A High-Tech Solution for the Low Resource Setting:

A Tool to Support Decision Making for Patients with Traumatic Brain Injury

by

Cyrus Elahi

Duke Global Health Institute
Duke University

Date:_____
Approved:

_____
Michael Haglund, Supervisor

_____
Joao Vissoci

_____
Catherine Staton

_____
Eric Green

Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in the Duke Global Health Institute
in the Graduate School of Duke University

2019

ABSTRACT

A High-Tech Solution for the Low Resource Setting:

A Tool to Support Decision Making for Patients with Traumatic Brain Injury

by

Cyrus Elahi

Duke Global Health Institute
Duke University

Date:_____

Approved:

_____
Michael Haglund, Supervisor

_____
Joao Vissoci

_____
Catherine Staton

_____
Eric Green

An abstract of a thesis submitted in partial
fulfillment of the requirements for the degree
of Master of Science in the Duke Global Health Institute
in the Graduate School of Duke University

2019

Abstract

Background. The confluence of a capacity-exceeding disease burden and persistent resource shortages have resulted in traumatic brain injury's (TBI) devastating impact in low- and middle-income countries (LMIC). Lifesaving care for TBI depends on accurate and timely decision making within the hospital. As result of technology and highly skilled provider shortages, treatment delays are common in low resource settings. This reality demands a low cost, scalable and accurate alternative to support decision making. Decision support tools leveraging the accuracy of modern prognostic modeling techniques represents one possible solution. This thesis is a collation of research dedicated to the advancement of TBI decision support technology in low resource settings. Methods. The study location included three national and referral hospitals in Uganda and Tanzania. We performed a survival analysis, externally validated existing TBI prognostic models, developed our own prognostic model, and performed a feasibility study for TBI decision support tools in an LMIC. Results. The survival analysis revealed a greater surgical benefit for mild and moderate head injuries compared to severe injuries. However, severe injury patients experienced a higher surgery rate than mild and moderate injuries. We developed a prognostic model using machine learning with a good level of accuracy. This model outperformed existing TBI models in regards to discrimination but not calibration. Our feasibility study captured the need for improved prognostication of TBI patients in the hospital. Conclusions. This work has provided a foundation for further investigation and implementation of TBI decision support technologies in low resource settings.

Contents

List of Tables

List of Figures

Acknowledgements

To create a thesis project of this scope, one requires the guidance, effort and endless support of many exceptional people. My completion of this thesis was not possible without me standing on the shoulders of some giants. I would be remiss if I did not take a moment to thank those that made this project possible and, more importantly, invested in me professionally and personally.

To Thiago Rocha, you are selfless individual, truly interested in helping others before yourself. You were a patient mentor for me walking me through the fundamentals of machine learning. Your positive demeanor is infectious; you were a joy to work with.

To Dr. Eric Green, while our meetings were few, your guidance was instrumental. You made time at key moments to shape this project. From business cards and speaking tips to coding and statistical support, I thank you for the breadth and depth of the teachings you provided me.

To Dr. Catherine Staton, despite me not being you research assistant, you somehow found the time for meticulous guidance as a researcher. Through active coaching and setting an example with your leadership, I have learned many of the tools needed to succeed as a future physician-researcher.

To Dr. Tony Fuller, I do not anticipate ever having another 24/7 mentor as I did in you. You were available at every step from an introductory skype conversation to introductions in Uganda, to advice on healthy daily habits. You are an example of the potential we all have when we embody the quote "success is the sum of small efforts, repeated day in and day out."

To Dr. Joao Vissoci, I would need another thesis to capture what all you have taught me. You have been caring mentor, carefully balancing healthy pressure with compassionate coaching. Professionally, you helped me find the edge of my comfort zone, and then taught me the tools to far exceed my perceived limitations. You applied the same approach of teaching me how to think through some of life's tough questions. I will carry these lessons for the rest of my life. Thank you.

To Dr. Michael Haglund, I do not know what compelled you to respond to a random email from a medical student in Texas, but I am forever grateful you did. From day one, you have opened doors I did not know existed, took a genuine interest in my development, and mentored me with research, neurosurgery and life. You established aggressive goals for my time at Duke but gave me every opportunity to succeed. Without exaggeration, this experience has completely altered my life trajectory. Thank you for the opportunity.

Finally, my motivation and dedication to this project was not possible without the unconditional support from my family. To my mother, Faye Elahi, the fervor with which you approach life no matter the obstacles you face provide an endless source of motivation for me. Thank you for encouraging my ceaselessly the last 28 years, these last two years in particular. To my father, Kamran Elahi, if I ever forget what hard work or discipline should look like I must look no further than you. The example you have set for Layla and me is more than a child could wish for. I appreciate the sacrifices you have made throughout your life to make the opportunities I reap possible. Finally, to my sister, Layla Elahi, despite being 10 years my junior, your maturity, discipline and approach to life have always been a source of inspiration. Your future is far brighter than mine; I am excited to cheer you on every step of the way.

1. Introduction

Over the last 25 years, the already devastating burden of traumatic brain injuries (TBI) has continued to increase.[1] This increase, while seen globally, is greatest in low and low-middle socio-demographic index (SDI) nations. [1] These nations must combat a high burden of disease while already suffering from human and technological shortages. These key resource limitations, which are needed to support the complex decision-making for triage of TBI patients, contribute to delays to treatment. With TBI, where change in survival is measured in minutes or hours, delays to care give rise to poor outcomes.[2] To reduce delays to care and improve outcomes, low resource settings could benefit from low cost, highly scalable solutions to the initial triage of TBI patients.

One tool which can support prognostication in low resource settings are TBI risk calculators. A risk calculator is based on a prognostic model which combines two or more clinical variables to calculate a patient's risk for a certain outcome. Prognostic models have been developed in high and low resource settings for a variety of neurosurgical pathologies including TBI.[3] The potential applications of TBI prognostic models include the support of real-time clinical decision making and clinical research design.[4,5] For research, TBI prognostic models have supported the design and analysis of randomized controlled trials. Clinically, the potential for these tools remains largely theoretically, with a persisting need for development, validation and feasibility research.

1.1 Current state of the literature

1.1.1 Traumatic brain injury

The deleterious effects of TBI are the result of primary and secondary brain injury, the latter of which is amenable to clinical intervention. Primary brain injury results from mechanical energy to the head from external forces at the time of injury. Secondary injury results from ensuing physiologic changes such as hypotension, hypothermia, and increased intracranial pressure (ICP). The Brain

Trauma Foundation (BTF) has established guidelines for management to mitigate these causes of secondary brain injury based on the best available evidence. For example, surgical decompression is an accepted intervention for select head injury patients to relieve increasing ICP.[6] BTF guidelines are developed by experts from well-resourced neurosurgical centers based on research from high income countries.[7] Health care centers in low resource settings, however, often lack the human and technical resources to fully implement these guidelines.[8]

One major cause of in- hospital death for TBI patients is secondary brain injury resulting from increasing ICP. Neurosurgical decompression, including burr-holes, craniectomies, and craniotomies, can relieve elevated intracranial pressure and prevent disability and death in certain TBI patients. The combination of limited neurosurgical resources and a capacity-exceeding TBI burden contribute to the increased mortality rates experienced in sub-Saharan Africa (SSA). In a comparison study, patients with a severe TBI in LMICs had over twice the odds of death compared to patients in high income countries (HIC).[9] TBI surveillance registries from sub-Saharan Africa (SSA) countries of Tanzania, Malawi, Uganda and Ethiopia reported mortality rates of 47%, 31%, 26% and 21% for severe TBIs. [10–13] These alarming mortality rates invite exploration of in-hospital care for TBI patients.

1.1.2 Decision making

Traumatic brain injury (TBI), "arguably the most complex disease in the most complex organ," is an intricate pathology with regards to etiology, mechanism, and severity.[14] [15] This heterogeneity poses a particular challenge to health care providers who must simultaneous predict which adverse events their patients may suffer and the risk associated with each complication. To estimate risk, providers in high income countries (HIC) rely on computed tomography scanners, intracranial pressure monitoring, electroencephalography and other key prognostic technologies [16]. These technologies are often limited or unavailable in the locations where a disproportionate burden

of TBIs occur; LMICs. [17] [18] Providers in low resource settings frequently lack the necessary

technologies to make key treatment decisions. This lack of technology forces providers to lean on

experience, intuition, vital signs and physical exam findings.

Poor TBI outcomes in LMIC which result from treatment and diagnostic delays are

influenced by limitations in health care infrastructure. These limitations include, but are not limited

to, highly skilled healthcare personnel, diagnostic and treatment resources. Specialized care,

particularly in emergency medicine, intensive care and neurosurgery, remains in its infancy in low

resources settings. In Africa, clinical care in most ICUs is provided by anesthesiologists who lack

formal critical care training. Aside from a few countries in the north and south, there are no formal

critical care fellowships for physicians in Africa. [19] In India, emergency medicine was recognized as a

separate specialty in 2009, and of the current 23,075 seats for medical education, 48 are dedicated to

emergency medicine.[20] The dearth of neurosurgery trained doctors is similar. The last comprehensive

review of neurosurgeons in Africa found 486 neurosurgeons in north and south Africa (1:358,000

people) and 79 neurosurgeons in Sub-Saharan Africa (1:3,600,000 people).[21] For comparison, the

ratio of neurosurgeons to total United States population is 1:65,580.[22]

1.1.3 Prognostic modeling

Prognostic models are a type of statistical model that predict a clinical outcome using two or

more input variables of patient information. [4] For TBI, the top predictive variables for outcome

include: age, Glasgow coma scale (GCS), arterial hypotension, computed tomography (CT) findings

and pupillary reactivity.[23] The most common outcomes predicted by TBI models are mortality or the

Glasgow Outcome Scale (GOS) at hospital discharge and at six months after injury.[24] [25]

An admission-based prognostic model for TBI could provide an objective estimate of

prognosis with readily available information upon patient presentation to the hospital.[26–28] Initial

management for TBI patients at the hospital is dependent on accurate categorization of injury

severity.[29] According to a survey of mostly LMIC doctors who routinely treat head injury patients, only 37% agreed that they currently assess patient prognosis accurately.[30] Limited access to diagnostic technologies and trained physicians in LMICs may contribute to a reduced capacity for complex decision making. With the capacity to influence patient, next-of-kin and physician decision making[25,31], prognostic models can support more rational use of limited resources.

Machine learning (ML), a branch of artificial intelligence, is a powerful approach to model development that can extract meaningful patterns from clinical data without the input of human programming.[3,32] This technique can reliably predict outcomes from vast numbers of predictors, provide real-time output, and reduce the chance for experts, using more rigid modeling approaches, to introduce bias.[33] In a comparison between artificial neural networks, an ML technique, and traditional regression modeling, Shi et al found the ML approach to achieve higher accuracy in predicting outcomes of severe TBI patients.[34] To our knowledge, no prognostic models for TBI in Sub Saharan Africa have used a ML approach.

1.1.4 Current models

Despite the potential for prognostic models to support TBI clinical decision making, there remains limited effort to develop these algorithms using data from LMICs.[4] The last major systematic review on prognostic models for TBI identified 102 prognostic models. [4] Only five of these models included populations from middle-income countries and two included populations from a low-income country. There is a need to develop prognostic models using populations from LMICs as models for HICs may perform worse when extrapolated to LMICs settings.[28] Another limitation of previous prognostic models is the use of clinical trial data for model development which could limit external validity.[35]

To date, the two most robust TBI prognostic models developed, CRASH and IMPACT, are developed on data primarily from randomized control trials (RCTs). [28,35] Despite providing the best

4

evidence of efficacy, data collected from an RCT can lack external validity and reduced the generalizability of a developed model.[36] Improved research infrastructure allowing quality data collection in low resource settings provides an opportunity to apply advanced statistical techniques such as machine learning based prognostic modeling to more robust datasets.[37]

1.2 Purpose of this study

Figure 1 outlines the steps required to implement a TBI prognostic model in LMICs. The objective of this study was to advance up the implementation triangle to the data for implementation level. Specifically, this thesis focused on "Assessment of Need", "Develop Prognostic Model", "Mobile App development", and "Data for Implementation". It is our hope that the findings generated from this thesis can lay a foundation for TBI decision support technology in LMICs.
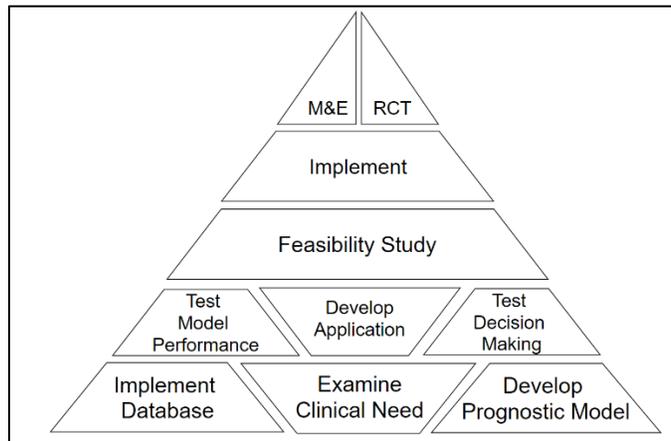
Figure 1: Implementation triangle for a decision support tool in LMICs

2. Methods: Framework of this project

We have provided the methods for each study completed for this project as indicated by the study number. This thesis includes the following studies:

- Survival analysis on acute in-hospital TBI outcomes (Study I)
- Prognostic model development (Study II)
- Three model validation (Study III)
- Vignette Experiment (Study IV)
- Feasibility study (Study V)

2.1 Study Design

2.1.1 Survival analysis (I)

We conducted a retrospective, cross-sectional observational study. In designing the analysis we followed the STROBE reporting guideline.[38] We incorporated best-practice recommendations published for time-to-event outcomes and survival analyses.[39–41]

2.1.2 Prognostic model development (II)

We conducted a retrospective analysis of clinical outcome data to determine key predictors of patient prognosis. In designing the analysis, we followed the Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research and the TRIPOD guideline.[42,43]

2.1.3 Three model comparison (III)

We conducted an external validation of the CRASH and IMPACT prognostic models using the Kilimanjaro Christian Medical Center (KCMC) TBI registry. We followed best practice recommendations proposed in the literature for validation studies.[44]

2.1.4 Vignette experiment (IV)

We created one detailed clinical vignette of a TBI patient (Appendix 1). A panel of emergency medicine doctors and neurosurgeons created the vignette to represent a realistic patient case for the low resource setting. The hypothetical patient case was a severe head injury with a CRASH calculator estimated risk of mortality of 51.4%. We chose this level of severity to provide near equal ranges to over or under estimate patient risk. In addition to the vignette, we administered a follow-up survey to capture participant background information and other considerations for decision making. We did not collect identifying participant information to preserve anonymity. The primary author (C.E) administered all surveys in-person between October and December 2018.

