration. RQ2 demonstrates that clinical progression patterns implicitly reflect underlying pathway dysregulation, validating GAMLSS-derived z-scores as capturing biologically meaningful deviation from normative aging trajectories rather than statistical noise.

RESEARCH QUESTION 3: TRANSLATIONAL PREDICTION MODEL

Question: Can integrated clinical staging and molecular pathway states improve levodopa-induced dyskinesia (LID) prediction accuracy?

Hypothesis: Incorporating SuStaIn stage and Boolean-predicted FOXO1 activity as model features will improve LID prediction from baseline AUC 0.72 to 0.80+, reflecting explicit capture of disease heterogeneity.

Rationale

Loo et al. (2024) demonstrated multi-cohort machine learning prediction of LID achieving AUC 0.67–0.72 using conventional clinical features including disease duration. However, duration represents a crude surrogate for biological progression, treating patients with identical diagnostic intervals as homogeneous despite potentially divergent underlying disease trajectories. Replacing duration with explicit biological staging and incorporating molecular pathway states should substantially improve predictive accuracy.

Methodological Assets

SuStaIn stage assignment captures biological progression rate independent of calendar time, providing superior information content compared to duration-based features. GAMLSS-derived severity z-scores quantify individual-level deviation beyond categorical stage membership, adding precision to risk stratification. Boolean pathway predictions from Dr. Hemedan's framework contribute molecular context absent from purely clinical models. LuxPark LID outcome data enables supervised model evaluation and comparative benchmarking.

Translational Significance for LIH

This analysis demonstrates clinical utility of integrated SuStaIn-Boolean frameworks. Improving AUC from 0.72 to 0.80+ corresponds to identifying approximately 13% additional at-risk patients before LID onset, enabling targeted pre-

ventive intervention in high-risk cases while avoiding unnecessary medication adjustment in low-risk patients - thereby operationalizing precision medicine principles.

Extension of Doctoral Research

RQ3 establishes that SuStaIn staging provides actionable clinical value beyond descriptive disease characterization. Demonstrating improved adverse outcome prediction validates patient stratification as enabling personalized intervention strategies, strengthening the translational impact of my methodological framework.

RESEARCH QUESTION 4: MOLECULAR PREDICTORS OF STAGE TRANSITIONS

Question: What clinical and molecular features predict stage transition events (disease progression)?

Hypothesis: Baseline miRNA profiles contain prognostic information predicting time-to-next-stage, enabling early identification of rapid progressors.

Rationale

Disease progression rates exhibit substantial inter-individual heterogeneity. Identifying patients destined for rapid progression at baseline - before clinical deterioration becomes evident - would enable early enrollment in neuroprotective trials and intensive monitoring protocols. This requires prospective validation that baseline molecular signatures predict future clinical trajectories, extending beyond retrospective staging.

Methodological Assets

Longitudinal SuStaIn stage assignments across four-year follow-up periods provide repeated stage transition events per patient. Baseline miRNA profiling (N=830 samples) captures molecular state at study entry. Cox proportional hazards survival analysis treats stage transitions as time-to-event outcomes, with miRNA signatures as predictive covariates. Dimensionality reduction via LASSO or principal components analysis identifies the most prognostically relevant miRNA subset. Fairness constraints ensure predictions maintain demographic parity across age and gender strata, preventing systematic bias.

Translational Significance for LIH

Dr. Hemedan's Boolean models predict which pathways drive accelerated progression. RQ4 tests whether baseline molecular states prospectively stratify progression risk, addressing the question: "Can pathway activity measurements

identify rapid progressors before symptom onset?" Affirmative findings would enable Dr. Hemedan's framework to directly support early risk stratification for clinical trial enrollment.

Extension of Doctoral Research

My doctoral work demonstrated SuStaIn's retrospective staging capability. RQ4 extends this to prospective prediction, transforming the methodology from descriptive characterization ("where is the patient now?") to prognostic forecasting ("where will the patient be?"), thereby establishing clinical decision support utility.

RESEARCH QUESTION 5: CROSS-COHORT REPLICATION AND GENERALIZABILITY

Question: Do LuxPark-derived subtypes generalize to independent PPMI and ICEBERG cohorts?

Hypothesis: External validation across geographically and methodologically diverse cohorts will yield AUC >0.75, indicating subtypes reflect fundamental disease biology rather than cohort-specific statistical patterns.

Rationale

Regulatory approval and clinical adoption require demonstrated generalizability across independent populations. LuxPark represents a European cohort with specific recruitment protocols and data collection standards. Replication in North American (PPMI) and German (ICEBERG) cohorts varying in population demographics, recruitment strategies, and assessment protocols would establish that identified subtypes capture universal disease mechanisms rather than population-specific artifacts.

