UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Development of a lesion localisation tool to improve outcome prediction in Traumatic Brain Injury patients

Afonso Santos Pourfarzaneh

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Engenharia Clínica e Instrumentação Médica

Dissertação orientada por:
Prof. David Menon
Prof. Alexandre Andrade

2021

Resumo

Esta dissertação é fruto de cooperação bilateral entre o Departamento de Física da Faculdade de Ciências da Universidade de Lisboa e a Divisão de Anestesia do Departamento de Medicina da Universidade de Cambridge. Cooperação esta efetuada através da supervisão e partilha de recursos associadas a este projeto, incluindo: o acesso a dados de pacientes, seminários mensais com uma equipa multidisciplinar e acompanhamento e comunicação diária no planeamento e desenvolvimento deste projeto.

As lesões cerebrais traumáticas representam um vasto e heterogéneo conjunto de patologias que causam numerosos problemas socioeconómicos, a uma escala global. Esta natureza heterogénea conjugada com a sua alta incidência torna as lesões cerebrais traumáticas a principal causa de mortalidade em jovens adultos e de incapacidade em todas as idades. Estas lesões são usualmente definidas como alterações da função cerebral e/ou danos estruturais no cérebro provocados por uma força externa. A fisiopatologia destas lesões está intrinsecamente ligada ao mecanismo físico da lesão, dado que a força externa pode danificar ambos o local de contacto como o local exatamente oposto, devido a forças inerciais. Adicionalmente, as acelerações rotacionais, comuns durante o impacto, apresentam um alto potencial para cisalhamento dos tecidos, enquanto que acelerações no plano lateral acarretam uma maior probabilidade de afetar as estruturas anatómicas mais profundas do cérebro humano. As principais causas das lesões cerebrais traumáticas são os acidentes de veículos motorizados, as quedas, os assaltos violentos e as lesões desportivas. Em suma, estas lesões podem acontecer a qualquer pessoa, em qualquer sociedade, o que dificulta a identificação de elementos comuns nestes pacientes. Todos estes fatores contribuem para a complexificação dos cuidados de saúde direcionados a estas lesões.

O avanço tecnológico das técnicas de imagem médica viabilizou a extração de uma maior quantidade de informação das imagens cerebrais e, consequentemente, o incremento da perícia nos cuidados de saúde nestes pacientes. Num contexto clínico, a implementação de um programa de reabilitação adequado ao paciente está intrinsecamente ligada ao seu prognóstico, que, por sua vez, depende frequentemente do tamanho e local das lesões do paciente. Esta dissertação tem como objetivo investigar o impacto que o volume, local e tipo de lesão têm no prognóstico do paciente. Para tal, foram tidas por base imagens médicas de um total de cento e vinte e cinco pacientes, com lesões cerebrais traumáticas, examinados em duas máquinas de ressonância magnética (3T SIEMENS *Trio* e *Verio*) e cujas lesões foram anotadas manualmente em múltiplos contrastes - *Fluid-Attenuated Inversion Recovery (FLAIR)*, *Susceptibility-Weighted Imaging (SWI)* e *Gradient Echo (GE)* - por médicos clínicos. As lesões foram localizadas por meio da comparação com um atlas do cérebro humano constituído por trinta e uma regiões de interesse distintas. A análise inicial das lesões incluiu a computação de várias características das lesões, como por exemplo, o volume lesionado e a conectividade (o grau de partição da lesão) de diferentes lesões. A localização das lesões, por sua vez, resultou na associação do volume lesionado a uma região cerebral e na computação de uma variável binária - *occupancy* - que descreve a presença de lesões em cada região.

As lesões anotadas manualmente foram atribuídas a um total de sessenta e três tipos de lesões, dos quais vinte foram considerados relevantes, tendo sido agrupados posteriormente, em cinco grupos de lesões: *Intraparenchymal Lesion Core (ILC)*, *Intraparenchymal Perilesional Oedema (IPO)*, *Petechial Haemorrhages (PH)*, *Extra-axial Haemorrhagic Lesions (EHL)* e *Ventricular and Cisternal Haemorrhages (VCH)*. Esta estratificação permitiu o estudo preliminar da cooccorrência das várias lesões, ao nível do tipo e grupo. Para tal, tipos e grupos de lesões foram emparelhados, e a cooccorrência de cada par foi definida pela percentagem de pacientes que demonstravam esse par, em função do número total de pacientes que demonstrava uma das lesões emparelhadas. Essencialmente, a coocorrência do par formado pelas lesões A e B é definida pela proporção de pacientes com a lesão A, que também mostram a lesão B. Esta análise permitiu revelar pares de lesões com uma coocorrência relativamente alta, que podem ser explicados pela natureza das lesões e/ou protocolo de anotação das mesmas. Estas hipóteses podem ser investigadas através da localização de lesões, capaz de elucidar a natureza das lesões, nomeadamente a proximidade espacial com que ocorrem. Tal averiguação estaria fora do âmbito principal da dissertação, no entanto, esta análise é escalável, podendo ser formalizada de forma a analisar padrões de lesões, em vez de pares. Eventualmente, investigações futuras poderão analisar o impacto que os padrões e a sua ocorrência têm no resultado funcional cognitivo do paciente.

Naturalmente, é provável que o uso de diferentes contrastes e/ou modelos de máquinas de ressonância magnética introduza variabilidade na medição das características das lesões. Esta variabilidade foi analisada no volume e conectividade das lesões. Os resultados sugeriram um aumento significativo (*Wilcoxon signed-rank (WSR)*, $p < 0.05$) do volume e número de *clusters* (partições da mesma lesão) das lesões anotadas em *SWI*, quando comparadas com as suas anotações em *FLAIR*. As imagens obtidas no modelo *Verio* também demonstraram um aumento de volume e contagem de *clusters*, quando comparadas com imagens obtidas no modelo *Trio*. Não obstante, a metodologia de comparação entre modelos foi aplicada aos mesmos grupos de lesões, mas em pacientes diferentes, dado que apenas um paciente foi examinado em ambos os modelos. Assim, a interpretação dos resultados é necessariamente incerta, uma vez que é impossível verificar se as diferenças observadas entre os volumes e números de *clusters* obtidos em cada modelo se devem á variação do modelo ou á variação entre os conjuntos dos pacientes examinados em cada modelo. Todavia, a comparação entre contrastes foi feita para as mesmas lesões, nos mesmos pacientes, e os resultados obtidos são coerentes com atributos conhecidos do contraste *SWI* que tendem a evidenciar artificialmente lesões hemorrágicas.

Os resultados dos pacientes, medidos através da *Glasgow Outcome Scale extended (GOSe)*, foram previstos através de quatro modelos lineares gerais, que incluíram o volume das lesões de quatro maneiras diferentes. O primeiro modelo considerou o volume total lesionado dos pacientes, independentemente dos tipos e locais das lesões. O segundo modelo integrou os volumes associados com cinco grupos de lesões, cada um constituído por vários tipos de lesões. Similarmente, o terceiro modelo associa os volumes com as diferentes regiões cerebrais identificadas. Por último, o quarto modelo combina o grupo e local da lesão, ambos quantificados em volume lesionado. A comparação entre modelos revelou que o volume das lesões poderá ser benéfico para o prognóstico do paciente, ou previsão do seu resultado, particularmente se ambos o tipo e local da lesão forem incluídos. Foram encontradas associações significativas entre preditores e resultado do paciente no primeiro (*Summed Volume (SV)*; $p = 0.047$) e quarto modelo (*ILC* com o lobo parietal esquerdo; $p = 0.047$), indicando que a combinação da região cerebral com o grupo da lesão aumenta a significância da sua associação com o resultado do paciente. No geral, esta metodologia mostra potencial para revelar o efeito que certos tipos e locais de lesões têm no resultado dos pacientes após uma lesão cerebral traumática.

Investigações futuras poderão beneficiar de segmentação automática de lesões, que provavelmente ultrapassará o processo lento e trabalhoso de anotação manual, possibilitando conjuntos de pacientes mais amplos, essenciais para a investigação de uma doença tão heterogénea como a lesão cerebral traumática. Avanços neste sentido aumentarão o poder estatístico da metodologia apresentada nesta dissertação.

Palavras chave: Traumatismo crânio-encefálico, Ressonância Magnética, Lesões Cerebrais Traumáticas, Localização de Lesões, Prognóstico

Abstract

Traumatic brain injury (TBI) is a highly heterogeneous pathology that poses severe health and socioeconomic problems on a global scale. Neuroimaging research and development has advanced its clinical care in numerous ways, as injured brains are being imaged and studied in greater detail. The size and location of TBI lesions are often necessary to accurately determine a prognosis, which is key in defining a patient-specific rehabilitation program. This dissertation aims to investigate the impact of lesion characteristics, such as volume and location, on outcome prediction in TBI patients. Lesion localisation was achieved by comparing annotated TBI lesions to a brain atlas. Furthermore, other lesion characteristics were examined across different Magnetic Resonance Imaging (MRI) sequences and scanners, with results suggesting that the use of different scanners or MRI contrasts introduced biases in said lesion characteristics. Patient outcome was predicted using four generalised linear models. Besides clinical variables, these models included lesion volume, group and location as predictors. Model comparison indicated that lesion volume could be beneficial for outcome prediction, but may be dependent on both lesion group and location. Overall, this methodology showed potential in uncovering the effect that certain lesion groups and/or locations have on patient outcome after TBI.

Keywords: Traumatic Brain Injury, Magnetic Ressonance Imaging, Lesion Volume, Lesion Localisation, Patient Outcome

Acknowledgments

I acknowledge the University of Cambridge and the University of Lisbon for the resources and supervision provided throughout the dissertation. Particularly, I acknowledge Professor Andrade and Professor Menon for accepting and fulfilling their roles as supervisors. I also thank Dr. Virginia Newcombe for her role in manually annotating the lesions in the data. I am very grateful to Stefan Winzeck, as he mentored me on a daily basis, throughout the dissertation, in a very accessible and didactic manner. His insight and support were essential in the making of this project.

This dissertation was done during less than usual times. Luckily, I had a network of friends and family to support me along the way. In this regard, I am very thankful for my parents and sister, who inspire and influence me everyday. I also appreciate all my friends who accompanied me during these university years, and wish them the best accomplishments.

Finally, I would like to give a special acknowledgment to my love, Mariana. Always the ray of sunshine in my cloudy day, you always push me to become a better person and student. Thank you for everything.

Table of Contents

Acronyms

BOLD Blood Oxygen Level-Dependent.

BP Brain Parcellation.

CNN Convolutional Neural Network.

CRF Conditional Random Field.

CT Computed Tomography.

DAI Diffuse Axonal Injury.

DSC Dice Similarity Coefficient.

DTI Diffusion Tensor Imaging.

DWI Diffusion-Weighted Imaging.

EDH Epidural Hematoma.

EHL Extra-axial Haemorrhagic Lesions.

FDR False Discovery Rate.

FEM Finite Element Model.

FLAIR Fluid-Attenuated Inversion Recovery.

fMRI Functional MRI.

GCS Glasgow Coma Scale.

GE Gradient Echo.

GOS Glasgow outcome scale.

GOSe Glasgow Outcome Scale extended.

ILC Intraparenchymal Lesion Core.

IPO Intraparenchymal Perilesional Oedema.

MAP Multi-Atlas Parcellation.

Acronyms

ML Machine Learning.

MR Magnetic Resonance.

MRI Magnetic Resonance Imaging.

MWU Mann-Whitney U.

PH Petechial Haemorrhages.

ROI Region of Interest.

SAH Subarachnoid Haemorrhage.

SDH Subdural Hematoma.

SV Summed Volume.

SWI Susceptibility-Weighted Imaging.

T1w T1-weighted.

T2w T2-weighted.

TAI Traumatic Axonal Injury.

TBI Traumatic brain injury.

VCH Ventricular and Cisternal Haemorrhages.

WSR Wilcoxon signed-rank.

List of Figures

List of Tables

Chapter 1

Introduction

The following chapter describes the context, motivation and structure within this dissertation.

1.1 Context and Motivation

Traumatic brain injury (TBI) is a major health and socioeconomic problem that impacts all societies, particularly in developing countries [1]. It is broadly defined as a brain injury that results from an impact or external force, thus, it comprises a vast collection of causes, lesion types and severities [1]. This heterogeneity combines with a high incidence rate, making TBI the leading cause of mortality in young adults and a major cause of disability across all ages [2, 3]. All in all, TBI can have a devastating impact on patients and their families, which is reflected in its socioeconomic cost estimates of up to $US400 billion each year [3]. These facts and figures have prompted experts and researchers to focus on the current shortcomings in prevention, health care and rehabilitation related to TBI [3]. One key aspect of TBI health care is prognosis, i.e. patient outcome prediction, as it is vital for tailoring patient-specific treatment and rehabilitation [4, 5]. Therefore, more accurate and reliable prognostic methods will improve patient health care whilst also alleviating the burden on families and socioeconomic cost [5, 6].

Multi-parametric magnetic resonance imaging (MRI) is a valuable tool for clinical management of TBI, as it enables complementary information on the damage caused by the injury, such as lesion size, type and location [7–10]. Researching efforts have been made to associate these features with patient outcome, with the objective of finding a potential pathway for more accurate and/or reliable TBI prognosis [6, 11–19].

In this context, this dissertation aims to investigate the relationship between patient outcome and lesion location, by developing a novel outcome modelling pathway based on multi-atlas lesion localisation. Put simply, this dissertation hypothesizes that features such as patient age, lesion type and location can be used to predict patient outcome, with hopes of improving or uncovering relevant factors related to TBI prognosis.

1.2 Structure

This dissertation is divided into seven chapters, including this introductory chapter. The following chapter elaborates on the principal concepts within this dissertation, starting with TBI and medical imaging related to this disease and ending with an overview of the statistical analysis and modelling methods used in this dissertation.

The third chapter provides an overview of lesion localisation methods and outcome prediction related to TBI, with examples of state-of-the-art studies and procedures.

The fourth chapter details the methodology used throughout the dissertation, from data collection and processing to bias and co-occurrence analysis, lesion localisation and patient outcome modelling.

The fifth chapter reports the results and findings associated with this dissertation. First, data demographics and lesion analysis are presented, setting the stage for the presentation of bias analysis results. Lastly, outcome-related findings are presented, initially associated to a qualitative analysis and then with respect to models and predictions.

The sixth chapter discusses the results previously mentioned and finishes with an overview of potential researching proposals related to this work, either by continuation, alteration or improvement.

Finally, the seventh chapter concludes this work, while an appendix with supplementary tables and figures is also included.

Chapter 2

Background Knowledge

This chapter introduces the major concepts present in this dissertation.

2.1 Traumatic Brain Injury

TBI is defined as a traumatically induced alteration of brain function and/or structural damage to the brain, as a result of an external force, such as an impact with an external object, foreign body penetration, acceleration/deceleration brain movements and forces generated by an explosion [20]. The pathophysiology of TBI is connected with the biomechanics of the injury, as external forces can induce a brain injury at both the site of contact and the opposite side of the brain, due to inertial forces [21]. Moreover, rotational accelerations, which are common during impact, have a high potential of resulting in shear-induced tissue damage, whereas lateral plane accelerations are more likely to damage the deep internal structures of the human brain [22].

The leading causes of TBI include motor-vehicle-related events, falls, assaults and sports-related injuries. These causes are also reflected in the risk factors for TBI which include sex, as motor-vehicle and violence-related TBIs lead to males being almost three times more likely to suffer a TBI, and age, as children and elders are more vulnerable to falls [23].