Randomization

We randomly assigned participants to one of two groups, the clinical vignette alone (control) or the clinical vignette plus the output from the CRASH risk calculator (experiment group). Before the experiment, we entered the patient information from the vignette into the CRASH risk calculator. The risk calculator estimated the patient risk for mortality at 14-days and for unfavorable outcome (death, vegetative state, or severe disability) at six months. We copied the output from the calculator and included it in the vignette for those in the experimental group. We also provided a brief written introduction to risk calculators for those in the experimental group. Other than the risk output and the description of risk calculators, the vignettes were identical between the two groups. A co-investigator (S.W.) randomized the order of clinical vignettes with or without the risk score using a computer program and placed them in envelopes. We kept the survey administrator (C.E.) blinded from the randomization process.

2.1.5 Feasibility study (V)

We implemented a mixed methods approach collecting quantitative data by direct observations and surveys and qualitative data through semi-structured interviews. The study design

was informed by Bowen et als "How We Design Feasibility Studies" and the mobile health evidence reporting and assessment checklist. [45,46]

2.2 Study Setting

2.2.1 Survival analysis, prognostic model, three model comparison (I-III)

The data for this portion of the study came from Kilimanjaro Christian Medical center (KCMC), a tertiary level hospital in northern Tanzania serving a catchment population of over 15 million people. KCMC treats about 1000 TBI patients annually.[10] At the time of data collection, there were no Emergency Medicine trained physicians, and two to four residents and interns staffing the Emergency Department at any time. KCMC has a training center for anesthesia; but there are no neurosurgical-trained physicians at KCMC, General surgery attendings and residents predominately perform emergency burr holes or more rarely hemicraniectomies for TBI patients.

In 2013, KCMC established a prospective TBI registry as part of a quality improvement process. [10] The registry consecutively enrolled all patients presenting to KCMC Emergency Department for treatment of their acute (<24 hours) TBI. Trained research nurses collected data on the enrolled patients' injury details, acute care, hospitalization care and condition at discharge. The research team and study PI performed quality analysis on data entering the registry at time of data entry and during weekly meetings.[10]

2.2.2 Vignette experiment and feasibility study (IV, V)

MRRH is a 350-bed referral hospital located 250 kilometers Southwest of the capital city Kampala. The hospital serves the surrounding 10 districts of Uganda and often treats patients from neighboring Rwanda and Tanzania. The annual volume of TBI patients is 600. There is one neurosurgery trained doctor at MRRH. There are no neurosurgery residents. Instead, general surgery residents complete six-week rotations in the neurosurgery department during their third year of

training. Interns, who are in their first year out of medical school, also rotate through the neurosurgery department. MRRH does have the country's first emergency medicine residency. The casualty department (emergency department), the location of first triage for TBI patients, is staffed by general surgery residents, emergency medicine residents and interns.

MNRH is a 1,500-bed national hospital located in the capital city of Kampala which treats around 1,500 cases annually including the casualty and neurosurgery departments. The hospital has a neurosurgery residency program staffed with 6 residents. The neurosurgery department, including residents and consultants, perform all the neurosurgical procedures at MNRH.

2.3 Study Population

2.3.1 Survival analysis, prognostic model, three model comparison (I-III)

The study included patients enrolled in the KCMC registry between May 5, 2014 and April 17, 2017. We excluded patients if they had missing data in the following key variables: admission date, discharge date, received surgery for TBI, and admission Glasgow coma scale (GCS) (Figure 2).

2.3.2 Vignette experiment and feasibility study (IV, V)

There are only 12 residency trained neurosurgeons in Uganda. To increase our sample size, we recruited doctors at different levels of training and in different specialties. We included doctors in other specialties which were closely involved in the care of TBI patients. This definition included emergency medicine doctors, general surgeons and neurosurgeons. The levels of training we included were interns, residents, and consultants. For an intern or general surgery resident to be included, they must have completed a rotation in the neurosurgery department or in the casualty department. Traditionally, doctors will work as a medical officer for two to four years after intern year before returning for residency training. As result, residents are about three to five years removed from medical school. After residency is complete, the doctor enters the workforce as a consultant (the

9

equivalent to an attending physician).

2.4 Variables

2.4.1 Survival analysis (I)

Explanatory variables

The analysis included the following variables: age, gender, self-reported consumption of alcohol before the injury, admission GCS, mechanism of the injury, if the patient had surgery for TBI, and if the patient was transferred from the surgical ward to the intensive care unit (ICU) representing a worsening in their condition or a missed triage from the Emergency Department. The TBI surgeries variable includes mostly emergency burr holes procedures and some craniotomies/craniectomies.

We stratified patients into three groups based on admission GCS : mild (14-15), moderate (9-13), severe (3-8) TBI. We used the classic GCS scale (GCS of 14-15 is mild rather than GCS of 13-15) to create more balanced groups for the survival analysis.[47] We transformed age from a continuous to an ordinal variable with the following categories: less than 18, 18-29, 30-39, 40-49 and older than 50. Mechanism of injury included road traffic injury, assault, falls, or other. Examples of other mechanisms of injury included accidental trauma to head with object.

Outcome

We selected the discharge GOS as the outcome measure for this study (Table 1). GOS ranges from one to five, each number representing a different level of recovery as described in Table 1. We dichotomized this variable into good outcome (GOS of 4 to 5) and poor outcome (GOS of 1 to 3). About half way through the data collection process, we switched from using GOS to using the Glasgow Outcome Scale extended (GOSe) to improve our descriptions of patients' outcome disabilities. This scale ranges from one to eight. We converted those with a GOSe outcome to a GOS as shown in Table 1. The registry did not include outcomes after discharge.

Outcome	GOSe score	GOS score
Death	1	1
Persistent vegetative state	2	2
Severe Disability	3 or 4	3
Moderate Disability	5 or 6	4
Good Recovery	7 or 8	5

2.4.2 Prognostic model development (II)

Prognostic variables

We initially selected the following prognostic variables: age, gender, mechanism of injury, intention of injury, presence of alcohol, initial vital signs collected in the ED (temperature, respiratory rate, pulse, systolic blood pressure, diastolic blood pressure, pulse oxygenation), glucose level, pupils size and reactiveness (pupils equal round reactive to light, accommodation - PERRLA), initial Glasgow coma score, if the patient had surgery for TBI, and if the patient had a TBI unrelated surgery. We selected the predictors based on emergency medicine and neurosurgery expert consensus. Additionally, we only considered predictor variables practical for the low resource setting. Practical for this project meant variable collection did not involve resource intensive equipment, variables were available shortly after patient presentation to hospital, and low and high-skilled personnel could collect the information. We did not include CT findings in the model as providers in this setting have inconsistent access to CT scanners.

Outcome selection

The endpoint for the prognostic model was the GOSe score at patient discharge. The discharging doctor or a research nurse calculated the GOSe score on all patients at hospital

discharge unless patient mortality during the hospitalization. The GOSe is an ordered outcome from one (worst outcome) to eight (best outcome), one represents patient death while eight represents upper good recovery. We dichotomized the GOSe scores into good and poor recovery. We classified scores from one to six as poor recovery and scores of seven or eight as good recovery. We chose a dichotomous outcome rather than ordinal because in our clinical setting, there are few patients with moderate outcomes. We selected scores less than seven, rather than less than five, as poor recovery to increase the number of patients with poor recovery. This step improved the balance between good and poor recovery in our dataset.

2.4.3 Three model comparison (III)

There were three models tested in this study: CRASH, IMPACT and a KCMC developed model.

CRASH

In 2008, a team of researchers developed the CRASH prognostic models from one multinational, multi-center RCT (n=10,008). The TBI patient population included were mild, moderate and severe injuries (GCS <= 14) with an age >= 16. The data included was from high income countries (HIC) and LMICs. The outcome measures used were mortality at 14 days and unfavorable outcome (GOS of 1-3) at six months after TBI. There are multiple versions of the CRASH model including the basic model and the CT model. The basic model predictors are age, GCS, pupil reactivity and major extracranial injury. The CT model includes the basic model predictors and additional findings from CT. Since the KCMC TBI registry does not include CT findings, we used the CRASH core model for model comparison. Additional information on the CRASH model can be found in the model's original publication.[35]

IMPACT

In 2008, a team of researchers developed the IMPACT prognostic models from the IMPACT database (n = 8509). This database includes prospectively collected data from eight RCTs and three observational studies. The TBI patient population included were moderate and severe injuries (GCS <= 12) with an age >=14. Additionally, all data came from HICs. The outcome measures used were mortality and unfavorable outcome (GOS 1-3) at six months after TBI. There are multiple versions of the IMPACT model including the IMPACT core model and the IMPACT extended model. The predictors for the core model are age, GCS motor score, and pupillary reactivity. The Extended model includes variables for hypoxia, hypotension and CT findings. As in the case with the CRASH model, we used the IMPACT core model for this study. Additional information on the IMPACT model can be found in Steyerberg et al's 2008 publication.[28]

KCMC TBI Model

In 2018, Rocha et al developed a machine-learning based prognostic model using the KCMC TBI registry (n=3138). We have previously described the details of the prospectively collected TBI registry.[11] The TBI patient population included were mild, moderate and severe injuries (GCS 3-15) with no age restrictions (pediatric and adult populations included). The outcome measure was GOS at discharge. The predictors included are injury characteristics, vital signs, physical exam findings and inpatient care processes. Additional information on this model can be found in Rocha et al's publication.

2.4.4 Vignette experiment (IV)

For the clinical vignette, we asked providers what was the best next step for the patient: surgery, admit to the intensive care unit (ICU), initiate comfort measures, or observe the patient. We then asked providers how many out of 100 patients similar to the vignette case would: survive the hospitalization, survive to six months, survive to 30 days, be able to perform activities of daily living within the next six months, and remain on a ventilator. We subtracted these numbers from 100 to

calculate the providers estimated risk for each outcome. We then divided the study population into experienced and inexperienced doctors. Experienced doctors were either in residency or had completed residency training. Inexperienced doctors were in their intern year.

To understand the influence of resources on the participants decision making, we asked providers what percentage of patients with an identical presentation would receive a CT within two days of admission and what percentage would receive surgery within two days of admission. Finally, we asked if patient prognosis, resources available or both influence their decisions.

The background information collected from each provider included medical specialty, gender, ethnicity, year of birth, religion, year of medical school graduation, country of primary practice, level of training (intern, resident, fellow, consultant-no residency, consultant-residency complete), do you supervise residents (yes, no), how important is religious faith in your life (scale including very important, important, neutral, unimportant, or very unimportant) and in which setting do you currently practice (private practice, employed by university, employed by government). We calculated age by subtracting 2018 from year of birth. We calculated years of experience by subtracting 2018 from year of medical school graduation.

The follow-up survey asked providers to select if they strongly agree, agree, neutral, disagree, or strongly disagree with the following statements: I have recommended surgery in patients that I should not have, I have not recommended surgery in patients that I should have, I am responsible if a patient dies within 30 days of any surgery I recommend, I am willing to not offer surgery even if it means a patient will die, I am obligated to offer surgery if there is a chance the patient will survive, I worry about incorrectly estimating prognosis, I use available literature and data to make a prognosis, I use prior experience with patients to make a prognosis and I feel that there is adequate quality data to base prognostication.

2.4.5 Feasibility study (V)

Participants completed a demographic questionnaire, a survey on potential uses of TBI

prognostic models, usability testing with closing user experience survey, and a semi-structured interview. There was overlap between data collection methods to capture the usability, acceptability and demand for a TBI risk calculator in this setting.

2.5 Statistical Analysis

2.5.1 Survival Analysis (I)

Data analysis

The data analysis compared the two study groups: TBI patients receiving or not receiving TBI surgery. We followed each patient from admission (hospital day [HD] zero) to discharge, inpatient death, or HD 14. We selected hospital day 14 because this is the cutoff to assess acute TBI outcomes. Additionally, survival plots should continue until 10-20% of the data remains.[39] In our study, 13% of patients remained without an outcome at HD 14. We right censored patient data without an outcome at HD 14 or with a good outcome prior to HD 14. We used R Language for Statistical Computing 3.4.1 for data management and statistical analyses.[48] We set significance at a p-value less than 0.05.

Descriptive Analysis

We calculated descriptive statistics for demographic and clinical care data including means, standard deviations, medians and IQR. We used chi-square tests to compare categorical variables.

Kaplan-Meier (KM) Survival Plots

We used the KM plots to present the survival curves and test the crude survival between the two study groups. One KM curve depicted the entire study population and a second set of KM curves depicted mild, moderate and severe TBI patients separately. We presented the statistical uncertainty with 95% confidence intervals (shaded region on plots). Our KM plots travel up with sliding y-axis scales to improve comparison between study groups as recommended in the literature.[39] We displayed the number of patients still at risk at a given time below the time axis.