Methodological Assets

My prior PPMI validation (N=1,358 patients) established a reproducible SuStaIn pipeline readily applicable to new datasets. LuxPark-derived models serve as the training framework for external testing. ICEBERG provides an independent European validation cohort with heterogeneous population characteristics. Standardized clinical assessments (UPDRS, MoCA) across all three cohorts enable harmonized cross-cohort analysis despite differing study designs.

Translational Significance for LIH

Regulatory bodies and high-impact journals require external validation demonstrating model robustness. Successful replication in PPMI (United States) and

ICEBERG (Germany) with AUC >0.75 elevates subtypes from statistical constructs to clinically credible biomarkers, enabling downstream precision trial design, biomarker-driven patient stratification, and Boolean model validation across diverse populations.

Extension of Doctoral Research

RQ5 completes my doctoral contribution. PPMI validation demonstrated North American applicability; European replication establishes cross-continental generalizability. This represents the methodological gold standard for validation studies and positions this work as foundational for international translational neuroscience applications.

INTEGRATED RESEARCH FRAMEWORK

The five research questions form a coherent translational arc progressing from mechanistic validation through clinical utility demonstration to generalizability confirmation.

Phase 1 (RQ1–RQ2): Mechanistic Foundation. RQ1 establishes that genetic variants cluster within specific SuStaIn subtypes, while RQ2 demonstrates these subtypes correlate with Boolean-predicted pathway dysregulation. Together, these validate the mechanistic hypothesis: GBA mutations drive FOXO1 dysregulation, manifesting as accelerated SuStaIn progression.

Phase 2 (RQ3–RQ4): Clinical Translation. RQ3 demonstrates practical utility through improved LID prediction when integrating staging and pathway data. RQ4 enables early risk detection via baseline molecular signatures predicting future stage transitions. These establish precision medicine capability - identifying at-risk patients prospectively using baseline biomarkers.

Phase 3 (RQ5): Clinical Credibility. RQ5 validates cross-cohort generalizability, confirming subtypes represent universal disease biology. This enables widespread clinical adoption and future precision trial applications.

Synergy with LIH Research. Dr. Hemedan's Boolean models explain why patients progress (pathway dysregulation mechanisms). My SuStaIn framework characterizes how patients progress (clinical stage transitions). Integration creates a unified framework linking molecular mechanisms to clinical phenotypes to outcome prediction - bridging systems biology discovery with clinical application.

FEASIBILITY AND TIMELINE

The proposed research is feasible within a 12-week internship period, with clearly defined milestones and deliverables.

Week	Research Focus	Primary Deliverable
1–2	Data access and harmonization	Integrated LuxPark dataset with Boolean scores
3–4	RQ1: Genetic stratification	Statistical report on genotype-subtype associations
5–6	RQ2: Pathway-stage correlation	Correlation matrices and validation statistics
7–8	RQ3: LID prediction refinement	Enhanced prediction model with comparative AUC analysis
9–10	RQ4: Transition prediction	Survival analysis results and miRNA signatures
11–12	RQ5: Cross-cohort validation	External validation report and manuscript draft

This timeline prioritizes methodological rigor and reproducibility over exhaustive publication preparation, while generating sufficient progress to support postdoctoral applications and future peer-reviewed manuscripts.

SIGNIFICANCE STATEMENT

For LIH: This work validates molecular systems biology models (Boolean networks) against longitudinal clinical staging (SuStaIn), bridging Dr. Hemedan's computational framework with Dr. Fritz's clinical innovation. It enables precision trial design, improves clinical prediction, and establishes LIH as a European leader in molecular-clinical integration for neurodegenerative disease.

For Parkinson's Medicine: Demonstrates that precision patient stratification improves clinical outcome prediction, enabling biomarker-driven trials and personalized intervention strategies. Identifies modifiable risk factors early in disease course when therapeutic intervention is most effective.

For Doctoral Research: Extends SuStaIn methodology from descriptive to predictive, incorporating genetic stratification and multi-omics integration beyond the current scope of my thesis. Establishes actionable clinical value through demonstrated improvement in adverse outcome prediction.

For Career Development: Positions me as a translational researcher bridging molecular systems biology and clinical data science - an ideal profile for competitive postdoctoral positions in precision medicine and computational neuroscience.

Prepared by:
Moad Hani, MSc
PhD Candidate, University of Mons (UMONS), Belgium
moad.hani@umons.ac.be
Expected Defense: September 2026