Furthermore, TBI is diagnosed with the emergence or worsening of at least one of the following clinical signs: loss or decreased level of consciousness, loss of memory for events immediately before or after the injury, alterations in mental state such as confusion and disorientation, neurological deficits (e.g. change in vision, lethargy and loss of balance) and intracranial lesions [24]. These clinical signs can be transient, long-lasting or permanent, and are usually accompanied by a variety of symptoms, such as, vomiting, headache, body numbness, speech difficulties and breathing problems [25]. Put simply, TBI is presented with vastly heterogeneous pathological patterns, making it difficult to identify single commonalities across patients [25]. Although TBI severity is best characterized as a continuum due to its strong heterogeneity [25], TBI is often classified based on the Glasgow Coma Scale (GCS), as mild (GCS 13-15), moderate (GCS 9-12) or severe (GCS$<$ 9) . This classification may be additionally guided by acute injury characteristics such as level of impaired consciousness and duration of post-injury amnesia [21].

2. BACKGROUND KNOWLEDGE

Having established the heterogeneity of TBI etiology and classification, i.e. causes and severity, it becomes transparent that TBI impacts a large demographic cohort. In fact, TBI has been reported as one of the most common neurological disorders [26], with estimates suggesting that there were 82,000 deaths and 2.1 million hospital discharges across Europe in 2012 [2], and that, each year, more than 50 million people worldwide suffer a TBI, with half of the entire world population having one or more TBIs during their lifetime [3]. Figure 2.1 illustrates additional statistics found in a TBI report made by the Centers for Disease Control and Prevention [27].

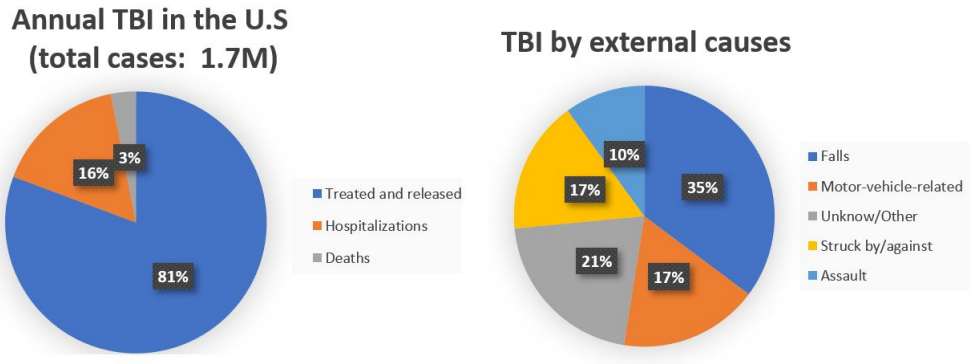

Figure 2.1: *Left: Annual TBI cases in the U.S with percentages of deaths, hospitalizations and treated and released from an emergency department. Right: External cause percentages related to TBI.*

Recovery and rehabilitation are also highly variable as they naturally depend on the severity of the injury. This combination of heterogeneity, incidence and morbidity contribute to an overwhelming socioeconomic cost associated with TBI, with estimates of the aggregate financial burden on U.S citizens ranging from 60.4 to 221 billion US dollars [28].

In summary, TBI occurs frequently, with a high degree of heterogeneity. Therefore, many research efforts have been (and are being) made to improve healthcare for patients that suffer from TBI, particularly through the development of faster and more accurate prognosis and diagnosis [29]. However, comprehensive reviews on this subject report a body of literature of varying quality, study design and objectives, with incorrect causal inferences deducted through cross-sectional studies [4, 29].

To understand these shortcomings, one must first analyze the current conventional and more advanced methods for TBI prognosis and diagnosis, starting with their general objectives [30], which are as follows:

- Identification of injuries that require emergency interventions, i.e. surgery;
- Detection of injuries that could benefit from early therapy and supervision;
- Accurate patient prognosis to enable personalized rehabilitative measures and to help family counseling and discharge planning.

2.2 Medical Imaging for Traumatic Brain Injury

One possible answer to these objectives was, and currently is, neuroimaging, as it enables unprecedented levels of information about human anatomy and physiology. When applied to TBI, neuroimaging methods such as MRI and Computed Tomography (CT), provide the means for detailed damage assessment in terms of location, categorization and severity, which is key for effective patient triage, injury classification and, ultimately, clinical trial design [31].

In the clinical setting of acute TBI, CT scans are usually the first imaging choice, and justifiably so, given their high sensitivity for blood, which can identify most urgent injuries. However, this is mainly due to the limitations of MRI, such as scanner availability, waiting time and compatibility [32]. As a matter of fact, when compared with CT scans, MRI has consistently shown better lesion detection, and is particularly useful in long-term TBI prognosis and management, because of its higher sensitivity for pathologies such as micro-haemorrhages, which are commonly presented with Diffuse Axonal Injury (DAI) [32, 33]. In the acute TBI context, MRI is usually requested if a CT scan is negative while the patient still presents abnormal neurologic behaviour, as reports have found that nearly one in four TBI patients admitted with a normal CT scan showed an abnormal MRI scan [33].

Moving forward, the focus turns to MRI in TBI, along with the different functionalities that comprise state-of-the-art procedures, as the next section provides a brief summary of MRI methods to highlight their relevance for clinical management of TBI.

2.2.1 Conventional Methods

In the reviewed literature, MRI methods are usually divided into two main categories: conventional and advanced methods. The three most relevant conventional MR contrasts, regarding their applicability to TBI, are T1-weighted (T1w), T2-weighted (T2w) and Fluid-Attenuated Inversion Recovery (FLAIR). These sequences display gross pathoanatomical features in the brain and have been shown to be beneficial when employed for TBI imaging. For instance, T1w has been extensively used to evaluate brain atrophy in chronic TBI, often being paired with automated or semiautomated software [34–37]. In fact, subacute haemorrhage products, as well as lipids and proteins, appear hyperintense on T1w scans, as a result of shortened T1 relaxation times [38].

T2-weighted scans, on the other hand, have consistently been shown to be useful in detecting various intracranial injuries that result in the increase of tissue water percentage, which increases T2 relaxation times. As such, examples of T2 hyperintense lesions are oedema, haematoma and axonal injury, and hypointense features - due to decreasing T2 relaxation time - include paramagnetic blood products [39, 40].

Lastly, FLAIR scans show contrast similarly to T2w scans while also suppressing normal ventricular and sulcal cerebral spinal fluid, which enhances the visibility of T2 hyperintense lesions like the ones aforementioned [41]. Examples of these MRI contrasts are shown in Figure 2.2.

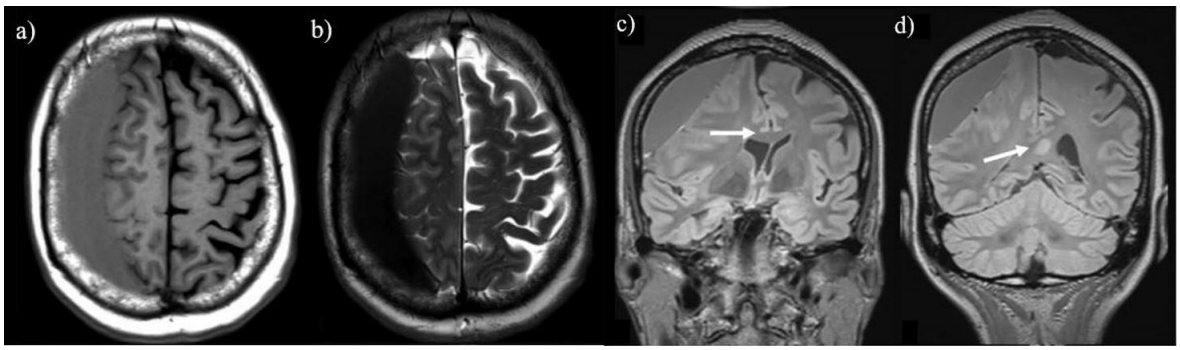

Figure 2.2: *Axial T1w (a) and T2w (b) imaging demonstrates a large acute subdural haematoma along the right cerebral convexity. Reformatted coronal T2w FLAIR imaging shows leftward subfalcine herniation (c) and a focus of axonal injury involving the splenium of the corpus callosum (white arrow) (d). Source: Kazam et al.[33].*

2.2.2 Advanced Methods

As for advanced methods, the ones usually referred to are Susceptibility-Weighted Imaging (SWI), Diffusion-Weighted Imaging (DWI) and functional MRI (fMRI). These techniques are of particular interest in situations where conventional MRI is not sufficiently sensitive, e.g. when imaging produces normal findings, but the patient is still symptomatic, and thus offer opportunity for better prognosis [42]. Similarly to the conventional methods mentioned before, the three advanced ones have also demonstrated well-documented applicability to TBI.

SWI combines magnitude and phase images with a 3-dimensional, velocity compensated GE sequence and has been reported as being highly sensitive for minute haemorrhages that usually occur with DAI [43, 44]. Diffusion weighted imaging captures the movement of water, i.e. Brownian diffusion of water molecules in a voxel of tissue, by quantifying it through diffusion coefficients that are mapped onto T2w images [45]. When applied to acute TBI, DWI has been shown to be very useful for assessing cerebral oedema and axonal injuries [46, 47]. Functional MRI (fMRI) enables noninvasive brain function evaluation, by assessing neuronal activation through a Blood Oxygen Level-Dependent (BOLD) signal. This signal quantifies magnetic field inhomogeneity caused by changes in the oxygenation state of hemoglobin [48]. This is perhaps the least reported technique, regarding its applicability to TBI, due to a lack of task-based studies within different behavioral/cognitive domains. In fact, only a small number of functional networks - mostly the default mode - have been researched with resting-state fMRI techniques, which limits insight into fMRI potential in the TBI context [42, 49].

Examples of SWI and DWI are presented in Figure 2.3, while examples of fMRI applications to TBI include the analysis of neuronal activation during functional tasks inferred from BOLD signal [48].

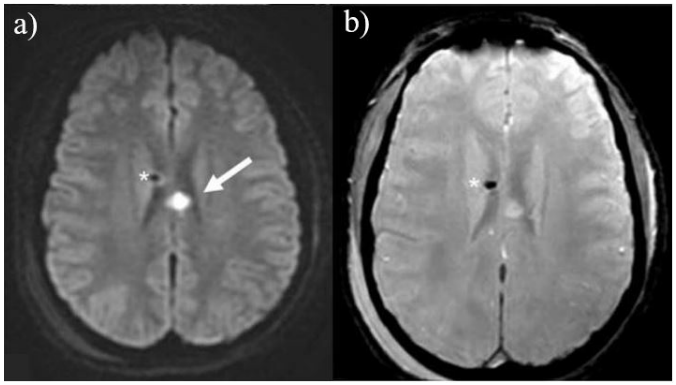

Figure 2.3: *DWI (a) clearly demonstrates axonal injury in the left body of the corpus callosum while SWI (b) shows it much more subtly. A ventricular catheterization (asterisk) was performed during examinations. Source: Kazam et al. [33]*

These imaging techniques exploit different contrast mechanisms to enhance our understanding of persistent symptoms and provide accurate prognosis. As such, thorough reviews have been made which feature quantitative and qualitative comparisons between them, enabling vital knowledge on these methods, e.g. parameter setting recommendations for optimal procedures and advantages and limitations of each technique [30, 31, 33, 42, 48, 50]. The key points in said reviews have been covered in this chapter for conventional and advanced MRI sequences. Nonetheless, the reviews are suggested for further reading.

2.3 Statistical Analysis and Modelling

Statistical analysis is a key component of biomedical data analysis, as descriptive and inferential statistics enables researchers to summarize findings and conduct hypothesis testing [51]. In this section, the inferential procedures used in this dissertation are introduced theoretically.

2.3.1 Mann-Whitney U Test

Quite often, t-tests are used to compare two populations means based on independent samples from the two populations or groups. To do so reliably, various assumptions need to be put in place, a particular one being that the data follows a normal distribution [52]. This constraint is eased in the Mann-Whitney U (MWU) test, enabling its application to non-normally distributed data or other cases in which this assumption is not accepted. Thus, the MWU test is often considered as the non-parametric alternative to a t-test.

Essentially, the MWU test ranks all the values from low to high, with no regard for the group each value belongs to. The minimum is attributed a rank of one while the largest value gets a rank of N, where N is the total number of values in the two groups. The ranks are then averaged, within each group, producing two average rank values to be compared. If these two averages are significantly different, the respective p-value will be small [53]. With a threshold α set at 0.05, a p-value below this value will indicate a statistically significant difference in group rank averages, with a confidence level of 95%.

This procedure, though less strict than the t-test, also carries two general assumptions: all observations from both groups are independent and the observed variable must be, at least, ordinal, to allow

comparison between observations. Additionally, with the assumption that the two groups have equally-shaped distributions, the MWU test is transformed in a test of medians [54].

Therefore, the null hypothesis that groups share the same distribution, is rejected if a p-value lower than α is obtained. Moreover, the test produces a U statistic value, which is computed in the following manner:

$$U = \sum_{i=1}^{n} \sum_{j=1}^{m} F(X_i, Y_j) \tag{2.1}$$

where

$$F(X,Y) = \begin{cases} 1, & \text{if } Y < X, \\ \frac{1}{2}, & \text{if } Y = X, \\ 0, & \text{if } Y > X. \end{cases} \tag{2.2}$$

The number of comparisons equals the product of n times m. If the null hypothesis is true, then the value of U should be approximately half the number of comparisons. Alternatively, if its value is much smaller than that, the p-value will likely be low, leading to rejection of the null hypothesis [55].

2.3.2 Wilcoxon Signed-Rank Test

The WSR test allows for the comparison of matched-pair data based on the distribution of their differences. Indeed, it accepts dependent observations and non-normally distributed data, serving as an alternative to the paired or dependent samples t-test [56].

The data must consist of samples $(X_1, Y_1)...(X_n, Y_n)$, so their differences are computed as $d = X_i - Y_i$, with $i = 1, 2, ..., n$. The assumption that d is symmetric around zero produces the two-sided null hypothesis that the median of the differences is zero against the alternative that it is different from zero. The one-sided test has the null hypothesis that the median is positive against the alternative that it is negative, or vice versa [57]. To compute the test statistic, one must first sort the differences by assigning ranks $R_1...R_n$, so that $0 < |X_{R_1}| < |X_{R_2}| < ... < |X_{R_n}|$. Next, let sgn denote the sign function, as $sgn(x) = 1$ if $x > 0$ and $sgn(x) = -1$ if $x < 0$. Lastly, the test statistic T is defined as:

$$T = \sum_{i=1}^{n} sgn(X_i) R_i \tag{2.3}$$

Subsequently, a p-value is produced by comparing the test statistic T to its distribution under the null hypothesis [57]. As before, a p-value below $\alpha = 0.05$ leads to a rejection of the null hypothesis, concluding that the median of differences is different from zero, i.e. there is a statistically significant difference between the paired observations.

2.3.3 Binary Logistic Regression

Regression analysis is usually described as the search for relationships among variables, mainly by estimating a mathematical relationship between a dependent variable, also termed outcome or response variable, and one or more independent variables, often called predictors or explanatory variables [58]. Logistic regression follows this same logic by modelling the probability of a discrete outcome variable, usually of binary nature, given one or more input variables. Because of this, logistic regression is often a valuable tool for classification problems, with many applications in the industry [59].

Let us consider a model with two predictors, x_1 and x_2, and a discrete binary outcome variable Y, which gives $p = \mathrm{P}(Y = 1)$, as the probability that Y equals 1. Next, a linear relationship between the predictors and the log-odds of $Y = 1$ is formulated as follows:

$$l = log_b \frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2, \tag{2.4}$$

where l is the log-odds, b is the base of the logarithm and β_i are the coefficients or parameters of the model. The probability p can then be retrieved via exponentiation of the log-odds and basic algebra, producing:

$$p = \frac{1}{1 + b^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}} \tag{2.5}$$

These coefficients are estimated by fitting 2.4 to adequate training data, and are key in interpreting the model. The y-intercept, β_0, is the log-odds of $Y = 1$ when $x_1 = x_2 = 0$. The other coefficients are interpreted in the same manner: obtaining a $\beta_i = k$ will imply that increasing x_i by 1, increases the log odds by k, or, alternatively, increasing x_i by 1 will increase the odds that $Y = 1$ by a factor of 10^k [60]. Along with the coefficients, each predictor is attributed a p-value, under the null hypothesis that there is no association between the predictor and the outcome variable. Therefore, a p-value below alpha rejects the null hypothesis, implying that there is a statistically significant association between said predictor and outcome variable [61].