15

Cox Proportional Hazard Model (Cox model)

We built Cox models to examine the association of TBI surgery on outcomes for the overall study population and by TBI severity. The Cox model is a powerful method to analyze survival data which can produce a hazard ratio (interpreted similar to an incident rate ratio or relative risk) for poor outcome while controlling for possible confounders.[41] For appropriate use and inference of a Cox model, the survival curves must be proportional throughout the study period. To assess for proportionality, we (1) checked if the survival curves crossed and (2) used the proportional hazards assumption of a Cox regression test (Coxph). If the proportional hazard assumption was violated, we used step functions to add time interaction terms[49]. We used the Coxph output to determine at which HD to add an interaction term. We produced unadjusted and adjusted effect estimates for surgery presented in tables 3A and 3B respectively.

2.5.2 Prognostic model (II)

Preprocessing of data for model building

All data were processed using the statistical language R. The first step was handling missing data. The outcome variable had no missing data and no patient had greater than 20% missing data. Input variables with more than 80% of cases missing were removed. The only variable removed during this first analysis was the glucose level. Next, we used multiple imputation using chained equations to impute missing values from variables with less than 20% of cases missing32. We separated each variable considering the measurement level (e.g. numeric, categorical) for different imputation approaches. We considered all the variables for the imputation process. We obtained the resulting dataset after 10 iterations.

A high correlation analysis was performed following the missing imputation steps to identify high correlation variables (i.e. greater than 0.9). No variable was dropped during this step. An analysis to exclude outliers among the numeric variables was performed and the following cases were

dropped: age greater than 75 years, respiratory rate greater than 75, and systolic blood pressure greater than 220 mmHg.

Several variables considered in the analysis were ordinal or categorical. As a result, the initial group of 21 variables expanded to 56 after dummy variable conversion. To maintain data integrity after this step, we investigated the data for the presence of near zero variance, high correlation and linear combos. Eighteen variables were removed after the near zero variance test, seven were removed after the analysis of high correlation, and the linear combo analysis did not exclude any variables. Ultimately, we used data from 3138 patients in the database and 31 variables after all preprocessing steps.

Internal validation strategy

We used a patient based cross validation with 10 folds and five repetitions for data splitting . Since the outcome of interest was imbalanced (14.4% had a poor recovery), we used a regularization for imbalanced procedure called Synthetic Minority Over-sampling Technique.[50]

Modeling techniques and performance metrics

We tested nine different models: k-Nearest Neighbor, Ridge Regression, Neural Network, Bagged Tree, Bayesian Generalized Linear Model, Bayesian Additive Regression Trees, Gradient Boosting Machine, Single C5.0 Ruleset and Random Forest. All methods were validated using an internal approach. The kappa statistic was the metric used to assess the prognostic models built. We did not have an external dataset so we could not test the precision of the model with unseen data. The metrics for comparison among the models were based on confusion matrix statistics: area under the ROC curve (AUC), sensitivity, specificity, positive predictive value, negative predictive value.[51–54] All AUC comparative analysis were performed using the R, pROC package.[48]

2.5.3 Three model comparison (III)

We calculated the distribution of predictors in the KCMC validation dataset and compared to the distribution in the CRASH and IMPACT development datasets. We then tested the performance of the CRASH and IMPACT models on the KCMC data set. We assessed the discrimination and calibration of these models. To test discrimination, we calculated the area under the receiver operating characteristic curve (AUC) and plotted the ROC curve. To test calibration, we created calibration plots. We also calculated the Brier score for all three models. We used the computing environment R for all statistical analyses and figure generation.[48]

2.5.4 Vignette experiment (IV)

We used a chi^2 testing to compare the participants' demographic and medical training background between the control and experiment study groups. This test allowed us to determine if our randomization was effective. We used an unpaired t-tests to compare the risk estimates between the two study groups. We repeated this technique after dividing the study population into experienced and inexperienced providers. For the follow-up clinical care questions, we reported counts, means and standard deviations. Finally, we presented the mean responses to the decision-making survey as a figure.

2.5.5 Feasibility study (V)

We totaled background demographic data on the participants and presented the findings in a tabular format. For the usability testing, we calculated the mean and standard deviation to complete an entry and the mean number of errors and questions. We presented the survey likert responses in a figure, totaling the score provided by the participants by each question. For the semi-structured interviews, we used an inductive content analysis to identify emerging themes. We present these emerging themes along with salient quotes to exemplify the theme.

2.6 Institutional review board

2.6.1 Survival Analysis, prognostic model, three model comparison (I, II, III)

Our study received ethical clearance from the Institutional Review Boards of Kilimanjaro Christian Medical Center, Moshi, Tanzania, as well as Duke University.

2.6.2 Feasibility study and vignette experiment (IV and V)

Our study received ethical clearance from the Institutional Review Boards of Makerere University, Uganda, as well as Duke University. We additionally received hospital clearance from Mbarara Regional Referral Hospital and Mulago National Referral Hospital.

3. Results

3.1 Survival analysis (I)

Demographics

We started with all 3209 patients enrolled in the TBI registry for this study. After the exclusion of cases with key missing data, 2506 TBI patients remained for inclusion in the analysis (Figure 2).
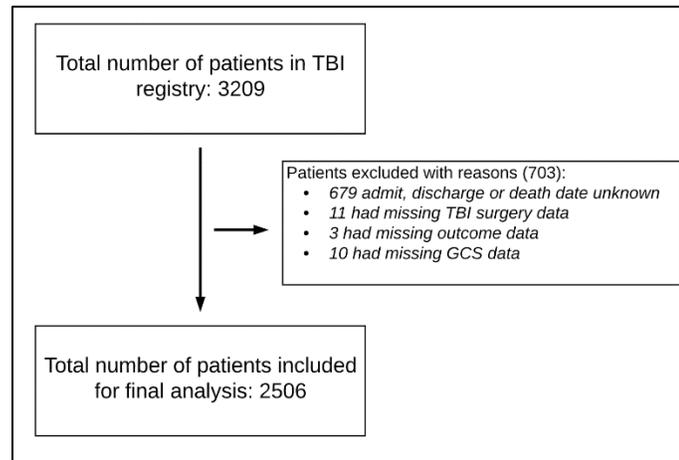
Figure 2: Diagram illustrating selection process for final study population

Out of the final 2506 patients, 628 (25%) received TBI surgery and 1878 (75%) did not receive TBI surgery. There were significantly less road traffic injuries and more violent causes of injury in those receiving surgery. Those receiving TBI surgery were also more likely to receive care in the ICU, have a poor outcome, have a moderate or severe TBI and stay in the hospital longer. There was no significant difference between the two study groups by age, gender, or presence of alcohol.

Of the 628 TBI surgeries performed, severe TBI patients received the highest proportion (123/316 or 39%), followed by moderate TBI patients (117/357 or 33%) and finally mild TBI patients (388/1833 or 21%).

Table 2: Study population for survival analysis

| Variables | Patient Population, n (%) | | | P-value |
	All patients n = 2506	Yes TBI Surgery, n = 628 (25%)	No TBI Surgery, n = 1878 (75%)	
Age				0.4019
<18	369 (15)	100 (16)	269 (15)	
18-29	883 (36)	202 (32)	681 (37)	
30-39	559 (22)	147 (23)	412 (22)	
40-49	325 (13)	88 (14)	237 (13)	
50+	343 (14)	88 (14)	255 (13)	
missing	27	3	24	
Gender				0.6175
Male	2087 (84)	526 (84)	1561 (83)	
Female	418 (16)	101 (16)	317 (17)	
missing	1	1	0	
Presence of Alcohol				0.1579
Yes	667 (27)	153 (25)	514 (27)	
No	1218 (49)	310 (49)	908 (49)	
Unknown	599 (24)	162 (27)	437 (24)	
missing	22	3	19	
Mechanism of Injury				<.001
RTI	1695 (68)	349 (56)	1346 (72)	
Assault	362 (14)	127 (20)	235 (12)	
Fall	252 (10)	72 (11)	180 (10)	
Other	189 (8)	75 (12)	114 (6)	
Surgical ward to ICU				
Yes	367 (16)	240 (41)	127 (7)	<.001
No	2024 (84)	347 (59)	1677 (93)	
Missing	115	41	74	
Outcome				.0035
Days to outcome median (IQR)	3.4 (1.8 - 7.2)	6.2 (3.6 - 11.2)	2.8 (1.5 - 5.7)	

Poor outcome (GOSe 1-4 or GOS 1-3)	257 (10)	90 (14)	167 (9)	
Good outcome (GOSe 5-8 or GOS 4-5)	2249 (90)	538 (86)	1711 (91)	
TBI				<.001
Mild (14-15)	1833 (73)	388 (62)	1445 (77)	
Moderate (9-13)	357 (14)	117 (19)	240 (13)	
Severe (3-8)	316 (13)	123 (20)	193 (10)	

Figure 3 is the KM plot for the entire study population. TBI patients not receiving surgery, the blue line, had more events of poor outcome compared to those receiving surgery. Around HD 11, the lines cross and those receiving surgery had more events of poor outcome. After HD one, the CI between the two lines overlap, and the p-value of .45 for the log rank test suggests the lines across the 14 days are not statistically different.
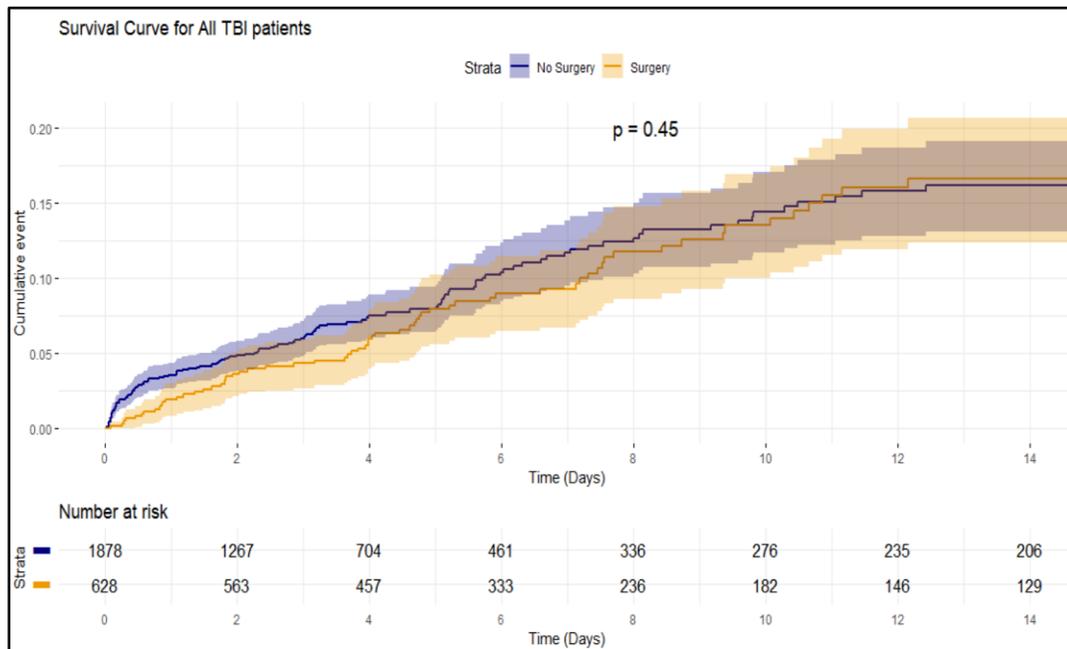
Figure 3: Kaplan Meier plot comparing outcomes between those receiving and not receiving TBI surgery for the entire study population

Figure 3 includes the KM plots for each TBI severity level separately. The slope of the survival curves increases from mild to severe reflecting an increase in poor outcome rates as severity

increases. In each of the three plots, the survival curves between patients receiving and not receiving surgery separate initially, draw near between HD five and 10, and then separate again. This fluctuation resulted in visible and statistical violation of the proportionality assumption.
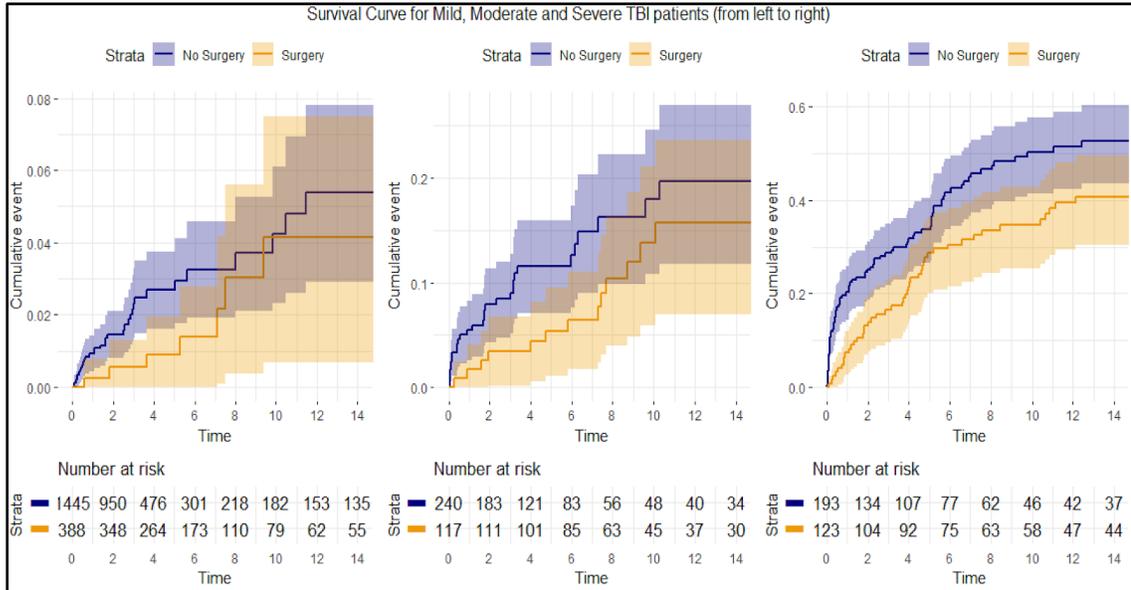
Figure 4: Kaplan Meier curves comparing outcomes between TBI patients receiving and not receiving surgery, stratified by GCS severity groups. We used a sliding y-axis scale.

