IMPACT OF PREOPERATIVE FRAILTY ON SIGNIFICANT POSTOPERATIVE DISABILITY IN OLDER PATIENTS FOLLOWING MAJOR SURGERY: A MULTICENTER PROSPECTIVE COHORT STUDY

Julian Daza (SSTP)[1], Peter Smith[2], Shabbir Alibhai[3], Erin Kennedy[1], Duminda Wijeysundera[4], FIT After Surgery Investigators

[1]Division of General Surgery, Department of Surgery, University of Toronto, Toronto, ON, Canada

[2]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[3]Department of Medicine, University of Toronto, Toronto, ON, Canada

[4]Department of Anesthesiology and Pain Medicine, University of Toronto, Toronto, ON, Canada

The authors have decided not to make the research results available at this time and will provide updates as soon as the results can be shared.

TURNING ASTROCYTES INTO NEURONS: DEVELOPING GENE THERAPY APPROACHES TO REPAIR THE STROKE INJURED BRAIN

Jack W. Hickmott [1], Gajeni Prabaharan [1], Tom Enbar [1], Ricky Siu [1], Varanan Vejeyathaas [1], Kriesha Eyer [1], & Cindi M. Morshead [1]

[1]Donnelly Centre for Cellular and Biomolecular Research, Division of Anatomy, Department of Surgery, University of Toronto, Toronto, ON

INTRODUCTION

Brain disorders are amongst the most serious health problems facing modern society. As the average human lifespan has increased, a concomitant increase in the incidence of death and disability due to brain disorders has negatively impacted life quality and is a major contributor to the Global Burden of Disease, with both social and economic implications[1,2]. Common to neurodegenerative disease and CNS injury, is the loss of neurons[3]. Since the adult brain does not replace lost neurons, one of the goals of therapeutic interventions to promote neuro-recovery is by generating new neurons. To this end, direct cellular reprogramming, specifically targeting glial cells to reprogram into neurons, is a particularly appealing technology to treat CNS injury and neurodegenerative disease as it addresses two issues: (1) neuronal loss and (2) cell toxicity. Many studies have targeted astrocytes as a cell source for reprogramming; their abundance, activation after injury and contribution to the pathology of neurodegenerative disease and stroke makes them attractive targets for neuronal conversion. Our goal is to use gene therapy approaches to directly convert astrocytes into neurons and enhance recovery of the stroke injured brain.

Gene therapy has gained increased traction with the advent of recombinant adeno-associated-viruses (AAVs)[4]: safe, non-integrating gene delivery vehicles that can be engineered to target specific cell types. We, and others, have demonstrated that AAV based reprogramming of astrocytes to neurons (AtN) with neurogenic transcription factors (TFs) improves recovery from stroke in an endothelin-1 mouse model of stroke in the primary motor cortex [5,6]. While promising, there is much to be learned regarding the underlying cellular mechanisms that lead to improved outcomes and understanding how best to optimize the gene therapy. For example, the diversity of glial and neuronal cells between brain regions may impact the efficacy of the approach. A number of considerations include: (1) Astrocyte heterogeneity: astrocytes are a heterogenous cell population and the efficacy of AtN conversion within the distinct pools has not been explored, (2) Regional heterogeneity: the degree to which AAV transduction varies across brain regions, and (3) Promoter promiscuity: the short GFAP promoter (shGFAP) which is commonly used to target TF expression in astrocytes, may not retain astrocyte specificity in an injured brain. Here we investigate regional and cellular heterogeneity in terms of AAV transduction and compare the fidelity of the shGFAP promoter versus the human *S100β* promoter in driving astrocyte versus neuron specific gene transcription towards the goal of enhancing AtN mediated neural repair.

METHODS

To trace transduced astrocytes over time, two approaches were taken: (1) a Cre-recombinase expressing vector was designed using the shGFAP promoter for use in a conditional reporter mouse and (2) EmGFP expressing vectors from either the shGFAP or the S100β promoters were designed for use in control mice. All vectors were cloned and packaged into AAV5 (VectorBuilder). AAV5 was stereotaxically injected into the primary motor cortex or medal prefrontal cortex of adult mice C57BL/6 or Ai14 (CAG-Lox-Stop-Lox-tdTomato) mice. Motor cortex stroke was induced by stereotaxic injection of ET-1 at RC +0.6 mm, ML -2.2 mm, DL -1.0 mm from bregma. Mice were harvested for histological analysis at various times post-AAV and/or stroke by cardiac puncture

and perfusion, followed by cryosectioning, immunofluorescent or RNAScope staining, and fluorescent microscopy.

Outcomes
Cell-type heterogeneity was assessed following AAV injection into the adult cortices and immunofluorescent antibody staining for neurons (NeuN) and astrocytes (S100β or GFAP). Brains were analyzed for the number of AAV transduced cells and *Gfap* expression by RNAscope. Bioinformatic analysis was performed on the DNA sequence of the shGFAP promoter to assess the presence of transcription factor binding sites.

Statistical analysis
Data was graphed and analysed using Prism (GraphPad Software, v10.1.2). Before performing parametric testing, assumptions of equal variance and gaussian distribution of residuals were verified using (F-test and Shapiro-Wilk, respectively). Data that met these assumptions of parametric testing were analysed using an unpaired t-test (for 2 group analysis) or using ANOVA followed by a Dunnett's Multiple Comparison Test (for >2 group analysis).

RESULTS
shGFAP promoter drives off target expression in medial prefrontal cortex neurons.

Brain injuries can occur at loci throughout the brain. Consequently, therapeutics treating brain injuries must be efficacious in different regions of the brain, despite considerable regional heterogeneity. Towards an AtN reprogramming based treatment we sought to test (1) AAV transduction from a ubiquitous promoter and (2) the specificity of the shGFAP promoter for astrocytes in different cortical regions. AAV5 encoding Cre-recombinase driven by shGFAP was injected into the medial prefrontal cortex or the motor cortex of Ai14 reporter mice. The expression of Cre in transduced cells results in permanent tdTomato (red) expression hence transduced cells can be identified and tracked over time. As predicted, the shGFAP promoter drove tdTomato expression in astrocytes at 21 days post AAV injection. Unexpectedly, we also observed tdTomato expression in resident cortical neurons, with significantly more tdTomato positive neurons in the medial prefrontal cortex than the motor cortex (52% and 11% of labelled cells, respectively; $p = 0.0105$). Consequently, use of the shGFAP promoter could result in off-target expression of reprogramming factors in cortical neurons, and this is more likely in the medial prefrontal cortex.

Gfap mRNA is endogenously transcribed in cortical neurons.

To better understand why the shGFAP promoter drives expression in cortical neurons, we explored the transcription patterns of the endogenous mouse *Gfap*. Using RNAScope, we assessed *Gfap* transcripts in $NeuN^+$ neurons of the medial prefrontal and motor cortices. Interestingly, *Gfap* mRNA was detected in neurons in a regionally distinct manner (51% vs 42% of neurons in the medial prefrontal and motor cortex, respectively, $p = 0.039$), revealing regional heterogeneity of *Gfap* RNA within neurons, which may account for off-target expression from the shGFAP promoter.

Neuronal Transcription Factor Binding Sites are present in the shGFAP promoter.

To help understand if the cortical neurons could drive expression from the shGFAP promoter, a bioinformatic analysis of the transcription factor binding sites present in shGFAP was conducted and cross-referenced to RNASeq data for the expression of transcription factors in neurons. Binding sites for 13 transcription factors modestly expressed in neurons were detected in the shGFAP promoter. The presence of these transcription factor binding sites is consistent with the

shGFAP promoter driven expression in neurons and could further contribute to the off target expression of the shGFAP promoter in cortical neurons.

Both the shGFAP and S100β promoters drive astrocyte-specific expression in stroke and non-stroke injured mouse brains.

Due to the expression driven by the shGFAP promoter in cortical neurons, and the need for astrocyte-specific expression for astrocyte to neuron reprogramming, we next explored an alternative astrocyte-specific promoter derived from the human *S100β* gene. Interestingly, using a direct EmGFP reporter system (*i.e.* expression is not Cre dependent) we observed high specificity for the GFAP$^+$ subpopulation of astrocytes with both promoters (55% vs 56% of EmGFP labelled cells for the shGFAP vs *S100β* promoters, respectively; **Fig. 1**). Similarly, we also so a high specificity for the S100β$^+$ subpopulation of astrocytes with both promoters (48% vs 40% of EmGFP labelled cells for the shGFAP vs *S100β* promoters, respectively). Finally, we only observed rare off-target expression in NeuN+ neurons (2% vs 3% of the EmGFP labelled cells from the shGFAP and *S100β* promoters, respectively). Most interesting, in stroke injured mice, the S100β promoter drove EmGFP expression in significantly more S100β$^+$ astrocytes than the GFAP promoter (36% vs 14%, respectively, $p = 0.0296$), while driving similar rare expression in NeuN$^+$ neurons (2.6% vs 2.99% for the shGFAP and S100β promoters, respectively). Further studies are needed to determine if this specificity translates to more sensitive reporter systems, such as those using Cre-recombinase, is retained in different brain regions, and if the S100β promoter provides sufficient specificity for AtN reprogramming.

CONCLUSIONS

Gene therapy for brain injuries such as stroke depend on specific targeting of cellular subtypes for their safety and efficacy. These findings reveal how the heterogeneity of the brain presents an important challenge that needs to be overcome as new brain-targeted gene delivery strategies are developed. We identified considerable heterogeneity even within the cortical regions of the brain, which may drive region specific expression patterns from genetic tools such as the shGFAP promoter. Importantly for therapeutic development, our studies also highlight how brain injuries such as stroke can influence the expression patterns driven by different promoters, as observed here with the performance of the shGFAP promoter changing in the context of stroke compared to the S100β promoter. Furthermore, our study also suggests astrocyte specificity may vary based on the gene cargo, and that the more sensitive Cre-Lox based reporter system may show higher 'off-target' effects compared to direct reporter systems. Together, these findings highlight neural cell heterogeneity as an important consideration when developing AtN reprogramming strategies for neural repair.

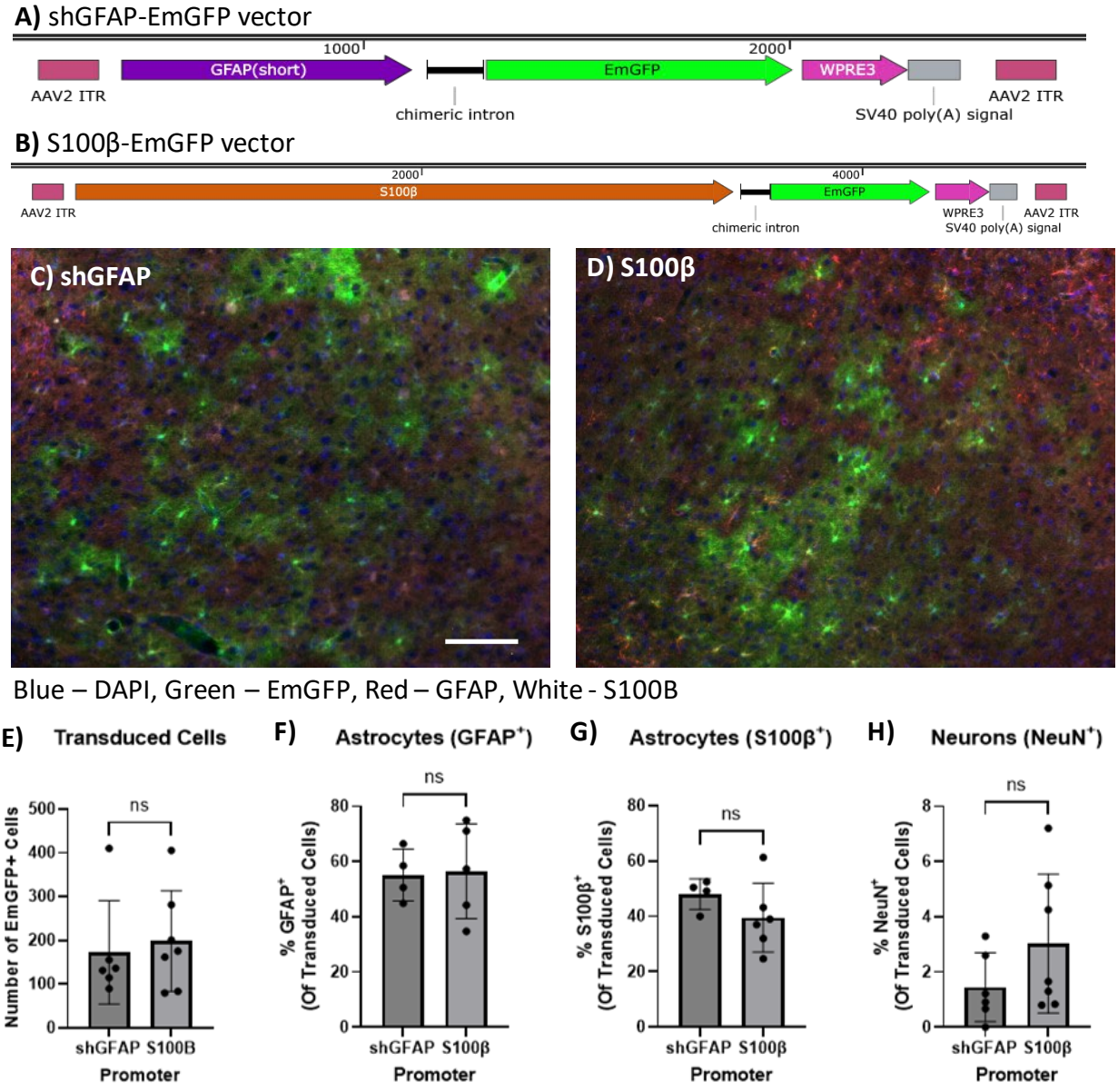


Figure 1. shGFAP and S100β promoters drive similar EmGFP expression patterns in non-stroke injured mouse motor cortices. A) & B) Vector schematics of the direct EmGFP reporter system for the shGFAP and S100β promoters. C& D) shGFAP-EmGFP and S100β-EmGFP expression (Green) in the mouse motor cortex 21 days following AAV5 injection. Sections were immunofluorescently stained with DAPI (blue), GFAP (red), and S100B (white). Scale bare = 100 µm. Both promoters drove expression in **E)** similar numbers of cells, **F)** GFAP$^+$ astrocytes, **G)** S100β$^+$ astrocytes, and **H)** NeuN$^+$ Neurons.

REFERENCES

1. Feigin VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, et al. World Stroke Organization (WSO): Global Stroke Fact Sheet 2022. Int J Stroke. 2022 Jan 1;17(1):18–29.

2. Feigin VL, Vos T, Nichols E, Owolabi MO, Carroll WM, Dichgans M, et al. The global burden of neurological disorders: translating evidence into policy. Lancet Neurol. 2020 Mar;19(3):255–65.

3. Campbell BCV, De Silva DA, Macleod MR, Coutts SB, Schwamm LH, Davis SM, et al. Ischaemic stroke. Nat Rev Dis Primer. 2019 Oct 10;5(1):1–22.

4. Pupo A, Fernández A, Low SH, François A, Suárez-Amarán L, Samulski RJ. AAV vectors: The Rubik's cube of human gene therapy. Mol Ther. 2022 Dec 7;30(12):3515–41.

5. Livingston J, Lee T, Daniele E, Phillips C, Krassikova A, Enbar T, et al. Direct reprogramming of astrocytes to neurons leads to functional recovery after stroke [Internet]. bioRxiv; 2020 [cited 2022 Nov 14]. p. 2020.02.02.929091. Available from: https://www.biorxiv.org/content/10.1101/2020.02.02.929091v1

6. Chen YC, Ma NX, Pei ZF, Wu Z, Do-Monte FH, Keefe S, et al. A NeuroD1 AAV-Based Gene Therapy for Functional Brain Repair after Ischemic Injury through In Vivo Astrocyte-to-Neuron Conversion. Mol Ther. 2020 Jan 8;28(1):217–34.

COST-UTILITY ANALYSIS OF TRAPEZIECTOMY AND LIGAMENT RECONSTRUCTION TENDON INTERPOSITION VS. SUTURE SUSPENSION ARTHROPLASTY FOR THUMB CARPOMETACARPAL JOINT OSTEOARTHRITIS

Chloe R. Wong (SSTP), Alice Zhu[2], Helene Retrouvey[3], Heather L. Baltzer[1], Christopher Witiw[4]

[1]University of Toronto, Department of Surgery, Division of Plastic, Reconstructive & Aesthetic Surgery, 149 College Street, Toronto, ON, Canada, M5T 1P5

[2]University of Toronto, Department of Surgery, Division of General Surgery, 149 College Street, Toronto, ON, Canada, M5T 1P5

[3]McMaster University, Department of Surgery, Division of Plastic, Reconstructive & Aesthetic Surgery, 50 Charlton Ave E, Hamilton, ON, Canada, L8N 4A6

[4]University of Toronto, Department of Surgery, Division of Neurosurgery, 149 College Street, Toronto, ON, Canada, M5T 1P5

The authors have decided not to make the research results available at this time and will provide updates as soon as the results can be shared.

GENOME EDITING TO IMMUNOLOGICALLY MODIFY DONOR LUNGS FOR TRANSPLANTATION USING CRISPR-CAS TECHNOLOGIES

Kumi Mesaki[1], Haruchika Yamamoto[1], Stephen Juvet[1,2], Jonathan Yeung[1,3], Zehong Guan[1], Akhi Akhter[1], Cameron Dickie[1], Henna Mangat[1], Aizhou Wang[1], Gavin W. Wilson[1,4], Andrea Mariscal[1], Jim Hu[5,6], Alan R. Davidson[7,8], Benjamin P. Kleinstiver[9,10,11], Marcelo Cypel[1,2], Mingyao Liu[1,4], Shaf Keshavjee[1,4]

[1]Latner Thoracic Research Laboratories, Toronto General Hospital Research Institute, University Health Network, Toronto, Ontario

[2]Division of Respirology, Department of Medicine, University of Toronto, Toronto, Ontario

[3]Division of Thoracic Surgery, Department of Surgery, University of Toronto, Toronto, Ontario

[4]Department of Surgery, University of Toronto, Toronto, Ontario

[5]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario

[6]Translation Medicine Program, the Hospital for Sick Children, Toronto, Ontario

[7]Department of Biochemistry, University of Toronto, Toronto, Ontario

[8]Department of Molecular Genetics, University of Toronto, Toronto, Ontario

[9]Center for Genomic Medicine, Massachusetts General Hospital, Boston, USA

[10]Department of Pathology, Massachusetts General Hospital, Boston, USA

[11]Department of Pathology, Harvard Medical School, Boston, USA

INTRODUCTION

Organ transplantation is a life-saving treatment rescuing over 140,000 patients with end-stage organ disease each year. In lung transplantation (LTx), life-long systemic immunosuppression, taken by recipients to prevent graft rejection, is one of the major impediments to post-transplant outcomes increasing the risk of complications, such as infection, malignancy, and adverse drug effects. We hypothesized that donor lungs with long-lasting immunomodulatory capacity could reduce the need for the systemic immunosuppression.

Genome editing[1] has the unique capability to induce sustained effects after one-time action of modifying the host genome, thereby holding the promise to engineer donor organs with beneficial traits for transplantation. We envisioned *genetically modifying donor lungs prior to implantation using our ex vivo lung perfusion (EVLP) platform to deliver immunologically optimized donor organs to recipients.* However, genome editing, particularly gain-of-function genetic modification, in the whole organ and its clinical translation has not been achieved - challenged by the limited modalities of editing, delivery and preclinical assessment models.

To realize our therapeutic vision, we first sought to develop a genome editing approach to upregulate an immunomodulatory gene IL-10 using a CRISPR-Cas construct.[2] We then developed optimized adenoviral vectors to achieve both early and sustained IL-10 induction by combining CRISPR genome editing and adenoviral vector mediated gene delivery. After achieving this in cell culture and in an in vivo rat model, we developed a method to evaluate genome editing preclinically in human donor lungs using a combination of the Toronto EVLP system and precision-cut lung slice (PCLS) technology. Collectively, we demonstrated the feasibility of donor organ modification using CRISPR-Cas genome editing and testing of such genome-editing interventions preclinically in human organs. Our findings address key challenges in the engineering of genetically modified organs towards the acceleration of safe and timely translation to overcome unmet clinical needs.

METHODS

Adenoviral vectors

DNA fragments containing a 2A self-cleaving peptide sequence, human IL-10 cDNA, a polyA sequence, and a guide RNA (gRNA) expression cassette were synthesized and cloned into the Cas9 expression plasmids. The E1/E3-deleted replication incompetent recombinant adenovirus serotype 5 was generated by transferring the whole transgene into a shuttle vector, generating primary virus, and amplification and purification.

Rat study

Rats were anesthetized by inhalation of isoflurane, orotracheally intubated and ventilated in the supine position. After guiding the intubation tube into the left bronchus, 500 µl of vector containing buffer was delivered to the left lung. The intubation tube was pulled back into the trachea and the rat was ventilated in the left lateral decubitus position until awakening. Standard clinical triple immunosuppression (cyclosporine, azathioprine, methylprednisolone) was injected intraperitoneally 2h before viral vector administration and on day 1 and day 2. Two hoursafter viral delivery, methylprednisolone alone was injected intraperitoneally. From day 3 to day 28, triple suppression was injected subcutaneously daily.

Ex vivo lung perfusion

Human donor lungs that were declined for clinical LTx with research consent, were used in this study. Donor lungs were retrieved using the same preservation protocol as in clinical transplantation in the Toronto Lung Transplant Program. The EVLP was performed according to the standard clinical Toronto EVLP protocol[3]. Briefly, cannulas are attached to pulmonary artery and left atrium. An endotracheal tube is inserted into the bronchus and secured with ties. Retrograde flushing was performed using low potassium dextran solution. The circuit was primed

with perfusate, supplemented with methylprednisolone, heparin and meropenem. During the initial 1h of EVLP, the flow was increased stepwise. The temperature was increased incrementally to 37 °C, and ventilation was initiated when the temperature reached 32 °C. A bronchial wash was performed by administering 20 ml of saline into the right or left main bronchus and aspirated using a bronchoscope. Perfusate and bronchial wash samples were snap-frozen and stored at -80 °C until analysis. Tissues were formalin fixed or snap frozen for downstream analysis.

Precision Cut Lung Slices (PCLS) of the human lung after EVLP

Upon completion of the EVLP, pieces of lung tissue were inflated with the DMEM/F-12 media containing 1.5% agarose, solidified on ice and sectioned at 500µm thickness using vibrating blade microtome and cultured in nutrient-enriched media. For the measurement of the secreted IL-10 protein, the slice was incubated with 1 ml of fresh media for 36h.

Statistical analysis

Data are displayed as mean ± SD and analyzed using the Prism software. An unpaired two-tailed Student's t-test or one-way ANOVA followed by Bonferroni's correction was used to analyze significance.

RESULTS

Genome editing of a regulatory element using Cas nucleases upregulated IL-10

We first explored genome editing approach to enhance the IL-10 gene, which encodes a key anti-inflammatory and immunomodulatory cytokine for graft immunomodulation.[4] Among various editing approaches, we sought to optimize CRISPR nuclease methods that can be achieved by simple delivery of the editing enzymes to the donor organ. We hypothesized that targeted mutagenesis of the IL-10 regulatory element could disrupt native repressor binding, leading to enhanced transcription through a *derepression mechanism*, as IL-10 expression is naturally suppressed in the native state.[5] We introduced mutations at different genetic loci in the regulatory element of the human IL-10 gene using Cas nuclease and assessed the impact on IL-10 expression in a human cell line (HEK 293). Targeted mutagenesis at regions 85- and 644-nucleotides upstream of the transcription start site significantly enhanced human IL-10 expression with an editing efficiency of 15.3% ± 0.64% and 10.3% ± 0.99%, respectively. This approach also induced the upregulation of endogenous IL-10 in rat and porcine cell lines, two species used for animal studies in the translational path.

A dual mode of IL-10 expression by combining Ad-hIL-10 delivery with genome editing to derepress native IL-10

Next, we sought a method that would allow for both early (for the innate ischemia reperfusion inflammatory response) and sustained IL-10 expression (for the later alloimmune acquired immune response) with a single intervention. Early ischemia-reperfusion-induced injury causes primary graft dysfunction, limiting short-term outcomes, and importantly is subsequently associated with chronic lung allograft dysfunction and rejection, which contribute to long-term mortality. Therefore, addressing both early and late graft injuries truly optimizes the therapeutic potential of graft immunomodulation. Although genome editing alone has unique advantages in inducing persistent effects, the achievement of editing and subsequent transcriptome alterations often take days to exert effect, exceeding the window available for pre-implantation donor organ modulation.