Having constructed the model, i.e. obtained its coefficients, one can use it to predict the probability that the outcome variable is 1, given certain values for the predictors. Subsequently, the model confusion matrix - a visual layout that summarizes the performance of the model - can help determine the correct and incorrect predictions [62]. Moreover, the model accuracy can be calculated as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.6}$$

where TP and TN are the true positives and negatives, respectively, and FP and FN are the false positives and negatives.

The error rate, err, can be inferred from the accuracy: $err = (1 - ACC) \times 100$. If the model is applied to the data it was trained on, this value represents the training error rate, whereas, if the model is applied to data, which was unknow during the training phase, then this value reports the test error rate. These calculations enable a key aspect in building a model: its validation [63].

2.3.4 Multiple Comparison Correction

Contrasting different situations or elements is often the basis for research projects, for instance in the form of group/treatment or product comparisons. In statistical analysis, this might occur as many tests of the same hypothesis, which leads to what is commonly referred as a multiple comparisons problem. Put simply, an increase in number of tests increases the likelihood of finding the effect or relationship under analysis. However, it also increases the likelihood of finding said effect by chance, thus producing a false claim [64].

Various procedures have been developed to solve this issue, for example by controlling the False Discovery Rate (FDR), i.e. the expected proportion of null hypothesis rejections that are incorrect. One method to control the FDR is the Benjamini-Hochberg p-value correction.

In short, p-values are ordered, and tested under the following inequality:

$$P_k \leq \frac{k}{m}\alpha, \tag{2.7}$$

where m is the number of p-values and $k = 1, 2, ..., m$.

The largest p-value that satisfies 2.7 is set as a threshold for significance, i.e. the largest p-value and all p-values below it are deemed significant. In addition to p-value corrections, researchers are also advised to plan before the analysis and perform reproducibility studies, to guard against multiple comparisons problems [65].

Chapter 3

Lesion Analysis in TBI

This dissertation is structured in two parts: lesion localisation and patient outcome prediction. Put simply, lesions were manually annotated and automatically located using a brain atlas, thus enabling analysis into the relationship between patient outcome and lesion type, size and/or location. This chapter provides a general overview of related work in lesion localisation and patient outcome prediction.

3.1 Lesion Localisation

As lesion segmentation identifies voxels associated with injury, it also enables the derivation of lesion features that provide further information for clinical assessments. One of the most obvious ones is the total lesion volume. However, lesion size alone might not provide a full picture of the injury, since lesion location within the brain also has an impact on the patients well-being, as the brainstem [19], cerebellum [66], corpus callosum [67] and several other locations [68] have been investigated with regards to their prognostic value as a lesion location. Indeed, lesions can be located in many ways, three of which are here described.

3.1.1 Visual Inspection

Lesions can be located by visual inspection, which often coincides with their manual annotation. For example, Williams et al. [19] investigated the prognostic value of brainstem lesions in acute CT scans. Initially, board-certified neuroradiologists evaluated the scans for the presence or absence of brainstem lesions. These lesions were then further categorized by location, size and type, with lesion localization being done manually by a fifth-year neurological surgery resident outlining and measuring regions of interest on appropriate software (OsiriX MD DICOM Viewer, v.11.0.2.). Figure 3.1 schematizes this procedure.

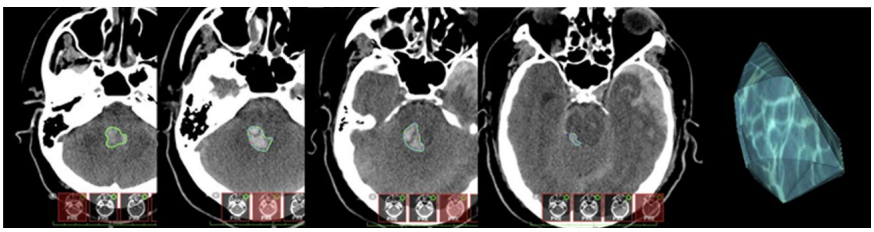

Figure 3.1: *Brainstem lesion segmentation on ascending axial head CT. Borders (light green) are hand-traced through the slices in Osirix MD software. These demarcated areas are merged to create a 3D region of interest for further quantitative analysis (last image on the right). Taken from William et al. [19].*

In this case, lesions were first located on the brainstem and then distinguished through three lesion features: size, type and sub-location. The study found that the addition of these lesion features into a patient outcome model did reveal significant associations between each feature and patient outcome, whereas the presence or absence of brainstem lesions were not significantly associated with outcome.

Inversely, a lesion type of interest can be defined first, and then analyzed within one or more regions: Moen et al [68] located traumatic axonal injuries across the hemispheres, cerebellum, corpus callosum, brainstem and thalamus/basal ganglia, as neuroradiologists characterized image findings on several MRI contrasts. These lesions were counted across all contrasts, and their volumes were estimated in FLAIR sequences. These features, i.e number of lesions and lesion volume, were then analyzed as potential predictors of patient outcome, with results suggesting that the number of DWI lesions and FLAIR lesion volumes in the corpus callosum, thalamus and brainstem were predictive of patient outcome. These findings are coherent with those produced by Cicuendez et al [69], as they also located traumatic axonal injuries in the corpus callosum and other regions, further categorizing these lesions across features such as the presence of haemorrhage, severity, sub-location, number of lesions and total lesion volume. Their findings suggest that lesion presence in the corpus callosum is associated with unfavourable patient outcomes, whereas the total volume is an independent prognostic factor for these same outcomes.

Matsukawa et al [67] employed a similar method, as two experienced neurosurgeons evaluated CT and MRI scans, and visually located lesions in the corpus callosum on MRI and intraventricular haemorrhage on CT. Lesions in the corpus callosum were distinguished by their presence or absence at three sub-locations, with the authors reporting that evidence of an intraventricular haemorrhage on CT may indicate lesion presence on the corpus callosum, which in turn is associated with unfavourable patient outcomes.

As seen above, visual inspection and neuroradiological expertise is commonly used as means to locate lesions, either by first selecting one or more brain regions and then analyzing the lesions within it [19], or by initially selecting a specific lesion type and then mapping its various occurrences across several regions [67–69]. These methods enable valuable information; however, visual inspection or manual lesion localisation is a laborious, time-consuming task, and may limit the sample size included in the analysis, which compromises TBI-related inference, due to its notorious heterogeneity, i.e. smaller sample sizes will result in fewer identifiable lesion types [7–10].

To surpass the time-consuming nature of manual lesion localisation, brain atlases and templates - which provide anatomical reference - can be used with the general objective of automatically mapping large collections of lesions [70, 71].

3.1.2 Brain Parcellation

This process of labelling distinct partitions across the spatially heterogeneous human brain is commonly described as Brain Parcellation (BP). In a historical perspective, it could be seen as the latest form of the century-old quest to develop an universal map of the brain [70]. This is challenging due to inter-individual variability and the multilevel complex topography of the human brain. However, there is now a wide range of MRI-based approaches for BP, focusing on several different features, from local properties of brain tissue to long-range connectivity patterns and functional and/or structural markers [70, 71].

Across research fields, BP has emerged as both a vital tool to advance knowledge of brain organization, and a practical step to perform biologically-based information compression by summarizing data from several voxels into more manageable groups that represent particular anatomical entities.

This reduction is critical in recent big-data approaches that aim to predict behavioral or clinical phenotypes from brain imaging data, as these predictions are usually related to distinct anatomical regions [72–75]. As such, efforts to automate BP have generally been successful, and several software tools are available [76–78]. However, these methods have mostly been optimized for healthy, anatomically normal brains, whereas brain pathologies, due to their heterogeneity, require more detailed solutions, depending on the lesion [79].

A popular approach in this setting has been atlas-based BP, which has been very useful for structural brain mapping, and widely used across neuroimaging [80]. It can also be achieved for individual or groups of patients, and many brain atlases are available as digital image files, containing multiple brain regions following region-specific image intensity values [80, 81]. Usually, brain atlases are defined as a template or reference brain image, be it a single subject image or an averaged image across a specific cohort. As such, they require image registration, a process that maximizes spatial similarity between source and target images. In other words, atlas-based brain mapping requires spatial matching between the atlas and the brain image being mapped. Image registration is in itself a nuanced topic, as many strategies have been developed to achieve it, including rigid-body and affine or non-affine registration [82, 83]. When applied to TBI, atlas-based BP allows for multi-modal registration and enables group comparisons with anatomical reference. Here, research efforts are divided across two main approaches: single-atlas and multi-atlas parcellation.

3.1.2.1 Single-Atlas Parcellation

For single-atlas parcellation, the common procedure is to register an image, that has yet to be segmented, to a single template image with delineated regions annotated by an expert. This allows the mapping of region labels from the template or atlas to the new image [84]. This approach has found various applications in neuroimaging [66, 85, 86].

In a prospective study, Daniels and Foundas [85] aimed to determine the relationship between neuroanatomical sites and increased risk of aspiration in acute stroke patients. In doing so, lesions were located using standardized templates of axial CT/MRI sections at various angles to the canthomeatal line: lesion features, such as sub-location and size, were mapped out within and across each CT/MRI image on the template. Once the lesion was mapped, its location was determined using standardized templates. The study found that various locations were injured in patients with risk of aspiration, but no specific cortical site was identified as a predictor of risk for aspiration.

In similar fashion, Kasahara et al. [86] combined voxel and atlas-based analysis to better distinguish between focal axonal injury and DAI, thus improving diagnosis on patients with mild TBI. To achieve this, diffusion maps, produced from Diffusion Tensor Imaging (DTI), were normalized to the International Consortium for Brain Mapping atlas.

This mapping allows for segmentation of Region of Interest (ROI)s and consequent region-specific data analysis, as demonstrated in figure 3.2.

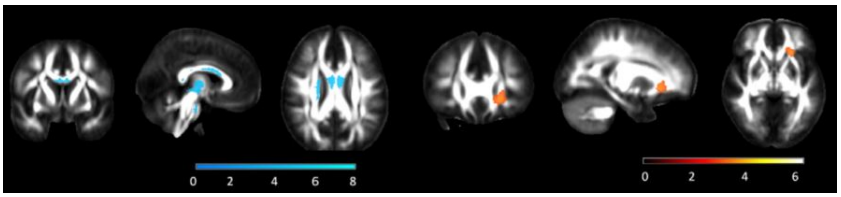

Figure 3.2: *Group analysis of DAI and mild TBI compared to control group. Atlas-based analysis enables the segmentation of ROIs in comparing different patient groups. Blue regions contain white matter diffusion measures that are lower in the DAI patient group while red regions contain white matter diffusion measures that are higher in the mild TBI patient group. Both representative values pertain to radial diffusivity. Retrieved from Kasahara et al. [86].*

Moreover, the authors further segmented and normalized all diffusion maps using a diffeomorphic registration tool in an effort to reduce misregistration. This is a common problem in single-atlas approaches, as deformation models used for novel image registration fail to capture the wide range of anatomical variation that comes with brain lesions [84, 86].

Another variation of single-atlas methodology is the introduction of probabilistic inference as novel images are registered to a a single atlas, producing additional information, such as the probability of observing a specific label at a given location [84, 87].

This procedure has also been applied to neuroimaging of brain lesions as Garfinkle et al [66] evaluated the impact of preterm cerebellar hemorrhage (CBH) in childhood development, with preterm MRI scans of 221 neonates being manually segmented for CBH quantification and location. These locations were then analyzed with respect to the likelihood of resulting in an adverse outcome.

To do this, the manually annotated CBH on each preterm scan was mapped to the common space of the early preterm template using a nonlinear transformation, which compensates the anatomical variability between patient and template images. Subsequently, probabilistic CBH maps were generated based on the cumulative number of CBH that occurred in homologous brain regions within the standard preterm template. These maps identify relevant regions and characterize the spatial distribution of CBH, whilst analogous outcome-based CBH maps enable visual assessment of difference in CBH location between different outcome groups. In this study, these maps demonstrated that CBH extending to deeper parts of the cerebellum predicted adverse motor, visuomotor and behavioral outcomes.

This probabilistic methodology can also be applied to multiple atlases, i.e. training images delineated by an expert, by summarizing them into a single probabilistic model. However, the consequential loss of information can lead to faulty region segmentation, particularly in the presence of lesions [84, 87].

3.1.2.2 Multi-Atlas Parcellation

Multi-Atlas Parcellation (MAP) has shown promising results for combating the issues plaguing single-atlas approaches [84, 88, 89]. Instead of summarizing atlases into probabilistic models or using a single standardized template, MAP takes into consideration all atlases as potential spaces for mapping novel images, e.g. pairwise registration between novel images and each atlas with label propagation for a specific region being done by label fusion [90]. This last process can be done by selecting the best atlas and using its corresponding labels, or by attributing the most frequent label across registrations to a given location [90, 91].

This methodology has been applied to neuroimaging in various forms, such as brain parcellation, i.e. whole-brain segmentation into a certain number of regions [88, 92], or with a focus on a specific set of regions or an individual ROI, which are often associated with pathologies [84, 93–95].

For example, the hippocampus has been the focus of several studies that use this methodology due to its association with Alzheimer's disease and dementia [84]. Other pathologies have been investigated in similar ways such as Parkinson's disease [93], multiple sclerosis [94] and epilepsy [95]. However, when compared to TBI, these pathologies produce much more subtle anatomical variation. Indeed, the mapping of TBI lesions is a challenging task even with the advantages of MAP, resulting in scarce literature regarding its application to TBI.

A published example of such application was produced by Ledig et al. [88], as the authors present a fully-automatic multi-atlas whole-brain segmentation method for MRI. The method was called Multi-Atlas Label Propagation with Expectation-Maximisation (MALP-EM) and it was based on a combination of a robust registration approach (MAPER), joint label fusion and intensity-based label refinement, which allowed for applicability to abnormal brain anatomy. To demonstrate this, authors applied this framework to segment 125 magnetic resonance images into 134 regions from patients with TBI, producing not only state-of-the-art segmentation, but also valuable insights on TBI disease progression and outcome prediction. Figure 3.3 shows the segmentation of an highly abnormal brain configuration.

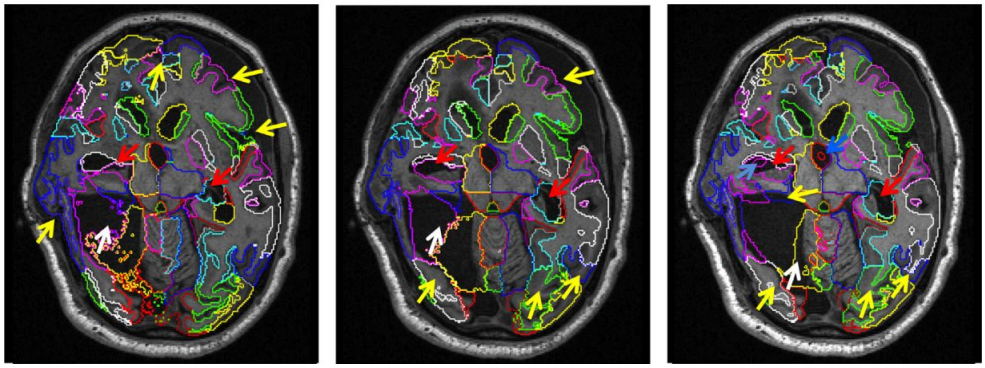

Figure 3.3: *Examples of segmentation of an highly abnormal brain configuration. Colored arrows point to errors that arise from oversegmentation and other inaccuracies. Retrieved from Ledig et al. [88].*

In short, this study showcases the potential and advantages of using MAP for abnormal brain anatomy, i.e. in presence of lesions. Automated lesion localisation has thus started to pave the way for larger and normative databases, as it scales visual inspection to much larger sample sizes. These tools not only facilitate data processing, but also produce a wealth of data features that can be associated and studied in various forms. Put simply, large collections of TBI patients clinical data coupled with automated lesion localisation, allow for unprecedented insights into various aspects of TBI. A key aspect in this regard is patient outcome analysis and prediction.