Table 3a and 3b provide the output for the unadjusted and adjusted Cox models testing the association of TBI surgery on outcomes. The overall study and severe TBI populations required time interaction terms at day three and day seven to ensure proportionality. Mild and moderate TBI populations needed a time interaction at day seven only.

Table 3: Unadjusted Cox proportional hazard ratios (HR) for TBI surgery by injury severity with time interactions.

Cox Proportional Hazards								
Hospital Day	Overall		Mild		Moderate		Severe	
	HR (CI)	p	HR (CI)	p	HR (CI)	p	HR (CI)	p
0-3	.71 (0.46, 1.08)	0.12	0.37 (0.13, 1.07)	0.06	**0.41 (0.18, 0.94)**	**0.03**	**0.51 (0.30, 0.85)**	**.01**
4-7	.90 (0.53, 1.53)	0.70					0.83 (0.44, 1.55)	0.55
8-14	1.66 (0.83, 3.33)	0.15	1.62 (0.36, 7.23)	0.53	1.86 (0.47, 7.47)	0.37	0.91 (0.34, 2.42)	0.84

In table 3a, surgery reduced the hazard for poor outcome for all study group before HD eight. This finding was only significant for moderate TBI patients before HD eight and severe TBI patients before HD four. Surgery increased the hazard for poor outcome for the overall, mild and moderate TBI populations after HD seven, however, this finding was insignificant. In table 3b, surgery significantly reduced the hazard for poor outcome for all study group before HD eight. Moderate TBI patients receiving surgery had the greatest reduction in hazard, an 83% decrease.

Table 4: Fully adjusted Cox proportional hazard ratios (HR) for TBI surgery by injury severity with time interactions. Covariates included, but not shown, are mechanism of injury, gender, alcohol, surgery to ICU, age, and GCS (for overall only).

Cox Proportional Hazards

Hospital Day	Overall		Mild		Moderate		Severe	
	HR (CI)	p	HR (CI)	p	HR (CI)	p	HR (CI)	p
0-3	**.32 (0.19, 0.54)**	**0.001**	**0.20 (0.06, 0.64)**	**0.006**	**0.17 (0.06, 0.49)**	**<0.001**	**0.47 (0.24, 0.89)**	**0.021**
4-7	**.46 (0.25, 0.86)**	0.015					0.65 (0.30, 1.39)	0.264
8-14	0.77 (0.36, 1.66)	0.505	0.69 (0.14, 3.31)	0.642	1.23 (0.23, 6.68)	0.808	0.70 (0.23, 2.13)	0.529

Table four provides the Cox model output for the fully adjusted cox model of the overall study population. Moderate TBIs, severe TBIs and transfer from the ward to the ICU significantly increased the hazard for poor outcome. Patients receiving surgical intervention and having an outcome by HD 3 had a 68% reduction in their hazard ratio. Similarly, patients with a surgery and outcome by HD seven had a 54% reduction in their hazard ratio.

Table 5: Fully adjusted Cox model results for entire study population.

Cox Model Output		
Variable	**Hazard Ratio (CI)**	**P-value**
Gender		
Male	1.51 (0.98, 2.35)	0.062
Age		

< 18	ref	ref
18-29	1.91 (1.04, 3.52)	0.037
30-39	2.45 (1.29, 4.64)	0.006
40-49	2.51 (1.28, 4.91)	0.007
50+	1.92 (0.99, 3.72)	0.053
MOI		
road traffic crash	ref	ref
assault	0.73 (0.41, 1.32)	0.302
fall	1.55 (0.96, 2.51)	0.072
Other	1.24 (0.67, 2.28)	0.493
Alcohol		
No	ref	ref
Yes	0.90 (0.59, 1.38)	0.633
Unknown	1.03 (0.70, 1.52)	0.878
TBI Surgery		
HD 0-3	**0.32 (0.19, 0.54)**	**<0.001**
HD 4-7	**0.46 (0.25, 0.86)**	**0.014**
HD 8-14	0.77 (0.36, 1.66)	0.505
Surgery to ICU Transfer		
Yes Transferred	**3.43 (2.32, 5.05)**	**<0.001**
GCS		
Mild	ref	ref
Moderate	**3.21 (1.99, 5.17)**	**<0.001**
Severe	**8.01 (5.18, 12.4)**	**<0.001**

3.2 Prognostic model (II)

Of the 3138 patients admitted to the KCMC, 2685 (85.6%) had a good recovery while 453 (14.4%) patients had a poor recovery (Table 1). There were 321 (71%) mortality cases in the poor recovery group. Both the good and poor recovery groups were predominantly male (82%, 83%) with unintentional injuries (81%, 87%). The mean age was also similar, 30.7 and 33.9 years for the good and poor recovery groups respectively. Seven-hundred and two (26.2%) of those in the good recovery group and 112 (24.7%) in the poor recovery group reported use of alcohol at time of injury.

The vital signs among good and poor recovery groups were not significantly different. However, this was not the case for clinical examination variables. The average GCS score of the poor recovery group was 8.7 while the good recovery group had a mean score of 14. In the good recovery group, 2641 (98.4%) patients had bilaterally reactive pupils compared to 318 (70.2%) in the poor recovery group. In the good recovery group, 554 (21%) patients were transferred from the surgical ward to the ICU compared to 242 (53%) in the poor recovery group. One-hundred-and-forty-nine (33%) patients in the poor recovery group received TBI surgery compared with 551 (21%) in the good recovery group.

Table 6: KCMC TBI patient profile

	All Patients GOSe 1-8 (n - 3138)	Good Recovery GOSe 7, 8 (n = 2685)	Poor Recovery GOSe 1-6 (n = 453)	pvalue
Age Mean (SD)	31.2 (15.3)	30.7 (15.2)	33.9 (15.7)	
Gender				.604
Male (n%)	2574 (82)	2198 (82)	376 (83)	
Intention of Injury				<.001
Unintentional (n%)	2584 (82)	2188 (81)	396 (87)	

Self Inflicted (n%)	1 (0)	0 (0)	1 (0)	
Inflicted by Other (n%)	533 (17)	487 (18)	46 (10)	
Unknown (n%)	20 (1)	10 (0)	10 (2)	
Alcohol				.111
Yes (n%)	814 (25.9)	702 (26.15)	112 (24.72)	
No (n%)	1518 (48.4)	1345 (50.09)	173 (38.19)	
Unknown (n%)	806 (25.7)	638 (23.76)	168 (37.09)	
Vital Signs				
Temp Mean (SD)	36.5 (0.8)	36.4 (0.7)	36.5 (1)	
Resp Rate Mean (SD)	21.7 (3.8)	21.6 (3.6)	22.5 (4.7)	
Heart Rate Mean (SD)	87.8 (18)	87.3 (16.8)	91.2 (23.6)	
Systolic BP Mean (SD)	121.3 (20.4)	121.5 (19.4)	119.7 (25.5)	
Pulse Ox Mean (SD)	95.5 (7.3)	96.4 (5)	89.9 (13.7)	
Clinical Exam				<.001
GCS Total Mean (SD)	13.2 (3.2)	14 (2.2)	8.7 (4.4)	
PERRLA				<.001
Yes (n%)	901 (29)	787 (29)	114 (25)	
No (n%)	22 (1)	8 (0)	14 (3)	
Unknown (n%)	2215 (71)	1890 (70)	325 (72)	
Disposition from the Emergency Department				<.001
ICU	92 (3)	28 (1)	64 (15)	
Surgery	2684 (86)	2322 (87)	362 (84)	
Operating Theatre	16 (1)	13 (0)	3 (1)	
Home	323 (10)	321 (12)	2 (0)	

Death	23 (1)	1 (0)	22 (0)	
TBI Surgery				<.001
Yes (n%)	700 (22)	551 (21)	149 (33)	
No (n%)	2107 (67)	1813 (68)	294 (65)	
Unknown (n%)	331 (11)	321 (12)	10 (2)	
Other Surgery				.439
Yes (n%)	377 (12)	312 (12)	65 (14)	

We tested nine different ML models on the dataset (Table 2). The AUC of the models varied from 66.2 % (K nearest neighbors) to 86.5% (Bayesian Generalized Linear Model). Despite this range, several models obtained similar AUC values. A confidence interval analysis of AUC identified the Bayesian Generalized Linear Model as the best approach to categorize the outcomes. For TBI patients receiving care at KCMC, this model produces a CI 95%: 85.6-87.4 and an AUC of 86.5 (Table 2).

Table 7:Comparison among machine learning models applied to predict TBI outcome.
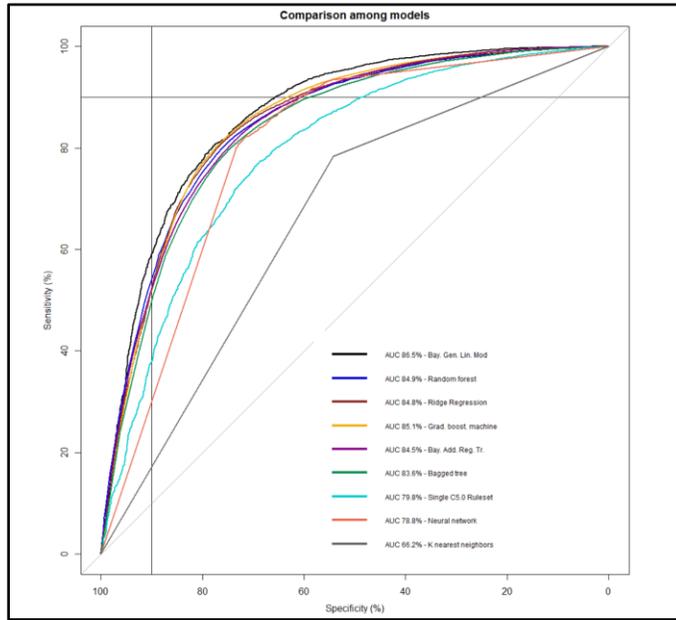
Table 8: Machine learning models and AUC with confidence intervals at 95%.

Method	AUC	95% CI - bottom limit	95% CI - upper limit
Bayesian Generalized Linear Model	86.5	85.6	87.4
Random forest	84.9	84.6	85.3
Ridge Regression	84.8	84.5	85.3
Gradient boosting machine	85.1	84.9	85.3
Bayesian Additive Regression Trees	84.5	84.3	84.8
Bagged tree	83.6	82.7	84.6
Single C5.0 Ruleset	79.8	78.8	80.9
Neural network	78.8	77.8	80.0
K nearest neighbor	66.2	66.1	65.5

The Bayesian generalized linear model had a sensitivity of 0.89 and a specificity of 0.67. The precision was 0.94 . The accuracy of the model was of 0.855 and the kappa statistic was 0.49. The best model achieved a moderate to high capability of predicting a good recovery and an intermediate ability of predicting a bad recovery.

We also extracted the predictive weight of each variable in the Bayesian generalized linear model (Figure 5). The top four predictors of a good outcome were an increasing GCS score, an increasing blood oxygen level, a foot or fist mechanism of injury (MOI) and if the injury was unintentional. The top four predictors of a poor outcome were a patient being sent to the intensive care unit (ICU) after surgery, a MOI by a gun, direct admission to the surgical floor, and a car MOI.
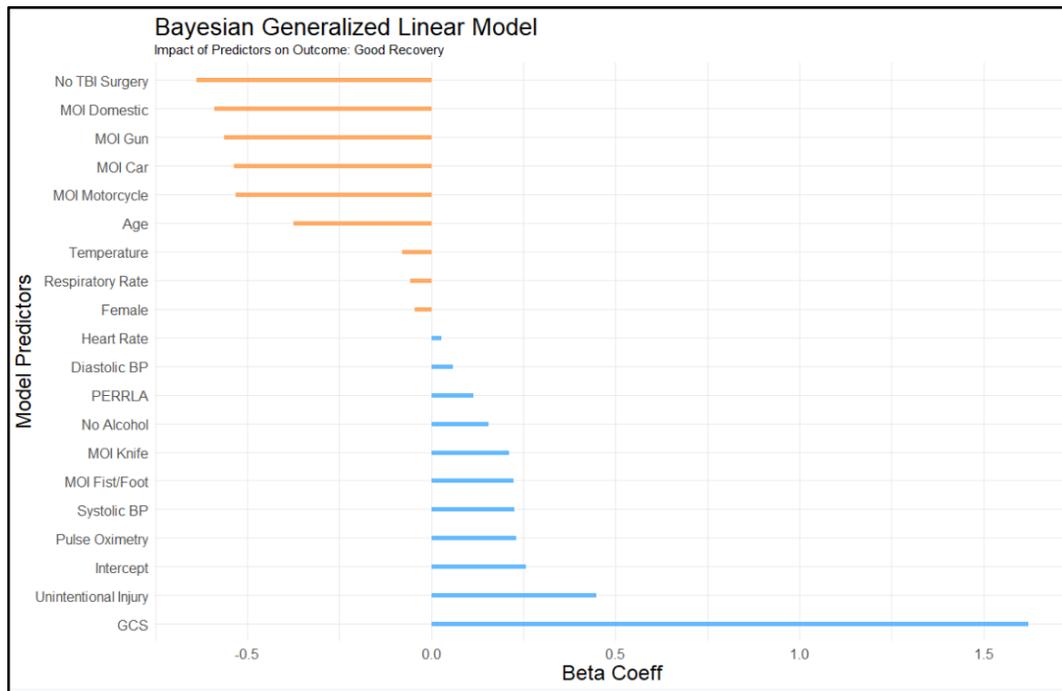
Figure 5: Beta coefficients for the highest performing predictive model.

3.3 Three model comparison (III)

The validation dataset from KCMC, after exclusion of cases with missing predictor variables, was 2972. The CRASH and KCMC datasets had a similar age mean and standard deviation with

more pediatric patients in the KCMC dataset and more elderly in the CRASH dataset. These two

datasets also had a similar gender and mechanism of injury distribution. The CRASH and KCMC

datasets had significantly different admission GCS, pupil reactivity and outcome distributions. The

IMPACT and KCMC datasets had significantly different admission GCS, pupil reactivity and

outcome distributions. The IMPACT data involved more severely injured patients compared to

CRASH and KCMC.