To achieve immediate intra-graft immunomodulation, we combined adenoviral vector mediated IL-10 gene delivery to achieve early IL-10 induction, with CRISPR genome editing for later and durable gene upregulation. [6] We developed an adenoviral vector (Ad-hIL10-CRISPR) carrying the human IL-10 cDNA, Cas9, and the gRNA, which targets the regulatory element of IL-10 in the genome, to induce dual mode of IL-10 expression (Fig. 1A). In vitro assessments of this combined

approach in a rat lung cell line (L2) indeed demonstrated early expression of human IL-10 expression by the cDNA delivery followed by genome editing mediated rat IL-10 upregulation as designed (Fig. 1B and 1C). The immediate human IL-10 expression from the exogenously delivered gene declined after peaking on day 3 (Fig. 1D) and the rat IL-10 upregulation induced by genome editing starting at ~24h was sustained for 28 days.

We further evaluated this approach in vivo in a rat model by transbronchially delivering the Ad-hIL10-CRISPR to the lung (Fig. 2A). Consistent with the findings in vitro, human IL-10 and Cas9 expression exhibited an early and transient expression (Fig. 2B). Human IL-10 protein levels in the lung tissue peaked between day 1 (124 ± 4.76 pg/mg protein) and day 3 (115 ± 4.56 pg/mg protein), followed by a decline afterward (Fig. 2C). Next-generation sequencing (NGS) of the genomic DNA in the lung tissue demonstrated a time-dependent increase in editing during the initial 5 days, reaching 2.52% ± 0.26% editing on day 5 ($p < 0.0001$ vs. the diluent control, Fig. 2D). The population of edited cells persisted for 28 days (3.84% ± 2.11%, $p = 0.0048$ on day 7, 1.78% ± 0.21%, $p < 0.0001$ on day 14, and 1.97% ± 0.46%, $p < 0.0001$ on day 28, Fig. 2D). In situ hybridization of lung tissue on day 14 revealed an enhanced signal of rat IL-10 mRNA in Cas9-expressing cells, supporting IL-10 upregulation in edited cells in the lung. Previous experiments have shown that this level of gene transduction is sufficient to achieve whole graft immunomodulation by IL-10 protein release working in a paracrine fashion. Importantly, we demonstrated that the combination of adenoviral gene delivery and genome editing using a single novel vector achieved both early and sustained IL-10 expression, which could be developed for pre-implantation donor organ engineering for graft immunomodulation.

CRISPR-based IL-10 modulation in ex-vivo perfused human donor lungs

We next sought to assess CRISPR genome editing in ex-vivo perfused human lungs, simulating pre-implantation donor organ modulation. We studied human lungs from two donors, that were declined for clinical transplantation. Using the Toronto EVLP system, we delivered Ad-hIL10-CRISPR to the lung through the airway (Fig. 3A). The lungs were then perfused at 37^0C for 12h (Fig. 3B) . In both cases, the treated lungs and showed an increase in IL-10 protein levels in the bronchial wash during the 12h post-intervention perfusion period (411.0 ± 285.4 pg/ml, change from baseline, Fig 3C) while the untreated lung showed no increase in IL-10 protein (18.4 pg/ml at 1 h vs. 17.6 pg/ml at 14 h). To investigate editing efficiency, we generated thin tissue slices using PCLS technologies at the end of EVLP and cultured them for 14 days. We demonstrated that the treated lungs exhibited higher production of hIL-10 protein on day 7, although IL-10 protein levels varied in different lung regions, likely due to heterogeneous vector distribution after non-aerosolized delivery (Fig. 3D). Analysis of the genomic DNA of the PCLS with high IL-10 production using NGS detected significant editing in the treated group on both days 7 and 14 compared with the untreated group (1.41% ± 0.35% on day 7, 1.00% ± 0.09% on day 14 vs. 0.43% ± 0.07% in the untreated group on day 14; $p = 0.020$ on day 14, Fig. 3E). These findings demonstrate that our combined approach can achieve IL-10 induction during 12h of EVLP and sustained editing in human donor lungs, underscoring the potential for application in clinical practice.

CONCLUSIONS

We developed a novel genome editing approach to enhance immunomodulatory gene expression in the donor lung prior and an optimized vector to achieve dual mode of IL-10 induction for early and sustained graft immunomodulation. Furthermore, we evaluated this genome editing intervention for the first time in human donor lungs ex vivo simulating organ modulation prior to implantation. Our findings lay the path towards a new frontier of organ immune-engineering and clinical translation to overcome challenges in patient care.

Figures

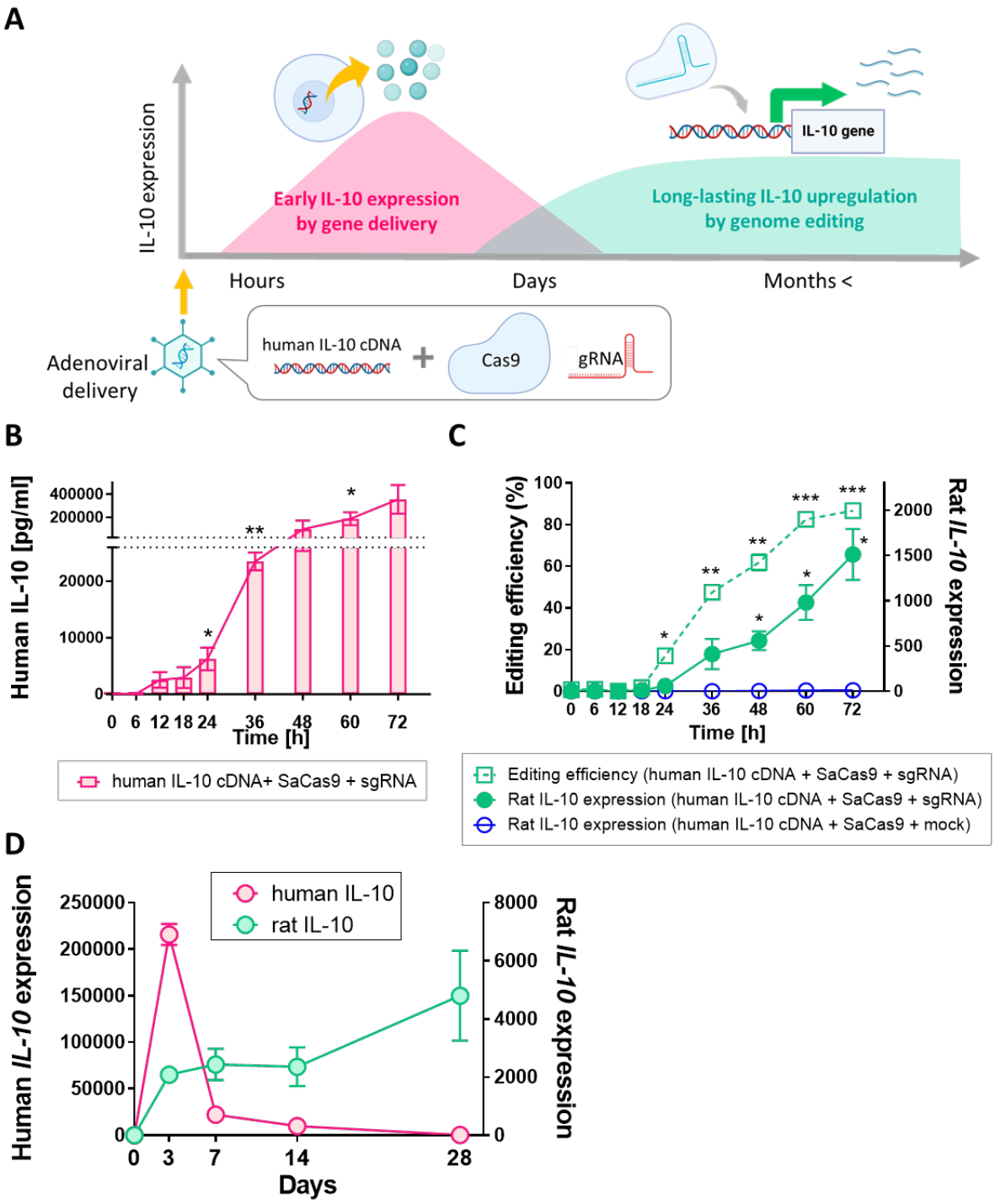


Figure 1. Temporal tuning of IL-10 by combining Ad IL-10 delivery with CRISPR genome editing. (A) Schematic representation of cellular events after the delivery of Ad-hIL10-CRISPR. (B, C and D) L2 cells were transduced with adenoviral vectors. The IL-10 protein levels in the media, mRNA levels in cell lysate, and genomic DNA sequence were measured by ELISA, qPCR, and Sanger sequencing, respectively. Significance compared to the levels at the start time (T0) was assessed. N=2 (B, C) or N=3 (D). Symbols indicate * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

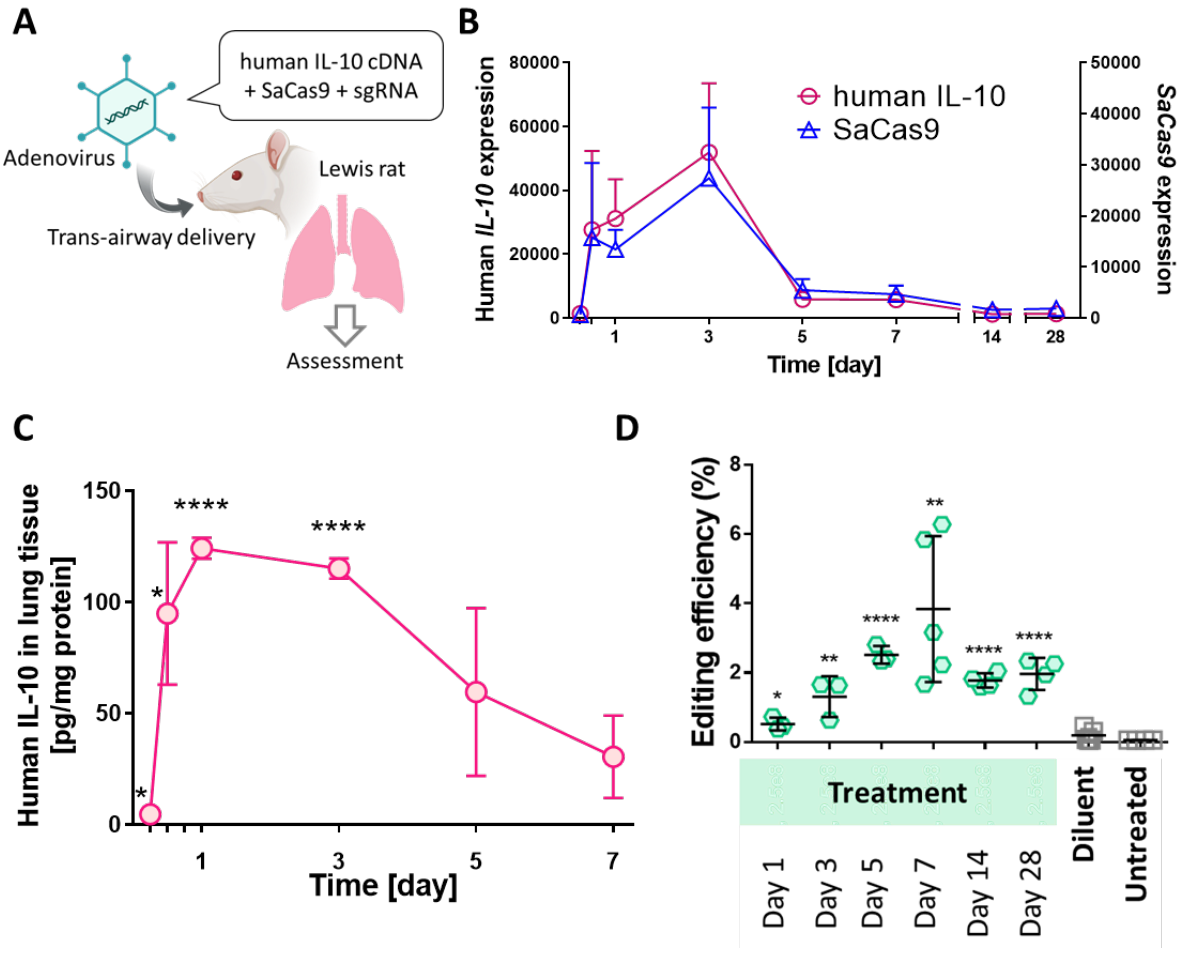


Figure 2. A dual mode of IL-10 expression after single dosing of a novel adenoviral vector in rat lungs in vivo. (A) Schematic representation of in vivo rat study. (B) The expression of human IL-10 and rat IL-10 in the lung tissues were measured by qPCR. Relative expressions to the levels in untreated rats are shown; N=3-5 for each time point. (C) Human IL-10 protein levels in the lung tissues were measured by ELISA. (D) Genome editing efficiency was analyzed by NGS. Significance was determined compared to the untreated cases. Symbols indicate * $p \leq 0.05$, ** $p \leq 0.01$, **** $p \leq 0.0001$.

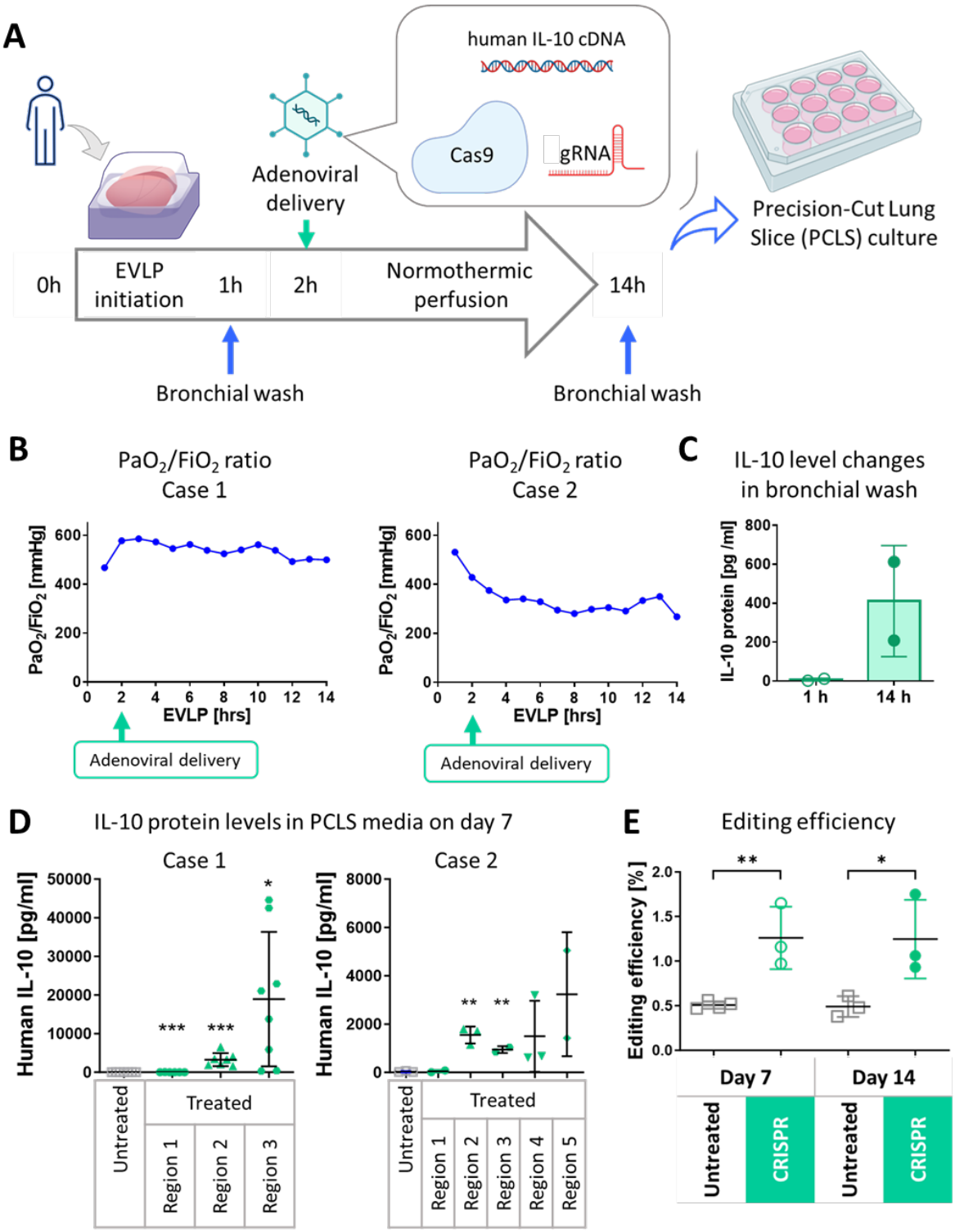


Figure 3. Evaluating genetic engineering in ex vivo perfused human donor lung. (A) Schematics of the declined human lung EVLP (ex vivo lung perfusion) study. (B) PaO_2/FiO_2 ratio during 14h of EVLP. (C, D and E) The levels of IL-10 in bronchial wash and PCLS media were measured using ELISA. Significance was determined compared to the untreated group. (F) The editing efficiencies in cultured PCLS were quantified by NGS. Symbols indicate * $p \leq 0.05$, **$p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

References

1. Doudna JA. The promise and challenge of therapeutic genome editing. *Nature*. Feb 2020;578(7794):229-236. doi:10.1038/s41586-020-1978-5

2. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. Aug 17 2012;337(6096):816-21. doi:10.1126/science.1225829

3. Cypel M, Keshavjee S. Ex Vivo Lung Perfusion. *Operative Techniques in Thoracic and Cardiovascular Surgery*. 2014/12/01/ 2014;19(4):433-442. doi:https://doi.org/10.1053/j.optechstcvs.2015.03.001

4. Machuca TN, Cypel M, Bonato R, et al. Safety and Efficacy of Ex Vivo Donor Lung Adenoviral IL-10 Gene Therapy in a Large Animal Lung Transplant Survival Model. *Hum Gene Ther*. Sep 2017;28(9):757-765. doi:10.1089/hum.2016.070

5. Saraiva M, O'Garra A. The regulation of IL-10 production by immune cells. *Nature Reviews Immunology*. 2010/03/01 2010;10(3):170-181. doi:10.1038/nri2711

6. Cypel M, Liu M, Rubacha M, et al. Functional repair of human donor lungs by IL-10 gene therapy. *Sci Transl Med*. Oct 28 2009;1(4):4ra9. doi:10.1126/scitranslmed.3000266

INTEGRATION OF CSF METHYLOME AND PROTEOME CAN OBVIATE THE NEED FOR SURGICAL BIOPSY IN CNS LYMPHOMA

Alex Landry (SSTP)[1,2], Jeff Zuccato[1,2], Vikas Patil[2], Mat Voisin[1,2], Justin Wang[1,2], Yosef Ellenbogen[1,2], Chloe Gui[1,2], Andrew Ajisebutu[2], Farshad Nassiri[1,2], Gelareh Zadeh[1,2]
[1]Division of Neurosurgery, Department of Surgery, University of Toronto, Toronto, Ontario
[2]Princess Margaret Cancer Center, University Health Network, Toronto, Ontario

INTRODUCTION

Liquid biopsy represents a major recent development in cancer research, with significant translational potential[1]. By sampling biofluids rather than tumour tissue itself, risks of surgery are mitigated and longitudinal sampling becomes feasible. This is of value for malignancies located in high-risk areas such as the central nervous system (CNS), wherein surgical risks are particularly high. While significant strides have been made in this field[2–5], improved model performance would profoundly reduce barriers to clinical implementation. In parallel, the increasing use of multi-platform (ie "multi-omic") approaches to cancer biology has been shown to help resolve heterogeneity and derive novel biologic insights[6]. Integrating multiple liquid biopsy platforms therefore holds promise as the next step in cancer diagnosis and surveillance. We hypothesize that the development of such multi-omic classifiers will improve the performance of existing models and allow for increased clinical utility.

We have created a composite score from cerebrospinal fluid (CSF) proteome and methylome to assess their synergistic power in common CNS malignancies. Using this novel approach, we have created the most accurate non-surgical CNS lymphoma classifier to date, able to completely discriminate this entity from glioblastoma (GBM) and brain metastases (BM) without invasive biopsy.

METHODS

Study cohort: Patients with histopathology confirmed GBM, BM, and PCNSL tumours with accompanying CSF were included in our study. CSF was sampled during placement of Ommaya reservoirs as part of their clinical care. All patients with GBM were confirmed to be IDH wild-type, WHO grade 4 tumours and all lymphoma cases were confirmed to be primary CNS lymphoma. This study received ethics approval from University Health Networks Institutional Review Board.

Data processing: Cell-free DNA (cfDNA) from each sample was extracted and processed according to our cfMeDIP-seq protocol as previously described[2–4]. Output sequencing data were aligned with the human genome and reduced to 300bp genomic windows which map to regulatory features using the *MEDIPS* package. Log 2-transformed counts per million (CPM) were used for downstream analysis. Proteomic sample processing was performed within 4 hours of CSF acquisition, the steps of the which are described in detail elsewhere[5]. Proteins were filtered to those detected by a minimum of two peptides and missing LFQ values were imputed with normalized iBAQ intensities. Output intensities were log2-transformed for downstream analysis.

Classifier construction and training: We use the *caret*[7] package in *R* (version 4.3.1)[8] for all model construction, training, and testing. For each configuration, the dataset was randomly split into training and testing cohorts in a ratio of 80% to 20%, respectively. All models were repeated

100 times to ensure robust summary statistics could be generated. Classifiers were binarized to each tumour, such that the outputs were of the form “tumour of interest” vs “other”.

To construct individual methylation and protein classifiers, binomial elastic net models were constructed and trained using 10-fold cross validation repeated 3 times. Input features were selected on the training cohort using pairwise moderated t statistics with up to 200 unique input features (differentially methylated regions or expressed proteins) per data type per model. Model performance was tested on the test set to generate an area under the receiver operating characteristic curve (AUROC) for each iteration. Median AUROCs and 95% confidence intervals were used to summarize model performance.

We subsequently constructed integrated models using three complimentary approaches (Figure 1A). In model 1 (*early integration with global tuning*), feature selection was performed separately on both methylation and protein data, and subsequently combined into an elastic net classifier with a single regularization parameter. In model 2 (*early integration with individualized tuning*), data were combined in the same way, but a data-type-specific regularization parameters were trained, this has previously been suggested to improve performance of similar models[9]. Finally, in model 3 (*late integration*), separate elastic net classifiers were trained on each data type and their posterior probabilities used to train a second layer integrated classifier.

Evaluating lymphoma model as a diagnostic tool: Given the clinical utility of identifying PCNSL without surgical biopsy, we further explored the “lymphoma vs other” classifier as a diagnostic tool in greater detail. For each patient, median class probabilities were extracted from each model iteration in which they were randomly allocated to the test set. Probabilities were binarized into predicted membership using thresholds of 0.3, 0.4, and 0.5 given the ‘1 vs 2’ ensemble construction of our classifiers. Median accuracy, sensitivity, and specificity were computed for all integrated classifiers as well as a consensus of all three, such that only patients with predictions consistent across all three integrated models were included and the rest considered to be unclassified. In similar fashion, we computed the same statistics using each tumour model together (i.e. “lymphoma vs other”, “GBM vs other”, and “BM vs other”), such that patients classified negatively in all models, or positively in more than one model, were deemed unclassified.

RESULTS

Our cohort included 20 patients with GBM, 17 with brain metastases, and 14 with PCNSL. All patients with GBM had IDH wild type, WHO grade 4 tumours, and patients with brain metastases had various primary malignancies breast (n=10), lung (n=4), esophageal (n=1), ovarian (n=1) and unknown (n=1). All patients with lymphoma were confirmed to be a result of PCNSL.

Overall, we found that the performance of parallel classifier paradigms was similar within each tumor type (Figure 1B). Remarkably, all versions of the integrated “lymphoma vs other” models demonstrated perfect performance with consistent median AUCs (and 95% CI) of 1 (1-1). By comparison, the median AUCs (95% CI) in the “GBM vs other” models were 0.88 (0.83-0.90), 0.85 (0.83-0.90), and 0.89 (0.85-0.90) for *early integration with global tuning*, *early integration with individualized tuning*, and *late integration* models, respectively. In the “BM vs other” models, median AUCs were 0.94 (0.89-0.94), 0.90 (0.89-0.94), and 0.94 (0.90-0.95). Comparing the median AUCs of these integrated models to single-platform models using pairwise, we found modest improvements in “GBM vs other” models, with higher median AUC in the *late*

integration model compared to both single-platform models (delta AUC = 0.06 and 0.09, p = 0.008 and 0.0001 for methylation and protein, respectively). Similarly, “BM vs other” models demonstrated improvement between the protein-only and *late integration* models (dAUC = 0.06, p = 0.0001). By contrast, there was marked improvement in performance when comparing all integrated classifiers to all single-platform classifiers across the “lymphoma vs other” models ($p<0.05$, Figure 1C, Supplementary Figure 2). These results were similar when comparing dAUCs computed at each iteration of the respective models (Figure 1D). Therefore, we show that combining methylation and protein data, regardless of method used, significantly increases the discriminative capacity of the constituent individual models. Importantly, this improvement in performance was far greater in PCNSL compared to GBM or BM, suggesting the epigenome and proteome are particularly synergistic in the biology of this disease.