3.2 Outcome Prediction

Traumatic brain injury prognosis, i.e. outcome prediction, is often described as a vital effort to address its heterogeneous incidence and consequential motor and mental disabilities, as faster, more accurate and reliable outcome predictions enable better rehabilitative planning. Indeed, acute prognostic measures play a fundamental role in clinical decision-making and enables patient counselling with regards to rehabilitation [96, 97].

Because of this, several variables have been studied as potential prognostic factors, such as age, sex, pupil reactivity, hypoxia and the presence of subarachnoid bleeding [11, 12]. In fact, recent prognostic models (CRASH and IMPACT) have been developed on large numbers of patients (10 008 and 8509, respectively), with studies showing that age, GCS and pupillary reactivity contain the largest amount of prognostic information, particularly when applied to moderate and severe TBI [98]. Their application to mild TBI, however, has been less successful [99]. One of the most established outcome measures after TBI is the Glasgow outcome scale (GOS), later extended to the GOSe, which ranges from 1 to 8. Hereby, the lowest value represents a fatal outcome and the highest value points to a favourable recovery, as the patient should show no symptoms due to TBI [13]. This measure is usually coupled with the GCS score, which ranges from 3 (completely unresponsive) to 15 (responsive) and is often a reliable measure to assess level of consciousness after TBI. Both measures have shown promising results in early prediction of long-term functional outcome, however, their limitations are increasingly recognizable, particularly when applied to large patient cohorts. To address these limitations, sets of guidelines have been developed to standardize the clinical interview used to assign outcome [6, 13].

Several studies have combined specific lesion patterns with the aforementioned variables and measures to model and predict patient outcome. Gauthier et al. [14] administered several language tests to a sample of 145 adult patients with TBI, as a means of evaluating cognitive-communication impairments in this cohort. GOSe and GCS, along with several other variables, were combined with lesion pattern analysis to acutely predict outcome. The authors found that older and less educated patients with a more severe TBI were at high risk of unfavourable outcome. Additionally, it was reported that left temporal, frontal and right parietal lesions may increase the risk for impairment.

Aldossary et al. [15] focused on DAI and its prognostic value, as neurological assessment of 251 patients with severe TBI was done according to the GOSe. The authors report an association between unfavourable functional outcome (GOSe below or equal to 4) and DAI, especially in patients with cerebral hemisphere and brainstem lesions. Moreover, other clinical variables such as duration of post-traumatic amnesia, hospital stay, DAI volume and loss of consciousness were all associated with unfavourable neurocognitive outcome.

Williams et al. [19] investigated the prognostic value of brainstem lesions in TBI patients across the severity spectrum. Patients with brainstem lesions were matched by age, sex and GCS to patients with cerebrum lesions only. Additionally, all patients undertook a head CT scan within the first 48 hours after the injury, and were attributed a GOSe score.

The report found no significant difference in GOSe scores for patients with brainstem lesions, compared with patients with similar scores but only cerebrum lesions. Nonetheless, further categorization of lesions into lesion types identified a proportion of patients with significantly less favorable outcomes. Lastly, a multivariate logistic regression model of favorable patient outcome showed that including lesion type and GCS produces a significant relationship with patient outcome.

Chastain et al. [16] investigated TBI outcome prediction across different imaging modalities and injury distributions. The imaging modalities considered were CT, T2w, FLAIR and SWI, which were obtained for each of 38 adults between 0.1 and 22 months after injury. Additionally, outcomes were assigned with respect to GOS. Figure 3.4 shows how dichotomized outcomes - unfavourable outcome (GOS 1-3) and favourable outcome (GOS 4-5) - vary according to median total lesion volume.

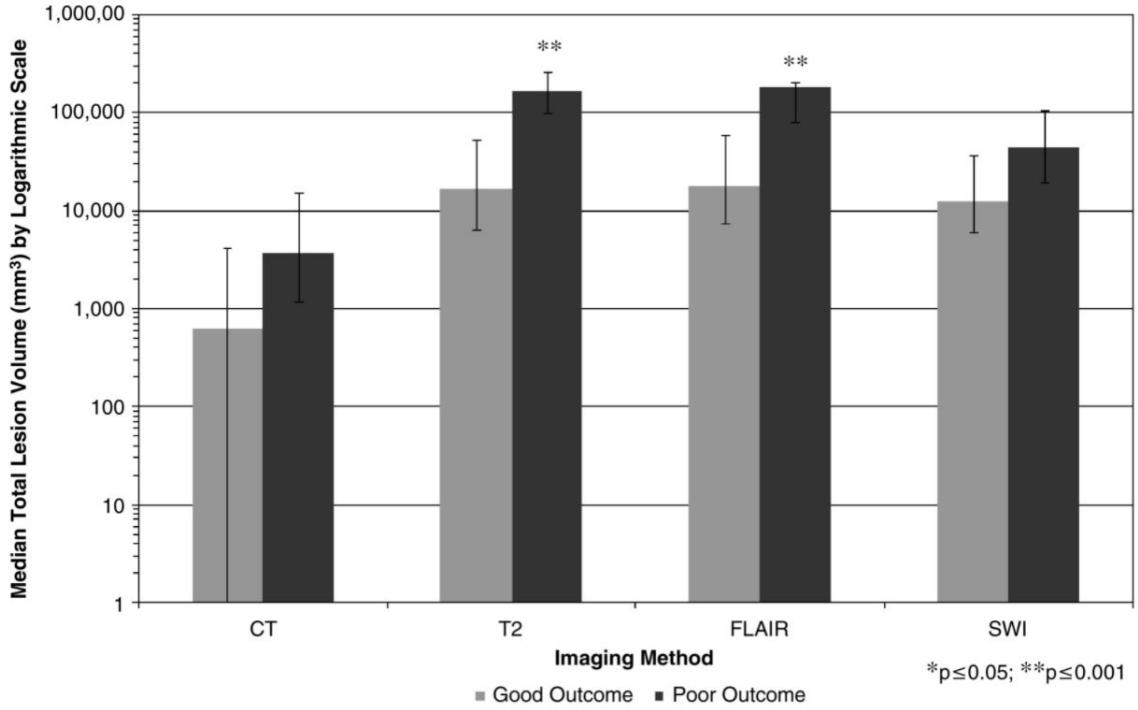

Figure 3.4: *Variation of median total lesion volume for each outcome, across different imaging modalities. Interquartile range is depicted by error bars for each outcome group, while a logarithmic scale with a base total lesion volume of 1 mm^3. Retrieved from Chastain et al. [16].*

Recent systematic reviews (2017-2018) analyzed a total of 90 studies to conclude that several lesion patterns are significantly associated with long-term functional outcome: DAI and lesions in the corpus callosum were associated with an unfavourable outcome. Additionally, while lesion locations and lesion count produced inconsistent results, lesion volume was predictive for outcome in the FLAIR sequence and apparent diffusion coefficient calculated from DWI [17]. Other studies relating patient outcome to lesion features such as lesion type, location, severity and size were in section 3.1, for instance, Garfinkle et al [66] have shown that the cerebellum holds significant prognostic value, while Matsukawa et al demonstrate the same for corpus callosum injuries [67]. Moen et al [68] specified lesions even further, by introducing lesion volume and lesion count, and found that the number of DWI lesions and FLAIR lesion volumes in the corpus callosum, thalamus and brainstem were predictive of patient outcome. Despite promising results, systematic reviews have pointed out that most studies were at high risk of methodological bias, as reviewers indicate that most studies did not report data adjusted for confounding prognostic variables and were at risk of selection bias due to frequent lack of consecutive patient recruitment [17]. In conclusion, reviewers a call for large standardized studies to produce more reliable and robust results [17, 18].

In conclusion, lesion localisation produces several lesion features, which can then be combined with patient clinical data to model and predict patient outcome. The next chapter explores this methodology

while presenting some of its implicit steps such as data reduction, lesion and bias analysis.

Chapter 4

Methods

Having presented an overview of state-of-the-art procedures in lesion localisation, as well as advances in patient outcome prediction, this chapter focuses on the methods used in this dissertation. Besides data collection and processing procedures, important tasks pertaining to lesion localisation and patient outcome prediction are described. It should be noted that the data collection and processing sections include tasks that were not done during, or as part of, the dissertation. Nonetheless, they are briefly described as they pertain to the acquisition and preparation of the data used throughout the dissertation.

4.1 Data Collection

4.1.1 MRI Acquisition & Demographics

The MRI data used for this dissertation were acquired either on a 3 Tesla SIEMENS Trio or a 3 Tesla SIEMENS Verio scanner, both located at the Wolfson Brain Imaging Centre (Addenbrooke's Hospital, Cambridge, UK). A total of 267 scans from 125 subjects that experienced TBI were collected, with an overview of the Magnetic Resonance (MR) sequence parameters being provided in Table 4.1. Furthermore, scans were collected within a large time window, ranging between 0 and 3422 days since injury, with most patients having multiple scans along this timeline.

Table 4.1: Overview of MR Sequence Parameters.TI is the inversion time, TE is the echo time and TR is the repetition time.

Parameters	T1w	FLAIR	GE	SWI
Acquisition Type	3D	2D	2D	3D
TE	2.98 ms	95 ms	20 ms	20 ms
TR	2300 ms	7840 ms	500 ms	28 ms
TI	900 ms	2500 ms	–	–
Flip Angle	9°	150°	20°	15°
Dimensions	$1\times1\times1$ mm^3	$0.7\times0.7\times4$ mm^3	$0.8\times0.8\times5$ mm^3	$0.5\times0.5\times1.2$ mm^3

In addition to MRI scans, demographics such as age, sex and education were also collected. Furthermore, the severity, mechanisms of injury, GCS score, and patient outcome (GOSe score at six months) were recorded (see section 3.2). An overview of clinical variables can be found in the Appendix (Table A.1).

4.1.2 Manual Lesion Segmentation

Experienced clinicians manually annotated multiple lesions types on FLAIR, SWI and GE scans. This resulted in five lesion groups: ILC, IPO, PH, EHL and VCH. Where SWI sequences were unavailable, lesions with the same imaging characteristics were annotated on GE sequences, as both show similar contrast. This was the case for four out of six classes of lesions that were annotated on SWI. Lesion frequency in subjects and scans varied within groups, as demonstrated in Table 4.2.

Table 4.2: Lesion groups accompanied by lesion type frequency across subjects and scans.

Lesion Group	Lesion Type	Subjects (%)	Scans (%)
ILC	SWI Contusion Core	112 (90.3)	218 (82.6)
	FLAIR Contusion Core	111 (89.5)	184 (69.7)
	SWI Callosal Blood not Petechial	43 (34.7)	70 (26.5)
	Basal Ganglia Contusion Core	23 (18.5)	35 (13.3)
	FLAIR Brainstem Contusion Core	24 (19.3)	34 (12.9)
	Intracerebral Haemorrhage	1 (0.81)	1 (0.38)
IPO	FLAIR Contusion Oedema	113 (91.1)	187 (70.8)
	FLAIR Callosal Oedema	57 (46.0)	85 (32.2)
	FLAIR Petechial Oedema	56 (45.2)	76 (28.8)
	FLAIR Brainstem Contusion Oedema	38 (30.6)	54 (20.4)
	Basal Ganglia Contusion Oedema	21 (16.9)	28 (10.6)
PH	SWI Petechial Haemorrhage	107 (86.3)	189 (71.6)
	SWI Brainstem Blood	48 (38.7)	82 (31.1)
	SWI Callosal Blood Petechial	47 (37.9)	67 (25.4)
EHL	Subdural Hematoma (SDH)	99 (79.8)	157 (59.5)
	Subarachnoid Haemorrhage (SAH)	55 (44.3)	73 (27.6)
	Post-surgical Subdural Debris	32 (25.8)	42 (15.9)
	Epidural Hematoma (EDH)	12 (9.7)	14 (5.3)
VCH	Intraventicular Haemorrhage	83 (66.9)	121 (45.8)
	SWI Blood in Cisterns	24 (19.3)	29 (11.0)

4.2 Data Processing

4.2.1 Lobe Atlas Construction

Lesion locations were computed using an atlas-based lesion localisation method. The lobe atlas was constructed through the parcellation of 652 MR T1w scans from the Cam-CAN study [100] via MALP-EM [88]. First, the T1w scans were non-linearly registered to the standard MNI template. The found transformation was then applied to the individual MALP-EM parcellations, to project them to the common MNI space. All parcellations were fused to generate an average ROI atlas comprised of 138 MALP-EM anatomical brain regions. Since MALP-EM segments white matter as a whole by design, each white matter voxel was associated to the closest cortical region within the atlas. The assignment of white matter voxels to specific ROIs via proximity, was elegantly solved via distance maps. Finally, adjacent regions were fused based on prior anatomical knowledge to reduce the number of ROIs. This resulted in 31 ROIs defining larger brain regions, such as the frontal or occipital lobe, but also ventricles or the brainstem. Figure 4.1 schematizes the atlas construction process. Figure 4.2 displays the final product from four different views.

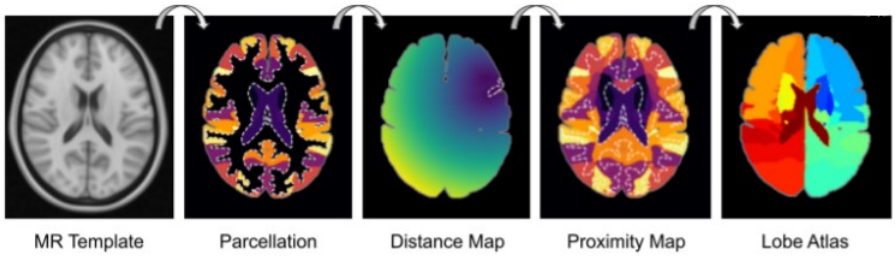

Figure 4.1: *Atlas construction pipeline for lesion localisation.*

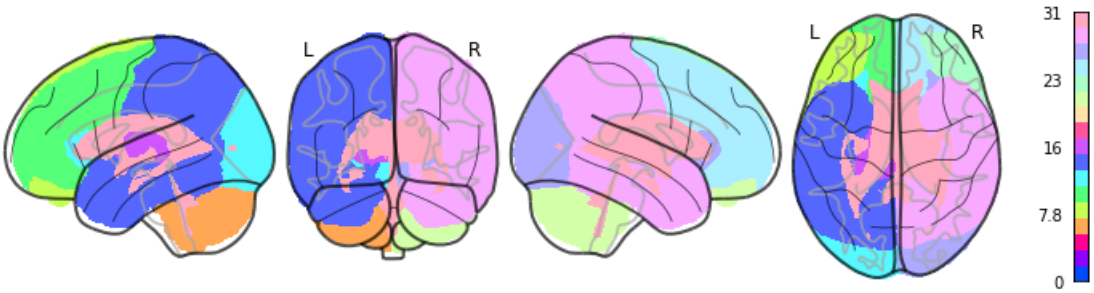

Figure 4.2: *Visual representation of the brain atlas from four different views. From left to right: sagittal left hemisphere, coronal, sagittal right hemisphere and axial.*

4.2.2 Image Preprocessing

All T1w scans were first skull-stripped [101], corrected for scanner field inhomogeneities [102], and spatially normalised to the MNI T1w template. The transformations for the affine registration and non-linear deformation were computed with ANTS (SyN) [103]. All other anatomical scans (FLAIR, SWI and GE) were co-registered to their respective T1w scan (acquired during the same training session). The rigid transformation, only allowing for translation and rotation, was also computed with ANTS. Both inverse transformations were concatenated and applied to the lobe atlas, to project it from MNI space to the native space of each of the scans with manually annotated lesions (FLAIR, SWI and GE) using nearest-neighbour interpolation.