Table 9: Patient characteristics of IMPACT, CRASH and KCMC datasets.

Characteristics	KCMC Data (N= 2972)	CRASH (N=10,008)	IMPACT (N=8509)
Development data			
Study design	Obs	RCT	RCT+Obs
Inclusion period	2013-2017	1999-2004	1984-1997
Age			
<20	19.4%	12.5%	NA
20-24	15.8%	17.0%	NA
25-29	14.9%	13.0%	NA
30-34	12.8%	10.7%	NA
35-44	18.5%	17.9%	NA
45-54	10.8%	12.5%	NA
>54	7.7%	16.7%	NA
Median (IQR)	29 (21-40)	NA	30 (21-45)
Mean (SD)	31.1 (15.2)	32 (17.1)	NA
Gender			

Male	82.5%	81.0%	NA
Mechanism of injury			
Road traffic crash	68.1%	65.1%	NA
Other	31.9%	35.0%	NA
Total Glasgow coma scale			
Mild (15)	66.3%	0.0%	NA
Mild (13-14)	13.5%	30.2%	NA
Moderate (9-12)	8.8%	30.4%	18.0%
Severe (3-8)	11.4%	39.5%	82.0%
Motor Score			
None	3.5%	NA	16%
Extension	2.9%	NA	12%
Abnormal flexion	3.4%	NA	13%
Normal flexion	4.7%	NA	23%
localizes/obeys (5/6)	85.5%	NA	30%
Untestable/missing	0%	NA	5%
Pupils			
Both reactive	95.5%	82.8%	63.0%
One reactive	2.0%	6.3%	12.0%
None reactive	2.6%	8.2%	25.0%
Unable to assess	0%	2.5%	16.3%

Outcome			
Death or severe disability*	10.1%	37.2%	48.0%

* Death or severe disability is calculated at 6 months for the CRASH and IMPACT while the KCMC data collected the outcome at hospital discharge or in-hospital death. Follow-up outcome data was not available for the KCMC dataset.

We first compared the discrimination of the three models. The ROC curves are depicted in figure 1 and the AUC for each curve is provided in table 2. The KCMC model had significantly higher AUC than the CRASH model. The CRASH model had a significantly higher AUC than the IMPACT model. Upon examination of figure 1, the KCMC model appear to have a higher sensitivity than the CRASH and IMPACT models.

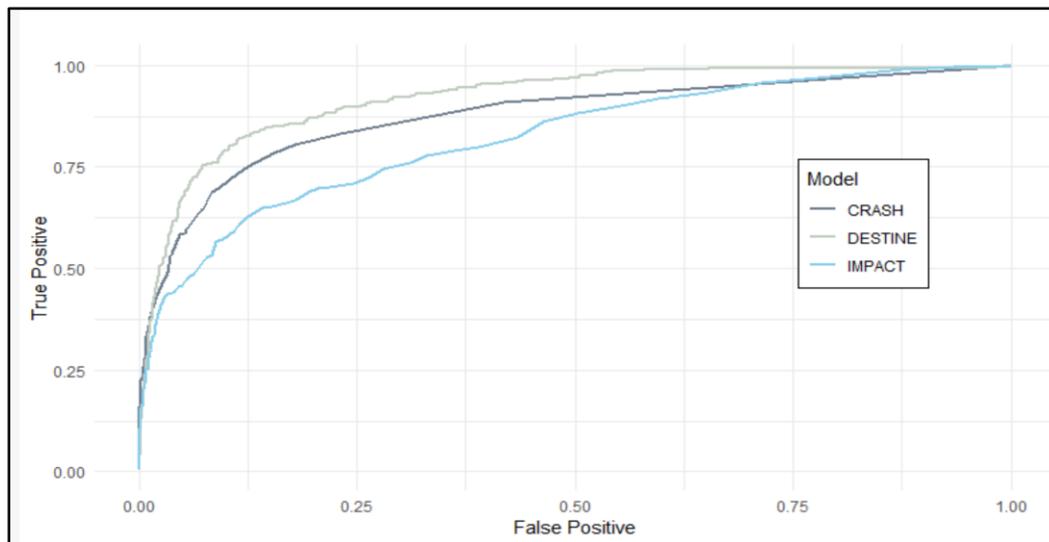Figure 6: A series of receiver operating characteristic curves comparing the discrimination of our model, CRASH, and IMPACT using the KCMC dataset.

We then tested the calibration of the three models using the KCMC dataset. The calibration was best for the CRASH model (Brier score = 0.061), followed by IMPACT (Brier score = 0.075) and the KCMC model (Brier score = 0.099). This pattern is reflected in the calibration plots. The

CRASH model calibration curve travels nearest to the y=x intercept (which signifies perfect

calibration) compared to IMPACT and KCMC models. The calibration plots also revealed the

CRASH model underestimates risk while the IMPACT and KCMC model overestimate risk.
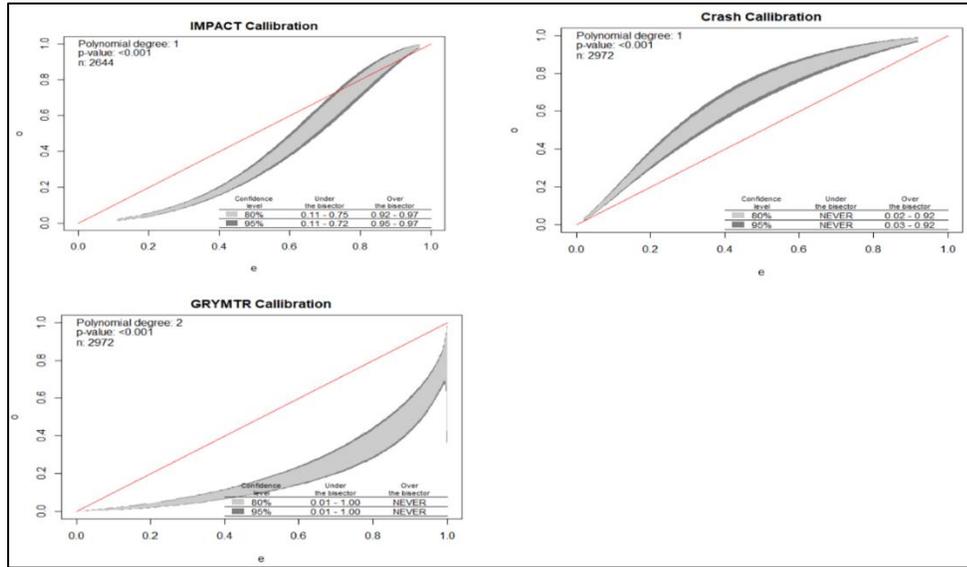
Figure 7: A series of calibration plots comparing the calibration of our model, CRASH, and IMPACT
using the KCMC dataset.

Table 10: Discrimination and calibration values for the IMPACT, CRASH and KCMC models.

Characteristics	KCMC Model	CRASH Model	IMPACT Model
Discrimination			
AUC (CI)	0.919 (0.902, 0.936)	0.876 (0.852, 0.90)	0.821 (0.793, 0.849)
Calibration			
Brier Score	0.099	0.061	0.075

Finally, we plotted the frequency of TBI patients presenting to KCMC by their predicted risk

based on the CRASH, IMPACT and KCMC models (Figure 3). The KCMC model appears to have

the greatest distribution and granularity of predicted risks. The CRASH model had the lowest

distribution of predicted risks.
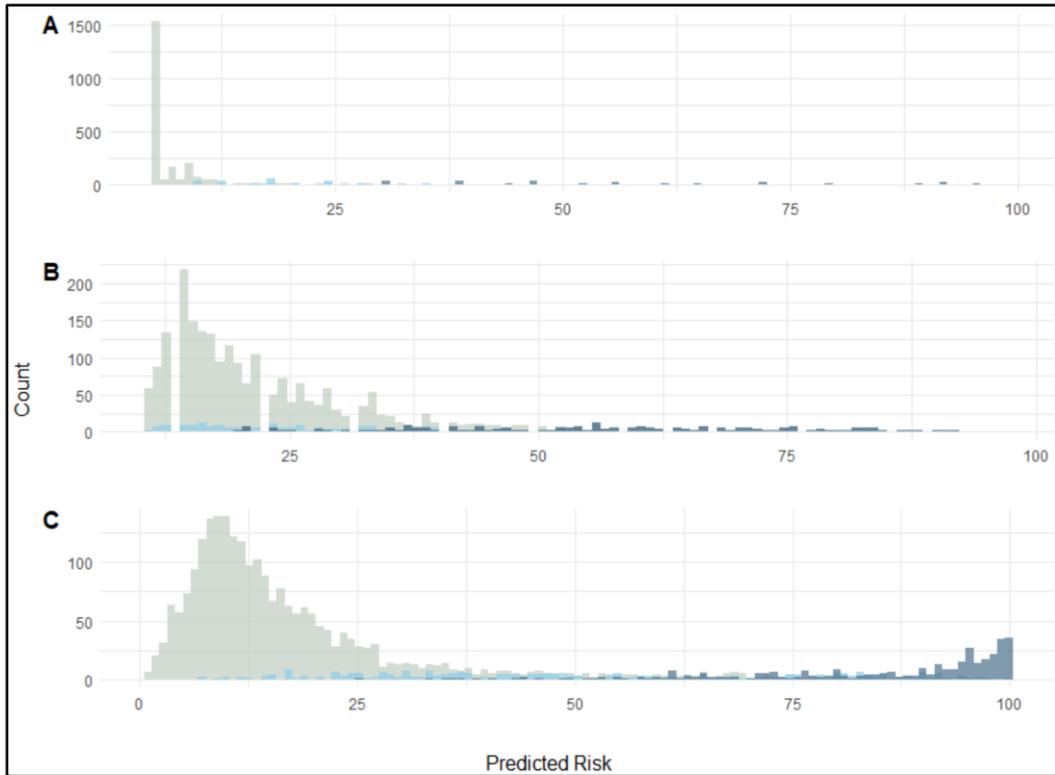
Figure 8: Histogram of the predicted for each patient in the KCMC dataset as calculated by (A) CRASH, (B) IMPACT and (C) our prognostic model.

3.4 Vignette experiment (IV)

Between the MNRH and MRRH locations, we enrolled 31 doctors to participate in this study (Table 10). Of the 31 participants, 16 completed the patient vignette with the CRASH risk calculator output, 15 received the patient vignette alone. The majority of doctors were male (N=24, 77%), residents (N=14, 45%), and working at MRRH (N=19, 61%). The most common medical specialty represented in the study, aside from intern doctors (N=11, 35%), was neurosurgery (N=8, 26%). We decided the randomization was successful as there was no significant difference between the groups in regard to gender, medical specialty, level of experience, importance of religion in life or location of experiment.

Table 11: Demographic and medical training background of participants in the study.

Category	Total	With Crash (Experiment)	No Crash (Control)	p-value
Participants				
Total participants	31	16	15	
Demographics				
Female	7	3	4	0.923
Age / sd	31 / 7.2	31.7 / 7	30.3 / 7.7	
Medical Specialty				0.272
Emergency Medicine	5	2	3	
General Surgery	6	5	1	
Intensivist	1	0	1	
Intern Doctor	11	4	7*	
Neurosurgery	8	5	3	
Level of Experience				0.372
Consultant	5	3	2	
Fellow	1	0	1	
Resident	14	9	5	
Intern	11	4	7	
Average experience (years)	5	5.4	4.5	
Importance of religious faith in your life				0.673
Very important	21	12	9	
Important	7	3	4	
Neutral	2	1	1	
Not important	0	0	0	
Not available	1	0	1	
Location of Experiment				0.821
Mbarara Regional Referral Hospital (MRRH)	19	9	10	
Mulago National Referral Hospital (MNRH)	12	7	5	
* Two interns in this group did not estimate prognosis: one was uncomfortable with the exercise and one did not understand.				

Table 11 reports the estimated risk for different clinical outcomes for the experiment (with

CRASH risk score) and control (without CRASH risk score) groups. Overall, participants were more

optimistic in estimating patient risk compared to the CRASH risk score for both acute and chronic outcomes. The experiment and control groups estimated the risk of inpatient mortality was 37% and 30% respectively while the CRASH estimate was 51.4%. The experiment and control groups estimated risk of unfavorable outcome at six months was 67% and 63% respectively compared to the CRASH estimate of 89.8%. When comparing the experiment and control groups, those with the clinical vignette plus the CRASH risk output reported higher risk estimates for each clinical outcome. These findings were not statistically significant. However, there were significant difference between the experiment and control groups when comparing across equal levels of training (Figure 8).

Table 12: The provider's estimated risk for each clinical outcome.