Given the striking performance of the integrated “lymphoma vs other” models, we further explored their performance as a stand-alone diagnostic tool (Figure 2). To do this, we examined each integrated “lymphoma vs other” model on its own and in combination by consensus, wherein patients were only classified if all three models agreed in their outputs. With this latter approach, 86% of patients were classified with an accuracy of 95%, specificity of 100% and sensitivity of 80% using a threshold of 0.3 (Figure 2A,C). We subsequently applied a more restrictive approach wherein patients were “successfully” classified if and only if a single tumour model (“GBM vs other”, “BM vs other”, or “lymphoma vs other”) yielded a positive output. Therefore, if a patient was classified negatively in all three models (indicating an absence of all three tumour subtypes) or was classified positively in more than one model (indicating an invalid prediction of multiple tumour types simultaneously), they were considered unclassified. In this paradigm, the *early integration with global tuning* model classified 73% of all patients with a sensitivity of 88%, specificity of 100%, and accuracy of 97%; similarly, the late integration model classified 75% of patients with a sensitivity of 89%, specificity of 100%, and accuracy of 97% (Figure 2B).

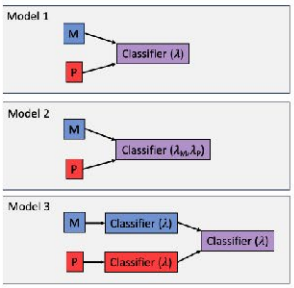
CONCLUSIONS

In this study, we show that combining the CSF methylome and proteome significantly improves classifier performance in CNS lymphoma. Importantly, patients with PCNSL currently undergo surgery purely for diagnosis and do not derive any therapeutic benefit from cytoreduction unlike many other CNS malignancies[10]. This is associated with an estimated mortality of 2.8%[11], healthcare costs in the vicinity $10,000[12], significant resource utilization, and perioperative delays to systemic therapy. While the use of genomic profiling for CNS tumours is not yet standard of care, it’s use in clinical management has steadily gained traction. Further, it is becoming increasingly understood that many common neurosurgical diseases are best represented by their multi-omic molecular fingerprints rather than more traditional classification paradigms[13]. As the clinical utility of these molecular techniques continues to increase, it seems inevitable that they will soon become commonplace in clinical neuro-oncology. Additionally, any economic barriers to implementing these genomic technologies are far outweighed by the cost savings they would generate in the case of CNS lymphoma.

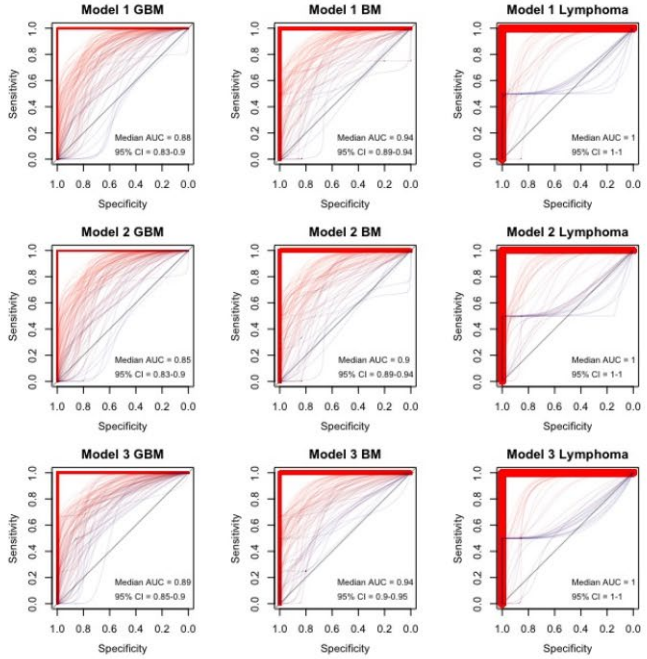
Overall, the development of multi-platform liquid biopsy classifiers represents a promising new avenue for cancer diagnosis and surveillance. Using this technique, we present the most specific and accurate CNS lymphoma classifier to date, filling a significant gap in the care of these patients. This has the potential to change current standard of care practice.

FIGURES

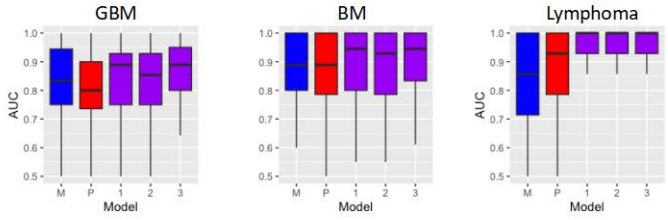
A



B



C



D

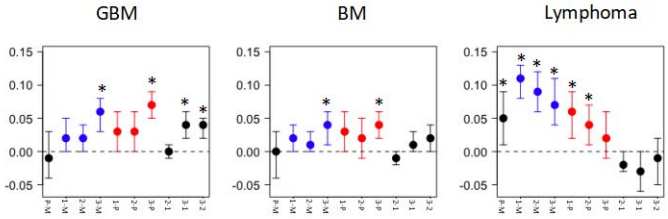


Figure 1: *Classifier performance across tumour types. A: Overview of integrated model design. Model 1 = early integration with global tuning, Model 2 = early integration with individualized tuning, Model 3 = late integration. M = methylation, P = protein. B: Receiver operating characteristic curves for each integrated model and tumour. Thickness of the AUC = 1 curve is directly proportional to the number of overlapping curves. C: AUCs by model. In the lymphoma model, each integrated model outperforms both individual classifiers and protein alone outperforms methylation alone (p <0.05). D: Mean (95% CI) of delta AUCs across all model iterations, pairwise model comparisons. Blue indicates comparison between methylation-only models and integrated models and red indicates comparison between protein-only models and integrated models. X-axis label defines comparison (i.e. P – M indicates that M is subtracted from P). * p<0.05.*

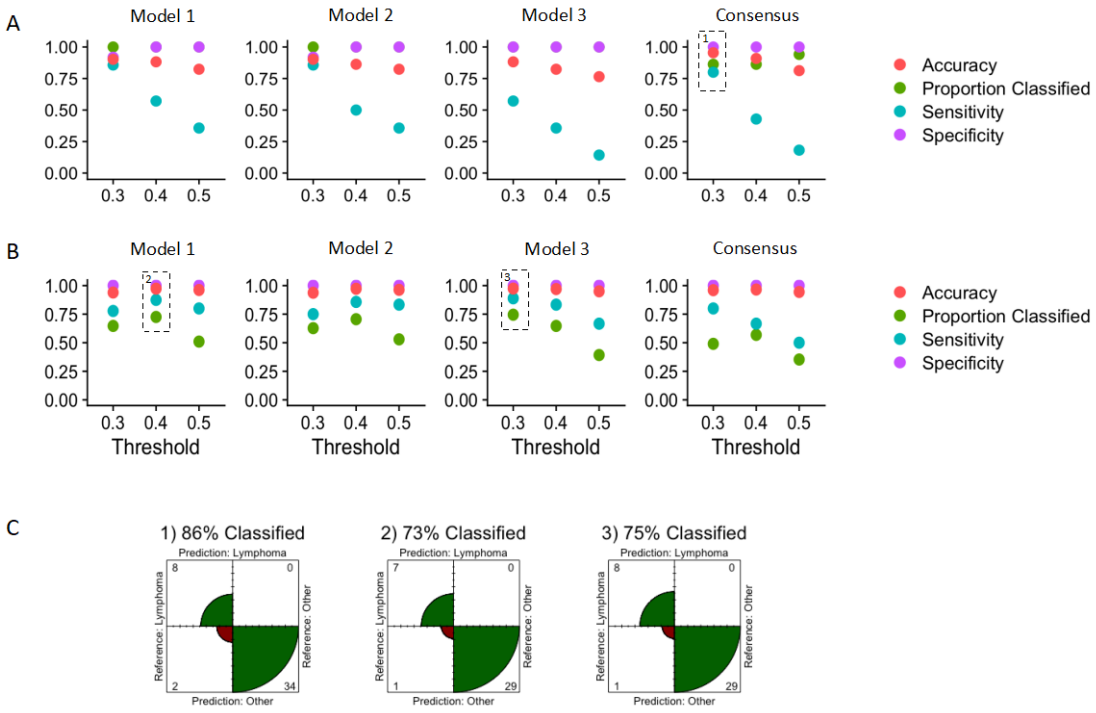


Figure 2: *Interrogation of the lymphoma classifier. A: Performance metrics (proportion of patients classified, accuracy, sensitivity, specificity) for each integrated "lymphoma vs other" model at various thresholds. In the consensus plot, patients are excluded if there is no consensus across all 3 models. B: Performance metrics of "lymphoma vs other" when layering "GBM vs other" and "BM vs other" models. Only patients with consistent outputs (1 positive and 2 negative outputs) are included in the performance metrics. Top performing configurations are boxed and labeled numerically. C: Confusion matrices of highest performing classifier configurations as labeled in A) and B). The area of the quarter circle is proportional to the number of patients in each quadrant.*

References

1. Lone SN, Nisar S, Masoodi T, et al. Liquid biopsy: a step closer to transform diagnosis, prognosis and future of cancer treatments. *Mol Cancer*. 2022;21(1):79. doi:10.1186/s12943-022-01543-7

2. Shen SY, Singhania R, Fehringer G, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature*. 2018;563(7732):579-583. doi:10.1038/s41586-018-0703-0

3. Nassiri F, Chakravarthy A, Feng S, et al. Detection and discrimination of intracranial tumors using plasma cell-free DNA methylomes. *Nat Med*. 2020;26(7):1044-1047. doi:10.1038/s41591-020-0932-2

4. Zuccato JA, Patil V, Mansouri S, et al. Cerebrospinal fluid methylome-based liquid biopsies for accurate malignant brain neoplasm classification. *Neuro Oncol*. 2023;25(8):1452-1460. doi:10.1093/neuonc/noac264

5. Mikolajewicz N, Khan S, Trifoi M, et al. Leveraging the CSF proteome toward minimally-invasive diagnostics surveillance of brain malignancies. *Neurooncol Adv*. 2022;4(1). doi:10.1093/noajnl/vdac161

6. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol*. 2017;18(1):83. doi:10.1186/s13059-017-1215-1

7. Kuhn M. Building Predictive Models in R Using the caret Package. *J Stat Softw*. 2008;28(5). doi:10.18637/jss.v028.i05

8. R Development Core Team. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing Vienna Austria*. 2016;0:{ISBN} 3-900051-07-0. doi:10.1038/sj.hdy.6800737

9. Liu J, Liang G, Siegmund KD, Lewinger JP. Data integration by multi-tuning parameter elastic net regression. *BMC Bioinformatics*. 2018;19(1):369. doi:10.1186/s12859-018-2401-1

10. Scheichel F, Pinggera D, Popadic B, Sherif C, Marhold F, Freyschlag CF. An Update on Neurosurgical Management of Primary CNS Lymphoma in Immunocompetent Patients. *Front Oncol*. 2022;12. doi:10.3389/fonc.2022.884724

11. Malone H, Yang J, Hershman DL, Wright JD, Bruce JN, Neugut AI. Complications Following Stereotactic Needle Biopsy of Intracranial Tumors. *World Neurosurg*. 2015;84(4):1084-1089. doi:10.1016/j.wneu.2015.05.025

12. Nassiri F, Li L, Badhiwala JH, et al. Hospital costs associated with inpatient versus outpatient awake craniotomy for resection of brain tumors. *Journal of Clinical Neuroscience*. 2019;59:162-166. doi:10.1016/j.jocn.2018.10.110

13. Nassiri F, Liu J, Patil V, et al. A clinically applicable integrative molecular classification of meningiomas. *Nature*. 2021;597(7874):119-125. doi:10.1038/s41586-021-03850-3

LONG-TERM OUTCOMES OF PERCUTANEOUS CORONARY INTERVENTION VERSUS CORONARY ARTERY BYPASS GRAFTING IN WOMEN WITH SEVERE CORONARY ARTERY DISEASE

Kevin R. An[1,2,3,4], Dominique Vervoort[1,4], Feng Qiu[4], Derrick Y. Tam[5], Rodolfo V. Rocha[1], Lamia Harik[2], Sameer Hirji[6], Mario F.L. Gaudino[2], Harindra C. Wijeysundera[4,7,8], Stephen E. Fremes[1,8]

[1]Division of Cardiac Surgery, Department of Surgery, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada

[2]Department of Cardiothoracic Surgery, Weill Cornell Medicine, New York, NY, USA

[3]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA

[4]Institute for Clinical Evaluative Sciences, Toronto, ON, Canada

[5]Department of Cardiac Surgery, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA.

[6]Division of Cardiac Surgery, Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, Harvard University, Boston, MA, USA

[7]Division of Cardiology, Schulich Heart Program, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada

[8]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

INTRODUCTION

Coronary artery disease (CAD) is the leading cause of death in women worldwide.[1] The treatment of severe CAD remains controversial, with several large randomized clinical trials (RCTs) of percutaneous coronary intervention (PCI) compared to coronary artery bypass grafting (CABG) demonstrating increased myocardial infarction (MI) and repeat revascularization with PCI, while evidence regarding mortality is less clear.[2] Sub-group analyses of women in RCTs have shown conflicting results due to lack of sufficient power, although, a meta-analysis of these RCTs suggested an increased composite outcome (all-cause death, MI, and/or stroke) in women undergoing PCI (hazard ratio (HR) 1.31 [95% confidence interval (CI), 1.05-1.63]).[3] Contemporary, large, observational studies comparing real world long-term outcomes after PCI and CABG in women are limited.[4] Women differ from men in the development of CAD in several ways. They develop CAD at an older age, they have smaller, more vasoreactive coronary arteries and grafts, less obstructive disease, and more microvascular disease.[5] Given the established sex-specific differences in the development and treatment of CAD, we examined and compared the real-world long-term outcomes of PCI and CABG in women with severe CAD.

METHODS

Study Design

We conducted a population level retrospective cohort study in Ontario, Canada, of women ≥18 and <80 years of age diagnosed with severe CAD who underwent an elective, first-time, isolated, coronary revascularization procedure between April 1st, 2012 and December 31st, 2021. Severe coronary artery disease was defined as left main disease, two vessel coronary artery disease involving the proximal left anterior descending artery, or three vessel coronary artery disease.[6] Clinical registry databases from CorHealth Ontario and the population-level administrative health databases housed at the Institute for Clinical Evaluative Sciences were used for the study. Coronary angiography records were obtained from CorHealth Ontario. Data accuracy has been validated by selective chart audits.[7] Baseline demographic variables were obtained from the CorHealth Registry, or the Canadian Institute for Health Information-Discharge Abstract Database (CIHI-DAD), using previously validated algorithms, where applicable.[8] In-hospital deaths were ascertained through the CIHI-DAD and the Registered Persons Database (RPD); and out of hospital deaths were ascertained through the RPD. Rehospitalizations were determined from the CIHI-DAD based on the main readmission diagnosis, as per ICD-10 definitions. Physician services were determined through the Ontario Health Insurance Plan billing database. Statistics Canada's census data was used to determine income quintile, based on median neighborhood income of individuals.

Outcomes

The primary outcome was a composite of all-cause mortality, myocardial infarction, stroke, and repeat revascularization (MACCE). Secondary outcomes included the individual components of the composite outcome and a composite readmission outcome (myocardial infarction, heart failure, and stroke). Tertiary outcomes included in-hospital outcomes (all-cause mortality, myocardial infarction, stroke).

Statistical analysis

Baseline characteristics were compared using the χ^2 for categorical variables and the Student's t-test for normally distributed continuous variables. Missing data for variables with <10% missing data were managed using multiple imputation, carried out 10 times with 10 iterations per imputation. Variables with >10% missing data were excluded from the propensity score model.[9]

Propensity scores were generated using a multivariable logistic regression model regressed on 33 baseline characteristics. Matching was performed 1:1 with greedy nearest-neighbor matching techniques and a caliper distance of 0.20 of the logit of the propensity score standard deviation. A standardized mean difference of less than 0.10 was used to determine success of matching for a given covariate.[10] Time-to-event survival analysis was conducted using Kaplan-Meier survival curves and log-rank testing stratified on the matched pairs to test the equality of the estimated survival curves. Cox proportional hazard models with robust variance estimators accounting for clustering within matched sets were used to regress survival status at latest follow-up on intervention status to estimate a hazard ratio for MACCE (all-cause mortality, myocardial infarction, stroke, and repeat revascularization) and all-cause mortality. Fine-Gray sub-distribution hazard models with robust variance estimators accounting for death as a competing risk and clustering within matched sets were fitted to the propensity score-matched sample and used to assess secondary outcomes (myocardial infarction, stroke, repeat revascularization, and the composite readmission outcome). Sensitivity analyses included patients with acute coronary syndrome, included patients over the age of 80, and excluded patients who received PCI after cardiac surgery consultation. A falsification end point analysis using pneumonia was also performed. All tests were 2-sided; p-values <0.05 were considered significant without adjustment for multiplicity. Effect estimates for Cox proportional hazards regression were reported using hazard ratios (HR) with a 95% confidence interval (95% CI). All analyses were carried out using R (Version 3.1.2; R Project for Statistical Computing, Vienna, Austria) or SAS (version 9.4; SAS Institute, Inc, Cary, NC, USA).

RESULTS

Primary Matched Analysis

In total, there were 7,499 patients in the PCI group and 7,845 patients in the CABG group. After excluding patients who were >80 years of age and patients who did not undergo elective revascularization procedures, 2,469 patients in the PCI group and 3,721 patients in the CABG group were included (**Table 1.**). Prior to matching, patients who underwent PCI were more likely to be frail, have heart failure, have had previous MI, and have end-stage renal disease on dialysis; however, PCI patients had fewer diseased coronary vessels. After matching on 33 baseline characteristics, 4,066 patients remained (2,033 PCI and 2,033 CABG). The patients were well-matched, with standardized mean differences <0.10 for all covariates. In-hospital death and in-hospital stroke were lower with PCI than with CABG (0.7% vs 1.5%, p=0.04) and (<0.2% vs 1.4%, p<0.001), respectively. There was no difference in perioperative MI between PCI and CABG (1.1% vs 0.6%, p=0.12) (**Table 2.**). In long-term follow-up, at a median (IQR) of 5.1 (2.9-7.5) years in the matched cohort, MACCE was higher with PCI compared with CABG (37.7% vs 23.3%, HR 1.79, [95% CI: 1.66-1.91], p<0.001) (**Figure 1A.**). All-cause mortality was higher with PCI compared with CABG (17.8% vs 13.4%, HR 1.33, [95% CI: 1.21-1.45], p<0.001) (**Figure 1B.**). Spontaneous MI and repeat revascularization were higher with PCI compared with CABG (8.7% vs 3.6%, HR 2.28, [95% CI: 2.16-2.42], p<0.001) and (21.9% vs 9.1%, HR 2.84, [95% CI: 2.72-2.97], p<0.001), respectively. However, stroke was lower with PCI compared with CABG (2.4% vs 3.0%, HR 0.85, [95% CI: 0.78-0.92], p<0.001). The composite readmission outcome (MI, heart failure, stroke) was higher with PCI compared with CABG (16.2% vs 11.2%, HR 1.42, [95% CI: 1.37-1.47], p<0.001) (**Figure 1C.**).

Sensitivity Analyses

In a sensitivity analysis including women undergoing inpatient or urgent revascularization, the results were consistent with the primary analysis; the occurrence of MACCE and all-cause mortality remained higher with PCI compared with CABG (HR 1.78, [95% CI: 1.66-1.91],

p<0.001) and (HR 1.33, [95% CI: 1.21-1.45], p<0.001), respectively. In a sensitivity analysis of women of all ages undergoing elective PCI or CABG, the results were similarly consistent with the primary analysis, with higher MACCE and all-cause mortality after PCI compared with CABG (HR 1.83, [95% CI: 1.67-2.01], p<0.001) and (HR 1.48, [95% CI: 1.31-1.67], p<0.001), respectively**.** When excluding patients with previous cardiac surgery consultation, the results remained consistent for MACCE (HR 1.83 [95% CI: 1.64-2.03], p<0.001) and death (HR 1.32, [95% CI: 1.14-1.53], p<0.001)**.** In a falsification end-point analysis, there was no significant difference in the incidence of pneumonia between PCI and CABG (HR 0.97, [95% CI: 0.84-1.13], p=0.73).

Crude Outcomes

In the unmatched cohort, both MACCE and all-cause mortality were higher in the PCI cohort compared to the CABG cohort (37.7% vs 24.5%, HR 1.83, [95% CI: 1.67-2.00], p<0.001) and (18.6% vs 14.4%, HR 1.42, [95% CI: 1.26-1.61], p<0.001), respectively.

CONCLUSIONS

In our analysis, we found that women under 80 years of age with stable severe CAD had significantly higher MACCE and long-term mortality after PCI compared to CABG. In addition, spontaneous MI, repeat revascularization, and readmission for either MI, heart failure, or stroke was higher in the PCI cohort compared to CABG, while stroke was lower after PCI compared to CABG. These results were consistent when including women who underwent inpatient or urgent revascularization, when including women over 80 years of age, and when excluding patients who received a cardiac surgery consultation prior to PCI.

Our findings are consistent with a retrospective analysis of the New York State database from 2012-2018 which evaluated 3,966 propensity score matched pairs of patients and found higher long-term mortality after PCI compared to CABG (HR 1.29, [95% CI 1.14-1.45]).[11] Additionally, our findings are consistent with a meta-analysis of six RCT sub-group analyses, where the composite of death, stroke, and MI was higher in PCI compared to CABG in women (HR 1.31 [95% CI 1.05;1.63]).[3]

Our results add to the evidence that women may derive long-term benefit from CABG compared to PCI despite the upfront higher risk of MACCE and mortality. Differences between women and men with regards to the biology of coronary disease may play a role. Women have a greater ratio of non-obstructive to obstructive CAD compared to men, leading to a greater overall plaque burden in women compared to men for a given amount of obstructive disease.[12] Non-obstructive plaques are not benign, as one study demonstrated, a majority of MIs occurred from non-obstructive plaques rather than obstructive plaques.[13] An analysis from the CONFIRM (Coronary CT Angiography Evaluation for Clinical Outcomes: An International Multicenter) registry showed that women with nonobstructive left main disease had a nearly 80% higher risk for MACCE compared with men with nonobstructive left main disease at 5-year follow-up.[14] As PCI primarily treats obstructive CAD whereas CABG treats both obstructive and non-obstructive CAD, it is plausible that the early operative risk of CABG in women is compensated by greater long-term reduction in MI and mortality in CABG compared to PCI.[15] Our findings provide evidence to suggest women with severe CAD may derive greater long-term benefit from CABG compared to PCI and suggest that the treatment recommendation should shift towards CABG in appropriately selected women.

Table 1. Baseline Characteristics before and after propensity score matching.