4.3 Lesion Characteristics

4.3.1 Lesion Volume & Connectivity

Manual lesion segmentation allows for the computation of lesion characteristics such as the total lesion volume and cluster count (i.e. counting the number of disconnected lesion parts for a particular lesion type). The total lesion volume is computed by multiplying the number of voxels annotated as lesion with the individual voxel volume. The cluster count is the number of voxel-wise disconnected regions that comprise the lesion.

4.3.2 Lesion Co-occurrence

The differentiation of lesions into several lesion types associates each patient with a set of lesions, enabling an assessment of lesion patterns within the data. This is done by pairing up lesion types and computing a lesion co-occurrence matrix, with co-occurrence defined as the proportion of patients with lesion A, that also show lesion B:

$$C = \frac{N(A,B)}{N(A)}, \tag{4.1}$$

where C is the co-occurrence value, $N(A)$ equals the number of patients with lesion A, and $N(A,B)$ the number of patients with lesion A and lesion B. This method was applied to all lesion types and lesion groups. To improve visualization, intracerebral haemorrhage was neglected since this lesion type occurs in only one patient.

4.3.3 Atlas Based Features

An established atlas-based mapping of annotated lesions allows for another level of insight into the data, particularly through feature extraction. In this dissertation, the following features were computed:

- **Lesion volume**, i.e. the lesion volume within a ROI;
- **Relative lesion volume**, which equals the lesion volume within a ROI, divided by the total lesion volume.
- **Occupancy**, a binary metric for lesion location, equal to 1 if the region is deemed occupied by the lesion, while it is equal to 0, if it is not. The region was considered occupied when at least 5% of the total lesion volume was within the region [104].

Both lesion volume and relative lesion volume are useful when characterizing specific lesions, however, when visualized across all lesion types, tend to marginalize smaller lesions. To address this issue, an occupancy feature was incorporated. Put simply, if at least 5% of the lesion volume is located within the ROI, then the occupancy for said ROI is equal to 1 [105].

4.4 Bias Analysis

Lesions types were attributed their respective volume and cluster count. This is extended for lesion groups, by adding volumes and cluster counts within each group, for each patient. Since the lesions were annotated on two different MR sequences and scanners (SWI and GE were used interchangeably - see section 4.1.2), the impact of this variation on lesion volumes and cluster counts was analyzed, to assess acquisition biases. To do this, scanner and MR sequence group comparisons were made, qualitatively and quantitatively, for each lesion group. Moreover, statistical analysis was employed to further inspect this hypothetical impact.

Practically all patients were associated with only one scanner, therefore lesion group volumes associated with each scanner can be considered independent observations. MR sequences, on the other hand, are usually both present in the same scan, i.e. nearly all patients present volumes in FLAIR and SWI/GE sequences. Put simply, while volumes are being compared for different MR sequences, they still pertain to the same lesion group, within the same patient. Because of this, these observations are deemed dependent. This differentiation is key in selecting adequate statistical tests. An additional factor

is that lesion volume and cluster count data failed standard normality tests, therefore their distributions are not assumed.

Following these restrictions, scanner group differences were assessed using MWU tests, with the null hypothesis that the two groups have equally-shaped distributions, i.e. medians. With alpha set at 0.05, a p-value below alpha will reject the null hypothesis in favor of the alternative: there is a statistically significant difference in medians, i.e. scanner variation impacts lesion volumes and/or cluster counts.

MR sequence comparisons, since they were made on dependent observations, were carried out with WSR tests, under the null hypothesis that the median of group differences is zero against the alternative that it is different from zero. A median of differences different from zero is equivalent to a difference in groups. Additionally, the one-sided version of the test was also computed, with the null hypothesis that the median is positive, against the alternative that it is negative. Here, a two-sided p-value below alpha will indicate that there is a significant difference between the paired observations, whilst a one-sided p-value below alpha will indicate which group holds bigger values. Essentially, this translates to a MR sequence producing significantly bigger volumes and/or cluster counts, for the same lesion group, within the same patient. These claims are made with a confidence level of $100(1-\alpha)\%$, i.e. 95%.

Furthermore, all p-values were corrected for multiple comparisons mistakes, by controlling the FDR using the Benjamin-Hochberg method described in section 2.3.4.

4.5 Patient Outcome Prediction

4.5.1 Data Selection

The combination of patient clinical and imaging data, i.e. lesion segmentation and localisation, can be used to model patient outcome in function of said data, using binary logistic regression.

The first step in building this model is data selection, particularly by grouping patients in similar conditions, starting with patient outcome. In this dissertation, much like in the aforementioned literature (see section 3.2), the GOSe was used to perform this stratification, as patients with a GOSe score above 4 were attributed to the favourable outcome group, while patients with a score below or equal to 4 were classified as having an unfavourable outcome. Additionally, a time window of interest, termed hyper-acute, was defined as containing scans collected within at most, three days since injury, resulting in a cohort of 74 scans (27.7%). Within this time frame, three patients had two scans, with the latter being discarded for each case, bringing the total to 71 scans and patients. When this cohort is intersected with the GOSe stratification and lesion selection, the total number of patients decreases to 58. Figure 4.3 illustrates the distribution of scan time post-injury up to 30 days, with an identification of hyper-acute scans and outcome availability.

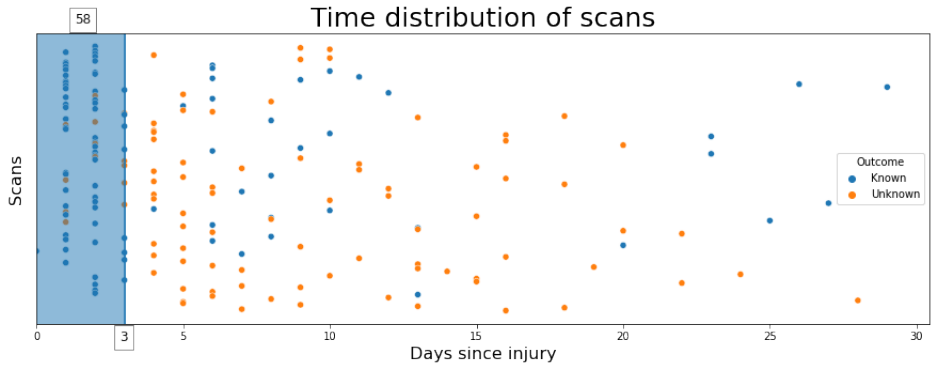

Figure 4.3: *Distribution of scan time post-injury up to 30 days. The vertical line at three days since injury timestamps the hyper-acute phase (blue area) that includes 74 scans from 71 patients. Blue dots represent scans pertaining to a patient with a known outcome, while orange dots stand for unknown outcomes. The final sample size stands at 58 patients/scans.*

An overview of the selected data is presented in Table 4.3, distributed across relevant variables, along with their respective categories.

Table 4.3: Overview of patients within outcome groups. Patients are distributed across relevant variables (top to bottom): sex, age, scanner, days since injury (DSI), Glasgow Coma Scale (GCS), injury mechanism and education.

Variable	Unfavourable outcome	Favourable outcome
Sex (M/F)	21/7	24/6
Age (mean [min, max])	40.9 [18, 71]	36.6 [17, 69]
DSI (mean [min, max])	1.7 [0, 3]	1.6 [1, 3]
Scanner (Trio/Verio)	18/10	19/11
GCS (%)		
3	12 (42.9 %)	7 (23.3 %)
4	2 (7.1 %)	1 (3.3 %)
5	2 (7.1 %)	1 (3.3 %)
6	2 (7.1 %)	3 (10.0 %)
7	3 (10.7 %)	6 (20 %)
8	4 (14.3 %)	2 (6.7 %)
9	-	2 (6.7 %)
10	-	1 (3.3 %)
11	-	2 (6.7 %)
12	1 (3.6 %)	2 (6.7 %)
13	1 (3.6 %)	-
14	-	3 (10.0 %)
15	1 (3.6 %)	-
Injury mechanism (%)		
Struck by object	1 (3.6 %)	1 (3.3 %)
Fall	7 (25.0 %)	7 (23.3 %)
Violence or assault	7 (25.0 %)	1 (3.3 %)
Road traffic collision	13 (46.4 %)	21 (70.0 %)
Education (%)		
GCSE/AS/A levels equivalent	4 (14.3 %)	13 (43.3 %)
Diploma	-	1 (3.3 %)
University degree	-	4 (13.3 %)
Below GCSE	-	-
Unknown	24 (85.7 %)	12 (40.0 %)

After data selection, the focus turns to the combination of lesion type and location, and how it can be

incorporated in the model. The rationale was to use lesion volume as a means to represent these features across a total of four models. Each model is built on top of the following base formula:

$$outcome = \beta_0 + \beta_1 \times age + \beta_2 \times GCS, \tag{4.2}$$

where β_0 is the y-intercept coefficient, *age* and *GCS* are numerical variables associated with each patient, and β_1 and β_2 are the coefficients for *age* and *GCS*, respectively. Since patient outcome is a binary variable, a favourable outcome is valued as 1 and an unfavourable outcome as 0, for all models.

4.5.2 Overview of Statistical Models

The first model ignores lesion type and location by attributing a SV value to each patient. This volume is equal to the sum of all lesion volumes within a patient, i.e. patient with lesion A and B will have a SV value of lesion volume A plus lesion volume B. Thus, the formula for the first model is as follows:

$$outcome = \beta_0 + \beta_1 \times age + \beta_2 \times GCS + \beta_3 \times SV \tag{4.3}$$

Importantly, for patients that show multiple lesion measurements, i.e. patient shows lesion A with two different volumes, one for a FLAIR sequence and other for SWI/GE, the FLAIR sequences were considered and others discarded, meaning that when a lesion is repeated on other sequence, only its FLAIR sequence volume measurement is considered. This is the rule for all but one patient, which presented repeated lesions but only had SWI/GE sequences. This convention is implemented across models.

The second model incorporates lesion types by extending SV to each lesion group and attributing them across patients. As such, each patient will be associated with their respective lesion group volumes, calculated by adding up all lesion volumes within the group. Therefore, the second model is computed with the following formula:

$$outcome = \beta_0 + \beta_1 \times age + \beta_2 \times GCS + \beta_3 \times LG_i, \tag{4.4}$$

where $i = (1, ..., N)$ and LG_i is the volume of lesion group i, pertaining to a patient with N lesion groups. This model is repeated for all lesion groups, producing five distinct formulas and sets of coefficients and p-values.

The third model introduces lesion location in the form of ROI-based lesion volumes by associating each patient with the lobe regions, each containing the sum of all lesion volumes within said region. Some regions lack representation in one of the outcome groups. To combat this imbalance, regions with median equal to zero, within outcome groups, were discarded, reducing the overall number of regions to 15. As in the second model, the third was repeated for each region, under the following formula:

$$outcome = \beta_0 + \beta_1 \times age + \beta_2 \times GCS + \beta_3 \times r_i, \tag{4.5}$$

where $i = (1, ..., N)$ and r_i is each region volume, out of a total of $N = 15$.

Lastly, the fourth model combines lesion groups with regions in a single formula, which is repeated for all possible group-lesion pairs:

$$outcome = \beta_0 + \beta_1 \times age + \beta_2 \times GCS + \beta_3 \times LS_i + \beta_4 \times r_k, \tag{4.6}$$

where $i = (1, ..., 5)$, $k = (1, ..., 15)$, LS_i is each group and r_k each region, with both being calculated as in the previous models.

Each model, or each formula repetition, produces a set of coefficients and p-values. The p-values were, as before, adjusted using the Benjamin-Hochberg multi-comparison correction. The coefficients are estimated by fitting the training data to the model formula, and can then be used to predict patient outcome given a set of new data.

Thus, the models were validated by making two sets of predictions. The first set was made with an input of the same data that the model was trained with; whereas the second set of predictions was made using a model trained only on Trio scanner data, given as input the Verio scanner data.

Both sets of predictions enable the construction of confusion matrices, which produce accuracy and error rate values for each set. The first set of predictions is associated with a training accuracy and training error rate, while the second set is associated with a test accuracy and test error rate, as defined in section 2.3.3. Having computed these values, the most promising models were compared with respect to their error rates.

All computations were made using Python on Jupyter Notebook, a web-based interactive computational environment that allows for coding in document or notebook format. Inside this enviornment, the statsmodels and scikit-learn Python packages were used to estimate the statistical models, while the Pandas and Seaborn packages were used to manage and visualize the data. The next chapter contains the results produced by this methodology.

Chapter 5

Results

This chapter describes the results produced throughout this dissertation.

5.1 Data Distribution

As one of the first stages in this project, a demographic characterization of the data offers an overview of data distributions across different variables. This begins by identifying the elements within said data, their respective count and distribution across relevant categories within the data. Table 5.1 shows an inventory of data elements: subjects, scans and lesions.

Table 5.1: Data overview. Lesion selection and grouping increases the quality of the data but decreases overall sample sizes.

Data	Count	Notes
Subjects	125	123 after lesion selection
Scans	267	253 after lesion selection
Lesions	63	20 lesion types were selected and grouped

Additionally, the data contains a substantial majority of male subjects and severe TBI cases, as 96 (76.8%) of all 125 subjects are male, and 70 (80%) experienced severe TBI. Nonetheless, the data presented less disparity with regards to age groups. Table 5.2 shows the sex and severity distributions, from left to right, while Table 5.3 contains the age distribution.

Table 5.2: Sex and severity distributions within subjects.

Sex	Number of Subjects (%)
Male	96 (76.8 %)
Female	29 (23.2 %)

Severity	Number of Subjects (%)
Severe	100 (80.0 %)
Moderate	14 (11.2 %)
Mild	11 (8.8 %)

Table 5.3: Age distribution within subjects.

Age	Number of Subjects (%)
10-19	18 (14.4 %)
20-29	34 (27.2 %)
30-39	22 (17.6 %)
40-49	22 (17.6 %)
50-59	14 (11.2 %)
60-69	13 (10.4 %)
70-79	2 (1.6 %)

5.2 Lesion Analysis

The following sections analyze the annotated lesion types within the database by first computing their co-occurrence, with the goal of identifying lesion pairs that were frequently observed together, and secondly, by locating lesion types within the previously described atlas (Section 4.2) to understand which brain regions were mostly affected by said lesions.

5.2.1 Lesion Patterns

Combining lesions together according to their group (see Table 4.2) produced high co-occurrence rates (Figure 5.1). For example, IPO co-occurred in 99% of the scans that also presented ILC. Similar rates were found for petechial and extra-axial haemorrhages, appearing frequently together with lesion cores and perilesional oedemas (co-occurrence rates all approximately 90%). The lowest values were found between VCH and all other lesion types (co-occurrence rates all approximately 73%).