Clinical Outcome	With Crash (Experiment)	No Crash (Control)	p-value
Risk of inpatient mortality	37	30	0.440
Risk of death at 30 days	53	37	0.115
Risk of unfavorable outcome at six months	67	63	0.730
Risk of death at six months	57	43	0.195
Risk of permanently needing ventilator	14	9	0.253
*CRASH estimated 14 day mortality was **51.4% (42.8 - 59.8)** CRASH estimated risk of unfavorable outcome at 6 months was **89.8% (86.0 - 92.6)***			

Intern doctors in the experiment group estimated the risk on inpatient mortality at 44% compared to 14% for intern doctors in the control group (p-value = 0.037). Similarly, interns in the experiment group estimated the risk of death at six months was 56% compared to 27% for intern doctors in the control group (p-value = 0.071). There was a similar difference in risk estimates for intern doctors for risk of unfavorable outcome at six months however this finding was not significant (p-value = 0.217). Figure 8 shows that risk estimates by intern doctors with the CRASH risk output more closely mirrored risk estimates by more experienced residents and consultants.
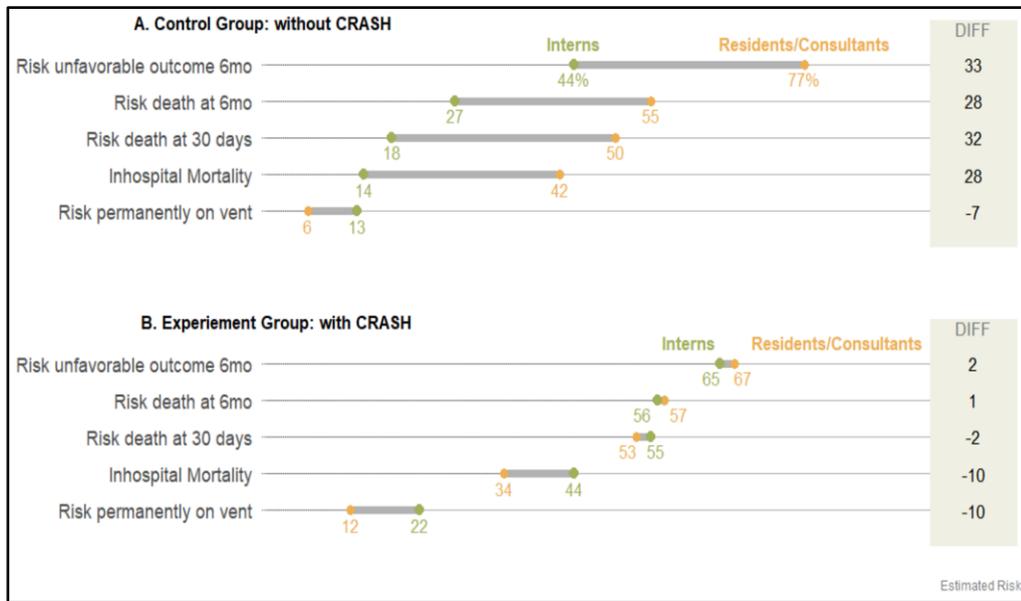
Figure 9: The participant's estimated risk for each clinical outcome if the participant had the CRASH risk output (A) and if the participant did not have the CRASH risk output (B).

The most common selected best next step for the patient was surgery (N=27, 87%). The participants estimated this hypothetical patient had a 60% chance of receiving a CT scan and 48% chance of receiving TBI surgical intervention within two days of admission. 29 participants reported considering patient prognosis and resources available when deciding the best course of action for the patient. Two participants reported considering resources only.

Table 13: Participant responses to follow-up questions pertaining to the clinical vignette. Responses are provided with all participants grouped together (total) and separated by experimental and control groups.

Responses to patient vignette			
Category	Total	With Crash (Experiment)	No Crash (Control)
Best next step			
Surgery	27	14	13
Intensive Care Unit	3	2	1
Observe	1	0	1
Comfort Measures	0	0	0
Estimate resources			

% get CT in 2 days	60	58	62
% get surgery in 2 days	48	47	49
What factors into decision			
Resources only	2	1	1
Prognosis only	0	0	0
Both	29	15	14

A total of 29 participants completed the follow-up survey on clinical decision making (Figure 9). The majority of participants agreed or strongly agreed with worrying about accurately assessing prognosis in head injury patients (N=22, 76%). The participants also agreed or strongly agreed with using patient data/literature (N = 26, 90%) and prior experience (N= 21, 72%) to estimate patient risk. The participants endorsed a strong sense of obligation to offer surgery if there was a chance at survival (N= 25, 86%). Only 10 participants (34%) agreed or strongly agreed that there was adequate data to estimate patient prognosis. Also, only 4 participants (14%) agreed or strongly agreed that they were willing to not offer surgery even if it meant the patient would die.
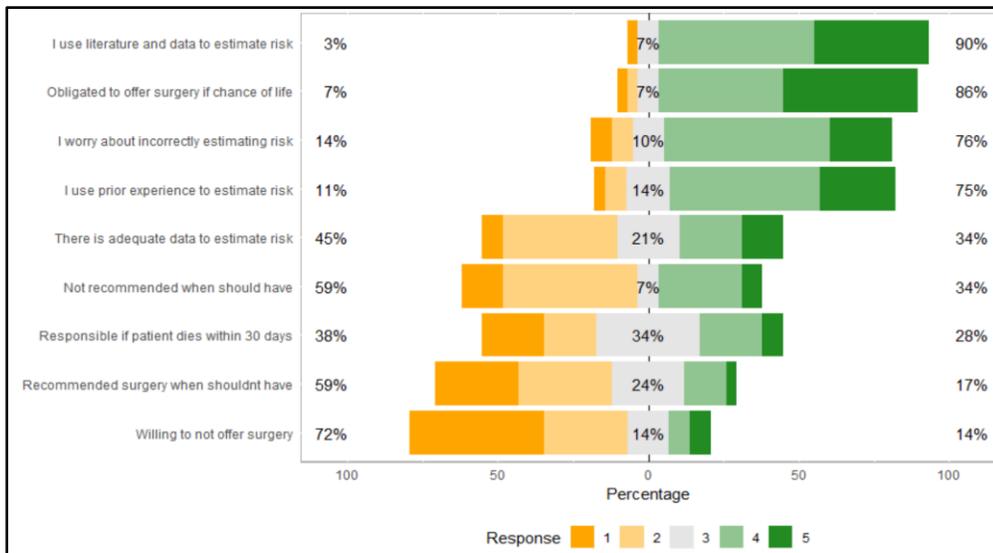
Figure 10: Participant responses to a follow-up survey on clinical decision making. The likert scale included: Strongly Disagree = 1, Disagree = 2, Neutral = 3, Agree = 4, Strongly Agree = 5.

3.5 Feasibility study (V)

In total, we included 29 health care providers in this study. Of the 29, 27 providers completed the potential use cases for prognostic modeling survey and 10 completed the application testing. The majority of participants completing the survey were male (N=21), resident doctors (N=14) and involved in neurosurgery (N=7). Emergency medicine residents (N=4) and intern doctors (N=4) were the most common health care professionals to complete the application testing.

Table 14: Demographic and medical training background for participants in the study. We provided data on those completing the survey and application testing.

Category	Overall	Use Cases Survey	Application Testing
Participants			
Total participants	29	27	10
Demographics			
Female	7	6	3
Age / sd	29.7 / 5.1	29.9 / 5.1	28.8 / 5.1
Medical Specialty			

41

Emergency Medicine	5	5	4
General Surgery	6	6	1
Intern Doctor	10	9	4
Neurosurgery	7	7	0
Nurse	1	0	1
Level of Experience			
Consultant	3	3	1
Fellow	1	1	0
Resident	14	14	4
Intern	10	9	4
Nursing School	1	0	1
Average experience (years)	4.3	4.5	3.1

Figure 1 illustrates the responses to the potential use cases of TBI prognostic modeling survey. The top four applications of prognostic modeling, endorsed as important or very important, were deciding which patient to send to the ICU (89% of respondents), which patient needed a decompressive craniotomy (93% of respondents), which patient needed surgery (93% of respondents) and in communicating with patients or families (93% of respondents). The least important use case of prognostic modeling was deciding which patient treatments should be withdrawn.
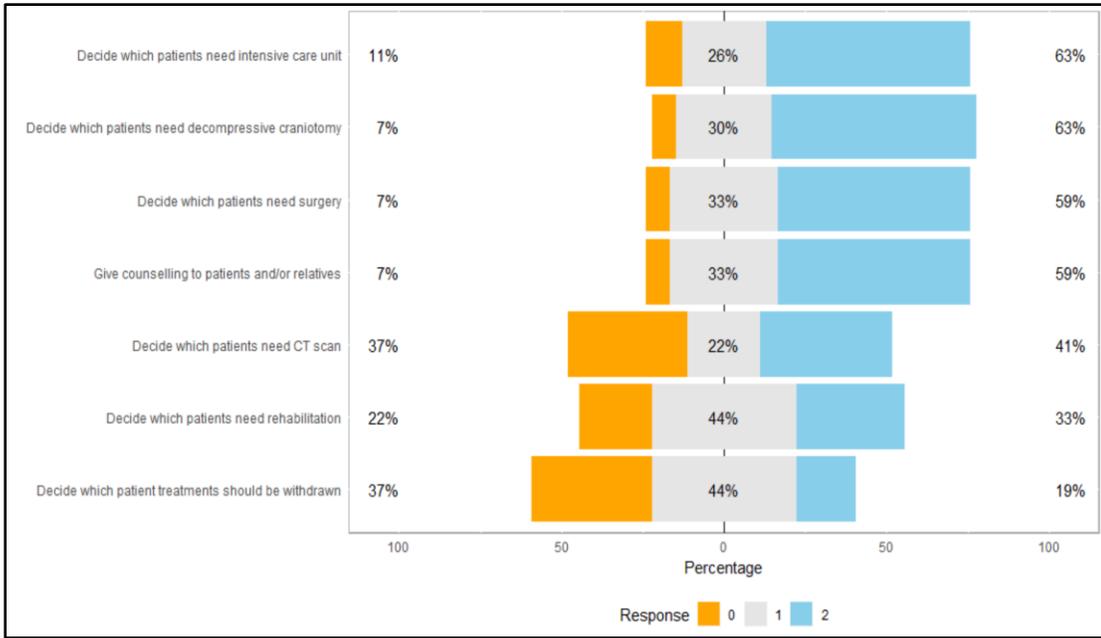
Figure 11: We asked participants "In what decisions could a TBI prognostic risk calculator provide value?" We provided a likert scale which included very important - 2, important - 1, not important - 0.

Next, we allowed the participants to test the application. The average time for one patient entry was two minutes and 22 seconds, slightly faster when using the application as a laptop compared to a tablet. There was considerable improvement in entry time between the first and second entry (23 second reduction. There were few errors committed or questions when using the application.

Table 15: Results from beta-testing of decision support application. We provided the data as overall, application testing using a laptop, and application testing using a tablet.

Variable	Overall	Laptop	Tablet
Entry Time (Seconds)			
Overall average	142	139	145
First entry	153	144	167
Second entry	130	131	130
User Experience			

Errors per entry	0.2	0.0	0.3
Questions per entry	0.2	0.1	0.3

After using the application, we asked the participants to complete the modified Computer System Usability Questionnaire. All participants were satisfied with the application and rated the application as quick and simple. One participant was not comfortable using the application as this participant disagreed with the estimated risk score.
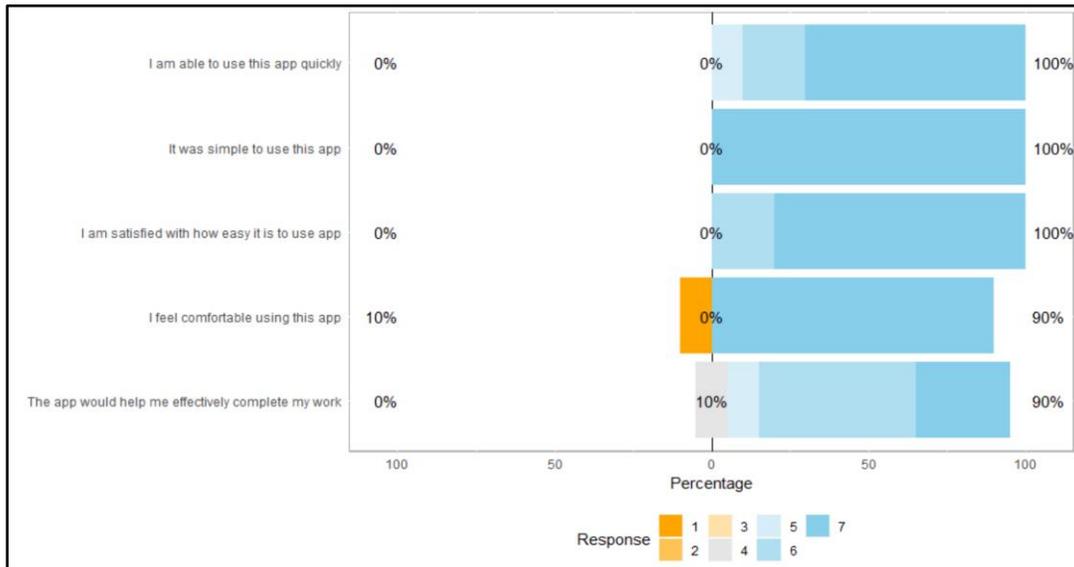

Figure 12: Participant responses to the Computer System Usability Questionnaire. We provided a likert scale ranging from strongly agree - 7 to strongly disagree - 1.

4. Discussion

4.1 Resource limitation

To ensure limited resources are appropriately distributed based on injury severity, providers must accurately assess patient risk. Providers assess risk by using data from diagnostic tests, literature, experience, and clinical findings. The shortcomings in risk assessment are magnified in settings with limited resources. The doctors in this study estimated a significant proportion of patients would not receive CT or surgeries within two days even if needed. Further, all study participants considered availability of resources when making treatment decisions. These findings underscore the importance of prudent resource allocation to maximize clinical impact.

We found surgery was associated with reduced HRs for poor outcome for all TBI severity groups. Patients with moderate TBIs receiving surgery had the greatest reduction in HR, followed by mild and then severe TBIs. When faced with multiple patients needing surgery, surgical teams in low resources settings must reconcile severity of injury with potential for good outcome. Previous studies present mortality rates for patients receiving or not receiving surgery, but direct comparisons were not possible due to data and method limitations.[11] Our finding of an 83% reduction in HR for moderate patients receiving surgery could support optimization of triage practices and prudent resource allocation for the low resource setting.

4.2 Current state of decision making

Participants in the study were overly optimistic in their risk estimation, particularly inexperienced doctors. However, we found that interns presented with the CRASH risk output estimated risk more closely to experienced doctors. Shortages of experienced doctors in LMICs has resulted in increased use of task shifting or the delegation of decision making to less specialized health workers. Intern doctors, despite being less than one year removed from medical school, are heavily involved in the inpatient care in LMIC hospitals. In Uganda, interns receive the patient, start

45

initial treatment plans, and activate care pathways by contacting their supervising resident. These decisions are informed by the perceived severity of injury. Our findings highlight the potential for risk calculators to augment risk assessment and empower inexperienced providers.