	Before Propensity Score Matching			**After Propensity Score Matching**		
	PCI (n=2469)	**CABG (n=3721)**	**SMD**	**PCI (n=2033)**	**CABG (n=2033)**	**SMD**
Age, yrs	66.7 (8.9)	66.5 (8.2)	0.0183	66.6 (8.9)	66.6 (8.1)	0.0030
Charlson (mean ±SD)	1.2 (1.6)	1.2 (1.4)	0.0113	1.2 (1.5)	1.2 (1.4)	0.0202
Frailty risk score	1.5 (3.5)	1.0 (2.4)	0.1514	1.3 (3.0)	1.2 (2.6)	0.0350
Creatinine level, mg/dl	90.6 (80.7)	83.6 (70.3)	0.0001	87.6 (73.3)	86.4 (79.1)	0.0147
Income quintile			0.0432			0.0197
1	607 (24.6)	847 (22.8)		497 (24.4)	472 (23.2)	
2	532 (21.5)	808 (21.7)		425 (20.9)	438 (21.5)	
3	498 (20.2)	763 (20.5)		410 (20.2)	415 (20.4)	
4	453 (18.3)	695 (18.7)		384 (18.9)	379 (18.6)	
5	379 (15.4)	608 (16.3)		317 (15.6)	329 (16.2)	
Rural status	354 (14.3)	544 (14.6)	0.0080	290 (14.3)	275 (13.5)	0.0210
Teaching hospital	1309 (53.0)	1827 (49.1)	0.0785	1054 (51.8)	1022 (50.3)	0.0315
Diabetes	1180 (47.8)	1949 (52.4)	0.0918	998 (49.1)	995 (48.9)	0.0030
Current smoker	413 (16.7)	646 (17.4)	0.0170	340 (16.7)	352 (17.3)	0.0158
Former smoker	519 (21.0)	818 (22.0)	0.0236	430 (21.2)	416 (20.5)	0.0169
History of MI	219 (8.9)	259 (7.0)	0.0671	170 (8.4)	156 (7.7)	0.0242
Hypertension	2008 (81.3)	3040 (81.7)	0.0095	1659 (81.6)	1651 (81.2)	0.0101
History of stroke	30 (1.2)	32 (0.9)	0.0324	19 (0.9)	21 (1.0)	0.0090
History of heart failure	443 (17.9)	442 (11.9)	0.1580	323 (15.9)	299 (14.7)	0.0308
Atrial arrhythmia	117 (4.7)	105 (2.8)	0.0902	77 (3.8)	73 (3.6)	0.0093
Cerebrovascular disease	63 (2.6)	63 (1.7)	0.0544	38 (1.9)	38 (1.9)	<0.0001
COPD	262 (10.6)	306 (8.2)	0.0775	194 (9.5)	188 (9.2)	0.0096
Hyperlipidemia	1068 (43.3)	1660 (44.6)	0.0274	884 (43.5)	909 (44.7)	0.0248
Peripheral vascular disease	42 (1.7)	58 (1.6)	0.0110	35 (1.7)	39 (1.9)	0.0152
History of cancer	92 (3.7)	95 (2.6)	0.0619	74 (3.6)	64 (3.1)	0.0260
Renal disease	94 (3.8)	77 (2.1)	0.0908	60 (3.0)	60 (3.0)	<0.0001
Dialysis	82 (3.3)	58 (1.6)	0.0983	48 (2.4)	44 (2.2)	0.0110
Depression	22 (0.9)	13 (0.3)	0.0576	13 (0.6)	11 (0.5)	0.0105
Liver disease	16 (0.6)	21 (0.6)	0.0104	15 (0.7)	14 (0.7)	0.0061
CCS Class						
0	554 (22.4)	691 (18.6)	0.0927	432 (21.2)	411 (20.2)	0.0071
1	340 (13.8)	519 (13.9)	0.0051	278 (13.7)	282 (13.9)	0.0014
2	834 (33.8)	1306 (35.1)	0.0279	695 (34.2)	703 (34.6)	0.0083
3	566 (22.9)	993 (26.7)	0.0895	498 (24.5)	500 (24.6)	0.0012

4	175 (7.1)	212 (5.7)	0.0542	130 (6.4)	137 (6.7)	0.0077
Number of diseased vessels	2.33 (0.70)	2.42 (0.82)	0.1294	2.37 (0.74)	2.39 (0.82)	0.0266
Disease by vessel						
Left main (≥50%)	367 (14.9)	1424 (38.3)	0.6578	360 (17.7)	420 (20.7)	0.0719
Proximal LAD (70-99%)	1607 (65.1)	2049 (55.1)	0.2102	1221 (60.1)	1184 (58.2)	0.0423
Proximal LAD occlusion (100%)	115 (4.7)	116 (3.1)	0.0731	85 (4.2)	84 (4.1)	0.0070
Mid LAD (70-99%)	1481 (60.0)	2404 (64.6)	0.0943	1297 (63.8)	1325 (65.2)	0.0141
Mid LAD occlusion (100%)	125 (5.1)	202 (5.4)	0.0167	108 (5.3)	123 (6.1)	0.0247
Left circumflex (70-99%)	1656 (67.1)	2841 (76.4)	0.1974	1445 (71.1)	1476 (72.6)	0.0209
Left circumflex occlusion (100%)	213 (8.6)	320 (8.6)	0.0010	176 (8.7)	187 (9.2)	0.0315
Right coronary artery (70-99%)	1780 (72.1)	2895 (77.8)	0.1272	1496 (73.6)	1509 (74.2)	0.0307
Right coronary artery occlusion (100%)	343 (13.9)	703 (18.9)	0.1446	306 (15.1)	316 (15.5)	0.0284

PCI – percutaneous coronary intervention; CABG – coronary artery bypass grafting; SMD – standardized mean difference, MI – myocardial infarction; COPD – chronic obstructive pulmonary disease; CCS – Canadian Cardiovascular Society; LAD – left anterior descending artery

Table 2. Early and late outcomes in propensity score matched cohort.

	Outcomes		
	PCI (n=2033)	**CABG (n=2033)**	**P-value**
In-hospital			
Peri-operative MI	22 (1.1)	12 (0.6)	0.121
Stroke	<5 (<0.2)	28 (1.4)	<0.001
Death	15 (0.7)	30 (1.5)	0.036
Long-term			
MACCE	767 (37.7)	473 (23.3)	<0.001
Spontaneous MI	177 (8.7)	73 (3.6)	<0.001
Stroke	48 (2.4)	61 (3.0)	0.244
Repeat revascularization	445 (21.9)	184 (9.1)	<0.001
Death	362 (17.8)	272 (13.4)	<0.001
Composite readmission (MI, heart failure, stroke)	330 (16.2)	228 (11.2)	<0.001

PCI – percutaneous coronary intervention; CABG – coronary artery bypass grafting; MI – myocardial infarction; MACCE – major adverse cardiovascular and cerebrovascular events

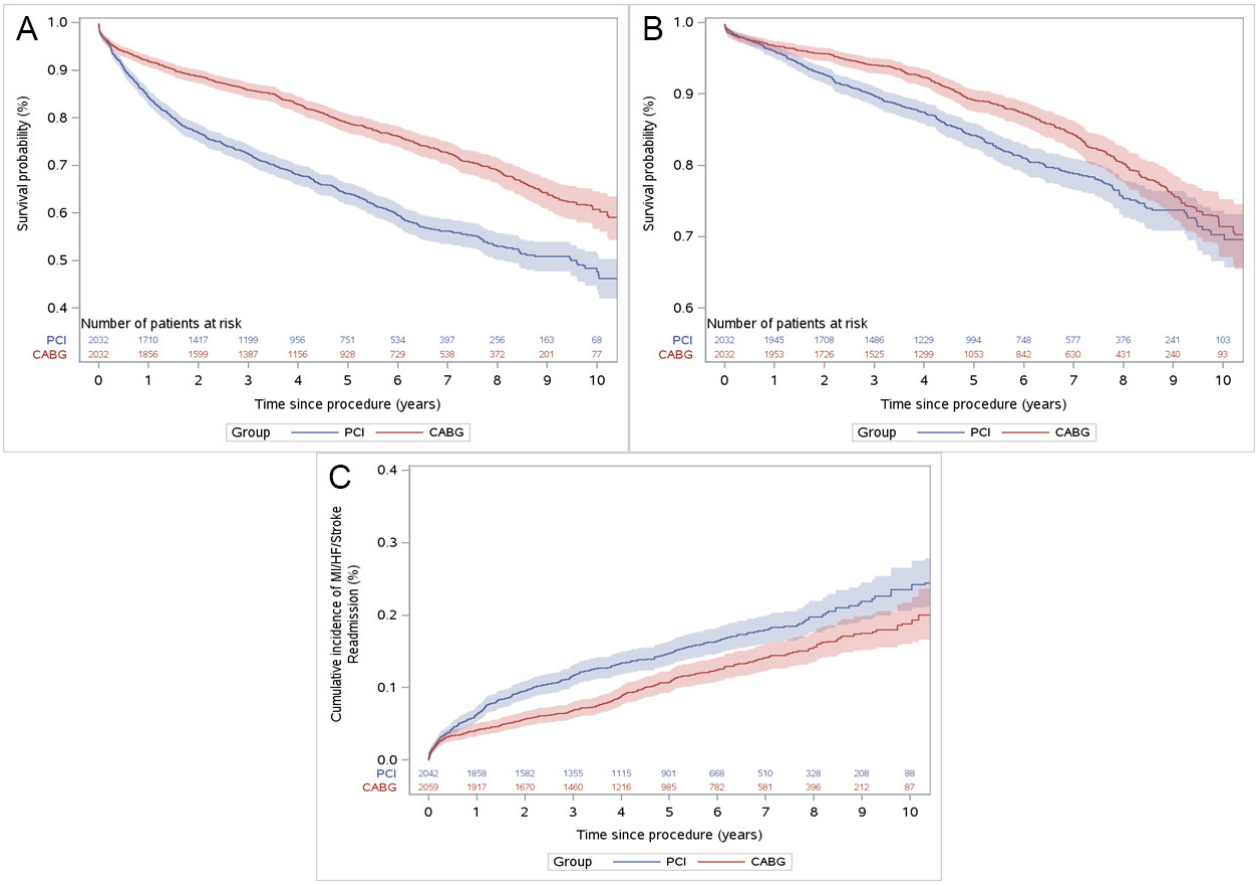


Figure 1A. Kaplan-Meier curve of freedom from major adverse cardiovascular and cerebrovascular events with PCI versus CABG in the propensity score matched cohorts. **1B.** Kaplan-Meier curve of all-cause mortality with PCI versus CABG in the propensity score matched cohorts. **1C.** Cumulative incidence function of readmission for myocardial infarction, stroke, or congestive heart failure with PCI versus CABG in the propensity score matched cohorts.

PCI – percutaneous coronary intervention; CABG – coronary artery bypass grafting; MI – myocardial infarction; HF – heart failure

REFERENCES

1. Vogel B, Acevedo M, Appelman Y, et al. The Lancet women and cardiovascular disease Commission: reducing the global burden by 2030. *The Lancet.* 2021;397:2385-2438.
2. Sabatine MS, Bergmark BA, Murphy SA, et al. Percutaneous coronary intervention with drug-eluting stents versus coronary artery bypass grafting in left main coronary artery disease: an individual patient data meta-analysis. *Lancet.* 2021;398:2247-2257.
3. Gul B, Shah T, Head Stuart J, et al. Revascularization Options for Females With Multivessel Coronary Artery Disease. *JACC: Cardiovascular Interventions.* 2020;13:1009-1010.
4. Moroni F, Beneduce A, Giustino G, et al. Sex Differences in Outcomes After Percutaneous Coronary Intervention or Coronary Artery Bypass Graft for Left Main Disease: From the DELTA Registries. *Journal of the American Heart Association.* 2022;11:e022320.
5. Harik L, Perezgrovas-Olaria R, Jr Soletti G, et al. Sex differences in coronary artery bypass graft surgery outcomes: a narrative review. *J Thorac Dis.* 2023;15:5041-5054.
6. Fihn SD, Gardin JM, Abrams J, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the Diagnosis and Management of Patients With Stable Ischemic Heart Disease. *Circulation.* 2012;126:e354-e471.
7. Wijeysundera HC, Stukel TA, Chong A, Natarajan MK, Alter DA. Impact of clinical urgency, physician supply and procedural capacity on regional variations in wait times for coronary angiography. *BMC Health Serv Res.* 2010;10:5.
8. Tam DY, Dharma C, Rocha R, et al. Long-Term Survival After Surgical or Percutaneous Revascularization in Patients With Diabetes and Multivessel Coronary Disease. *J Am Coll Cardiol.* 2020;76:1153-1164.
9. Lee JH, Huber JC, Jr. Evaluation of Multiple Imputation with Large Proportions of Missing Data: How Much Is Too Much? *Iran J Public Health.* 2021;50:1372-1380.
10. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med.* 2009;28:3083-3107.
11. Hannan EL, Wu Y, Harik L, et al. Coronary Artery Bypass Surgery vs. Percutaneous Interventions for Women with Multivessel Coronary Artery Disease. *J Thorac Cardiovasc Surg.* 2023.
12. Plank F, Beyer C, Friedrich G, Wildauer M, Feuchtner G. Sex differences in coronary artery plaque composition detected by coronary computed tomography: quantitative and qualitative analysis. *Neth Heart J.* 2019;27:272-280.
13. Stone GW, Maehara A, Lansky AJ, et al. A Prospective Natural-History Study of Coronary Atherosclerosis. *New England Journal of Medicine.* 2011;364:226-235.
14. Xie JX, Eshtehardi P, Varghese T, et al. Prognostic Significance of Nonobstructive Left Main Coronary Artery Disease in Women Versus Men. *Circulation: Cardiovascular Imaging.* 2017;10:e006246.
15. Gaudino MFL, An KR, Calhoon J. Mechanisms for the Superiority of Coronary Artery Bypass Grafting in Complex Coronary Artery Disease. *Ann Thorac Surg.* 2023;115:1333-1336.

BARRIERS TO BLACK MEDICAL STUDENTS AND RESIDENTS PURSUING AND COMPLETING SURGICAL RESIDENCY IN CANADA: A QUALITATIVE ANALYSIS

Edgar Akuffo-Addo[1*], Jaycie Dalson[1*], Kwame Agyei [1], Samiha Mohsen[1], Safia Yusuf[1], Clara Juando-Prats[2,3], Jory S Simpson[1,3]

[1]Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[2] Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[3] Division of General Surgery, St. Michael's Hospital, Toronto, ON, Canada
*Edgar Akuffo-Addo and Jaycie Dalson have contributed equally to this study and should be considered as co-first authors.

Introduction: Within the medical field in Canada, there is a persistent and disproportionate underrepresentation of racial and ethnic minority groups in surgical disciplines.[1 2] The limited available data suggest that the North American medical workforce does not reflect the racial diversity of the patient population it serves especially within surgery.[3 4 5 6 7 8 9 10 11] There is a growing impetus to improve diversity in the surgical force. The American College of Surgeons (ACS) recognizes the importance and need of increasing diversity within the field of surgery to improve patient outcomes. [12] The health care benefits of patient-provider race concordance are well established. Black patients were more receptive to a surgeon's recommendation for surgery when their physicians were of the same race.[13] One can appreciate the importance of this especially when it comes to the surgical management malignancies. In addition, patients rated their visits with a physician as more satisfying when race concordant.[14] To improve patient satisfaction and the quality-of-care patients receive, the surgical workforce must reflect the variety of patients for which they provide care.

Prior research has exclusively focused on exploring the challenges underrepresented minorities who are training in postgraduate surgical training programs in Canada face.[2] They show that postgraduate residents who are visible minorities experience mistreatment, increased imposter syndrome, patient dismissal of medical expertise because of race/ethnicity, and stressors related to the recurrence of reported discriminatory incidences.[2] However, to our knowledge and from our literature search, there have been no studies to date that characterize the challenges faced by racial/ethnic minority medical students in undergraduate education in accessing and entering surgical specialties in a Canadian context. However, as the number of Black medical students in Canada increases,[15] there is a greater need to understand the unique barriers and challenges Black medical students face in pursuing surgical residency.

Negative perceptions and stereotypes about surgeons likely discourage many Black medical students from pursuing a surgical career.[16] Surveys among medical students show that they hold uniform stereotypes of surgeons as self-confident and intimidating; surgery as competitive, masculine, and requiring sacrifice. Students felt they had to fit these stereotypes - often shaped by gendered constructs - if they were to succeed in surgery, excluding those unwilling, or who felt unable to conform.[17] These negative stereotypes could disproportionately affect Black medical students who already question their sense of belonging in the Canadian medical system.[18] Interventions aimed at improving these negative perceptions and stereotypes have shown immediate positive results. [19 20 21 22 23] However, there are likely unique factors that contribute to Black medical students' negative perceptions about the field of surgery. Interviewing current Black surgical residents may also provide insights on barriers encountered during surgical training and difficulties faced to in pursuing surgical careers. Illuminating the factors that differentially impact Black medical students and residents is required to develop more targeted interventions to address these factors with the overall goal of increasing the number of practicing Black surgeons in Canada. As such, our research aims to explore the experiences of Black medical students in applying for a surgical residency in Canada. In addition, we sought to examine the experiences of Black surgical residents who have gone through the application process and their current encounters in residency.

Methods: Using critical qualitative inquiry, purposive sampling pursuing heterogeneity for gender, geographical location, and student/trainee year, we recruited self-identifying Black medical students and surgical residents across Canada. Online in-depth semi-structured interviews were conducted and transcribed verbatim. Transcripts were analyzed through an inductive reflexive narrative thematic process by four analysts.

Results: The authors interviewed 27 participants including 18 medical students and 9 residents (Table 1). They ranged in age from 20 to 40 years of age. 16 identified as females (59%) and 11 as males (41%). Almost half of the participants (41%) were from a perceived low SES background. More than half (59%) were born outside Canada. The results showed three major themes that characterized the experiences: journey to and through medicine, perceptions of the surgical culture, and recommendations to improve the student experience.

Theme 1: Journey to and through medicine

Each participant described their initial interest in medicine and surgery, as well as how this interest evolved based on experiences both within and outside the clinical environment. Intersectionality is defined as multiple interlocking identities with relative sociocultural power and privilege that shape people's individual and collective experiences. This plays a large role in the lives of our participants, who are all black, but with unique cultures, genders, sexual orientations, socioeconomic status etc. and may be additive, multiplicative, or intersectional in different situations. [24] With this in mind, we identified 4 subthemes within participants' journey to and through medicine: 1) Motivation 2) Admissions Criteria 3) Finances 4) Social Capital

Motivation

Among students interested in surgical specialties, most were drawn to the immediate gratification that came with the hands-on work associated with surgery. Some students cited the anticipated satisfaction they would derive from being able to provide their patients with near-instantaneous relief for their medical conditions through surgery. In exploring specialties for postgraduate training, Black medical students were more likely to consider primary care and public health specialties. They shared the burden of expectation and moral obligation to pursue primary care practice to fill care gaps in their local communities. Some non-OBGYN surgical specialties were perceived as having minimal community impact.

Admission Criteria: Research vs Advocacy

Black medical students were more likely to have engaged in advocacy and other community-building initiatives that are perceived not to be formally credited in the surgical residency admission process. This made several candidates hesitant to apply considering many of their extracurricular activities were centered around those initiatives.

Socio-economic Status

While most students from higher SES were financially supported by their parents with the cost of application and test registration fees, students from lower SES often had to seek part-time employment to cover these costs. Moreover, Black applicants from lower SES were more likely to limit their applications to a few medical schools as a cost-saving method, thereby significantly reducing their chances of admission. These applicants were also less likely to see medicine as a viable career path after one unsuccessful application cycle. Some applicants shared that the cost of the MCAT and application fees were a large deterrent for reapplying.

Finances and compensation did not singularly influence Black medical students' choice of postgraduate specialty. Students were assured that a career in medicine will provide adequate compensation irrespective of the specialty they choose but some shared their concerns around job security in surgical specialties.

Social Capital

Another common barrier shared by most participants, irrespective of SES, included an admission system that favors candidates in a more privileged social position than those in lower positions (e.g., visible minorities, newcomers, English as a second language, or gender minorities).

Students felt they were lacking the social capital seemingly required to get into medicine. Most participants were the first in their family to pursue medicine and shared their struggles with finding research opportunities and networking. Participants also described significant challenges finding mentors once they matriculated into medical school.

Theme 2: Sense of Belonging
In exploring their journeys through medical school and postgraduate medical training, most Black learners cited feelings of isolation and lack of belonging. We identified 3 subthemes: 1) Isolation 2) Microaggressions & Verbal abuse and 3) Gender.

Isolation
For many of the students, their experience of being the only Black learner in their respective programs contributed to their isolation, anxiety, and depression. Black medical students struggled with the idea of being the only black surgical resident or staff in their institutions and the possible impact to their mental well-being.

Microaggressions and Verbal Abuse
Black medical students and residents continue to experience discrimination in the form of microaggressions. Microaggressions are brief and commonplace daily verbal or behavioral indignities, whether intentional or unintentional that communicate hostile, derogatory, or negative racial slights and insults toward people of color. [24] Some medical students who were considering surgical specialties received unsupportive remarks from their supervising staff. They were made to question if they could handle the rigors of surgical specialties. Students found these comments discouraging, unsupported and were left wondering if they received these comments because of their race.

Gender
The intersection of race and gendered experiences also influenced the specialty choice of black female medical students. Medical students who identified as female and who were interested in surgical specialties were more likely to consider Obstetrics and Gynecology over other subspecialties. The underrepresentation of women in surgical specialties such as Orthopedics, Neurosurgery and Vascular surgery was a major deterrent for female medical students.

Theme 3: Surgical Lifestyle and Culture
The field of surgery was deemed as mentally, emotionally, and physically demanding by participants. It was seen as a specialty that requires significant dedication and sacrifice. Through our analysis we identified two subthemes including 1) Surgical Lifestyle & Sacrifice of Personal Needs 2) Duty Hours & Sleep Deprivation.

Conclusion: In summary, medical students identified lack of mentorship and representation, and experiences with racism as the main barriers to pursuing surgical training. Surgical trainees cited systemic racism, lack of representation and poor safe spaces as the main deterrents to program completion. The intersection with gender exponentially increased these identified barriers.

Except for a few surgical programs, medical schools across Canada do not offer a safe space for Black students and trainees to access and complete surgical training. An urgent change is needed to offer diverse mentorship that is transparent and acknowledges the real challenges related to systemic racism and biases, as well as being inclusive of different racial and ethnic backgrounds. We provide a table of recommendations for programs to consider to optimize the experiences of Black medical students and residents pursuing and completing surgical residencies in Canada.

Table I: Demographic characteristics of 27 Black medical students and residents interviewed to explore their experiences in pursuing and completing surgical residency in Canada

Characteristics	Undergraduate		Postgraduate		Full sample	
	n	%	*n*	%	*n*	%
Gender						
Female	12	60	4	57	16	59
Male	8	40	3	43	11	41
Age range						
20 - 24	11	55	1	14	12	44
25 - 29	9	45	3	43	12	44
30 - 34	0	0	2	29	2	7
35 - 40	0	0	1	14	1	4
Socioeconomic status						
Low	9	45	2	29	11	41
Mid	10	50	4	57	14	52
High	1	5	1	14	2	7
Location of training						
Atlantic Region	2	10	0	0	2	7
Central Canada	16	80	3	42	19	70
Prairie Provinces	2	10	2	29	4	15
West Coast	0	0	2	29	2	7
Country of Origin						
Canada	7	37	4	57	11	41
Nigeria	7	11	0	0	7	26
Ghana	2	37	1	14	3	11
Sudan	1	5	1	14	2	7
Russia	0	0	1	14	1	4
UK	1	5	0	0	1	4
St Lucia	1	5	0	0	1	4
Unknown	1	5	0	0	1	4
Specialty Interest						
Surgery	13	65	N/A		N/A	
Medicine	5	25	N/A		N/A	
Psychiatry	1	5	N/A		N/A	
Pediatrics	1	5	N/A		N/A	

Atlantic Region: Newfoundland and Labrador, Prince Edward Island, Nova Scotia, New Brunswick
Central Canada: Quebec, Ontario
Prairie Provinces: Manitoba, Saskatchewan, Alberta
West Coast: British Columbia

Table II. Table of Recommendations

	Barrier(s)	Recommendation(s)
Undergraduate	Mentorship	1. Actively support and promote the nationwide mentorship program established by the Black Physicians of Canada Group. This vital program facilitates connections between students and potential mentors, playing a crucial role in fostering concordant mentorship. It is imperative to increase the involvement of Black physicians, residents, and students in the program, ensuring that mentorship opportunities are abundant, relevant, and reflective of the diverse experiences and challenges faced within the medical community. [25] 2. Establish National surgery interest groups to enable students from Canadian regions with limited access to Black surgeons to network with staff and faculty from other universities.
	Admissions	1. Consider holistic review of applicants' files and recognize the complexity of lived experiences that applicants may have in their path to applying to medicine. 2. De-emphasize GPA as a metric for admission as some minority students face financial barriers that may require them to work part-time during their undergraduate education.
Postgraduate	Training Environment	1. Promote a collegial and supportive work environment for residents, ensuring that it actively counters racism and fosters inclusivity. 2. Implement mandatory training on implicit bias and anti-racism for faculty and staff to create an environment where all learners feel equally valued and understood. 3. Cultural change to encourage eating, drinking water, using the washroom and sleep. 4. Adjustment of duty hours and increase in residency spots with consideration of the number of residents per surgical specialty. 5. Transparent process for requesting accommodations including the establishment of a standardized maternity leave process across all surgical specialties.

	Barrier(s)	Recommendation(s)
	Admissions	1. Demonstrate commitment to promoting diversity, equity, and inclusion by establishing pathway programs. Pathway programs signal to applicants that a program is ready to provide holistic and multifaced support to minority students. 2. Highlight programs' commitment to providing medical care to diverse populations by rewarding applicants' advocacy efforts. 3. Permit applicants to substitute advocacy initiatives to fulfill research requirements. This will encourage minority applicants who may have spent their most of their efforts on advocacy at the expense of research to apply. 4. Consider mandatory implicit bias training for all interviewers and file reviewers.