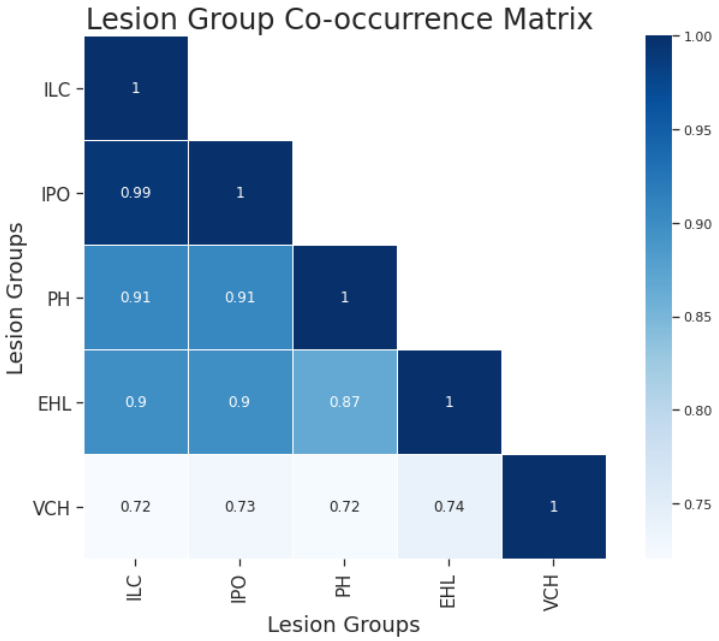

Figure 5.1: *Lesion group co-occurrence matrix. As lesion types are grouped, the resulting group co-occurrence values naturally increase as group pairs are much more frequent.*

The same process was done to investigate lesion type co-occurrence, with the respective matrix in Figure 5.2. This is an unbiased approach in terms of MR sequence, location or scanner, thus, as expected, the increase in number of lesions results in lower co-occurrence rates overall, as the information provided by the matrix also increases.

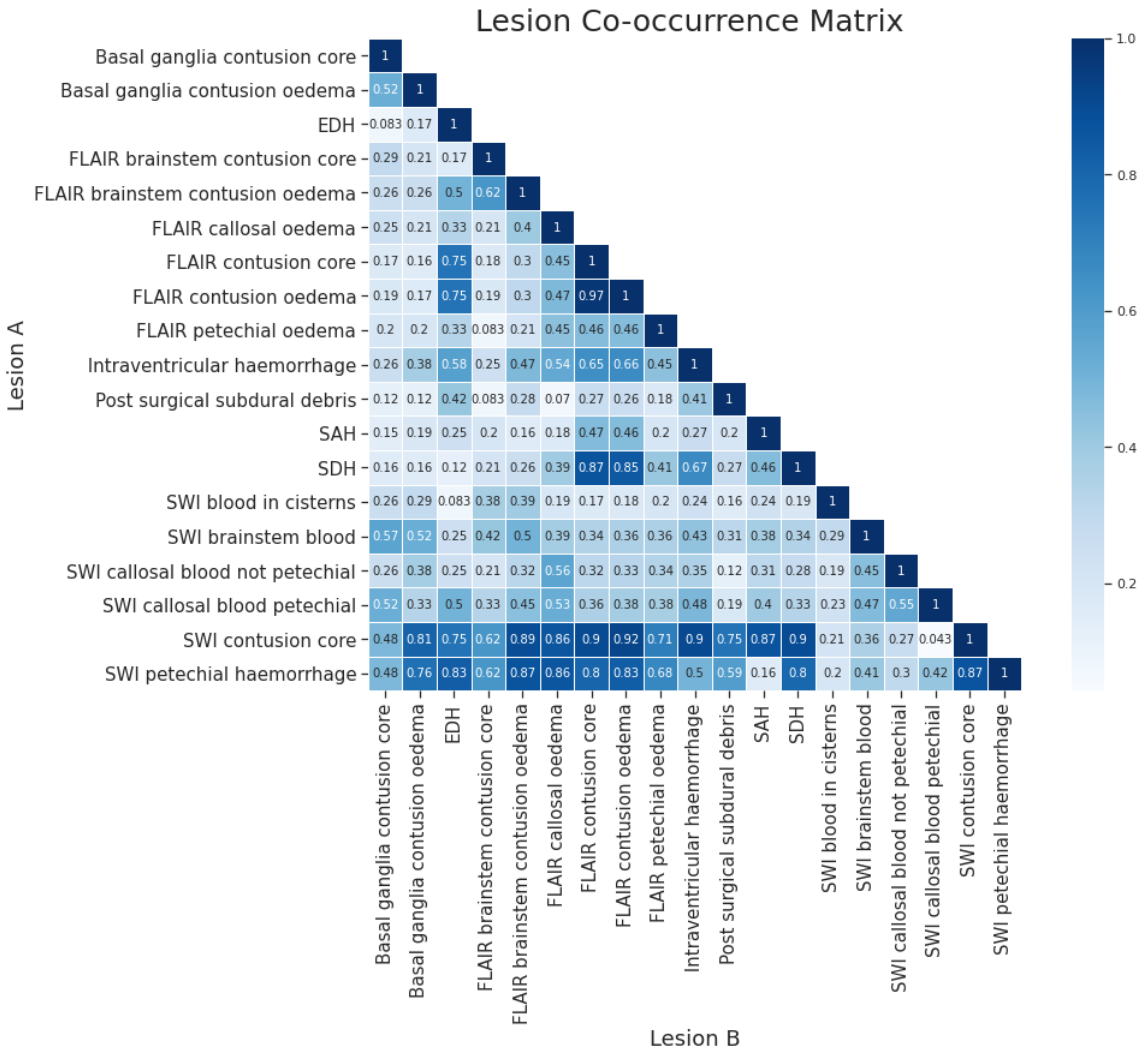

Figure 5.2: *Lesion co-occurrence matrix. Each square holds the ratio between patients with Lesion A and B, and patients with Lesion A. The matrix format coupled with colored mapping allows for an easier identification of high pair-specific co-occurrences.*

The pairwise co-occurrence matrix highlights lesion types that were often found together in patient scans. FLAIR contusion oedema and FLAIR contusion core exhibit the highest co-occurrence rate at 97% and share very similar values across other lesion types, e.g. both have a co-occurrence rate of 75% with EDH, 30% with FLAIR brainstem contusion oedema and 46% with FLAIR petechial oedema. However, the 97% co-occurrence rate between FLAIR contusion core and FLAIR contusion oedema decreases when the same lesion pair is attributed to a specific brain region: FLAIR brainstem contusions (core and oedema) have a co-occurrence rate of 62% and basal ganglia contusions have 52%.

Additionally, FLAIR brainstem contusion core and FLAIR contusion core did not appear together often (co-occurrence rate at 18%). This is analogous for contusion oedema despite an increase to 30%. Similar values were also found for basal ganglia contusions (17% co-occurrence rate between FLAIR contusion core/oedema and basal ganglia contusion core/oedema).

As expected, the highest co-occurrence values seem to concentrate around the most frequent lesion types - SWI contusion core and petechial haemorrhage and FLAIR contusion core and oedema (see Table 4.2). Intraventricular haemorrhage and SDH also appear in most subjects (66.9% and 79.8% of subjects, respectively - Table 4.2), although their respective co-occurrence values across the other lesion types seem less elevated. Interestingly, the lowest co-occurrence rate (4.3%) was found between the most frequently annotated lesion - SWI contusion core - and SWI callosal blood petechial.

5.2.2 Localisation Features

Along with pairing frequency, lesion types were also analyzed with regards to their location. The atlas described in section 4.2 allows for the computation of features like lesion volume and occupancy (Section 4.3) across 31 distinct regions, which can then be visualized in a lesion localization matrix (Figure 5.3).

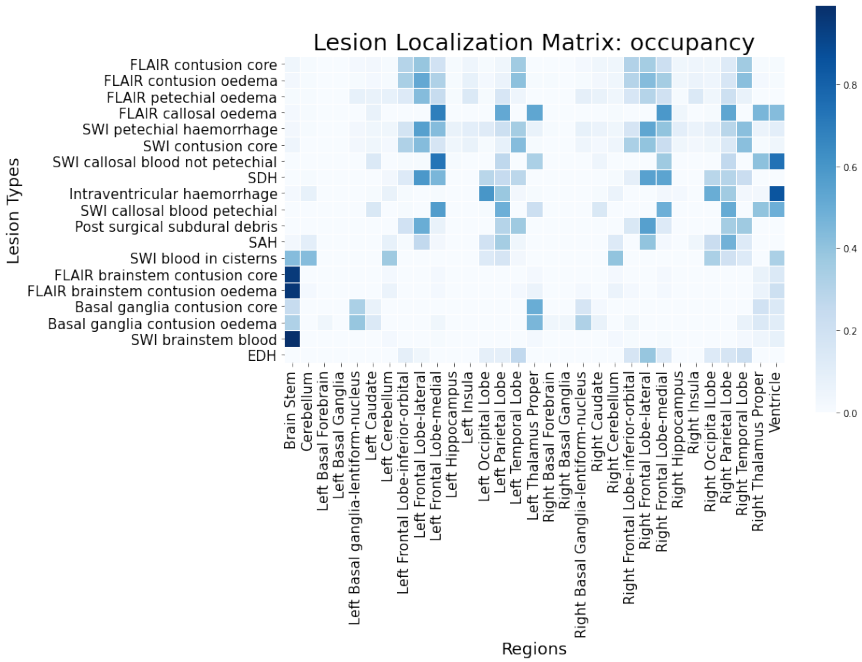

Figure 5.3: *Lesion Localization matrix. Each square holds the lesion-specific (y-axis) average occupancy value for that ROI (x-axis). Color-mapping helps identify high-occupancy zones.*

As expected, contusions annotated within the brainstem (i.e. FLAIR brainstem contusion core and oedema) show highest occupancy values in the brainstem (occupancy = 0.95, 0.96, respectively), however, these same lesions also appear in other regions such as the ventricles (occupancy = 0.14, 0.21, respectively). This is analogous with intraventricular haemorrhage lesions, which naturally have their highest occupancy value associated with the ventricles (occupancy = 0.85), yet also appear in the left and right occipital (occupancy = 0.6, 0.5) and parietal (occupancy = 0.38, 0.36) lobes. Importantly, the presence of ventricular haemorrhage seemed to be frequently associated with all three callosal lesions, i.e. FLAIR callosal oedema and SWI callosal blood petechial or not petechial, with occupancy values of 0.43, 0.49 and 0.74, respectively. Additionally, SWI blood in cisterns are similarly distributed across the brainstem (0.43), cerebellum (0.43 for cerebellum, 0.37 for left cerebellum and 0.4 for right), the ventricles (0.33) and the right occipital lobe (0.33).

Overall, the most occupied regions coincide with the largest regions: left and right frontal lobe lateral/medial and left and right parietal and temporal lobes (compare Figure 5.3 and Figure A.1 in the Appendix). After these major regions, and the already mentioned ventricles and brainstem, regions such as the left and right basal ganglia-lentiform-nucleus, thalamus proper and cerebellum are less but still visibly occupied. Lastly, the left and right basal forebrain, basal ganglia, caudate, insula and hippocampus are barely occupied, with a maximum occupancy value among these regions of 0.15, for SWI callosal blood petechial in the left and right caudate.

Additionally, these findings are further explored in Figure 5.4, which show 3D glass brain templates with overlaying lesion occupancy maps.

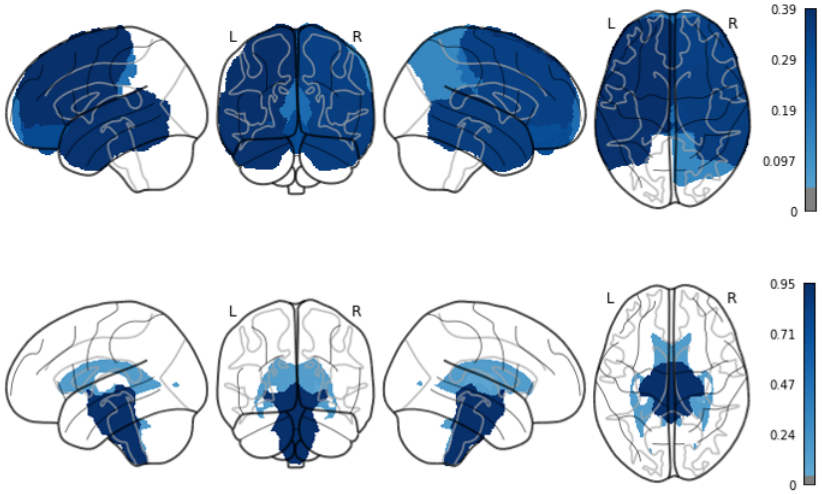

Figure 5.4: *Color-mapping of FLAIR contusion core (top) and FLAIR brainstem contusion core (bottom) average occupancy values for each ROI on a 3D brain template across 4 different views: sagittal left hemisphere, coronal, sagittal right hemisphere and axial.*

The top maps show the brain regions affected by FLAIR contusion core whilst the bottom ones show them for FLAIR brainstem contusion core. These lesion types are clearly differentiated by the size of the regions in which they were found while also mirroring the matrix shown before, as the brainstem lesion is mapped across the same two regions - the brainstem and, more subtly, the ventricles. In contrast, FLAIR contusion core is mapped across the temporal and frontal lobes. This lesion comparison enables the clear distinction between FLAIR brainstem contusion core and FLAIR contusion core.

5.3 Bias Analysis

The lesions were annotated on scans collected on two different scanners and MR sequences (Section 4.1). This section aims to investigate how these factors affected lesion volumes and number of lesion clusters (cluster count), by comparing lesion groups across scanner and sequences. While scans were distributed more evenly across both scanners (Table 5.4), some lesion groups were mostly annotated on either FLAIR or SWI/GE scans. For example, IPO were predominantly delineated on FLAIR ($n = 117$) and rarely on SWI/GE ($n = 6$) scans. On the contrary, petechial haemorrhagic lesions were primarily segmented on SWI/GE scans ($n = 112$) and only on a single FLAIR scan. VCH, despite being mostly annotated on SWI/GE ($n = 86$), were also segmented in several FLAIR scans ($n = 25$). The biggest disparity in number of scans across scanners was found for ILC (Trio: $n = 66$, Verio: $n = 53$).

Table 5.4: Number of scans containing each lesion group across scanners (Trio and Verio) and MR sequences (FLAIR and SWI/GE).

Lesion group	Number of Scans			
	Trio	Verio	FLAIR	SWI/GE
ILC	66	53	113	115
IPO	64	53	117	6
PH	61	51	1	112
EHL	59	50	100	91
VCH	44	44	25	86

5.3.1 Scanners

Figure 5.5 shows clear differences in total lesion volumes and cluster counts across scanners for each of the five lesion groups.

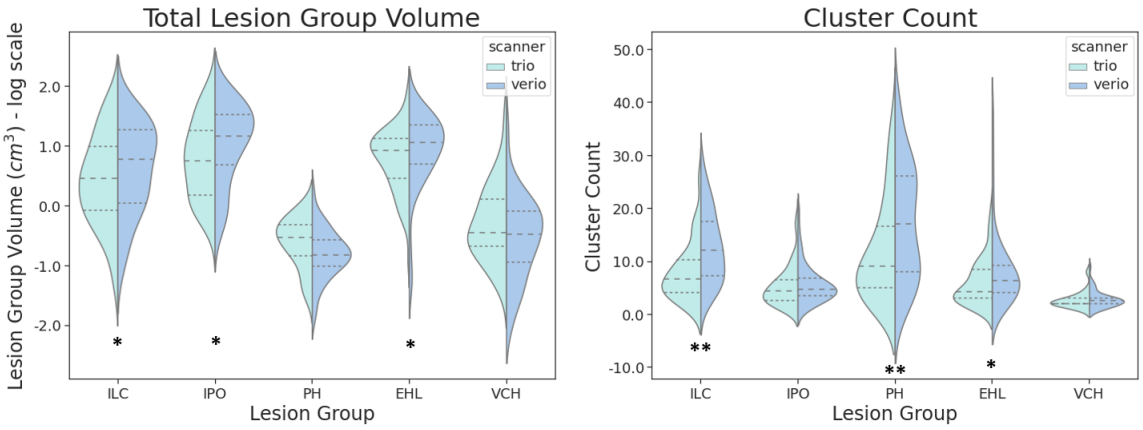

Figure 5.5: *Violin plots displaying the differences between scanner groups in total lesion volume (left) and cluster count (right). The dashed lines represent the third quartile, the median and the first quartile (from top to bottom). Cluster count values above 50 were removed before plotting the data to improve visualization. Asterisks indicate statistically significant differences (* = $p < 0.05$; ** = $p < 0.005$)*

In fact, the figures seem to indicate a trend in which Verio scans are associated with larger lesion group volumes and cluster counts. This trend was in most cases supported by quantitative differences and statistical testing: for example, ILC lesions showed significantly larger volumes and cluster counts

(MWU; $p = 0.0095$, $p < 0.0001$, respectively) on Verio (average ILC volume of 15.8 cm^3, average ILC cluster count of 17.1 clusters) than on Trio (9.4 cm^3, 8.0 clusters) (Table 5.5).