Machine learning based prognostic models have potential to increase efficiency and precision of decision making for neurosurgical conditions including TBI.[3] In a survey of HIC and LMIC physicians who treat patients with head injury, accurate prognostic information was considered very important for the following: deciding to undertake a decompressive craniotomy, deciding who should receive intensive care, communication with patient families and which patients' treatment should be withdrawn.[29] A previous study exploring how a prediction system influences intensity of management based on injury severity supports the potential for a more rational use of hospital resources[17]. The researchers found more intensive management for patients with moderate or good prognosis. Conversely for those with a worse prognosis, the frequency of using resources for intensive management decreased by 39%. These findings also support the prediction of poor outcome rather than prediction of good outcome. In resource poor settings, prediction of poor outcome may result in more prudent use of allocation of limited resources. Conversely, prediction of good outcome may increase resources dedicated to a patient, problematic in a setting devoid of resource.

4.3 LMIC based model (pros and cons)

This study used ML to produce a TBI prognostic model which can identify a patient risk for inpatient poor recovery upon hospital admission. The importance of successfully applying ML to TBI prognostic modeling is two-fold. First, data constraints have limited prognostic models using ML to predominately high income and upper middle income countries. Our access to the TBI patient registry collected by Staton et al. provided a sufficiently robust dataset to apply ML techniques.[10] Second, previous studies have found ML to outperform traditional statistical methodologies and, at times, clinical experts.[34] The best performing model in this study had an AUC 86.5 (CI: 85.6, 87.4).

As comparison, the only ML-based TBI prognostic model identified in Sender's et als review had an AUC of 85.527.[3] Their model used retrospectively collected data to predict in-patient mortality compared to our model which used prospectively collected data to predict poor recovery.

The Bayesian generalized linear model, the model with the best performance in this study, is not considered a black box model since the weight of each model predictor can be extracted (Figure 5). The GCS total score, oximetry, mechanism of injury, as well as the availability of surgery and ICU beds were the main predictors of a good recovery for TBI patients admitted to KCMC. Our findings are aligned with the previous works reported in the literature about the importance of GCS, vital signs, and characterization of injury.[55,56]

4.4 Potential implementation

For risk calculators to achieve meaningful use in this setting, the implementation must be culturally and contextually relevant. In high resource settings, providers endorse the potential for clinical decision support systems to improve communication with families and justify treatment decisions.[57] Conversely, providers have expressed concern with these technologies as a final decision maker, particularly with end of life care.[57] Our study captured a sense of duty among providers to perform surgery if there was a chance of survival and an opposition to withholding care. These are important contextual considerations when designing an implementation of decision support technology in this setting. Future research should explore which decisions on the care continuum can a risk calculator support, what happens when providers disagree with the risk calculator score, and how they would communicate the risk score to patients and families.

Physician buy-in is paramount for technology adoption. In a survey of physicians about prognostic models, most physician (67%) reported a more accurate prognostic model would change their current patient management.[29] We can encourage physician acceptability toward prognostic modeling by maintaining transparency when possible. Some ML approaches, however, generate "black box models" meaning the modeling techniques used to build relationships among the

47

variables examined is not easily interpretable. A black box model is likely to be met with skepticism from clinicians, slowing clinical adoption.[58] Models based on neural networks, deep learning, and random forest, for example, create intermediary steps hard for the end-user to understand.

An example application of a prognostic model for the hospital in this study is the decision to send a patient to the ICU (intensive management) or surgical ward (less intensive management). We found 53% of patients who ultimately were classified as having poor outcome at discharge had been transferred from the surgical ward to the ICU prior to discharge compared to only 21% of patients classified as having a good outcome at discharge. This finding means patients were assessed in the ED, deemed sufficiently stable for care on the surgical ward, and later required transfer to the ICU for more aggressive care. The significant difference between the good and poor recovery group requiring this transfer represents an opportunity to improve initial triaging decisions in the ED. This example is a potential application for prognostic models to support complex clinical decision making in hospital settings with a limited number of highly-skilled healthcare providers.

4.5 Statistical consideration and lessons learned

Considerations for prognostic model development

A major strength of this study was the quality of data used to develop the prognostic model. The registry data used in this study was prospectively collected, had little missing values, and included over 3,000 patients. The missing values which were present were estimated statistically using the MICE package in R. Perel et al.'s systematic review on TBI prognostic models found three quarters of the models included less than 500 patients and less than 15% of models handled missing data statistically.[59]

There was a low volume of patients with a poor recovery in the dataset (14.4%). Our use of the SMOTE technique to handle the imbalance during the training phase of the models was intended to overcome this limitation. Finally, we included the receipt of TBI surgery as a prognostic variable. The decision to operate is not a neutral decision and thus may introduce bias into our model. Also,

we do not know if patients receiving TBI surgery, and having a good outcome, would have had a good outcome without surgical intervention. Despite these potential limitations, we included this variable in order to estimate patient prognosis with or without surgical intervention. This novel approach to including downstream treatment variables may produce a more understandable risk estimate by allowing providers to see the change in outcome depending on receipt of an intervention.

Considerations for survival analysis

The lessons learned by applying a survival analysis, including KM plots and Cox models, to our data has important clinical and statistical considerations. First, our data exhibited significant non-proportionality, the effect of surgery varied depending on day of hospital discharge and TBI severity. Statistically, this finding required the use of time interaction terms in the Cox model to avoid erroneous inferences from the data.[49,60] For example, severe TBI patients receiving surgery, compared to moderates and mild, had the lowest HR when a time interaction term was not included. This finding was the opposite of the association seen when we used time interaction terms. Clinically, the non-proportionality of our data present an interesting finding warranting further investigation. For moderate TBIs, the positive association of surgery on outcomes flips at HD eight. This finding was not seen in mild and severe TBI patients. Patients with a moderate TBI may be more prone to clinical deterioration compared to mild and severe TBI patients with more confidently expected good and poor outcomes respectively. This potential conclusion necessitates further research surrounding inpatient care for this specific patient population.

4.6 Next steps

Our long-term vision is improve clinical care for TBI especially in LMIC settings. To achieve this ambitious goal, there remains a number of statistical, clinical and implementation next steps. Statistically our machine-learning based decision support tool will be adapted to and learn from multiple settings at varying levels of care. This step will both externally validate our model using data from different hospitals and countries and enrich the dataset from which the model is trained. This

enrichment will further improve the performance of the model.

While this work was novel by exploring the feasibility to implement a TBI decision support tool in the low resource setting, additional implementation research is needed. We plan on having providers test the application while rounding on their patients to replicate a more realistic test environment.

Clinically, we must continue to explore how decision support tool influence provider decision making at the patient bedside. We will repeat the vignette experiment from this study with a larger study population. Additionally, we will perform effectiveness testing to investigate the impact of decision support tools on risk estimation and care of real TBI patients.

The work of this thesis, as well as previous work produced by the Duke Global Neurosurgery and Neurology division, will go towards an NIH RO1 submission. This funding would support the execution of these next steps in multiple LMICs.

4.7 Limitations

4.7.1 Survival analysis (I)

A robust statistical analysis on registry data from a low resource setting warrants careful consideration of limitations. First, the TBI registry used in this study did not include the GCS at the time of surgery because collecting this information was not always readily available to the research team. The type of surgery was also not documented in the database, however, the general surgery staff described performing mostly burr holes and some hemicraniectomies. The registry included less patients who suffered severe and moderate TBI compared to mild injuries. The smaller sample size limited our ability to draw inferences from the survival analyses at later HDs. We did not have data on the actual injury (e.g. computed tomography [CT] results) to understand severity of TBI other than admission GCS. We also did not have information on other injuries which might have caused poor outcomes other than TBI. Finally, since our study included over three years of data, there is

potential for patient outcomes to change over time if the quality of surgical care significantly changed. We tested for this potential limitation by comparing survival curves for patients enrolled early and late in the study period. The curves appeared similar suggesting no basis for this limitation.

4.7.2 Prognostic model (II)

This study did have some limitations. First, there was a low volume of patients with a poor recovery in the dataset (14.4%). This imbalance was also the reason for classifying GOSe scores of 1-6 as poor recovery. Previous work has used GOSe scores of 1-4 for poor recovery[15,29]. The incorporation of more data, and more patients with poor recovery should help increase the specificity levels achieved by our models. Next, the generalizability of our results is also limited due to our inability to externally validate our model. The model presented has potential to be suitable to other contexts but more data from different sources are necessary to strengthen the potential presented here.

4.7.3 Three model comparison (III)

The outcome measure used for the external validation of CRASH and IMPACT was discharge GOS. This outcome measure was different the outcome used in the development of CRASH and IMPACT. As result, the performance metrics obtained in this study may underestimate the performance of the two models had the outcome been the same. However, the limited research and clinical infrastructure in many low resource settings makes obtaining six month follow-up outcome data a challenge.

4.7.4 Vignette experiment (IV)

The limitations of this study include the sample size and heterogeneity of participants in the study. The final sample size was 31 which decreased further when subsetting the data by level of experience. Although the findings were significant when comparing the experimental and control groups for interns, the small sample size warrants further research to confirm these findings in a

larger population. In an effort to increase sample size, we included providers at two different hospital and from different medical specialties. Providers from different medical specialties likely weigh clinical findings differently when estimating patient risk. Also, these providers may use different decision algorithms when determining best next steps for patients. We believed the risk of this limitation were mitigated by the randomization procedure and the level of task sharing in this setting.

4.7.5 Feasibility study (V)

The study experienced several limitations. First, we performed the testing of the application in quite, private settings (e.g. doctors office, empty patient room, etc). Thus the captured entry time and errors likely underestimate the actual values when used as intended at patient patient bedside. Due to time constraints, we were unable to test the application with neurosurgeons. The majority of application testers were emergency medicine residents and interns. At these hospitals, however, the doctors who mostly triage TBI patients initially are the interns and emergency medicine resident. These providers then call the neurosurgeons or general surgeons. Finally, due to connectivity challenges, we were unable to test the application as a mobile application on smartphone. Testing should be repeated with smartphones since these devices were universally purported as the device of choice to access the application.

5. Conclusions

While fundamental efforts to increase hospital capacity and prevent TBIs are needed, the devastating burden of TBI in LMICs demand innovative solutions which can have immediate impact with the resources in place. One such solution is a TBI decision support tool leveraging the strengths of modern modeling techniques. While a decision support tool is not a panacea for TBI hospital care, it can help providers root life or death decisions in objective data. It is my hope that this work can help propel the field of decision support technologies in low resources settings from an academic exercise to a clinical reality.

Appendix. Survey tool for studies IV and V

A 56-year old male presents to the casualty unit as the unrestrained driver in a motor vehicle collision just prior to arrival.

PMH: Unknown, no known medications.
Patient vitals: RR: 16, BP: 142/86, pulse: 105 bpm, SpO2: 91% on room air
Exam: The patient has surface abrasions to the head and lower extremities. GCS: M-2, V-2, E-2, pupils: Lt- non reactive and dilated, Rt- normal. Airway is patent, gag reflec intact.
CT Brain: Demonstrates a 1 cm acute subdural hematoma with 8 mm of midline shift. There is also scattered traumatic subarachnoid hemorrhage.
Labs: Glucose is 90mg/dL, coagulation studies within normal limits, and hemoglobin is 9g/dL.

TBI risk calculators (like pictured below) can estimate a patient prognosis upon patient presentation to the hospital. The output below is meant to support a doctor's decision. The risk estimate below is **with the hematoma evacuated**.

If asked by a trainee, what is the best next step?

☐ Take to the theatre (operating room)
☐ Admit to the ICU for aggressive medical management
☐ Initiate comfort measures
☐ Observe patient

If you proceed with surgery, what is your best guess that given 100 similar cases, how many would...

a. survive the hospitalization? _____

b. survive to 6 month _____

c. survive 30 days? _____

d. be able to communicate and
independently perform activities of daily _____
living (ADLs) within the next 6 months?

e. remain permanently on a ventilator _____

If needed, and considering costs and availability, how many patients out of 100 will receive a CT within 2 days of admission? _____

If needed, and considering costs and availability, how many patients out of 100 will receive surgery within 2 days of admission? _____

Which of the following are you considering when making your decision?

☐ Patient prognosis
☐ Resources available
☐ Both

Demographic Data

Please provide the following:
1. Medical Specialty_____
2. Gender   _____
3. Race/Ethnicity   _____
4. Year of Birth   _____
5. What is your religion, if any? _____
6. Year of medical school graduation   _____
7. In which country is your primary practice? _____

Please check one for each question below:
8. Level of training ☐ intern ☐ resident ☐ fellow ☐ consultant-no residency
   ☐ consultant- residency complete

9. Do you supervise residents/fellows? ☐ Yes ☐ No

10.   How important is religious faith in your life?
☐ Very important ☐ Important ☐ Neutral ☐ Unimportant ☐ Very
unimportant

11.   Which best describes your current employer (Check one)?
   ☐ Independent (Private), physician-owned practice
   ☐ employed by a university hospital/health system
   ☐ employed by a non-university hospital/health system (government)

Follow up Questions
How would you rate the following (specific to severe TBI you have treated in your lifetime)?

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Looking back on patients with severe TBI, I have **recommended** surgery in patients that I should not have.					
Looking back on patients with severe TBI, I have **not recommended** surgery in patients that I should have.					
I am responsible if a patient dies within 30 days of any surgery I recommend.					
As a physician, I am willing to **not offer** surgery even if it means a patient will die.					
Regardless of prognosis, I am obligated to offer surgery if there is a chance the patient will survive.					
I worry about incorrectly estimating prognosis.					
I use available literature and data to make a prognosis.					
I use prior experience with patients to make a prognosis.					
I feel that there is adequate quality data to base prognostication.					