References

1. Khan R, Apramian T, Kang JH, Gustafson J, Sibbald S. Demographic and socioeconomic characteristics of Canadian medical students: a cross-sectional study. *BMC Medical Education.* 2020;20(1):1-8

2. Mocanu V, Kuper TM, Marini W, et al. Intersectionality of Gender and Visible Minority Status Among General Surgery Residents in Canada. *JAMA Surg.* 2020;155(10):e202828. doi:10.1001/jamasurg.2020.2828

3. Canadian Medical Association. Addressing gender equity and diversity in Canada's medical profession: A review. *Canadian Medical Association.* 2018

4. Talamantes E, Henderson MC, Fancher TL, Mullan F. Closing the gap—making medical school admissions more equitable. *N Engl J Med.* 2019;380(9):803-805

5. Attiah MA. The new diversity in medical education. *N Engl J Med.* 2014;371(16):1474-1476

6. Pories SE, Turner PL, Greenberg CC, Babu MA, Parangi S. Leadership in American surgery: women are rising to the top. *Ann Surg.* 2019;269(2):199-205

7. West MA, Hwang S, Maier RV, et al. Ensuring equity, diversity, and inclusion in academic surgery: an American Surgical Association white paper. *Ann Surg.* 2018;268(3):403-407

8. Poon S, Kiridly D, Mutawakkil M, et al. Current trends in sex, race, and ethnic diversity in orthopaedic surgery residency. *JAAOS-Journal of the American Academy of Orthopaedic Surgeons.* 2019;27(16):e725-e733

9. Aibana O, Swails JL, Flores RJ, Love L. Bridging the gap: holistic review to increase diversity in graduate medical education. *Academic Medicine.* 2019;94(8):1137-1141

10. Peter TY, Parsa PV, Hassanein O, Rogers SO, Chang DC. Minorities struggle to advance in academic medicine: a 12-y review of diversity at the highest levels of America's teaching institutions. *J Surg Res.* 2013;182(2):212-218

11. Cohen JJ, Gabriel BA, Terrell C. The case for diversity in the health care workforce. *Health Aff.* 2002;21(5):90-102

12. Williams-Karnesky RL, Kashyap M, Courtney C, Park C, Ritter KA, Hanke R. Shoring up the pipeline: Increasing diversity in surgery by enhancing equity and inclusion in the surgical learning environment. *Bull Am Coll Surg.* 2021

13. Saha S, Beach MC. Impact of physician race on patient decision-making and ratings of physicians: a randomized experiment using video vignettes. *Journal of general internal medicine.* 2020;35:1084-1091

14. Takeshita J, Wang S, Loren AW, et al. Association of racial/ethnic and gender concordance between patients and physicians with patient experience ratings. *JAMA network open.* 2020;3(11):e2024583

15. Collie M. Black medical students accepted to U of T Medicine—the most in canadian history. *Global News.* 24;2

16. Schmidt LE, Cooper CA, Guo WA. Factors influencing US medical students' decision to pursue surgery. *J Surg Res.* 2016;203(1):64-74

17. Hill EJ, Bowman KA, Stalmeijer RE, Solomon Y, Dornan T. Can I cut it? Medical students' perceptions of surgeons and surgical careers. *The American Journal of Surgery.* 2014;208(5):860-867

18. Mathieu J, Fotsing S, Akinbobola K, et al. The quest for greater equity: a national cross-sectional study of the experiences of Black Canadian medical students. *Canadian Medical Association Open Access Journal.* 2022;10(4):E937-E944

19. McKinley SK, Sell NM, Saillant N, et al. Enhancing the formal preclinical curriculum to improve medical student perception of surgery. *Journal of surgical education.* 2020;77(4):788-798

20. Hicks KE, Doubova M, Winter RM, Seabrook C, Brandys T. Surgical exploration and discovery program: early exposure to surgical subspecialties and its influence on student perceptions of a surgical career. *Journal of Surgical Education.* 2019;76(5):1248-1257

21. Kim S, Farrokhyar F, Braga LH. Survey on the perception of urology as a specialty by medical students. *Canadian Urological Association Journal.* 2016;10(9-10):349

22. Peel JK, Schlachta CM, Alkhamesi NA. A systematic review of the factors affecting choice of surgery as a career. *Canadian Journal of Surgery.* 2018;61(1):58

23. Smith AA, Duncan SF, Esparra BC. Can brief interventions by hand surgeons influence medical students toward a career in hand surgery? *J Hand Surg.* 2007;32(8):1267-1270

24. Sue DW, Capodilupo CM, Torino GC, et al. Racial microaggressions in everyday life: implications for clinical practice. *Am Psychol.* 2007;62(4):271

25. Egbedeyi O, El-Hadi H, Madzima TR, Semalulu T, Tunde-Byass M, Swaleh R. Assessing the need for Black mentorship within residency training in Canada. *CMAJ.* 2022;194(42):E1455-E1459

DEEP MULTIPROTEOMIC ANALYSIS OF HUMAN THORACIC AORTIC ANEURYSMS AND DISSECTION: RESULTS FROM THE MULTITAAD

Malak Elbatarny(SSTP)[1,2], Uros Kuzmanov[2,5], Daniella Eliathamby[3], Vivian Chu[4], Rashmi Nedadur[2,4], Cristine Reitz[2,5], Raymond Kim[6], Craig Simmons[3], Jennifer CY Chung[1], Bo Wang[4], Maral Ouzounian[1], Anthony O Gramolini[2,5], on behalf of the MultiTAAD[7]

[1]Division of Cardiovascular Surgery, University of Toronto, Canada
[2]Department of Physiology, University of Toronto, Canada
[3]Department of Biomedical and Molecular Engineering, University of Toronto, Canada
[4]Vector Institute of Artificial Intelligence, University of Toronto, Canada
[5]Translational Biology & Engineering Program, Ted Rogers Centre, University of Toronto, Canada
[6]Cardiac Genome Clinic, Hospital for Sick Children, Toronto, Canada
[7]MultiTAAD: Multiomics of Thoracic Aneurysm & Dissection, University of Toronto, Canada

The authors have decided not to make the research results available at this time and will provide updates as soon as the results can be shared.

FEMALE MICE DISPLAY SEX-SPECIFIC DIFFERENCES IN CEREBROVASCULAR FUNCTION AND SUBARACHNOID HAEMORRHAGE-INDUCED INJURY

Hoyee Wan[1,2,3 *], Danny D Dinh[1,2 *], Darcy Lidington[1,2 +] and Steffen-Sebastian Bolz[1,2,4 +]

[1]Department of Physiology, University of Toronto, Toronto, Canada

[2]Toronto Centre for Microvascular Medicine at *The Ted Rogers Centre for Heart Research* Translational Biology and Engineering Program, University of Toronto, Canada

[3]Plastic, Reconstructive and Aesthetic Surgery, Department of Surgery, University of Toronto, Toronto, Canada

[4]Heart & Stroke / Richard Lewar Centre of Excellence for Cardiovascular Research, University of Toronto, Toronto, Canada

***** denotes equal first-author contribution

\+ denotes equal senior author contribution

Introduction

Aneurysmal subarachnoid haemorrhage (SAH) is a devastating type of stroke with high mortality and morbidity [1]. As with other brain injuries [2], the time of day that an SAH occurs significantly impacts the extent of injury that subsequently ensues [3,4]. Defining the diurnal/circadian mechanisms that drive these time-of-day differences is vital since, in addition to impacting injury severity and outcome, these mechanisms are also likely to influence the efficacy of therapeutic interventions [2].

Our recent work in male mice demonstrates that: (i) cerebral resistance arteries possess a circadian rhythm in myogenic reactivity and (ii) the level of myogenic tone at the time of SAH strongly associates with the degree of SAH-induced injury [4]. At the molecular level, circadian oscillations in cystic fibrosis transmembrane conductance regulator (CFTR) channel expression appear to drive the myogenic reactivity rhythm in male cerebral arteries and, by association, the level of SAH-induced injury [4]. CFTR is a prominent modulator of cerebrovascular reactivity: in addition to its role as a chloride channel, its sphingolipid transporter function [5] sequesters the myogenic signalling mediator sphingosine-1-phosphate away from its receptors, thereby preventing the activation of the pro-constrictive signalling pathways that augment myogenic vasoconstriction [6–9].

Given that our previous work focused solely on male mice [4], one key question to resolve is whether the male phenotype is generalizable to females. Remarkably, there are several reasons to propose that females will possess distinctly different vascular and SAH injury phenotypes compared to their male counterparts: (i) biological sex differences in cerebrovascular function, cerebral blood flow autoregulation and cerebrovascular pathologies are well-documented [10]; (ii) female sex hormones are known to modulate CFTR expression [11,12], inflammatory processes [10] and cellular stress responses [10]; and (iii) females appear to have a higher incidence of delayed cerebral ischemia at 3-5 days post-SAH [13]. Circadian time is not frequently incorporated into experimental plans and thus, most studies addressing sex differences have been conducted during daylight hours. Unfortunately, this hampers the generalizability of the reported observations, as any identified difference (or their lack thereof) could be sensitive to the circadian time of data collection.

This investigation examined whether biological sex differences exist in cerebral artery myogenic reactivity and how this influences injury following experimental SAH. To understand the role of sex hormones in females, we conducted assessments in both naïve and ovariectomized female mice. Our previous work establishes a strong association between cerebral resistance artery myogenic reactivity and SAH-induced injury [4]: thus, we hypothesized that this association persists in females. Indeed, our data show that this association is preserved in females, although naïve and ovariectomized female mice possess a distinctly different vascular phenotype compared to males.

Methods

SAH was modelled by pre-chiasmatic blood injection. Olfactory cerebral resistance arteries were functionally assessed by pressure myography; these functional assessments were related to brain histology (Fluoro-J, anti-active caspase-3, anti-Iba1 immunostaining) and neurobehavioral assessments.

Results

Female mice do not display circadian variations in myogenic tone and SAH-induced injury.

In contrast to our previous observations from male cerebral arteries [4], we did not identify a 24-hour rhythm in female cerebral artery myogenic vasoconstriction **(Figure 1a**). Strikingly, when the female tracing is qualitatively compared to our previously published male data [4], the female profile has the appearance of a “dampened rhythm”, with the female tone level aligning with the trough level of tone observed in male mice. Phenylephrine-stimulated vasoconstriction is not identified as rhythmic in female cerebral arteries, which matches our previous observations in male arteries (**Figure 1b**) [4]. Consistent with the lack (or very low amplitude) of rhythmicity, a comparison of female cerebral arteries isolated at ZT11 and ZT23 shows that Zeitgeber time has a very small effect of on myogenic tone and agonist-stimulated vasoconstriction (**Figure 1c**); in contrast, a ZT11/ZT23 myogenic tone

comparison for male cerebral arteries assessed in parallel (**Figure 1d**) shows data compatible with a much larger effect size.

Figure 2a shows representative images of activated and non-activated microglia cells. In contrast to our previous work in male mice [4], the data in **Figure 2a** and **Supplemental Figure 4** are not compatible with the hypothesis that there are meaningful differences in microglia number and/or their activation state when SAH is induced at ZT23 versus ZT11. Representative images of cortical cells stained for the cell injury markers Fluoro-Jade and activated caspase-3 are shown in **Figure 2b**, with quantifications in **Figures 2c-d**. Since the female myogenic tone tracing reasonably aligns with the low ZT11 level of tone observed in male arteries, we expected that the SAH-induced injury in females would align with the ZT11 injury level previously described in male mice [4]. Indeed, the caspase-3 positive staining and neurofunctional scores observed in females largely match the previously described ZT11 injury levels in male mice. However, the data do not support the hypothesis that SAH-induced injury in female mice is more severe when SAH is induced at ZT23 versus ZT11, as previously reported in males [4]. Modified Garcia scores are slightly higher (indicating better neurofunctional outcome) when SAH is induced at ZT23 versus ZT11.

Collectively, the data presented in **Figures 1-2** strongly reinforce the association that we previously described between myogenic tone at the time of SAH and the subsequent SAH-induced injury level [4]. Of note, we did not compare Fluoro-Jade staining across different experiments: we have found that the method is highly-sensitive to experimental variations and consequently, are not confident that separate data sets are directly comparable.

Ovariectomy does not unmask a cerebrovascular rhythm in females.

We hypothesized that ovariectomy would “unmask” a hidden cerebrovascular rhythm in female mice that drives higher levels of myogenic tone and SAH-induced injury at ZT23: data collected from ovariectomized female mice, however, do not support the hypothesis. Both myogenic reactivity (**Figure 3a**) and phenylephrine-stimulated vasoconstriction in cerebral arteries isolated from ovariectomized female mice remain non-rhythmic, according to cycle analysis. Intriguingly, the new myogenic reactivity profile displays a prominent “isolated peak” in myogenic tone at ZT19 (**Figure 3a**). When qualitatively compared to the naïve female data, the myogenic tracing from ovariectomized females displays reasonable alignment with the naïve myogenic tone level, except for the aforementioned ZT19 peak. Vascular reactivity data from arteries isolated from ovariectomized female mice at ZT11 and ZT23 (**Figure 3b**) indicate that Zeitgeber time has a very small effect on vascular reactivity; contrary to our hypothesis, the data are more compatible with ZT11 possessing higher tone compared ZT23 than vice versa. In terms of SAH-induced injury in ovariectomized female mice at 2 days post-SAH cortical Fluoro-Jade (**Figure 3d**) and activated caspase-3 (**Figure 3e**) positive cell counts were more compatible with either a non-substantive effect or reduced injury when SAH is induced at ZT23 versus ZT11 than vice versa; likewise, modified Garcia score data (**Figure 3f**) are more compatible with an improved outcome when SAH is induced at ZT23, rather than a detrimental outcome. Taken together, we conclude that ovarian sex hormones do not explain the differences between males and females with respect to circadian variations in cerebral artery myogenic tone and time-of-day differences in SAH-induced injury.

Discussion

Our previous work demonstrated that (i) myogenic reactivity in olfactory cerebral arteries isolated from male mice oscillates with a circadian rhythm and (ii) male mice display a diurnal difference in SAH-induced injury, with the extent of injury associating with the level myogenic tone at the time of SAH [4]. The present study demonstrates that female mice possess a distinctly different phenotype: there is no statistical rhythm in myogenic reactivity, nor is there a diurnal difference in SAH-induced injury.

Circadian rhythms are present in virtually all living organisms: their entrainment to the solar day/night cycle allows organisms to anticipate recurring environmental patterns and to adapt their

physiology and behaviours accordingly [15,16]. Time of day is a crucial biological variable that, unfortunately, is rarely incorporated into experimental designs, despite a clear influence on both behaviour and experimentally-induced pathologies [4,15,16]. Within the circadian research field, female subjects remain underrepresented, despite clear evidence of physiological sex differences in circadian timing and animal behaviour [16]. Incorporating these variables into scientific research is essential for generalizability, translation and ensuring safety, efficacy and benefit across the entire population.

In this context, the critical observation of the present study is that the cerebrovascular myogenic reactivity rhythm previously described in males [4] is dampened to the extent that it is *functionally absent* in females. To our knowledge, this is the first time such a profound difference between the sexes has been demonstrated at the circadian level. Female sex hormones have well-known acute effects on vascular contractility [17] and thus, we were surprised to observe that ovariectomy has no effect on myogenic reactivity. Similarly, oestrogens are widely considered to be neuroprotective [10], and it was therefore surprising that ovariectomy did not enhance SAH-induced activated caspase-3 staining and elicited only a tendency for worse modified Garcia scores following SAH. Although classically considered an ovarian sex hormone, oestrogen is synthesized by astrocytes and neurons, providing a brain-specific source of oestrogen that serves neuroprotective and anti-inflammatory roles [18]. While ovariectomy undoubtedly reduces circulating oestrogen levels, it is not clear how ovariectomy affects brain oestrogen levels in our SAH model. Thus, the results stemming from our ovariectomy model must be viewed cautiously.

Consistent with the observations from our ovariectomized mouse model, several studies in eumenorrheic human female subjects also indicate that ovarian sex hormones minimally influence autoregulatory function [19–21]. Within the field, however, studies evaluating the influence of sex hormones on vascular function have generated heterogeneous and frequently contradictory results [20,21]: these discrepancies arise from many factors, including human subject characteristics (e.g., eumenorrheic, menopausal, use of contraception, etc.) and the interventional challenges utilized to perturb the system (e.g., hypercapnia, postural changes). Additionally, the *in vitro* system utilized in the present study to assess vascular reactivity washes away diffusible factors and therefore, likely eliminates any reversible effect during functional assessment. Therefore, caution must be applied when comparing our myogenic data to clinical measurements *in vivo*.

Only a handful of experimental studies have assessed the effect of sex on SAH-induced injury [22-25]. A commonality across these previous studies is the use of vascular perforation / aneurysm rupture models that, compared to the blood injection model used in the present study, possess variable bleed severity, extended periods of elevated intracranial pressure and higher mortality. Additionally, these previous studies utilized vastly different time points for injury assessment (minutes to 21 days post-SAH) and none statistically compared male and female injury levels at 48 hours post-SAH. Given these confounds, it is not practical to directly compare our study results to this previous body of work. However, the primary goal or our investigation was not to address sex differences in SAH-induced injury *per se*, but rather to identify potential sex differences in the cerebrovascular phenotype and assess its association to SAH-induced injury. In this regard, our vascular data predicted that female mice would display less brain injury than males when the SAH ictus occurs during the waking/active phase of the circadian cycle: a comparison of our current and previously published histological/neurofunctional data substantiate this prediction [4].

In summary, this investigation shows that female mice possess a distinctly different cerebrovascular phenotype than their male counterparts. Female mice lack the diurnal difference in SAH-induced injury previously observed in males, fairing equal or better depending on SAH ictus. Notably, ovariectomy does not unmask a rhythmic phenotype. Thus, translational studies clinically assessing therapeutics in cerebrovascular pathologies must be prepared to identify and quantify potential sex differences in efficacy.

Figures

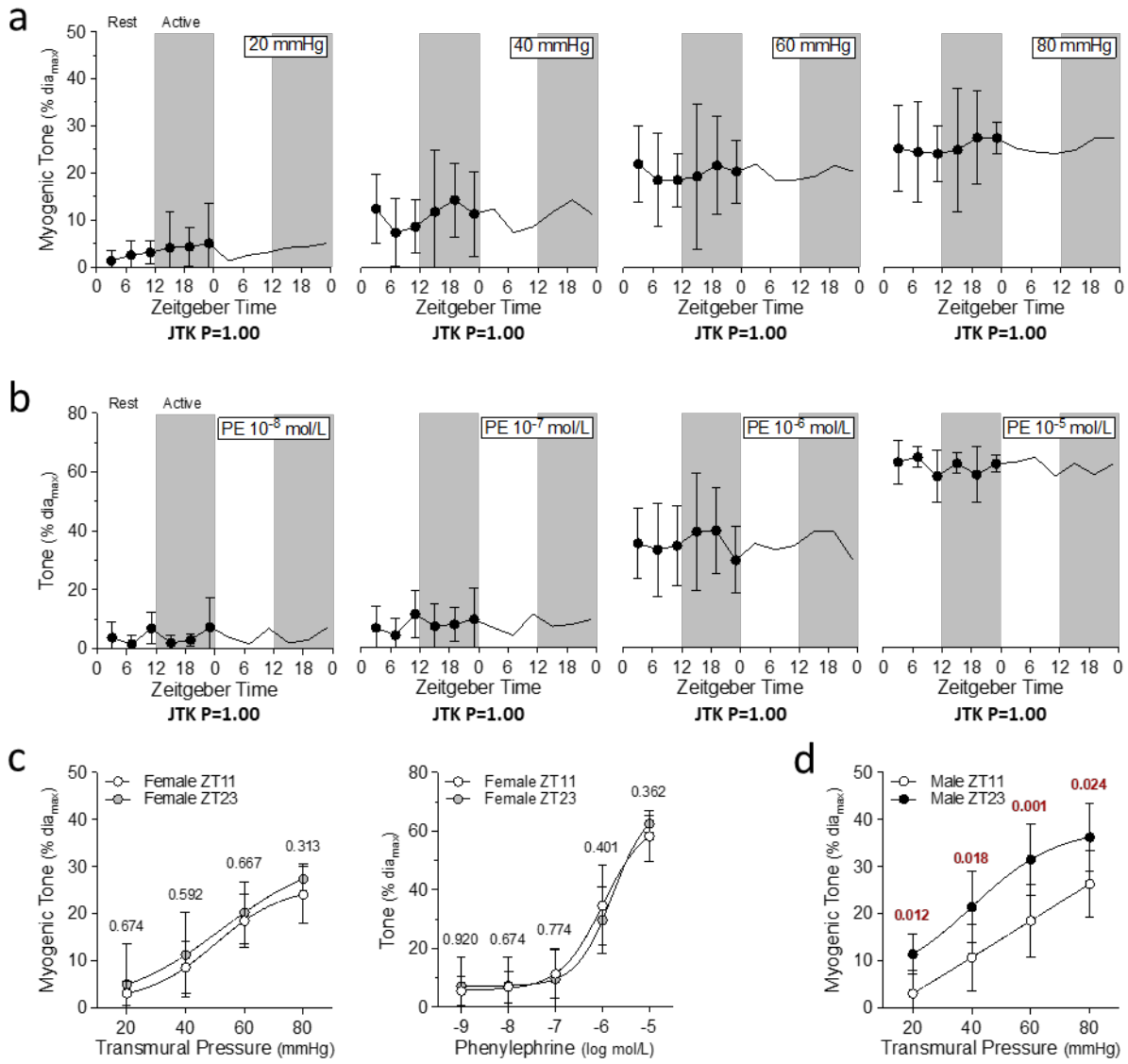


Figure 1: Female mice do not display circadian variations in myogenic tone.
Panels a and b display wild-type female olfactory cerebral artery (a) myogenic tone and (b) phenylephrine-stimulated vasoconstriction plotted over Zeitgeber time (n=5 arteries from 3 mice at Zeitgeber times 3, 11, 15 and 23, n=6 arteries from 3 mice at Zeitgeber time 7 and n=6 arteries from 4 mice at Zeitgeber time 19). Data are double-plotted for visualization purposes; white shading indicates "lights on" and dark shading indicates "lights off". Neither myogenic tone nor phenylephrine-stimulated vasoconstriction display a statistically significant circadian rhythm according to JTK_CYCLE analysis. (c) Female olfactory cerebral artery myogenic tone and phenylephrine-stimulated vasoconstriction measured in arteries isolated at Zeitgeber time 11 (ZT11; n=5 arteries from 3 mice) and Zeitgeber time 23 (ZT23; n=5 arteries from 3 mice). (d) Male olfactory cerebral artery myogenic tone measured in arteries isolated at ZT11 (n=8 arteries from 4 mice) and ZT23 (n=6 arteries from 4 mice). All data are means ± standard deviation. In panels a-b, the JTK_CYCLE p value for the given transmural pressure or PE concentration is shown below the respective graph. In panels c-d, comparison test p-values are shown above the error bars.

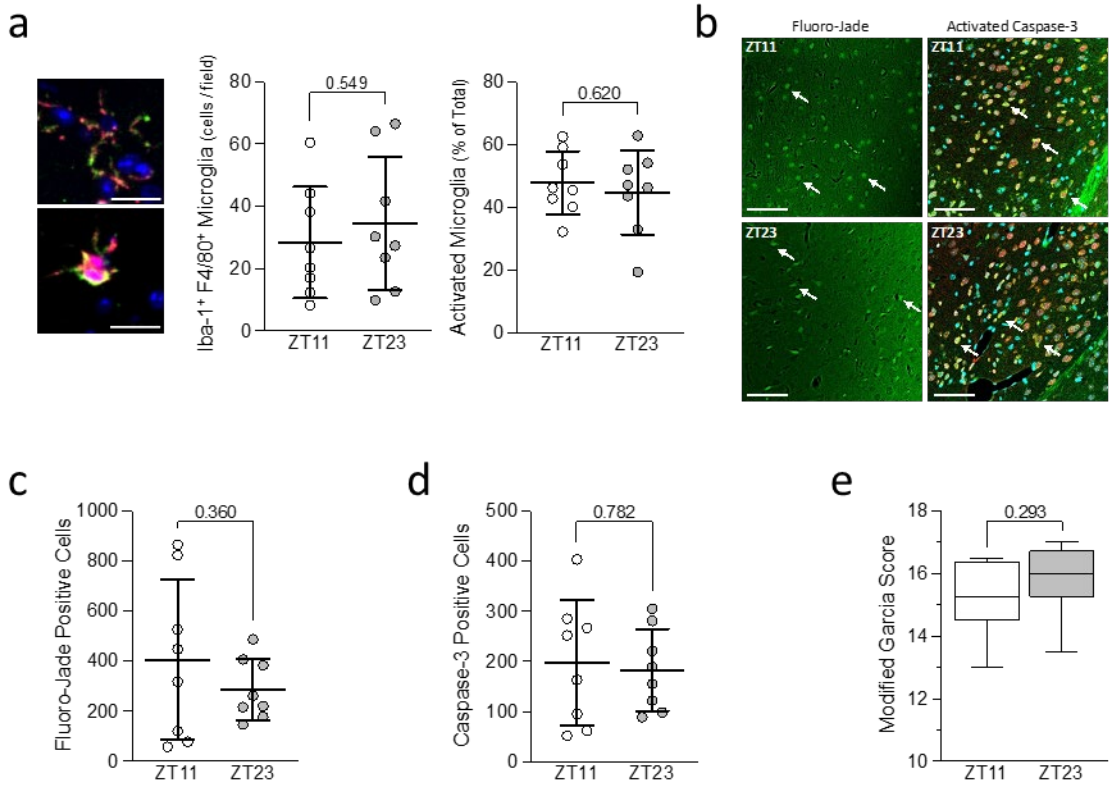


Figure 2: Female mice do not display circadian variations SAH-induced injury.