The EHL group showed similar behavior (MWU; $p = 0.0150$ for volumes and $p = 0.0298$ for cluster counts), despite having more similar cluster counts across scanners, with an average of 6.7 clusters in Trio and 8.5 in Verio. IPO lesions showed significantly larger volumes on Verio (MWU; $p = 0.0114$), however, cluster counts were not significantly impacted by scanner variation (MWU; $p = 0.1414$). Inversely, PH show a significant increase in number of clusters (MWU; $p < 0.0001$) when annotated on Verio scans, yet, their volumes were not significantly impacted by the scanner used (MWU; $p = 0.1438$). Lastly, this trend was not observed for VCH lesions, where both volumes and cluster counts were not significantly impacted by scanner variation (MWU; $p = 0.1543$, $p = 0.1249$, respectively).

Table 5.5: Quantitative differences between scanner Trio and Verio in total lesion volume and cluster count for each lesion group. MWU p-values regarding scanner comparison are also presented, with an asterisk identifying p-values below alpha = 0.05.

Lesion group	**Total Lesion Volume (cm^3) (mean [median])**			**Cluster Count (mean[median])**		
	Trio	Verio	MWU (p)	Trio	Verio	MWU (p)
ILC	9.4 [2.8]	15.8 [8.6]	0.0095*	8.0 [6.6]	17.1 [12.3]	<0.0001*
IPO	16.2 [5.5]	23.8 [14.5]	0.0114*	5.3 [4.4]	5.7 [4.7]	0.1414
PH	0.5 [0.6]	0.4 [0.3]	0.1438	17.9 [9.2]	25.7 [22.3]	<0.0001*
EHL	10.1 [8.0]	14.5 [11.7]	0.0150*	6.7 [4.2]	8.6 [6.5]	0.0298*
VCH	1.7 [0.3]	1.0 [0.3]	0.1543	2.5 [2.0]	2.8 [2.6]	0.1249

Overall, scanner variation seemed to have impacted four out of five lesion groups, with Verio scans generally producing larger lesion volumes and cluster counts. Importantly, in the case of petechial haemorrhagic lesions, there was no indication that either scanner significantly impacted total lesion volume (MWU; $p = 0.1543$), despite Figure 5.5 suggesting that Trio scans show increased volumes for this group.

5.3.2 MR Sequences

In Figure 5.6, the disparity in number of scans (see Table 5.4) is again demonstrated in IPO and PH, which were mainly annotated on FLAIR or SWI/GE, respectively.

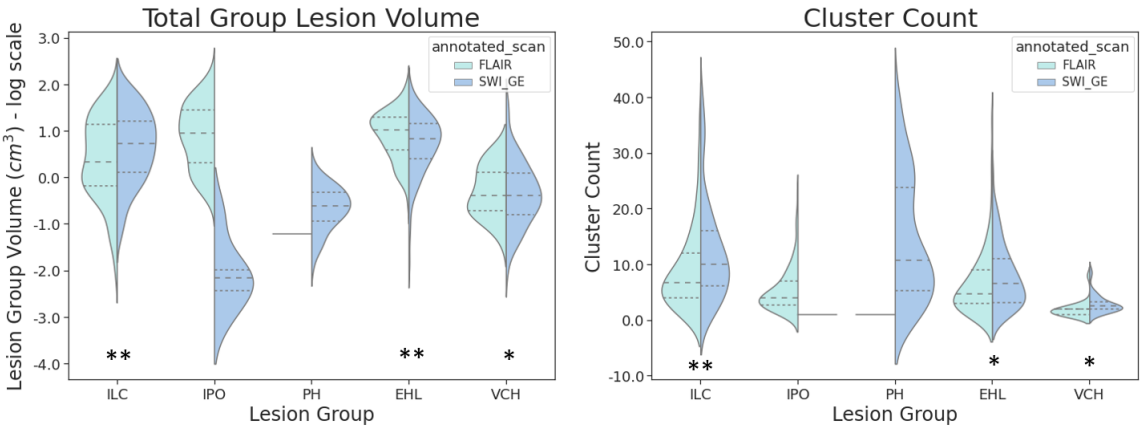

Figure 5.6: *Violin plots displaying the differences between MR sequence groups in total lesion group volume (left) and cluster count (right). The dashed lines represent the third quartile, the median and the first quartile (from top to bottom). Sample size imbalance disrupts visualization of blood petechial and contusion oedema groups. Cluster count values above 50 were removed to improve visualization. Asterisks indicate statistically significant differences (* = $p < 0.05$; ** = $p < 0.005$)*

Lesion volumes and cluster counts, appeared to be affected by the MR sequence used for manual segmentation. Indeed, ILC lesions showed significantly larger volumes and more clusters (WSR; $p < 0.001$ in both cases) on SWI/GE scans (median volume of 5.1 cm^3 and 10.4 median cluster count) than on FLAIR (2.8 cm^3, 6.2 clusters) (Table 5.6). Similarly, VCH lesions also appeared significantly bigger and with more clusters on SWI/GE scans (WSR; $p = 0.0093$, $p = 0.0016$, respectively). Interestingly, EHL lesions showed significantly bigger volumes (WSR; $p = 0.0002$) on FLAIR scans (median of 11.2 cm^3) than on SWI/GE (median of 6.1 cm^3), despite showing significantly more clusters on SWI/GE (WSR; $p = 0.0197$). It should be noted that two-sided WSR tests indicate there is a significant difference in volumes/clusters for each MR sequence, whereas the one-sided WSR tests the hypothesis that one specific MR sequence produces bigger volumes/clusters (see Section 4.4). Table 5.6 shows the two-sided WSR p-values, as the one-sided test produced similar p-values, all below alpha.

Table 5.6: Quantitative differences between MR sequences in total lesion volume and cluster count for each lesion group. WSR p-values regarding each comparison are also presented, with an asterisk identifying p-values below alpha = 0.05.

Lesion group	**Total Lesion Volume (cm^3) (mean [median])**			**Cluster Count (mean[median])**		
	FLAIR	SWI/GE	WSR (p)	FLAIR	SWI/GE	WSR (p)
ILC	12.0 [2.8]	11.9 [5.1]	<0.0001*	12.4 [6.7]	12.3 [10.1]	<0.0001*
EHL	12.6 [11.2]	12.9 [6.1]	0.0002*	7.6 [4.8]	7.1 [6.5]	0.0197*
VCH	1.3 [0.4]	1.4 [0.3]	0.0093*	2.4 [2.0]	2.7 [2.9]	0.0016*

In this case, MR sequence variation impacts all lesions considered, despite average total lesion volume and cluster count values appearing fairly similar (Table 5.6). Median values are less balanced for ILC and EHL, though they remain relatively similar for VCH lesions (FLAIR: 0.4 cm^3, SWI/GE: 0.3 cm^3).

5.4 Outcome Analysis

This section investigates patient outcome, as outcome groups (Table 4.3; Section 4.5) were analyzed and modelled, thus enabling outcome predictions.

5.4.1 Lesion Volume

When inspected across lesion groups, patients with unfavourable outcomes (GOSe ≤ 4) tended to have increased lesion volumes (Figure 5.7).

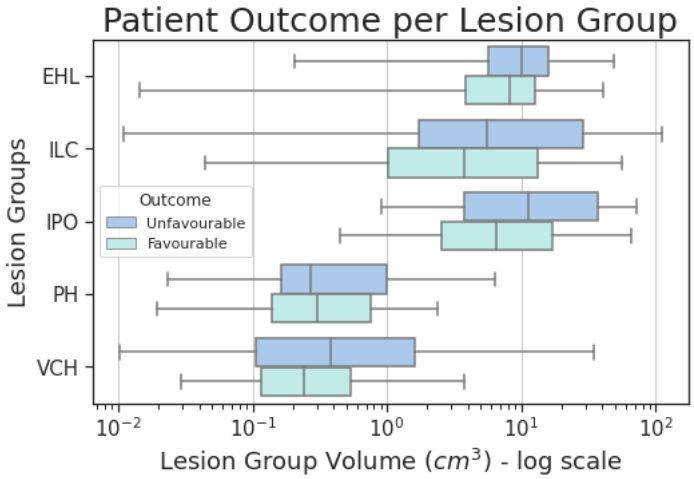

Figure 5.7: *Outcome groups boxplots for each lesion group, across lesion volume (x-axis). Volumes were averaged for each lesion group. Patient outcome is stratified in unfavourable (blue) and favourable outcome (cyan blue), as indicated in the legend.*

This volume difference was relatively pronounced in IPO, where patients with favourable outcome displayed lower lesion volumes (median = 6.5 cm^3) than patients with unfavourable outcome (median = 11.3 cm^3) (See Table A.2 in Appendix). ILC showed a similar pattern (favourable outcome: 3.6 cm^3, unfavourable outcome: 7.1 cm^3), which is coherent with previous findings of high co-occurrence between these groups (99%; Section 5.2.1). The association between higher lesion volumes and unfavourable outcomes was much more subtle for EHL (favourable: 8.9 cm^3; unfavourable: 9.2 cm^3) and VCH (0.2 cm^3; 0.3 cm^3). Moreover, this trend is not seen in petechial lesions, where the volumes associated with unfavourable outcomes appear to be bigger, yet the median volume increases by a small margin (approximately 0.03 cm^3) for favourable outcomes.

The same trend (higher volumes in patients with unfavourable outcome) was observable across most lesion types (Figure A.2).

5.4.2 Outcome Modelling

To further substantiate the relationship between lesion volume and patient outcome, four binary logistic regression models were fitted (Section 4.5.2). The first model neglects lesion type and location, and instead attributes a SV value to each patient. Both the SV and a GCS scores were found to be significantly associated with patient outcome ($p = 0.0466$ for both; Table 5.7): patients with unfavourable outcome had on average a SV of 94.3 (median: 65.9) cm^3 in contrast to 45.0 (27.1) cm^3 in patients with favourable outcome.

Table 5.7: Coefficients, p-values and corrected p-values for each predictor in model 1, including the intercept. GCS and SV present p-values below alpha ($\alpha = 0.05$), with or without multi-comparison correction.

Model 1: Outcome ~Age + GCS + SV			
Predictors	**Coefficients**	**P-values**	**Corrected p-values**
Intercept	0.2781	0.7320	0.7320
Age	-0.0205	0.3290	0.4386
GCS	0.2368	0.0233	0.0466
SV	-0.0153	0.0171	0.0466

In fact, a larger SV significantly decreased (coefficient: -0.0153, $p = 0.0466$) the likelihood of having a favourable outcome (favourable and unfavourable outcome was represented with 1 and 0 for the binary prediction, respectively; see Section 4.5). Conversely, higher GCS scores significantly increased the likelihood of having a favourable outcome (0.2368, $p = 0.0466$). This is expected as higher GCS scores indicated that the patient is suffering from milder TBI. Interestingly, age was not a significant predictor of outcome, despite its coefficient suggesting that older patients may have a less favourable (unfavourable) outcome (-0.0205, $p = 0.4386$).

Additionally, the model was fitted to the training data and correctly predicted patient outcome with an accuracy of 70.7%, which equals a training error rate of 29.3% (100% − 70.7%). The model was then trained using only Trio scanner data, and tested on Verio scanner data, producing an accuracy of 66.7%, and subsequently, a test error rate of 33.3% (Figure 5.8).

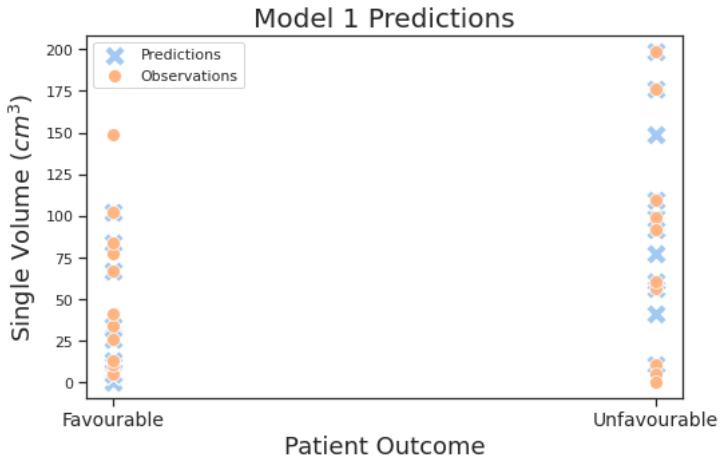

Figure 5.8: *Model 1 predictions for Verio scanner data. Outcome groups (x-axis) are plotted against SV (y-axis). Model predictions are represented by blue crosses, while orange spheres represent observed outcomes.*

The second model (Equation 4.4; Section 4.5) aims to reveal the relationship between lesion volumes of specific lesion groups and patient outcome (Table 5.8).

Uncorrected p-values showed a trend of patient outcome affected by volumes of ILC and IPO ($p = 0.0394$, $p = 0.0464$, respectively). However, this effect vanished after correcting p-values for multiple comparisons and no significant associations between patient outcome and either lesion group were found (both: $p = 0.1160$). EHL ($p = 0.2164$; corrected $p = 0.2992$), VCH ($p = 0.2393$, corrected $p = 0.2992$) and petechial ($p = 0.7602$, corrected $p = 0.7602$) lesions were not found to be significantly linked to patient outcome, neither before nor after correcting the p-values for multiple comparisons. As for the coefficients, all lesion groups show negative values, suggesting that an unitary increase in the volume of any of these groups would lower the likelihood of a favourable outcome.

Table 5.8: Coefficients, p-values and corrected p-values for each lesion group, obtained through model 2. Contusion groups show initial promising p-values, that increase well above alpha ($\alpha = 0.05$) when corrected with multi-comparison tests.

Model 2: Outcome ~Age + GCS + Lesion group			
Lesion group	**Coefficients**	**P-values**	**Corrected p-values**
EHL	-0.0149	0.2164	0.2992
ILC	-0.0452	0.0394	0.1160
IPO	-0.0217	0.0464	0.1160
PH	-0.0680	0.7602	0.7602
VCH	-0.4845	0.2393	0.2992

The absolute highest coefficient was associated with VCH (-0.4845), whereas EHL (-0.0149), ILC (-0.0452), IPO (-0.0217) and PH (-0.0680) were smaller. Each model was used to predict patient outcome, enabling a comparison between training accuracy and test accuracy across lesion groups (Figure A.3). As expected, model test accuracy is lower than training accuracy for all lesion groups. Furthermore, all model test accuracy values are above 50%, i.e. test error rates are all, inversely, below 50%. The ILC model stands out with the highest test accuracy (approximately 66.7%), while models trained on VCH and PH volumes provided the lowest test accuracy (52.4%). Interestingly, the EHL model exhibits the highest training accuracy (70.7%) and the ILC model shows a relatively small variation between training and test accuracy (approximately 2.3%). Overall, the ILC model seems to be the most promising, which is why it was selected for later model comparison.