Please circle your response:
Have you previously heard of TBI prognostic risk calculators such as CRASH or IMPACT?   Yes or No

A TBI prognostic risk calculators provide the chance of morbidity or mortality for head injury patients. The calculator uses data such as vital signs and physical exam findings from prior studies to estimate the patient's risk for a bad outcome. It is not meant to substitute clinical judgment however is intended to support it.

In what decisions could a TBI prognostic risk calculator provide value?

	Very important	Important	Not Important
To decide which patients need decompressive craniotomy			
To decide which patients need Intensive Care Unit			
To give counselling to patients and/or relatives			
To decide in which patients treatments should be withdrawn			
To decide which patients need surgery			
To decide in which patients CT scan should be done			
To decide which patients need rehabilitation			

Comments: _____

Mobile Application Testing

Please enter the following into the application. Please think out loud as you use the application.

Age: 45

Gender: Female

Intention of Injury: Unintentional

Alcohol Involved: Yes

Temperature: 39

Resp Rate: 24

Pulse: 105

Systolic BP: 132

Diastolic BP: 85

Pulse Oxygen: 93

Glasgow Coma Score, Eye: To pain

Glasgow Coma Score, Verbal: Inappropriate words

Glasgow Coma Score, motor: Withdraws from pain

Pupils equal: No

Casualty Department Disposition: Surgery Ward

Went from Surgery to ICU: Yes

Mechanism: Car

Can you explain the output for me?

Mobile Application Testing

Please enter the following into the application. Please think out loud as you use the application.

Age: 21

Gender: Male

Intention of Injury: Unintentional

Alcohol Involved: No

Temperature: 37

Resp Rate: 20

Pulse: 92

Systolic BP: 122

Diastolic BP: 93

Pulse Oxygen: 93

Glasgow Coma Score, Eye: To speech

Glasgow Coma Score, Verbal: Confused

Glasgow Coma Score, motor: Localizes

Pupils equal: No

Casualty Department Disposition: Surgery Ward

Went from Surgery to ICU: No

Mechanism: Motorcycle

Can you explain the output for me?

Mobile Application Feedback (Please circle one number for each of the following)

Overall, I am satisfied with how easy it is to use this app: Strongly agree 1   2   3   4   5   6   7 Strongly disagree

It was simple to use this app: Strongly agree 1   2   3   4   5   6   7 Strongly disagree

The app would help me effectively complete my work: Strongly agree 1   2   3   4   5   6   7 Strongly disagree

I am able to use this app quickly: Strongly agree 1   2   3   4   5   6   7 Strongly disagree

I feel comfortable using this app: Strongly agree 1   2   3   4   5   6   7 Strongly disagree

Mobile Application Testing

Tablet or laptop _____

Entry#1 Time to complete one entry _____

Entry#1 number of errors _____

Entry#1 number of questions _____

Tablet or laptop _____

Entry#2 Time to complete one entry _____

Entry#2 number of errors _____

Entry#2 number of questions _____

Mobile Application Feedback

What are the strengths of using this tool in clinical practice?

What are the weaknesses of using this tool in clinical practice?

Can you think of anything that can be added or removed to improve the application?

What would be the method to use this tool at the hospital? Computer, laptop, smart phone, tablet, SMS?

Who is the best user of this tool?

Any other suggestions?

References

1. James, S. L. *et al.* Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology* **18**, 56–87 (2019).

2. Vaca, S. D. *et al.* Temporal Delays Along the Neurosurgical Care Continuum for Traumatic Brain Injury Patients at a Tertiary Care Hospital in Kampala, Uganda. *Neurosurgery* doi:10.1093/neuros/nyy004

3. Senders, J. T. *et al.* Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. *World Neurosurgery* **109**, 476–486 (2018).

4. Perel, P., Edwards, P., Wentz, R. & Roberts, I. Systematic review of prognostic models in traumatic brain injury. *BMC Medical Informatics and Decision Making* **6**, (2006).

5. Young, N. H. & Andrews, P. J. D. Developing a Prognostic Model for Traumatic Brain Injury—A Missed Opportunity? *PLOS Medicine* **5**, e168 (2008).

6. Stone, J. L. *et al.* Acute subdural hematoma: direct admission to a trauma center yields improved results. *J Trauma* **26**, 445–450 (1986).

7. Carney, N. *et al.* Guidelines for the Management of Severe Traumatic Brain Injury, Fourth Edition. *Neurosurgery* **80**, 6–15 (2017).

8. Smart, L. R. *et al.* Severe Traumatic Brain Injury at a Tertiary Referral Center in Tanzania: Epidemiology and Adherence to Brain Trauma Foundation Guidelines. *World Neurosurgery* **105**, 238–248 (2017).

9. Silva, D. *et al.* Patient outcome after traumatic brain injury in high-, middle- and low-income countries: analysis of data on 8927 patients in 46 countries. *Int J Epidemiol* **38**, 452–458 (2009).

10. Staton, C. A. *et al.* A prospective registry evaluating the epidemiology and clinical care of traumatic brain injury patients presenting to a regional referral hospital in Moshi, Tanzania: challenges and the way forward. *International Journal of Injury Control and Safety Promotion* **24**, 69–77 (2017).

11. Eaton, J., Hanif, A. B., Mulima, G., Kajombo, C. & Charles, A. Outcomes Following Exploratory Burr Holes for Traumatic Brain Injury in a Resource Poor Setting. *World Neurosurgery* **105**, 257–264 (2017).

12. Aenderl, I., Gashaw, T., Siebeck, M. & Mutschler, W. Head Injury-A Neglected Public Health Problem: A Four-Month Prospective Study at Jimma University Specialized Hospital, Ethiopia. *Ethiopian Journal of Health Sciences* **24**, 27 (2014).

13. Tran, T. M. *et al.* Burden of Surgical Conditions in Uganda: A Cross-sectional Nationwide Household Survey. *Annals of Surgery* **266**, 389–399 (2017).

14. Wheble, J. L. C. & Menon, D. K. TBI—the most complex disease in the most complex organ: the CENTER-TBI trial—a commentary. *Journal of the Royal Army Medical Corps* **162**, 87–89 (2016).

15. Maas, A. I., Stocchetti, N. & Bullock, R. Moderate and severe traumatic brain injury in adults. *The Lancet Neurology* **7**, 728–741 (2008).

16. Lingsma, H. F., Roozenbeek, B., Steyerberg, E. W., Murray, G. D. & Maas, A. I. Early prognosis in traumatic brain injury: from prophecies to predictions. *The Lancet Neurology* **9**, 543–554 (2010).

17. Fuller, A., Tran, T., Muhumuza, M. & Haglund, M. M. Building neurosurgical capacity in low and middle income countries. *eNeurologicalSci* **3**, 1–6 (2016).

18. Ramesh, A. *et al.* Challenges and Solutions for Traumatic Brain Injury Management in a Resource-Limited Environment: Example of a Public Referral Hospital in Rwanda. *Cureus* (2014). doi:10.7759/cureus.179

19. MacLeod, J. B. A., Kirton, O. C. & Maerz, L. L. Surgical intensivist and global critical care: is there a role? *Trauma Surg Acute Care Open* **1**, e000058 (2016).

20. Jain. Development of post graduate program in emergency medicine in India: Current status, scope and career pathways. Available at: http://www.astrocyte.in/article.asp?issn=2349-0977;year=2014;volume=1;issue=3;spage=218;epage=221;aulast=Jain. (Accessed: 6th November 2018)

21. Khamlichi, A. E. African Neurosurgery: Current Situation, Priorities, and Needs. *Neurosurgery* **48**, 1344–1347 (2001).

22. Rosman, J. *et al.* Is There a Shortage of Neurosurgeons in the United States? *Neurosurgery* **73**, 354–366 (2013).

23. Bullock, R. *et al.* Guidelines for the management of severe head injury. Brain Trauma Foundation. *Eur J Emerg Med* **3**, 109–127 (1996).

24. Teasdale, G. M., Pettigrew, L. E. l., Wilson, J. t. L., Murray, G. & Jennett, B. Analyzing Outcome of Treatment of Severe Head Injury: A Review and Update on Advancing the Use of the Glasgow Outcome Scale. *Journal of Neurotrauma* **15**, 587–597 (1998).

25. Honeybul, S., O'Hanlon, S. & Ho, K. M. Decompressive Craniectomy for Severe Head Injury: Does an Outcome Prediction Model Influence Clinical Decision-Making? *Journal of Neurotrauma* **28**, 13–19 (2011).

26. Subaiya, S., Roberts, I., Komolafe, E. & Perel, P. Predicting intracranial hemorrhage after traumatic brain injury in low and middle-income countries: A prognostic model based on a large, multi-center, international cohort. *BMC Emergency Medicine* **12**, (2012).

27. Zador, Z., Sperrin, M. & King, A. T. Predictors of Outcome in Traumatic Brain Injury: New Insight Using Receiver Operating Curve Indices and Bayesian Network Analysis. *PLoS One* **11**, (2016).

28. Steyerberg, E. W. *et al.* Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics. *PLoS medicine* **5**, e165 (2008).

29. Frattalone, A. R. & Ling, G. S. F. Moderate and Severe Traumatic Brain Injury: Pathophysiology and Management. *Neurosurgery Clinics of North America* **24**, 309–319 (2013).

30. Perel, P. *et al.* Prognosis following head injury: a survey of doctors from developing and developed countries. *Journal of Evaluation in Clinical Practice* **13**, 464–465 (2007).

31. Bandiera, G. *et al.* The Canadian C-Spine rule performs better than unstructured physician judgment. *Annals of Emergency Medicine* **42**, 395–402 (2003).

32. Obermeyer, Z. & Emanuel, E. J. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine* **375**, 1216–1219 (2016).

33. Reddy, C. K. & Aggarwal, C. C. *Healthcare data analytics.* **36**, (CRC Press, 2015).
\
34. Shi, H.-Y., Hwang, S.-L., Lee, K.-T. & Lin, C.-L. In-hospital mortality after traumatic brain injury surgery: a nationwide population-based comparison of mortality predictors used in artificial neural network and logistic regression models. *Journal of Neurosurgery* **118**, 746–752 (2013).

35. MRC CRASH Trial Collaborators. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* **336**, 425–429 (2008).

36. Britton, A. *et al.* Threats to Applicability of Randomised Trials: Exclusions and Selective Participation. *J Health Serv Res Policy* **4**, 112–121 (1999).

37. Vissoci, J. R. N., Rocha, T. A. H., Elahi, C., Haglund, M. M. & Staton, C. PW 2468 Development of a predictive modeling approach to analysis of patient outcomes related to traumatic brain injury at a emergency center in a low income country. *Injury Prevention* **24**, A191–A192 (2018).

38. Vandenbroucke, J. P. *et al.* Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. *PLOS Medicine* **4**, e297 (2007).

39. Pocock, S. J., Clayton, T. C. & Altman, D. G. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *The Lancet* **359**, 1686–1689 (2002).

40. Stel, V. S., Dekker, F. W., Tripepi, G., Zoccali, C. & Jager, K. J. Survival Analysis I: The Kaplan-Meier Method. *NEC* **119**, c83–c88 (2011).

41. van Dijk, P. C., Jager, K. J., Zwinderman, A. H., Zoccali, C. & Dekker, F. W. The analysis of survival data in nephrology: basic concepts and methods of Cox regression. *Kidney International* **74**, 705–709 (2008).

42. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC medicine* **13**, 1 (2015).

43. Luo, W. *et al.* Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. *Journal of Medical Internet Research* **18**, e323 (2016).

44. Vergouwe, Y., Steyerberg, E. W., Eijkemans, M. J. C. & Habbema, J. D. F. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *Journal of Clinical Epidemiology* **58**, 475–483 (2005).

45. Agarwal, S. *et al.* Decision-support tools via mobile devices to improve quality of care in primary healthcare settings. *Cochrane Database of Systematic Reviews* (2018). doi:10.1002/14651858.CD012944

46. Bowen, D. J. *et al.* How We Design Feasibility Studies. *American Journal of Preventive Medicine* **36**, 452–457 (2009).

47. Kortbeek, J. B. *et al.* Advanced Trauma Life Support, 8th Edition, The Evidence for Change: *The Journal of Trauma: Injury, Infection, and Critical Care* **64**, 1638–1650 (2008).

48. R Core Team. *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.* (2018).

49. Therneau, T., Crowson, C. & Atkinson, E. Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model. 27

50. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic Minority Over-sampling Technique. *J Artif Intell Res* **16**, 321–357 (2002).

51. Fawcett, T. An Introduction to ROC Analysis. *Pattern Recogn. Lett.* **27**, 861–874 (2006).

52. Carpenter, J. & Bithell, J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine* **19**, 1141–1164 (2000).

53. DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **44**, 837–845 (1988).

54. Kuhn, M. Caret: Building Predictive Models in R Using the caret Package. *J Stat Softw* **28**, 1–26 (2008).

55. Sadaka, F., Patel, D. & Lakshmanan, R. The FOUR Score Predicts Outcome in Patients After Traumatic Brain Injury. *Neurocrit Care* **16**, 95–101 (2012).

56. Kalpakis, K. *et al.* Permutation entropy analysis of vital signs data for outcome prediction of patients with severe traumatic brain injury. *Computers in Biology and Medicine* **56**, 167–174 (2015).

57. Weber, S. Impacts of Clinical Decision Support Technology on Nursing and Medical Practice in U.S. Critical Care. *Canadian Journal of Nursing Informatics* **5**, (2011).

58. Holmes, D. E., Jain, L. C. & Beaulieu-Jones, B. *Machine Learning for Structured Clinical Data.* **137**, (Springer International Publishing, 2018).

59. Perel, P., Edwards, P., Wentz, R. & Roberts, I. Systematic review of prognostic models in traumatic brain injury. *BMC Med Inform Decis Mak* **6**, 38 (2006).

60. Bellera, C. A. *et al.* Variables with time-varying effects and the Cox model: Some statistical concepts illustrated with a prognostic factor study in breast cancer. *BMC Med Res Methodol* **10**, 20 (2010).