Subarachnoid haemorrhage (SAH) was induced in wild-type female mice at Zeitgeber Time 11 (ZT11) or Zeitgeber Time 23 (ZT23); behavioural assessments and tissue collection were conducted 48 hours afterwards. **(a)** Representative images of microglia double-stained with ionized calcium binding adaptor molecule 1 (Iba) and F4/80 are shown on the left. The top panel shows a highly fenestrated cell, while the bottom panel shows a rounded morphology (bar = 20 μm, pink = Iba, green = F4/80, blue = DAPI). To the right of the images are quantifications of double-positive cell counts and percent activated cells (based on morphology) when SAH is induced at ZT11 (n=8 brain slices from 8 mice) or ZT23 (n=8 brain slices from 8 mice). **(b)** Representative images of cortical cells stained for/with Fluoro-Jade (left) and activated caspase-3 (right); arrows point to positively stained neuronal cells (red = NeuN, green = activated caspase-3, cyan = DAPI). For all images, bar = 60 μm.
Positive neuronal cell counts for **(c)** Fluoro-Jade staining and **(d)** activated caspase-3 when SAH is induced at ZT11 (n=8 brain slices from 8 mice) or ZT23 (n=8 brain slices from 8 mice). **(e)** Modified Garcia scores when SAH is induced at ZT11 (n=8 mice) or ZT23 (n=8 mice). Data in *Panel e* (modified Garcia scores) are ordinal data and are presented as box and whisker plots displaying median, upper and lower interquartile range and upper and lower extremes; all other data are presented as means ± standard deviation. In panels a and c-e, comparison test p-values are shown above the error bars.

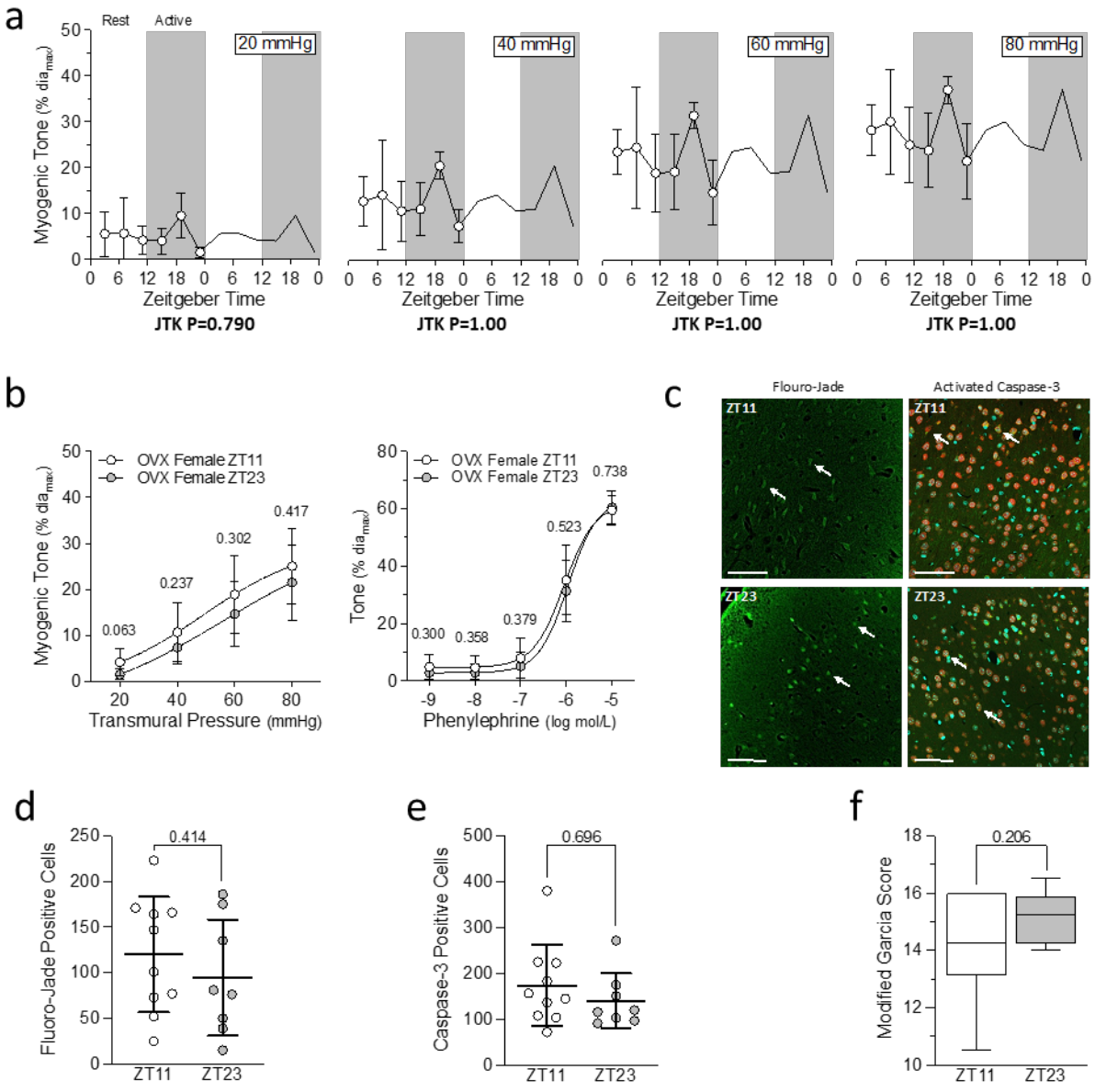


Figure 3: Ovariectomy does not unmask circadian variations in myogenic tone or SAH-induced injury.

(a) Myogenic tone in cerebral arteries isolated from ovariectomized female mice (OVX) is plotted over Zeitgeber time (n=5 arteries from 3 mice at Zeitgeber times 7 and 19, n=6 arteries from 3 mice at Zeitgeber times 3 and 15, n=7 arteries from 4 mice at Zeitgeber time 11 and n=8 arteries from 5 mice at Zeitgeber time 23). Data are double-plotted for visualization purposes; white shading indicates "lights on" and dark shading indicates "lights off". No statistically significant circadian rhythm is present according to JTK_CYCLE analysis. **(b)** Myogenic and phenylephrine-stimulated vasoconstriction in olfactory cerebral arteries isolated from OVX mice at Zeitgeber time 11 (ZT11; n=7 arteries from 4 mice) and Zeitgeber time 23 (ZT23; n=8 arteries from 5 mice). In *Panels c-f*, subarachnoid haemorrhage (SAH) was induced in OVX mice at ZT11 or ZT23; behavioural assessments and tissue collection were conducted 48 hours afterwards. **(c)** Representative images show cortical cells stained for Fluoro-Jade (left) and activated caspase-3 (right); arrows point to positively stained neuronal cells

(Bar = 60 μm; red = NeuN, green = activated caspase-3, cyan = DAPI). Positive neuronal cell counts for **(d)** Fluoro-Jade staining and **(e)** activated caspase-3 when SAH is induced at ZT11 (n=10 brain slices from 10 mice) or ZT23 (n=8 brain slices from 8 mice). **(f)** Modified Garcia scores when SAH is induced at ZT11 (n=10 mice) or ZT23 (n=8 mice). Data in *Panel f* (modified Garcia scores) are ordinal data and are presented as box and whisker plots displaying median, upper and lower interquartile range and upper and lower extremes; all other data are presented as means ± standard deviation. In panel a, the JTK_CYCLE p value for the given transmural pressure is shown below the respective graph. In panels b and d-f, comparison test p-values are shown above the error bars.

References

1. Solenski NJ, Haley EC, Kassell NF, Kongable G, Germanson T, Truskowski L, et al. Medical complications of aneurysmal subarachnoid hemorrhage: a report of the multicenter, cooperative aneurysm study. Participants of the Multicenter Cooperative Aneurysm Study. *Crit Care Med*. 1995;23:1007–1017.
2. Hetman M, Slomnicki LP, Hodges ER, Saraswat Ohri S, Whittemore SR. Role of circadian rhythms in pathogenesis of acute CNS injuries: Insights from experimental studies. *Exp Neurol*. 2022;353:114080.
3. Schallner N, Lieberum J-L, Gallo D, LeBlanc RH, Fuller PM, Hanafy KA, et al. Carbon Monoxide Preserves Circadian Rhythm to Reduce the Severity of Subarachnoid Hemorrhage in Mice. *Stroke*. 2017;48:2565-2573.
4. Lidington D, Wan H, Dinh DD, Ng C, Bolz SS. Circadian Rhythmicity in Cerebral Microvascular Tone Influences Subarachnoid Hemorrhage–Induced Injury. *Stroke*. 2022;53:249–259.
5. Boujaoude LC, Bradshaw-Wilder C, Mao C, Cohn J, Ogretmen B, Hannun YA et al. Cystic fibrosis transmembrane regulator regulates uptake of sphingoid base phosphates and lysophosphatidic acid: modulation of cellular activity of sphingosine 1-phosphate. *J Biol Chem*. 2001;276:35258–35264.
6. Lidington D, Fares JC, Uhl FE, Dinh DD, Kroetsch JT, Sauvé M, et al. CFTR Therapeutics Normalize Cerebral Perfusion Deficits in Mouse Models of Heart Failure and Subarachnoid Hemorrhage. *JACC Basic Transl Sci*. 2019;4:940–958.
7. Lidington D, Wan H, Bolz SS. Cerebral Autoregulation in Subarachnoid Hemorrhage. *Front. Neurol.* 2021;12:688362.
8. Meissner A, Yang J, Kroetsch JT, Sauvé M, Dax H, Momen A, et al. Tumor Necrosis Factor-α-Mediated Downregulation of the Cystic Fibrosis Transmembrane Conductance Regulator Drives Pathological Sphingosine-1-Phosphate Signaling in a Mouse Model of Heart Failure. *Circulation*. 2012;125:2739–2750.
9. Yagi K, Lidington D, Wan H, Fares JC, Meissner A, Sumiyoshi M, et al. Therapeutically Targeting Tumor Necrosis Factor-alpha/Sphingosine-1-Phosphate Signaling Corrects Myogenic Reactivity in Subarachnoid Hemorrhage. *Stroke*. 2015;46:2260–2270.
10. Robison LS, Gannon OJ, Salinero AE, Zuloaga KL. Contributions of sex to cerebrovascular function and pathology. *Brain Res*. 2019;1710:43–60.
11. Ismail N, Giribabu N, Muniandy S, Salleh N. Estrogen and progesterone differentially regulate the levels of cystic fibrosis transmembrane regulator (CFTR), adenylate cyclase (AC), and cyclic adenosine mono-phosphate (cAMP) in the rat cervix. *Mol Reprod Dev*. 2015;82:463–474.
12. Jin H, Wen G, Deng S, Wan S, Xu J, Liu X, et al. Oestrogen upregulates the expression levels and functional activities of duodenal mucosal CFTR and SLC26A6. *Exp Physiol*. 2016;101:1371–1382.
13. Dienel A, Ammassam Veettil R, Hong S-H, Matsumura K, Kumar T P, Yan Y, et al. Microthrombi Correlates With Infarction and Delayed Neurological Deficits After Subarachnoid Hemorrhage in Mice. *Stroke*. 2020;51:2249-2254.
14. Kloske CM, Gearon MD, Weekman EM, Rogers C, Patel E, Bachstetter A, et al. Association between APOE genotype and microglial cell morphology. *J Neuropathol Exp Neurol*. 2023;82:620–630.
15. Nelson RJ, Bumgarner JR, Walker WH, DeVries AC. Time-of-day as a critical biological variable. *Neurosci Biobehav Rev*. 2021;127:740–746.
16. Walton JC, Bumgarner JR, Nelson RJ. Sex Differences in Circadian Rhythms. *Cold Spring Harb Perspect Biol*. 2022;14:a039107.

17. Khalil RA. Sex Hormones as Potential Modulators of Vascular Function in Hypertension. *Hypertension*. 2005;46:249–254.
18. Korad S, Mündel T, Fan J-L, Perry BG. Cerebral autoregulation across the menstrual cycle in eumenorrheic women. *Physiol Rep*. 2022;10:e15287.
19. Abidi S, Nili M, Serna S, Kim S, Hazlett C, Edgell H. Influence of sex, menstrual cycle, and oral contraceptives on cerebrovascular resistance and cardiorespiratory function during Valsalva or standing. *J Appl Physiol (1985)*. 2017;123:375–386.
20. Favre ME, Serrador JM. Sex differences in cerebral autoregulation are unaffected by menstrual cycle phase in young, healthy women. *Am J Physiol Heart Circ Physiol*. 2019;316:H920–H933.
21. Skinner BD, Davies RJ, Weaver SR, Cable NT, Lucas SJE, Lucas RAI. A Systematic Review and Meta-Analysis Examining Whether Changing Ovarian Sex Steroid Hormone Levels Influence Cerebrovascular Function. *Front Physiol*. 2021;12:687591.
22. Friedrich V, Bederson JB, Sehba FA. Gender Influences the Initial Impact of Subarachnoid Hemorrhage: An Experimental Investigation. *PLoS One*. 2013;8:e80101.
23. Guo D, Wilkinson DA, Thompson BG, Pandey AS, Keep RF, Xi G, et al. MRI Characterization in the Acute Phase of Experimental Subarachnoid Hemorrhage. *Transl Stroke Res*. 2017;8:234-243.
24. Yanagisawa T, Zhang H, Suzuki T, Kamio Y, Takizawa T, Morais A, et al. Sex and Genetic Background Effects on the Outcome of Experimental Intracranial Aneurysms. *Stroke*. 2020;51:3083–3094.
25. Kagerbauer SM, Kadera V, Gordan LM, Blobner M, Török E, Schmid S, et al. Influence of sex and hormonal status on initial impact and neurocognitive outcome after subarachnoid haemorrhage in rats. *Behav Brain Res*. 2019;363:13–22.

USING MACHINE LEARNING TO PREDICT OUTCOMES FOLLOWING CAROTID ENDARTERECTOMY

Ben Li (SSTP)[1,2,3,4], Derek Beaton[5], Naomi Eisenberg[6], Douglas S Lee[7,8,9], Duminda N Wijeysundera[8,9,10,11], Thomas F. Lindsay[1,3,6], Charles de Mestral[1,2,8,9,11], Muhammad Mamdani[3,4,5,8,9,11,12], Graham Roche-Nagle[1,6], Mohammed Al-Omran[1,2,3,4,11,13]

[1]Department of Surgery, University of Toronto, Canada
[2]Division of Vascular Surgery, St. Michael's Hospital, Unity Health Toronto, Canada
[3]Institute of Medical Science, University of Toronto, Canada
[4]Temerty Centre for Artificial Intelligence Research and Education in Medicine (T-CAIREM), University of Toronto, Canada
[5]Data Science & Advanced Analytics, Unity Health Toronto, University of Toronto, Canada
[6]Division of Vascular Surgery, Peter Munk Cardiac Centre, University Health Network, Canada
[7]Division of Cardiology, Peter Munk Cardiac Centre, University Health Network, Canada
[8]Institute of Health Policy, Management and Evaluation, University of Toronto, Canada
[9]ICES, University of Toronto, Canada
[10]Department of Anesthesia, St. Michael's Hospital, Unity Health Toronto, Canada
[11]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Canada
[12]Leslie Dan Faculty of Pharmacy, University of Toronto, Canada
[13]Department of Surgery, King Faisal Specialist Hospital and Research Center, Kingdom of Saudi Arabia

INTRODUCTION

Carotid artery stenosis is responsible for ~30% of all ischemic strokes, a leading cause of mortality/morbidity worldwide[1,2]. Carotid endarterectomy (CEA) is the surgical management option for patients with severe stenosis; however, the procedure carries significant perioperative risks, with a pooled rate of postoperative stroke/death over 10% in the highest risk patients[3]. To balance the benefits and risks of CEA, the Society for Vascular Surgery (SVS) recommends maintaining a perioperative stroke/death rate of less than 6% and 3% in symptomatic and asymptomatic patients, respectively[4]. Therefore, accurate prediction of outcomes following CEA is critical for guiding clinical decision-making, such as deciding which patients to operate on[4].

Currently, there are no standardized tools to predict outcomes following CEA. A recent systematic review of 37 studies reporting outcome prediction models for carotid revascularization demonstrated significant methodological limitations, incomplete reporting, and inadequate performance[5]. Furthermore, these tools use traditional modelling techniques that require manual input of clinical variables, which deters routine use in busy clinical settings[6]. Therefore, there is an important need to develop better and more practical risk prediction tools for patients being considered for CEA.

Machine learning (ML) is a rapidly advancing technology that enables computers to learn from large amounts of data and make accurate predictions[7]. The field has been driven by the explosion of electronic data combined with increasing computational power[8]. In this study, we used data from a large clinical registry to develop ML models that accurately predict 1-year stroke or death following CEA at the pre-, intra-, and post-operative stages.

METHODS

Study approval

The Research Advisory Council of the SVS Patient Safety Organization (PSO) approved this project and provided the blinded data.

Design

We conducted a ML-based prognostic study and reported our findings based on the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement[9].

Dataset

The Vascular Quality Initiative (VQI) database is a clinical registry maintained by the SVS PSO with the goal of improving the delivery of vascular care[10]. Over 1,000 academic and community hospitals worldwide prospectively submit demographic, clinical, and outcomes data on consecutive eligible vascular surgery patients, including information from their initial hospitalization up to 9-21 months of follow-up[11]. Annual audits with comparison to hospital claims are performed to ensure accuracy of the submitted information[12].

Patient Cohort

All patients who underwent CEA for carotid artery stenosis from January 1, 2003 to January 6, 2022 were included. Patients with unreported presenting symptom status (i.e., asymptomatic or symptomatic [defined as amaurosis fugax, transient ischemic attack, or stroke within 180 days prior to CEA]) were excluded.

Features

Predictor variables (features) used in the ML models were divided into the pre-, intra-, and post-operative stages. Pre-operative features (n = 43) included demographics, comorbidities, functional status, investigations, medications, and symptom status. Intra-operative features (n = 21) included type of anesthesia, endarterectomy technique, use of adjuncts, procedural hemodynamics, and procedure time. Post-operative features (n = 7) included in-

hospital complications: myocardial infarction (MI), congestive heart failure (CHF) exacerbation, cranial nerve injury, reperfusion syndrome, and surgical site infection (SSI).

Outcomes

The primary outcome was stroke or death at 1 year following CEA. Stroke was defined as neurological deficits persisting for > 24 hours in the ipsilateral or contralateral ocular, cortical, or vertebrobasilar territory[13]. Death was defined as all-cause mortality. Secondary outcomes were individual components of the primary outcome.

Model development

Six ML models were trained to predict 1-year stroke or death following CEA: Extreme Gradient Boosting (XGBoost), random forest, Naïve Bayes classifier, support vector machine, multilayer perceptron artificial neural network (MLP ANN), and logistic regression. These are widely used in the literature and demonstrate the best performance for predicting surgical outcomes[14–16].

Our data was randomly split into training (70%) and test (30%) sets[17]. Ten-fold cross-validation and grid search were performed on the training set to find optimal hyperparameters for each ML model[18,19]. To improve class balance, Random Over-Sample Examples (ROSE) was applied to the training set[20]. The models were then evaluated on unseen data in the test set and ranked based on discriminatory metrics, primarily area under the receiver operating characteristic curve (AUROC). We compared models at the pre-operative stage because predictions at this timepoint are the most important in guiding a patient's care and provide the most opportunity to mitigate adverse events[21]. Our best performing model was XGBoost, which had the following optimized hyperparameters for our dataset: number of rounds = 200, maximum tree depth = 3, learning rate = 0.3, gamma = 0, column sample by tree = 0.6, minimum child weight = 1, subsample = 1.

Once we identified the best performing ML model at the pre-operative stage, we further trained it using intra- and post-operative data. At the pre-operative stage, only pre-operative characteristics were considered. At the intra-operative stage, both pre- and intra-operative features were used. At the post-operative stage, all pre-, intra-, and post-operative features were inputted into the model. The reason for building models in this fashion is to allow clinicians to understand a patient's risk throughout their surgical course, thereby guiding decision-making before, during, and after intervention[22].

Statistical analysis

Pre-, intra-, and post-operative characteristics for patients with vs. without stroke or death at 1 year following CEA were compared using independent t-test for continuous variables or chi-square test for categorical variables. Bonferroni correction was used to set statistical significance to account for multiple comparisons.

The primary metric for assessing model performance was AUROC (95% CI). Secondary performance metrics were accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). To further assess predictive performance of our models, we plotted calibration curves and calculated Brier scores[23]. In the final model, feature importance was determined by ranking the top 10 predictors based on the variable importance score (gain)[24]. To assess model robustness on various populations, we performed subgroup analysis of predictive performance based on age, race, ethnicity, sex, insurance status, symptom status, and urgency.

Based on a validated sample size calculator for clinical prediction models, to achieve a minimum AUROC of 0.7 with an outcome rate of ~5% and 43 pre-operative features, the minimum sample size required was 15,142 patients with 758 events[25]. Our cohort of 166,369 patients with 7,749 events met this sample size requirement. There was less than 5% missing data for variables and outcomes of interest; therefore, complete-case analysis was applied whereby only non-missing covariates for each patient were considered[26]. All analyses were performed in R version 4.2.1[27].

RESULTS

Patients, events, and follow-up

Overall, we included 166,369 patients and 7,749 (4.7%) had the primary outcome of stroke or death at 1 year. Of the patients who had an event, 2,357 (1.4%) developed a stroke and 5,392 (3.3%) died. Most events occurred after discharge (n = 5,320 [68.7%]) and the median time to event was 90 (IQR 11 – 228) days. Mean follow-up was 14.2 (SD 4.1) months.

Pre-, intra-, and post-operative characteristics

Pre-operatively, compared to patients who did not develop stroke or death at 1 year, those with an event were older and more likely to be Black, receive Medicare insurance, and have cardiovascular comorbidities. Functionally, a greater proportion of patients with an event required assistance with ambulation or were wheelchair dependent. Despite having more comorbidities, patients who developed stroke or death at 1 year were less likely to receive cardiovascular risk reduction medications. Patients with an event were more likely to have symptomatic carotid stenosis, undergo urgent/emergent surgery, and have an American Association of Anesthesiologists (ASA) class ≥ 4. Intra-operatively, patients with an event were less likely to receive a patch and neuromonitoring. A greater proportion of patients with an event received a drain but were less likely to receive protamine. Patients with an event were more likely to receive a completion angiogram, undergo surgical re-exploration, and have a longer procedure time. Post-operatively, compared to patients who did not have stroke or death at 1 year, individuals with an event were more likely to have in-hospital complications including MI, dysrhythmia, CHF exacerbation, cranial nerve injury, reperfusion syndrome, and SSI (Table 1).

Model performance

Of the 6 ML models evaluated at the pre-operative stage using test set data, XGBoost had the best performance in predicting 1-year stroke or death (AUROC 0.90 [95% CI 0.89 – 0.91]). The secondary performance metrics of XGBoost were the following: accuracy 0.82 (95% CI 0.81 – 0.83), sensitivity 0.82, specificity 0.83, PPV 0.83, NPV 0.81 (Table 2). We then built additional predictive models at the intra- and post-operative stages using XGBoost. The addition of intra-operative features did not significantly change performance, with the AUROC (95% CI) remaining at 0.90 (0.89 – 0.91). However, adding post-operative features improved performance to AUROC 0.94 (95% CI 0.93 – 0.95). There was good agreement between predicted and observed event probabilities, with Brier scores of 0.15 (pre-operative), 0.14 (intra-operative), and 0.11 (post-operative). For secondary outcomes, XGBoost predicted 1-year stroke with AUROC's of 0.85-0.90 and 1-year death with AUROC's of 0.91-0.96. The top 10 predictors of 1-year stroke or death following CEA in the final XGBoost model included 8 pre-operative features (dialysis, prior major amputation, pre-operative living status, existing CHF, symptomatic carotid stenosis, pre-operative hemoglobin, existing hypertension, and prior ipsilateral CEA), 1 intra-operative feature (surgical re-exploration), and 1 post-operative feature (in-hospital MI).

Subgroup analysis

Model performance remained robust on subgroup analyses of all demographic and clinical populations, with AUROC's ranging from 0.88 – 0.92 and no significant differences between majority and minority populations.

CONCLUSIONS

In this study, we used a large clinical registry (VQI) to develop automated and robust ML models that predict 1-year stroke or death following CEA with excellent performance (AUROC's ≥ 0.90). Our models can be applied at the pre-, intra-, and post-operative stages to guide clinical decision-making regarding strategies to mitigate the risk of adverse outcomes (Figure 1). Our models perform better than traditional logistic regression, and therefore, have potential for important utility in the care of patients with carotid artery stenosis.

Table 1. Pre-, intra-, and post-operative characteristics of patients undergoing carotid endarterectomy with and without stroke or death at 1 year

	Absence of stroke or death at 1 year (n = 158,620)	**Presence of stroke or death at 1 year (n = 7,749)**	**P**
Pre-operative characteristics			
Demographics			
Age, years, mean (SD)	70.4 (9.4)	72.6 (9.6)	< 0.001
Female	62,241 (39.2)	3,091 (39.9)	0.26
BMI, kg/m^2, mean (SD)	28.8 (6.1)	27.9 (6.1)	0.46
Race			
American Indian or Alaskan Native	377 (0.24)	22 (0.28)	< 0.001
Asian	1,590 (1.0)	65 (0.84)	
Black	7,086 (4.5)	438 (5.7)	
Native Hawaiian or other Pacific Islander	121 (0.08)	5 (0.06)	
White	143,817 (90.7)	6,983 (90.1)	
More than 1 race	199 (0.13)	6 (0.08)	
Unknown/other	5,332 (3.4)	227 (2.9)	
Hispanic ethnicity	4,825 (3.0)	228 (2.9)	0.65
Insurance status			
Medicare	78,060 (49.2)	4,110 (53.0)	< 0.001
Medicaid	5,493 (3.5)	277 (3.6)	
Commercial	49,148 (31.0)	2,120 (27.4)	
Medicare Advantage	1,677 (1.1)	91 (1.2)	
Military/Veterans Affairs	1,097 (0.7)	50 (0.6)	
Non-US Insurance	1,982 (1.3)	92 (1.2)	
Self-pay (uninsured)	6,370 (4.0)	167 (2.2)	
Unknown/other	14,793 (9.3)	842 (10.9)	
Comorbidities			
Smoking status			
Never	40,403 (25.5)	1,680 (21.7)	< 0.001
Prior	76,376 (48.2)	3,909 (50.4)	
Current	41,564 (26.2)	2,154 (27.8)	
Hypertension	141,257 (89.1)	7,115 (91.8)	< 0.001
Diabetes	56,464 (35.6)	3,374 (43.5)	< 0.001
Coronary artery disease	43,119 (27.2)	2,704 (34.9)	< 0.001
Congestive heart failure	16,930 (10.7)	1,708 (22.0)	< 0.001
Chronic obstructive pulmonary disease	35,328 (22.3)	2,485 (32.1)	< 0.001
Dialysis	1,780 (1.1)	285 (3.7)	< 0.001

	Absence of stroke or death at 1 year (n = 158,620)	Presence of stroke or death at 1 year (n = 7,749)	P
Previous procedures			
Prior ipsilateral carotid endarterectomy	2,679 (1.7)	208 (2.7)	< 0.001
Prior ipsilateral carotid stent	400 (0.3)	28 (0.4)	0.08
Prior aortic aneurysm repair	4,252 (2.7)	308 (4.0)	< 0.001
Prior bypass for peripheral artery disease	8,555 (5.4)	689 (8.9)	< 0.001
Prior endovascular intervention for peripheral artery disease	14,317 (9.0)	972 (12.5)	< 0.001
Prior major amputation	1,215 (0.8)	146 (1.9)	< 0.001
Functional status			
Living status			
Home	156,520 (98.7)	7,471 (96.4)	< 0.001
Nursing home	1,613 (1.0)	248 (3.2)	
Homeless	143 (0.09)	12 (0.16)	
Pre-operative ambulatory status			
Independent	128,712 (81.1)	5,381 (69.4)	< 0.001
With assistance	13,293 (8.4)	1,248 (16.1)	
Wheelchair-dependent	1,591 (1.0)	240 (3.1)	
Bedridden	103 (0.06)	29 (0.04)	
Investigations			
Hemoglobin, g/L, mean (SD)	132 (40.6)	124 (20.9)	< 0.001
Creatinine, umol/L, mean (SD)	98.3 (63.5)	104.0 (54.0)	0.45
Cardiac stress test			
Not done	107,212 (67.6)	5,382 (69.5)	< 0.001
Normal	38,816 (24.5)	1,578 (20.4)	
Abnormal	12,105 (7.6)	770 (9.9)	
Medications			
Acetylsalicylic acid	132,762 (83.7)	6,181 (79.8)	< 0.001
P2Y12 antagonist	53,308 (33.6)	2,649 (34.2)	0.30
Statin	131,833 (83.1)	6,245 (80.6)	< 0.001
Beta blocker	89,513 (56.4)	4,948 (63.9)	< 0.001
ACE-I/ARB	76,220 (48.1)	3,460 (44.7)	< 0.001
Anticoagulant	15,657 (9.9)	1,090 (14.1)	< 0.001
Anatomy			
Prior neck radiation	2,065 (1.3)	142 (1.8)	< 0.001
High-risk anatomic features			
Prior neck surgery	5,387 (3.4)	332 (4.3)	< 0.001
Tracheostomy or esophagostomy	59 (0.04)	8 (0.1)	< 0.001
Contralateral laryngeal nerve palsy	51 (0.03)	7 (0.09)	< 0.001

	Absence of stroke or death at 1 year (n = 158,620)	**Presence of stroke or death at 1 year (n = 7,749)**	**P**
Lesion extends below clavicle	160 (0.1)	12 (0.2)	< 0.001
Multiple high-risk features	186 (0.1)	17 (0.2)	< 0.001
Carotid stenosis percentage			
0 – 49%	4,101 (2.6)	190 (2.5)	< 0.001
50 – 69%	14,964 (9.4)	777 (10.0)	
70 – 79%	34,031 (21.5)	1,506 (19.4)	
80 – 99%	63,022 (39.7)	3,188 (41.1)	
Occluded	2,149 (1.4)	174 (2.3)	
Other pre-procedural characteristics			
Symptomatic carotid stenosis	48,872 (30.8)	3,102 (40.0)	< 0.001
Urgency			
Elective	137,841 (86.9)	6,078 (78.4)	< 0.001
Urgent	19,621 (12.4)	1,543 (19.9)	
Emergent	904 (0.6)	118 (1.5)	
ASA Class			
1	900 (0.6)	23 (0.3)	< 0.001
2	7,112 (4.5)	176 (2.3)	
3	113,854 (71.8)	4,661 (60.1)	
4	31,498 (19.9)	2,597 (33.5)	
5	53 (0.03)	11 (0.1)	
Intra-operative characteristics			
Antibiotics prior to incision	135,068 (85.2)	6,558 (84.6)	0.24
Anesthesia			
Local	2,425 (1.5)	109 (1.4)	0.65
Regional	9,494 (6.0)	457 (5.9)	
General	146,455 (92.3)	7,174 (92.6)	
Endarterectomy technique			
Conventional	140,095 (88.3)	6,884 (88.8)	0.17
Inversion	17,878 (11.3)	834 (10.8)	
Patch			
None	18,571 (11.7)	1,023 (13.2)	< 0.001
Prosthetic	137,038 (86.4)	6,542 (84.4)	
Autogenous	2,588 (1.6)	162 (2.1)	
Neuromonitoring			
EEG	46,497 (29.3)	2,195 (28.3)	0.06
Stump pressure	17,058 (10.8)	747 (9.6)	0.002
Other	34,489 (21.7)	1,602 (20.7)	0.002
Shunt			
None	77,885 (49.1)	3,423 (44.2)	< 0.001
Routine	67,815 (42.8)	3,396 (43.8)	
Selective	12,469 (7.9)	907 (11.7)	

	Absence of stroke or death at 1 year (n = 158,620)	**Presence of stroke or death at 1 year (n = 7,749)**	**P**
Drain	58,981 (37.2)	3,026 (39.1)	< 0.001
Heparin	156,870 (98.9)	7,661 (98.9)	0.67
Protamine	110,975 (70.0)	5,264 (67.9)	< 0.001
Highest heart rate, beats/min, mean (SD)	67.2 (37.9)	67.5 (38.9)	0.56
Completion testing			
Doppler	89,536 (56.4)	4,395 (56.7)	0.61
Duplex	38,584 (24.3)	1,861 (24.0)	0.55
Angiogram	6,451 (4.1)	396 (5.1)	< 0.001
Surgical re-exploration	2,191 (1.4)	290 (3.7)	< 0.001
Procedure time, min, mean (SD)	119.0 (50.8)	128.0 (65.0)	< 0.001
Post-operative in-hospital complications			
Myocardial infarction	520 (0.3)	165 (2.1)	< 0.001
Dysrhythmia	2,392 (1.5)	445 (5.7)	< 0.001
Congestive heart failure exacerbation	559 (0.4)	195 (2.5)	< 0.001
Cranial nerve injury	4,298 (2.7)	4379 (4.9)	< 0.001
Reperfusion syndrome	128 (0.08)	144 (1.9)	< 0.001
Surgical site infection	79 (0.05)	21 (0.3)	< 0.001
Length of stay, days, median (IQR)	1 (1-2)	3 (1-8)	< 0.001

Values are reported as No. (%) unless otherwise indicated.
Abbreviations: BMI (body mass index), ACE-I (angiotensin converting enzyme inhibitor), ARB (angiotensin II receptor blocker), ASA (American Association of Anesthesiologists), EEG (electroencephalography), SD (standard deviation), IQR (interquartile range).

Table 2. Performance of machine learning models in predicting 1-year stroke or death following carotid endarterectomy on test set data at the pre-operative stage

	AUROC (95% CI)	**Accuracy (95% CI)**	**Sensitivity**	**Specificity**	**PPV**	**NPV**
XGBoost	0.90 (0.89 – 0.91)	0.82 (0.81 – 0.83)	0.82	0.83	0.83	0.81
Random forest	0.87 (0.86 – 0.88)	0.79 (0.78 – 0.81)	0.80	0.78	0.77	0.81
Naïve Bayes	0.81 (0.80 – 0.83)	0.81 (0.80 – 0.83)	0.80	0.82	0.82	0.79
Support vector machine	0.79 (0.77 – 0.80)	0.72 (0.70 – 0.74)	0.71	0.73	0.73	0.71
Artificial neural network	0.74 (0.73 – 0.77)	0.71 (0.69 – 0.73)	0.70	0.72	0.75	0.68
Logistic regression	0.65 (0.63 – 0.67)	0.62 (0.60 – 0.63)	0.61	0.62	0.65	0.58

Abbreviations: XGBoost (Extreme Gradient Boosting), AUROC (area under the receiver operating characteristic curve), CI (confidence interval), PPV (positive predictive value), NPV (negative predictive value).

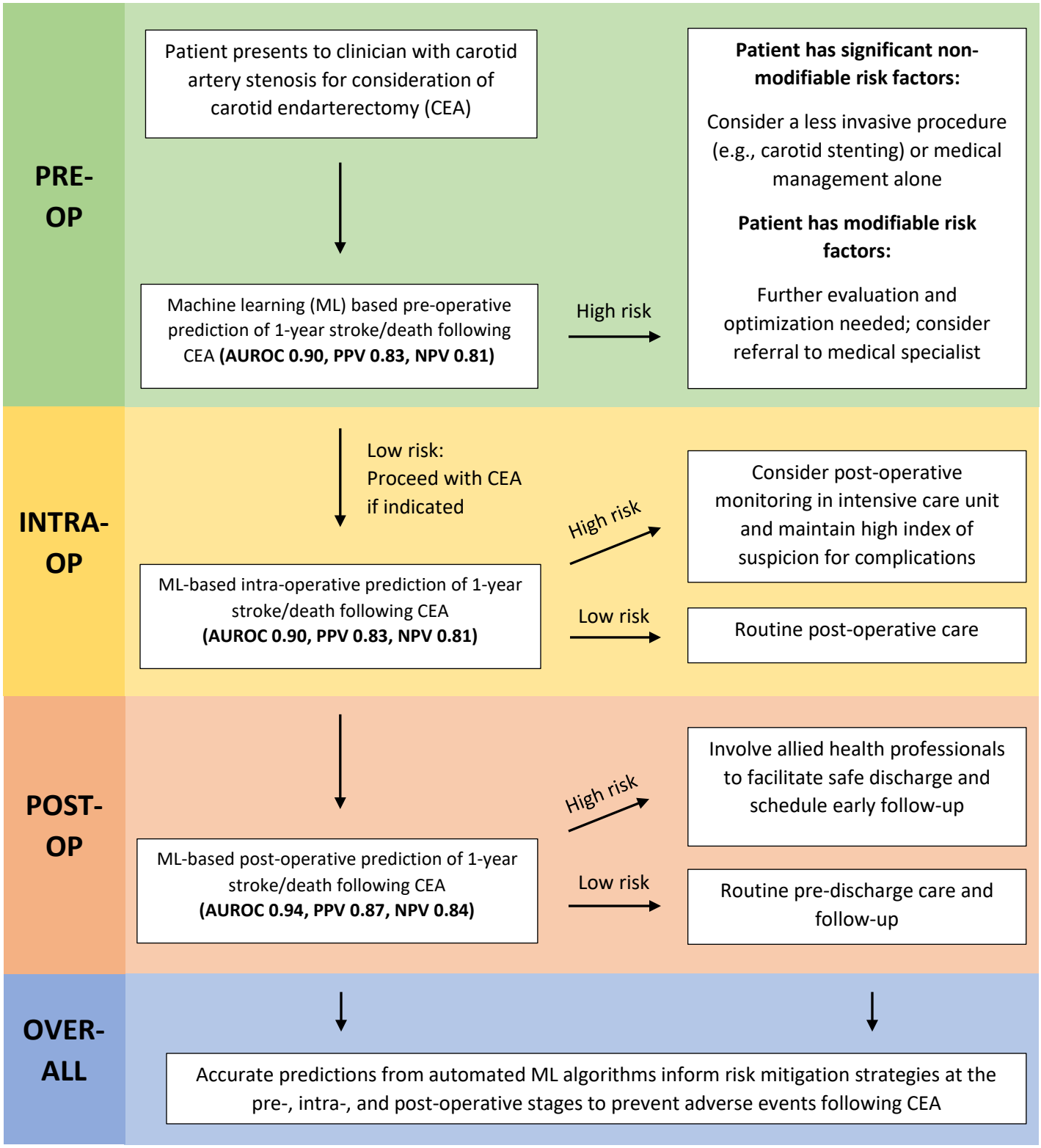


Figure 1. Clinical workflow for the use of our machine learning algorithm to guide clinical decision-making at the pre-, intra-, and post-operative stages for patients being considered

for carotid endarterectomy. Abbreviations: AUROC (area under the receiver operating characteristic curve), PPV (positive predictive value), NPV (negative predictive value). High risk defined as a model prediction positive for 1-year stroke/death. Low risk defined as a model prediction negative for 1-year stroke/death.

ACKNOWLEDGEMENTS

None.

SOURCES OF FUNDING

This research was funded by the Canadian Institutes of Health Research, Ontario Ministry of Health, and PSI Foundation (BL). The funding sources did not play a role in the design or conduct of the research.

DISCLOSURES

The authors have no conflicts of interest.

CODE AVAILABILITY STATEMENT

The complete code used for model development and evaluation in this project is publicly available on GitHub: https://github.com/benli12345/CEA-ML-VQI.

DATA AVAILABILITY STATEMENT

The data used for this study comes from the Vascular Quality Initiative Database, which is maintained by the Society for Vascular Surgery Patient Safety Organization. Access and use of the data requires approval through an application process available at https://www.vqi.org/data-analysis/.

FULL-TEXT PUBLICATION

Li B, Beaton D, Eisenberg N, Lee DS, Wijeysundera DN, Lindsay TF, de Mestral C, Mamdani M, Roche-Nagle G, Al-Omran M. Using machine learning to predict outcomes following carotid endarterectomy. Journal of Vascular Surgery. 2023 Oct 1;78(4):973-987.e6. https://pubmed.ncbi.nlm.nih.gov/37211142/

REFERENCES

1. Qaja E, Tadi P, Theetha Kariyanna P. Carotid Artery Stenosis. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 [cited 2022 Oct 5]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK442025/

2. Donkor ES. Stroke in the 21st Century: A Snapshot of the Burden, Epidemiology, and Quality of Life. Stroke Res Treat. 2018 Nov 27;2018:3238165.

3. Rerkasem K, Rothwell PM. Systematic review of the operative risks of carotid endarterectomy for recently symptomatic stenosis in relation to the timing of surgery. Stroke. 2009 Oct;40(10):e564-572.

4. AbuRahma AF, Avgerinos ED, Chang RW, Darling RC, Duncan AA, Forbes TL, et al. Society for Vascular Surgery clinical practice guidelines for management of extracranial cerebrovascular disease. J Vasc Surg. 2022 Jan;75(1S):4S-22S.

5. Volkers EJ, Algra A, Kappelle LJ, Greving JP. Prediction models for clinical outcome after a carotid revascularisation procedure: A systematic review. Eur Stroke J. 2018 Mar;3(1):57–65.

6. Sharma V, Ali I, van der Veer S, Martin G, Ainsworth J, Augustine T. Adoption of clinical risk prediction tools is limited by a lack of integration with electronic health records. BMJ Health Care Inform. 2021 Feb 19;28(1):e100253.

7. Baştanlar Y, Özuysal M. Introduction to machine learning. Methods Mol Biol. 2014;1107:105–28.

8. Shah P, Kendall F, Khozin S, Goosen R, Hu J, Laramie J, et al. Artificial intelligence and machine learning in clinical development: a translational perspective. NPJ Digit Med. 2019;2:69.

9. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015 Jan 6;162(1):55–63.

10. Society for Vascular Surgery Vascular Quality Initiative (VQI) [Internet]. [cited 2022 Jul 11]. Available from: https://www.vqi.org/

11. Vascular Quality Initiative [Internet]. [cited 2022 Jun 16]. Available from: https://www.vqi.org/

12. Cronenwett JL, Kraiss LW, Cambria RP. The Society for Vascular Surgery Vascular Quality Initiative. J Vasc Surg. 2012 May;55(5):1529–37.

13. Tadi P, Lui F. Acute Stroke. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 [cited 2022 Jul 1]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK535369/

14. Elfanagely O, Toyoda Y, Othman S, Mellia JA, Basta M, Liu T, et al. Machine Learning and Surgical Outcomes Prediction: A Systematic Review. J Surg Res. 2021 Apr 10;264:346–61.

15. Bektaş M, Tuynman JB, Costa Pereira J, Burchell GL, van der Peet DL. Machine Learning Algorithms for Predicting Surgical Outcomes after Colorectal Surgery: A Systematic Review. World J Surg. 2022 Sep 15;

16. Senders JT, Staples PC, Karhade AV, Zaki MM, Gormley WB, Broekman MLD, et al. Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. World Neurosurg. 2018 Jan;109:476-486.e1.

17. Dobbin KK, Simon RM. Optimally splitting cases for training and testing high dimensional classifiers. BMC Med Genomics. 2011 Apr 8;4:31.

18. Jung Y, Hu J. A K-fold Averaging Cross-validation Procedure. J Nonparametric Stat. 2015;27(2):167–79.

19. Adnan M, Alarood AAS, Uddin MI, Ur Rehman I. Utilizing grid search cross-validation with adaptive boosting for augmenting performance of machine learning models. PeerJ Comput Sci. 2022;8:e803.

20. Wibowo P, Fatichah C. Pruning-based oversampling technique with smoothed bootstrap resampling for imbalanced clinical dataset of Covid-19. J King Saud Univ - Comput Inf Sci. 2022 Oct;34(9):7830–9.

21. Lee SG, Russ A. Predicting and Preventing Postoperative Outcomes. Clin Colon Rectal Surg. 2019 May;32(3):149–56.

22. Ruan X, Fu S, Storlie CB, Mathis KL, Larson DW, Liu H. Real-time risk prediction of colorectal surgery-related post-surgical complications using GRU-D model. J Biomed Inform. 2022 Sep 24;135:104202.

23. Redelmeier DA, Bloch DA, Hickam DH. Assessing predictive accuracy: how to compare Brier scores. J Clin Epidemiol. 1991;44(11):1141–6.

24. Loh WY, Zhou P. Variable Importance Scores. J Data Sci. 2021 Sep 16;19(4):569–92.

25. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020 Mar 18;m441.

26. Schafer JL. Multiple imputation: a primer. Stat Methods Med Res. 1999 Mar;8(1):3–15.

27. Download R-4.2.1 for Windows. The R-project for statistical computing. [Internet]. [cited 2022 Oct 6]. Available from: https://cran.r-project.org/bin/windows/base/

DEPRESSIVE-LIKE PHENOTYPE INDUCED BY AAV-MEDIATED OVEREXPRESSION OF HUMAN A-SYNUCLEIN IN MIDBRAIN DOPAMINERGIC NEURONS

Laura Kondrataviciute,[1,2,9] Ignas Kondratavicius,[3] Minesh Kapadia,[2] Hien Chau,[2] Jimmy George,[2] Erdost Yildiz,9 Taufik Valiante,[1,2,4,5] Luka Milosevic,[1,2,4,5] Lorraine V. Kalia,[2,5,6,7] Suneil K. Kalia[4,9]

[1]Institute of Biomedical Engineering, University of Toronto, Canada
[2]Krembil Research Institute, Toronto Western Hospital, University Health Network, Toronto, ON, Canada
[3]Kauno Technologijos Universitetas, Kaunas, Lithuania
[4]KITE, Toronto Rehabilitation Institute, University Health Network,
[5]CRANIA, University Health Network and University of Toronto, Toronto, Canada
[6]Division of Neurology, Department of Medicine, University of Toronto, Toronto, ON, Canada;
[7]Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, ON, Canada
[8]Division of Neurosurgery, Department of Surgery, University of Toronto, Canada
[9]Physical Intelligence Department, Max Planck Institute for Intelligent Systems, 70569, Stuttgart, Germany

INTRODUCTION

Parkinson's disease (PD), as the most prevalent neurodegenerative motor disorder, presents a complex spectrum of symptoms encompassing both motor impairments (bradykinesia, resting tremor, freezing of gait) and non-motor manifestations (depression, anosmia, anxiety). Robust animal models that faithfully replicate this multifaceted pathology are indispensable for comprehensive PD research. This study aims to investigate if depressive-like behaviour, prominent in up to 40% of PD patients (Laux, 2022), emerges in the bilateral human mutated alpha-synuclein (A53T) rat model of PD.

METHODS

Study Design

All procedures were conducted under guidelines and regulations set by the Canadian Council on Animal Care (AUP #3818). Thirty-nine adult female Sprague-Dawley rats bilaterally injected into substantia nigra (SN) with either empty AAV1/2 vector (EV) or AAV1/2-expressing human mutated A53T-alpha synuclein (A53T) were used in the study. Baseline sucrose preference and novelty suppressed feeding tests were conducted to divide animals into groups with similar means and standard deviations. The same behaviour tests were repeated on weeks 3 and 6 at the same time on subsequent days. Post-mortem immunofluorescence staining for TH+, alpha-synuclein, DAPI was done to confirm neurodegeneration on week 6.

Surgical Procedure

The surgical procedure was adapted from a protocol described in previous literature (Koprich et al., 2010)(Nim et al., 2023). The surgical procedure was performed under isofluorane/oxygen anesthesia and ketoprofen (5 mg/kg) analgesia. Adult rats with an average weight of 270 g were stereotaxically bilaterally injected into SN with either 2.0 μL of empty AAV1/2 vector (EV) or 2.0 μL of AAV1/2-expressing human mutated A53T-aSyn (A53T) at the concentration of $3.4x10^{12}$ genomic particles (gp)/mL using coordinates from bregma: –5.2 mm anteroposterior (AP); either –2.0 mm (right) or +2.0 mm (left) mediolateral (ML); –7.5 mm dorsoventral (DV) (rat brain atlas of (Paxinos & Watson, 2013)). The rat's body temperature was stabilized at 37°C – 38°C throughout the surgery with a heating pad. Daily monitoring was performed one week after surgery.

Sucrose Preference Test

The sucrose preference test was used to evaluate depressive-like behavior according to a modified protocol (Liu et al., 2018). Before testing, rats underwent a habituation protocol during which they were water deprived for 1-hour and then placed in a standard cage with two bottles,

one containing water and the other containing a 20% sucrose solution, for 1hr. On test days, two bottles were given to the rats, one containing water and the other the 20% sucrose solution, with their weights measured before and after the 1-hour test period. Sucrose preference was calculated by dividing the amount of sucrose solution consumed by the total fluid intake from both bottles.

Novelty Suppressed Feeding Test

The modified novelty-suppressed feeding protocol was conducted to evaluate anxiogenic-like behavior (Blasco-Serra et al., 2017) and responsiveness to palatable stimulation. The test was performed in a cage (26cm W x 48cm D x 20cm H) with an illuminated dish containing 20 crisps. Rats were subjected to 12hr food deprivation before the test to enhance their motivation for food. The latency to feeding, defined as the time needed to start eating the first crisp, was manually recorded using a stopwatch. Additionally, the number of crisps consumed within 10 minutes was measured.

Statistical analysis

Data in Fig. 1 is expressed as the mean ± standard error of the mean (SEM) and were analyzed with JMP Pro v16 and GraphPad Prism software. The normality of the data was assessed using the Shapiro-Wilk W test. If the p-value from the Shapiro-Wilk W test was greater than 0.05, indicating normal distribution, two-way ANOVA was used to detect significant effects. If significant effects were detected, pairwise analyses were performed using t-test. When data was not normally distributed, non-parametric Wilcoxon matched-pairs signed rank test was performed.

RESULTS

The brief sucrose preference test was used to assess responsiveness to palatable stimulation (Scheggi et al., 2018). Both groups of rats showed a preference for the 20% sucrose solution. Comparison of A53T expressing and control group 3 weeks post injection revealed comparable preference for the sweetened solution ($p=0.313$). However, significantly lower preference for sucrose was found in alpha-synuclein expressing animals on week 6 compared to control group ($p=0.0099$). Results suggest a significantly higher preference for sucrose on week 6 compared to week 3 for EV group ($p=0.0004$) and no difference between timepoints for A53T group ($p=0.611$). No difference between timepoints was seen in total fluid consumption for both groups.

Anxiogenic-like behavior in the AAV-A53T rat model of PD was assessed through the latency to eating in novelty suppressed feeding test (Francois et al., 2022). No significantly elevated latency to eating in A53T animals was seen on any timepoint, suggesting no intensified anxiety in alpha-synuclein expressing animals ($p=0.515$ and $p=0.123$ on week 3 and week 6, respectively). In agreement with sucrose preference test findings, reduced responsiveness to highly palatable food consumption was observed in novelty suppressed feeding test, but statistically significant difference between A53T and EV groups was not reached at any timepoint. Study design and results are summarised in Fig. 1.

CONCLUSIONS

Our findings suggest that the A53T-alpha synuclein rat model manifests depressive-like behavior, evidenced by diminished responsiveness to palatable stimuli. This model holds promise for investigating non-motor pathologies associated with PD.

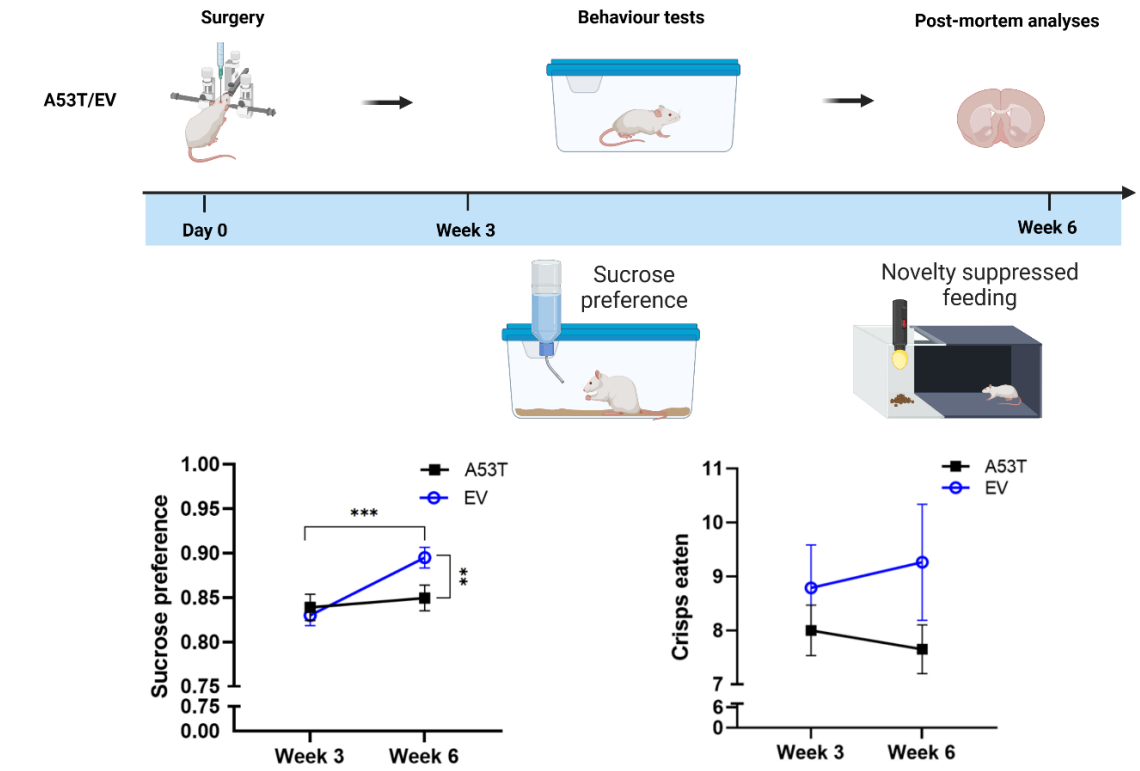


Fig. 1. Study design and results

REFERENCES

Blasco-Serra, A., González-Soler, E. M., Cervera-Ferri, A., Teruel-Martí, V., & Valverde-Navarro, A. A. (2017). A standardization of the Novelty-Suppressed Feeding Test protocol in rats. *Neuroscience Letters*, *658*, 73–78. https://doi.org/10.1016/j.neulet.2017.08.019

Francois, M., Canal Delgado, I., Shargorodsky, N., Leu, C.-S., & Zeltser, L. (2022). Assessing the effects of stress on feeding behaviors in laboratory mice. *ELife*, *11*. https://doi.org/10.7554/eLife.70271

Koprich, J. B., Johnston, T. H., Reyes, M. G., Sun, X., & Brotchie, J. M. (2010). Expression of human A53T alpha-synuclein in the rat substantia nigra using a novel AAV1/2 vector produces a rapidly evolving pathology with protein aggregation, dystrophic neurite architecture and nigrostriatal degeneration with potential to model the pathology of Parkinson's disease. *Molecular Neurodegeneration*, *5*(1), 43. https://doi.org/10.1186/1750-1326-5-43

Laux, G. (2022). Parkinson and depression: review and outlook. *Journal of Neural Transmission*, *129*(5–6), 601–608. https://doi.org/10.1007/s00702-021-02456-3

Liu, M.-Y., Yin, C.-Y., Zhu, L.-J., Zhu, X.-H., Xu, C., Luo, C.-X., Chen, H., Zhu, D.-Y., & Zhou, Q.-G. (2018). Sucrose preference test for measurement of stress-induced anhedonia in mice. *Nature Protocols*, *13*(7), 1686–1698. https://doi.org/10.1038/s41596-018-0011-z

Nim, S., O'Hara, D. M., Corbi-Verge, C., Perez-Riba, A., Fujisawa, K., Kapadia, M., Chau, H., Albanese, F., Pawar, G., De Snoo, M. L., Ngana, S. G., Kim, J., El-Agnaf, O. M. A., Rennella, E., Kay, L. E., Kalia, S. K., Kalia, L. V., & Kim, P. M. (2023). Disrupting the α-synuclein-ESCRT interaction with a peptide inhibitor mitigates neurodegeneration in preclinical models of Parkinson's disease. *Nature Communications*, *14*(1), 2150. https://doi.org/10.1038/s41467-023-37464-2

Paxinos, G., & Watson, C. (2013). *The Rat Brain in Stereotaxic Coordinates* (7th ed.).

Scheggi, S., De Montis, M. G., & Gambarana, C. (2018). Making Sense of Rodent Models of Anhedonia. *International Journal of Neuropsychopharmacology*, *21*(11), 1049–1065. https://doi.org/10.1093/ijnp/pyy083

IMPROVING ELECTIVE ARTHROPLASTY PROCEDURE SCHEDULING WITH MACHINE LEARNING PREDICTION AND OPTIMIZATION

Johnathan R. Lex (SSTP)[1,2], Aazad Abbas[1,2], Jay Toor[3], Elias B. Khalil[4], Bheeshma Ravi[1,5], Cari Whyne[1,2,5]

[1]Division of Orthopaedic Surgery, University of Toronto, 149 College Street Room 508-A, Toronto, ON, M5T 1P5, Canada.

[2]Orthopaedic Biomechanics Lab, Sunnybrook Research Institute, 2075 Bayview Avenue, Toronto, ON, M4N 3M5, Canada.

[3]Department of Orthopaedic Surgery, University of Manitoba, Winnipeg, MB, Canada.

[4]Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, ON, M5S 3G8, Canada.

[5]Division of Orthopaedic Surgery, Sunnybrook Health Science Centre, 2075 Bayview Avenue, Toronto, ON, M4N 3M5, Canada.

INTRODUCTION

Demand for primary and revision total hip and knee arthroplasty (THA and TKA, respectively) is increasing with a commensurate demand for improved access, efficiency, and value of care.[1-3] This has led to growing research and development of new models of care for arthroplasty delivery including centralization, transitioning to outpatient delivery, and bundled payment models.[4-8]

Operating room (OR) time is costly and is often the main constraint on surgeons to work through their patient waiting list.[9, 10] Despite this, poor estimations of duration of surgery (DOS) leads to a large proportion of OR time going underutilized or requiring overtime, associated with greater expenses due to supplemental personnel pay.[11] Despite the association of total joint arthroplasty DOS with several factors such as patient sex, age, body mass index, American Society of Anaesthesiologist (ASA) score and surgeon experience, few efforts have been made to use this information in order to optimize surgical scheduling.[12-15] Most institutions rely on a surgery-specific or surgeon-surgery-specific rolling average to generate OR schedules from manually selected cases.

More accurate estimates of operative time can lead to improved operating room utilization by allowing more precise surgical scheduling.[11, 16] The appropriate preoperative allocation of operative time is closely linked to OR efficiency, including greater total OR utilization, overtime and underutilization.[17-19] To date, research utilizing machine learning (ML) to predict DOS and operations research attempting to optimize elective surgery scheduling has only been performed in isolation.[20-23] Advances in ML techniques and computational power has now offered the opportunity to combine these techniques to predict and optimize surgical scheduling. The aim of this study was to evaluate the efficiency of a surgical schedule generated using ML and optimization to current practice (surgeon-surgery-specific rolling mean).

METHODS

Setting and Data

Ethics for this study was approved by local research ethics board review (project identification number: 5306). All primary and revision TKA and THA cases completed between April 2012 to February 2022 at the largest volume arthroplasty centre in Canada were included. Data was retrospectively analyzed from two prospectively maintained databases, an OR information management database and an arthroplasty specific database. Bilateral and all revision arthroplasty cases were included.

Model Development

ML models for each of the four procedures (primary TKA, revision TKA, primary THA, revision THA) were developed. Dataset features (predictors) were analyzed and visualized to identify the distribution and proportion of missing values. Features with greater than 15% missing data were excluded. Operative time was defined as patient total time in the operating room in minutes. As models were developed to inform future scheduling, data was split based on year of surgery to mimic forecasting onto unseen, future data. Surgeries performed in years 2012 to 2020 (n=13508, 88.5%) were used for training and validation and 2021 to 2022 for testing (n=1759, 11.5%). Data was trained and validated using five-fold cross-validation and tested on the operative times from the out-of-sample test set. Overall study design is depicted in Figure 1.

Feature selection was performed using recursive feature elimination with cross-validation using a linear regression estimator. The most important features for each procedure were then selected to train and tune the models. The ML models evaluated were linear regression with lasso regularization, AdaBoost and a neural network (multilayer perceptron). These three machine learning models and one “dummy” regressor model, using the surgery-specific mean

operative time, were then compared. Models were trained to optimize the score of the mean squared error (MSE). Other scoring metrics included the R^2 value, 15-minute buffer accuracy, and classification accuracy (<60, 60-90 or >90 minutes). The latter two scoring metrics involve regression prediction then binning, as previously defined.[24, 25] This improves interpretability of the results without binning *a priori,* therefore, maintaining the ability to predict a continuous variable using regression rather than classification.

Optimized Schedule Generation

Following model development, the realizable effect that improved operative time prediction can have on operating room utilization was simulated. Three schedules were created using different predicted operative times: 1) a “smart” schedule using ML-predicted times, 2) historic scheduling practice using surgeon-surgery-specific rolling mean times, and 3) surgery-specific mean times. The predictive ML model that performed the best was selected to inform the inputted operative times of the “smart” schedule. The schedules using the various predicted times were then evaluated based on the schedule that would have occurred in that OR based on the known actual operative times. Schedules were evaluated based on OR overtime, underutilization (idle time), and cases completed (Figure 1).

Weekly daily surgical schedules were then simulated over 80 weeks (20 months). Cases were sampled from the test subset. Based on the comparison of previously solved optimization problems in the literature, a mixed integer linear programming, specifically the multi-subset sum problem (MSSP), was chosen. This formulation optimizes a single surgeon’s waitlist for each of their assigned OR days. It maximizes the overall utilization of an operating room with a penalty for overtime. Notably, it does not aim to maximize the total throughput as this generates bias towards selecting cases of shorter duration.[26]

The following constraints and times on the optimization problem were reproduced from those at the local institution. Cases were performed by nine surgeons and assigned to the surgeon that performed the case. There were three operating rooms designated to arthroplasty each day from Monday to Friday. ORs were scheduled from 08:00 to 17:00 with a penalty incurred for overtime. An average OR turnover time of 30 minutes was scheduled between each case. Scheduling was performed with a granularity of 10-minute blocks rounded up to the nearest block. A waitlist size of 500 randomly selected cases from the test subset were chosen for each week. Revision TKA and THA surgeries were allocated higher priority and must have been completed within the planning horizon (one week).

Statistical Analysis

Statistical analysis, machine learning and optimization were all performed using Python (Python Software Foundation, version 3.9). Gurobi optimization package was installed to generate schedule optimization problems.[27] Models were compared and selected for based on the lowest average (median) MSE score on the test set across all four procedures. Most important features were determined based on the highest impurity-based feature importance scores. The overtime, underutilization and completed cases of the mean and rolling average schedules were independently compared to the ML-predicted schedule using Wilcoxon signed-rank test. Threshold for statistical significance was set at the p-value of 0.05.

RESULTS

There were 13529 (88.7%) primary TKA and THA procedures and 1728 (11.3%) revision procedures. There were 5799 male (38.0%) and 9468 females (62.0%) included. For TKA and THA, revision procedures were longer and had a wider standard deviation of time than their respective primary counterparts ($p<0.001$).

The accuracy of the historic modality (surgeon-surgery-specific rolling mean) to predict operative time was evaluated. Retrospectively evaluation revealed fair performance, with a median MSE across all four procedures of 909.7 and within a 15-minute buffer accuracy in 52.7% of cases.

Model Performance

The performance of all models is presented in Table 1. The lasso regression, AdaBoost and neural network model outperformed the historic modality of operative time prediction. In terms of MSE value, the neural network model outperformed the historic mean operative time, lasso regression and AdaBoost model when averaged across all procedures. The neural network model also generated the best median R^2 value across procedures (0.62) compared to lasso regression and AdaBoost models.

The neural network buffer accuracy was an improvement compared to the historic rolling mean time by 7.1% (59.8% vs 52.7%). The surgery-specific mean time is often still used at hospitals and is currently used at the local institution until a surgeon has performed more than 10 of a specific case, which is when the surgeon-surgery-specific rolling mean is started. This mean time was the worst predictor of operative times across all procedures and scoring metrics.

Schedule Comparison

The median MSE was the lowest using the neural network model, therefore, those models were selected to predict the operative times of the “smart” schedule.

The “smart” schedule using the ML-predicted operative times had the least amount of OR underutilization and the greatest number of completed cases. The mean weekly OR underutilization was 155.4 minutes (SD: 61.7) using ML-predicted times compared to 354.6 (SD: 108.2) and 687.7 (SD: 156.9) when using the historic practice of rolling and surgery-specific means, respectively. This represents a 56.2% and 77.4% reduction in underutilization. Over the three operating rooms over five days this equates to 10.4 minutes of idle operating room time per OR day. Compared to historic scheduling practices this is an improvement of 13.3 minutes per OR day and 35.3 minutes to using the surgery-specific mean operative time (Table 2).

The increase in OR utilization with the “smart” schedule was associated with an increased amount of overtime requirement; 3.6 and 3.8 minutes more per OR day compared to using the historic rolling and surgery-specific mean operative times. A 17.2% and 17.9% increase in overtime, respectively. The decreased OR underutilization with the “smart” schedule conferred an increase in the average number of completed cases by 1.6 and 3.1 cases per week compared to using historic rolling mean and surgery-specific mean operative times (Table 2).

To ensure the “smart” scheduling system was not biased towards selecting cases of shorter or longer duration, the proportion of cases scheduled that were greater than and less than 1 standard deviation from the mean were counted. Compared to using the historic surgeon-surgery-specific rolling mean operative times, the “smart” schedule selected more cases at both extremes, 1.4% more shorter cases ($p=0.069$) and 3.5% more longer cases ($p<0.0001$).

CONCLUSIONS

The current approach is novel, combining both improvement in prediction and optimization to quantify the realizable improvements on OR utilization for primary and revision TKA and THA. This is the first step towards automation and improvement of scheduling practices from current, largely manual, practice. The benefits have significant implications for patients, hospitals, and healthcare payers to maximize OR utilization and patient throughput without expanding resources. Future work of prospectively evaluating and then implementing this system is required to evaluate the true accuracy, impact, and fairness of these models.

Figures and Tables:

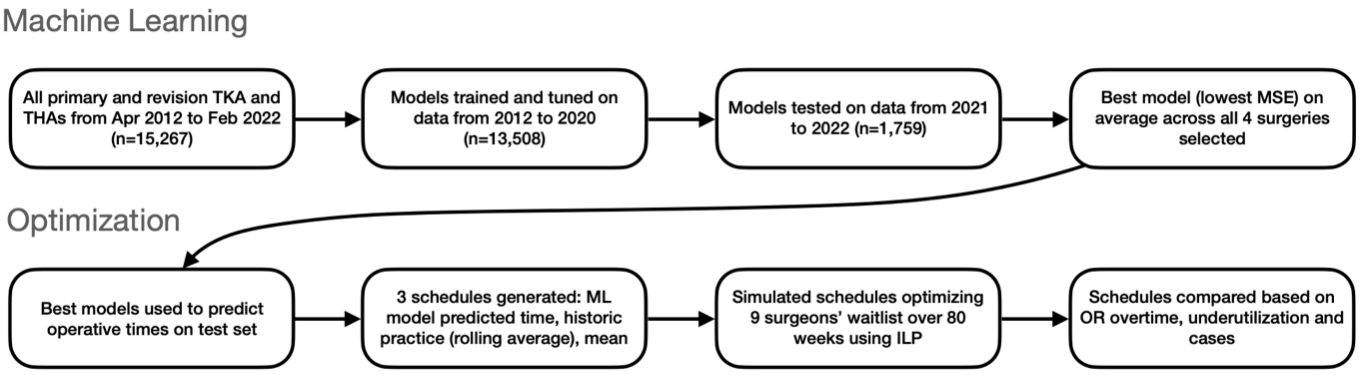


Figure 1. Flowchart of the overall study design.

TKA = total knee arthroplasty; THA = total hip arthroplasty; MSE = mean squared error; ML = machine learning; ILP = integer linear programming; OR = operating room

Model	Scoring metric	Primary TKA	Revision TKA	Primary THA	Revision THA	Median Score
Mean	MSE	407.7	2687.3	512.4	4817.1	1599.9
	R^2	-0.02	0.00	0.00	-0.02	-0.01
	15-minute BA (%)	65.1	26.9	49.7	12.0	38.3
	<60, 60-90, >90 CA (%)	55.7	89.2	52.0	92.8	72.5
Lasso Regression	MSE	224.2	1023.3	214.4	1394.8	623.8
	R^2	0.44	0.62	0.58	0.70	0.60
	15-minute BA (%)	79.3	51.5	74.7	26.5	63.1
	<60, 60-90, >90 CA (%)	77.0	90.8	78.5	87.9	83.2
AdaBoost	MSE	247.8	950.9	248.3	2454.8	599.6
	R^2	0.38	0.65	0.51	0.48	0.50
	15-minute BA (%)	77.1	46.9	74.2	32.5	60.6
	<60, 60-90, >90 CA (%)	74.6	93.1	75.2	92.8	84.0
Neural Network	MSE	225.4	963.8	202.9	1403.7	594.6
	R^2	0.44	0.64	0.60	0.70	0.62
	15-minute BA (%)	78.9	43.1	76.4	30.1	59.8
	<60, 60-90, >90 CA (%)	75.9	91.5	77.2	87.9	82.6

Table 1. Model performance on the test subset of data.

TKA = total knee arthroplasty; THA = total hip arthroplasty; MSE = mean squared error; BA = buffer accuracy; CA = classification accuracy

	“Smart” ML-predicted schedule	Surgeon-surgery-specific rolling mean schedule	P-value	Surgery-specific mean schedule	P-value
Overtime, mean (SD)	315.4 (111.2)	261.0 (109.6)	0.0001	258.9 (130.2)	0.0030
Underutilization, mean (SD)	155.4 (61.7)	354.6 (108.2)	<0.0001	687.7 (156.9)	<0.0001
No. Cases, mean (SD)	60.9 (1.8)	59.3 (1.6)	<0.0001	57.8 (0.8)	<0.0001

Table 2. Mean weekly surgical schedule outcome metrics following 80 weeks of simulated schedule generation using the three different modalities to predict operative time.

ML = machine learning; SD = standard deviation

References:

1. Sloan M, Premkumar A, Sheth N, Projected Volume of Primary Total Joint Arthroplasty in the U.S., 2014 to 2030. 2018; 100:17

2. Singh JA, Yu S, Chen L, Cleveland JD, Rates of Total Joint Replacement in the United States: Future Projections to 2020–2040 Using the National Inpatient Sample. 2019; 46:9

3. Culliford D, Maskell J, Judge A, Cooper C, Prieto-Alhambra D, Arden NK, Future projections of total hip and knee arthroplasty in the UK: results from the UK Clinical Practice Research Datalink. Osteoarthritis and Cartilage 2015; 23:4

4. Marlow NE, Barraclough B, Collier NA, Dickinson IC, Fawcett J, Graham JC, Maddern GJ: Centralization and the relationship between volume and outcome in knee arthroplasty procedures. Centre for Reviews and Dissemination (UK), 2010

5. Ramirez G, Myers TG, Thirukumaran CP, Ricciardi BF, Does Hypothetical Centralization of Revision THA and TKA Exacerbate Existing Geographic or Demographic Disparities in Access to Care by Increased Patient Travel Distances or Times? A Large-database Study. 2022; 480:6

6. Pincus D, Jenkinson R, Paterson M, Leroux T, Ravi B, Association Between Surgical Approach and Major Surgical Complications in Patients Undergoing Total Hip Arthroplasty. JAMA 2020; 323:11

7. Rozell JC, Ast MP, Jiranek WA, Kim RH, Della Valle CJ, Outpatient Total Joint Arthroplasty: The New Reality. J Arthroplasty 2021; 36:7S

8. Finch DJ, Pellegrini VD, Franklin PD, Magder LS, Pelt CE, Martin BI, The Effects of Bundled Payment Programs for Hip and Knee Arthroplasty on Patient-Reported Outcomes. J Arthroplasty 2020; 35:4

9. Childers CP, Maggard-Gibbons M, Understanding Costs of Care in the Operating Room. JAMA Surg 2018; 153:4

10. Moody A, Gurnea T, Shul C, Althausen P, True Cost of Operating Room Time: Implications for an Orthopaedic Trauma Service. 2020; 34:5

11. Hans E, Wullink G, van Houdenhoven M, Kazemier G, Robust surgery loading. European Journal of Operational Research 2008; 185:3

12. Cowley RJ, Frampton C, Young SW, Operating time for total knee arthroplasty in public versus private sectors: where does the efficiency lie? ANZ J Surg 2019; 89:1-2

13. Duchman KR, Pugely AJ, Martin CT, Gao Y, Bedard NA, Callaghan JJ, Operative Time Affects Short-Term Complications in Total Joint Arthroplasty. J Arthroplasty 2017; 32:4

14. Kosashvili Y, Mayne IP, Trajkovski T, Lackstein D, Safir O, Backstein D, Influence of sex on surgical time in primary total knee arthroplasty. Can J Surg 2010; 53:4

15. Lau RL, Perruccio AV, Gandhi R, Mahomed NN, The role of surgeon volume on patient outcome in total knee arthroplasty: a systematic review of the literature. BMC Musculoskelet Disord 2012; 13:

16. Lehtonen J, Torkki P, Peltokorpi A, Moilanen T, Increasing operating room productivity by duration categories and a newsvendor model. 2013; 26:2

17. Strum DP, Vargas LG, May JH, Surgical subspecialty block utilization and capacity planning: a minimal cost analysis model. Anesthesiology 1999; 90:4

18. Dexter F, Traub RD, Macario A, How to release allocated operating room time to increase efficiency: predicting which surgical service will have the most underutilized operating room time. Anesth Analg 2003; 96:2

19. Dexter F, Epstein RH, Operating room efficiency and scheduling. Curr Opin Anaesthesiol 2005; 18:2

20. Yeo I, Klemt C, Melnic CM, Pattavina MH, De Oliveira BMC, Kwon Y, Predicting surgical operative time in primary total knee arthroplasty utilizing machine learning models. 2022

21. Bartek MA, Saxena RC, Solomon S, Fong CT, Behara LD, Venigandla R, Velagapudi K, Nair BG, Lang JD, Improving Operating Room Efficiency: A machine learning approach to predict case-time duration. J Am Coll Surg 2019; 229:4

22. Huang CC, Lai J, Cho D, Yu J, A Machine Learning Study to Improve Surgical Case Duration Prediction. 2020

23. Entezari B, Koucheki R, Abbas A, Toor J, Wolfstadt JI, Ravi B, Whyne C, Lex JR, Improving Resource Utilization for Arthroplasty Care by Leveraging Machine Learning and Optimization: A Systematic Review. Arthroplast Today 2023; 20:

24. Abbas A, Mosseri J, Lex JR, Toor J, Ravi B, Khalil EB, Whyne C, Machine learning using preoperative patient factors can predict duration of surgery and length of stay for total knee arthroplasty. Int J Med Inform 2022; 158:

25. Witvoet S, de Massari D, Shi S, Chen AF, Leveraging large, real-world data through machine-learning to increase efficiency in robotic-assisted total knee arthroplasty. 2023

26. Guido R, Conforti D, A hybrid genetic approach for solving an integrated multi-objective operating room planning and scheduling problem. 2017; 87:

27. Gurobi Optimization L, Gurobi Optimizer Reference Manual. 2023