The third model introduces ROI-based lesion volumes by associating each subject with their 15 lobe regions, each containing the sum of all lesion volumes within said region (Section 4.5). Figures 5.9 and A.4 show the corresponding p-values and model accuracies for each region.

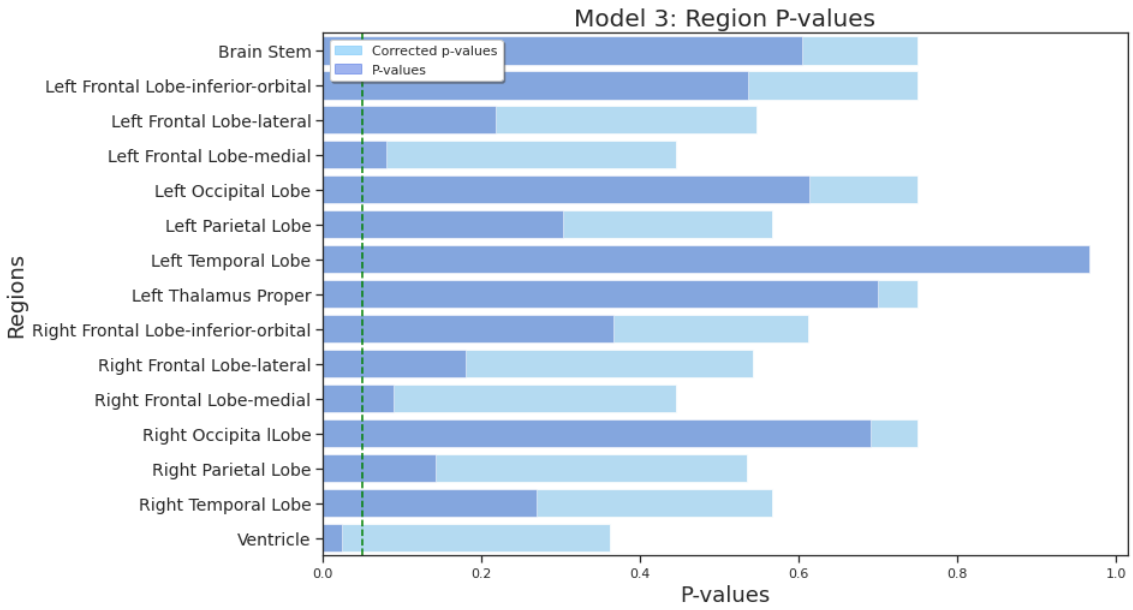

Figure 5.9: *Region (y-axis) p-values (x-axis) obtained through model 3. Corrected p-values are represented by light cyan colored bars. The dotted green line marks alpha ($\alpha = 0.05$).*

After correcting for multiple comparisons, no region was found to be significantly associated with patient outcome. However, before said correction, the ventricles region showed a significant association ($p = 0.0241$) with an unitary increase in volume within this region resulting in the biggest decrease (coefficient: -1.0519) of the likelihood of having a favourable outcome.

Similarly, left and right fronto-medial lobes also showed fairly low p-values (left: $p = 0.0809$, right: $p = 0.0890$), as did the right parietal lobe ($p = 0.1424$). However, multi-comparison correction increased all p-values, with the resulting lowest value still being attributed to the ventricles ($p = 0.3621$). Some regions, such as the left or right occipital lobe, brainstem and left thalamus proper show relatively high p-values, irrespective to the correction. The left temporal lobe exhibited the same high p-value before and after correction ($p = 0.9657$) with a coefficient closest to zero (< -0.001). Additionally, the left occipital lobe was the only region with a positive coefficient (0.0859, $p = 0.7497$).

The highest test accuracy (approximately 62%) was found in two regions: the left parietal lobe and the ventricles. On the other hand, the left temporal lobe and brainstem were tied for the lowest test accuracy (approximately 47.6%). For all but the previous four regions, the test accuracy was either 52.4% or 57.1%, as both training and test accuracy values were often repeated. Since the ventricles region model showed the most promising prediction accuracy,it was selected for model comparison.

Lastly, the fourth model includes lesion volume in a combination of lesion groups and locations (Equation 4.6, Section 4.5). Figure 5.10 shows the corrected lesion group p-values for each region, i.e. combination, while Figure A.5 shows the model test accuracy for each combination. Region p-values were also analyzed, with findings shown in the appendix (see Figure A.6).

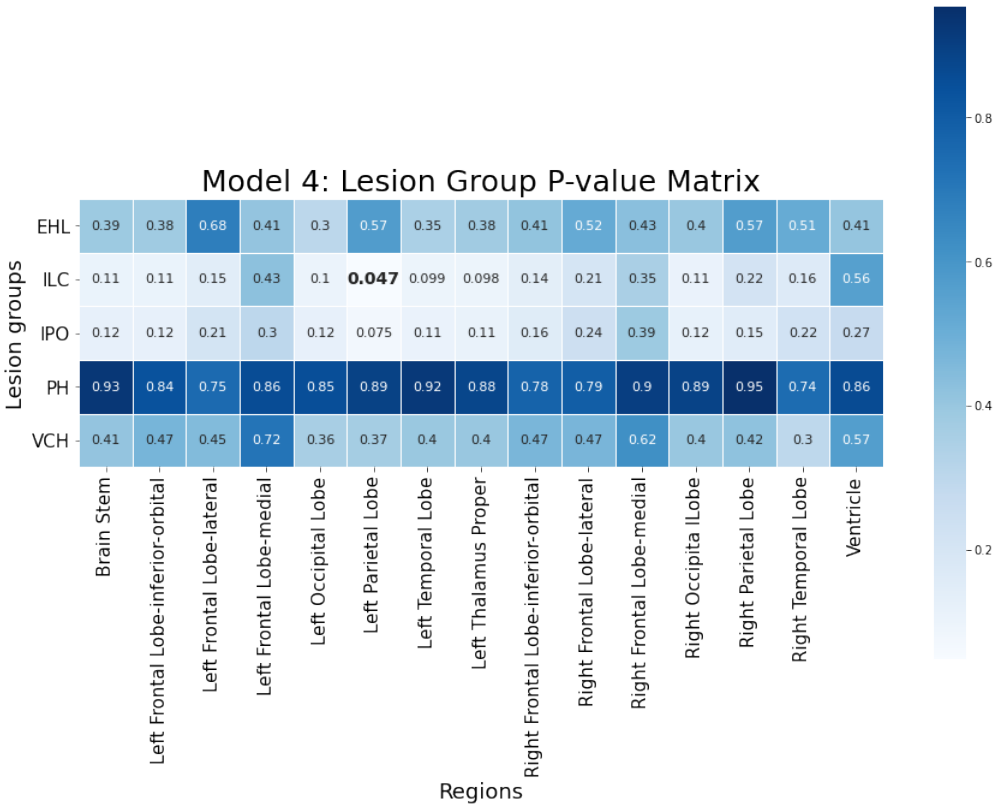

Figure 5.10: *Lesion group corrected p-values matrix. Each square holds the lesion group (y-axis) p-value produced through a modelled combination with a singular region (x-axis). Each p-value was corrected with multi-comparison tests.*

ILC lesions show both the lowest p-value ($p = 0.047$), when combined with the left parietal lobe, and the maximum model test accuracy at 67%, when combined with this same region, the left occipital lobe and the left thalamus proper. These results suggest that there is a significant association between patient outcome and ILC, when these are combined with the left parietal lobe.

5. RESULTS

Overall, ILC and IPO lesions seem to result in lower p-values (lowest: approximately $p = 0.047$, $p = 0.075$, respectively) while PH (approximately, $p = 0.74$) shows higher. EHL and VCH (approximately $p = 0.30$ in both) also show moderately high p-values. In fact, no other region or group p-values were found to be below the significance threshold ($\alpha = 0.05$). Model test accuracy varied between 48% and 67%, as the former - 48% - was repeated eight times against five for the latter. Three out of these five were associated with ILC, while the other two were associated with petechial and EHL. Overall, ILC appear to have the highest accuracy values out of the other lesion groups. As for regions, the left parietal (average accuracy across lesion groups of approximately 62%) and ventricles (61%) also seemed to stand out in terms of accuracy. As such, the chosen model to proceed for model comparison was the combination of contusion core lesions and the left parietal lobe region.

As a concluding analysis, all constructed models were compared with regards to their training and test error rate (Table A.3; See Section 4.5). The group or region which produced the lowest p-values and/or highest test accuracies, were chosen to represent the second to fourth model. In the second model, ILC stood out as the most promising, whereas the third and fourth model are represented by the ventricles and the left parietal lobe (combined with ILC), respectively.

The third model shows a relative high test error rate (approximately 38.1%) in comparison with the three other models, that showed lower training error rates (min: 27.6%, max: 31.0%). The test error rate did not continue to increase in the fourth model (approximately 33.0%), as it instead decreases to slightly less than in the first model (approximately 33.3%). These results suggest that the inclusion of lesion group and location slightly increase model test accuracy, thus providing an improved patient outcome prediction.

Chapter 6

Discussion

This chapter contextualizes the lesion localisation method and its potential for outcome prediction. Moreover, future research directions are suggested.

6.1 Data Distribution

The data were found to be mostly comprised by male patients ($n = 96$, 76.8%) and severe TBI cases ($n = 100$, 80.0%) (See Table 5.2). A pronounced male majority is not unusual, given that literature has suggested that males are nearly three times more likely to suffer a TBI [23]. The slightly higher prevalence of severe TBI cases was also expected, as the selected cases had manually annotated scans, which shows an inherent bias towards severe cases (mild cases often do not have lesions visible to the naked eye). In contrast, age groups were relatively more uniform (See Table 5.3), as their distribution seems skewed towards younger subjects (See Figure A.8 in the Appendix).

6.2 Lesion Co-occurrence

Lesion co-occurrence was analyzed first among lesion groups - Figure 5.1 - and then across lesion types (Figure 5.2), which resulted in an overall decrease in co-occurrence rates. This is likely caused by the transition from lesion groups to lesion types. Although the latter provide a more fine grained analysis of specific lesion types, it decreased the overall number of samples per lesion category. For example, splitting the ILC category into several lesion types (i.e. FLAIR contusion core, SWI contusion core or FLAIR brain stem contusion core) inevitably lowered the lesion appearance in those sub-categories. Lower occurrences of a lesion has a direct effect on the co-occurrence with other lesions types, thus leading to lower numbers observed. This has to be considered when interpreting the results. Additionally, lesion types were considered irrespective to MR sequence, location or scanner, which limits the interpretations stemming from such analysis. Nonetheless, it enables some insight into the nature of each lesion pair. For example, FLAIR contusion core and contusion oedema showed the biggest co-occurrence rate, at 97%. This indicates that these lesions were annotated together most of the time, as they are different lesion parts of the same contusion. It is unlikely that one of these lesions is identified without the other. However, this same pair of lesions was associated with the brainstem and basal ganglia, where it showed lower co-occurrence rates (brainstem: 62%; basal ganglia: 52%).

A potential explanation for this decrease in co-occurrence lies in the size of said regions, as the brainstem and basal ganglia are some of the smallest regions in the atlas (See Figure A.1). Put simply, small specific regions are potentially less likely to contain lesions than the rest of the brain. This may lead to lower total number of lesions within these regions, which directly affects the co-occurrence (as mentioned above).

As demonstrated, the co-occurrence matrix can provide valuable insight for investigating specific lesion pairs. However, disparity in lesion frequencies appeared to impact the matrix as the biggest rates concentrated around the most frequent lesion types. This could limit the aforementioned insight, as it produces ambiguity in the interpretation of the matrix: a high co-occurrence rate could be due to the nature of the lesions (e.g. lesions occur on adjacent regions) or to one of the lesions being highly frequent in the data. Future research efforts may combat this limitation by normalizing lesion frequencies as much as possible and cooperating with clinicians to determine reasons for high or low co-occurrence values.

6.3 Lesion Localisation

Lesion localisation was summarized through an occupancy matrix (Figure 5.3), with results demonstrating the advantages and limitations in using an automated, atlas-based localisation method. Regarding said limitations, some lesions were seemingly located in the wrong region. For example, brainstem lesions were located mostly in the brainstem (occupancy = 0.95), as expected, but also, in the ventricles (0.21). This can be explained by a misalignment between the atlas and the scan of the injured brain, as the ventricles are directly adjacent to the brainstem (Figure 5.4). Indeed, spatial normalisation of lesioned brains is a notoriously error prone task [79], thus it is unlikely that falsely located lesions could be avoided altogether. Another limitation inherent to atlas-based brain parcellation is volume heterogeneity among ROIs, as the largest regions contained most of the lesion types. Different atlases highlight different regions and will affect lesion localisation findings. For instance, the atlas used in this dissertation does not include the corpus callosum, a region that has been shown to impact patient outcome [17]. Therefore, different atlases could be tested in the future and a choice should be discussed closely with clinicians to answer questions relevant to TBI.

Despite these limitations, the localisation method was successfully used to contrast FLAIR brainstem contusion core and FLAIR contusion core (Figure 5.4) Indeed, this lesion comparison pathway provides a clear sanity check for the lesion localisation method and can be used to investigate interesting pairs found in the aforementioned co-occurrence matrix (Figure 5.2).

6.4 Scanner and MR Sequence Biases

Lesions showed different size and connectivity, i.e. number of clusters, when compared across the two scanners (Table 5.5). However, the measurements were independent, so this impact can be purely coincidental, i.e. patients that were scanned on Verio just happened to have bigger lesions and/or more clusters. Moreover, MR sequence may be impacting both comparisons, particularly since SWI sequences were nearly all attributed to the Verio scanner. Nonetheless, these results suggest that scanner bias could significantly impact results and should be expected in multi-scanner databases. Analysing the true bias introduced by different scanners would require scanning patients on two (or more) scanners withing a short time frame. This is, however, infeasible due to the patient discomfort and care management. Furthermore, different MR sequences seemingly affected volumes/clusters in all lesion groups considered (Table 5.6).

Particularly, SWI/GE displayed bigger volumes/more clusters in all cases except for EHL volumes, which appeared bigger on FLAIR. This effect of larger TBI lesions on SWI has been reported before [16, 106, 107], with authors suggesting that SWI is inherently more sensitive towards certain TBI lesions. Indeed, SWI, unlike the more conventional sequences (e.g. FLAIR, T2w), relies on both phase and magnitude images (see Section 2.2). Therefore, SWI is inherently more sensitive to compounds that affect the local magnetic field, such as the paramagnetic blood products associated with haemorrhagic TBI lesions [108]. However, Geurts et al. found contrary evidence, as SWI and FLAIR showed similar lesion volumes at a median of seven weeks post-injury [50]. In response to this, Spitz et al. conceded that the SWI superiority in detecting TBI lesions over FLAIR may be more pronounced at the chronic stage, as the time window associated with their study was on average 17.9 months post-injury [106]. In this dissertation, this time-related hypothesis does not seem applicable, since our TBI group was scanned at a median of nine days post-injury, approximately 40 days before Geurts et al. (49 days [50]), and SWI still exhibited larger lesion volumes. Therefore, the fact that SWI inherently magnifies haemorrhagic lesions seems more likely to have contributed to our findings, though, this can only be assumed with caution, since neither this hypothesis, nor the haemorrhagic nature of the annotated lesions, were explicitly evaluated in this dissertation. Such pathway could be explored in future research investigations: since the current comparison is already being made on dependent lesion volumes, i.e. the same patient is scanned on both sequences, an adjusted lesion differentiation based on their haemorrhagic nature could provide more evidence to explain the larger volumes seen on SWI scans.

To address scanner and MR sequence biases, future research efforts may start by quantifying them and attempting a balancing of their impact on lesion features. Moreover, there may be some ambiguity if lesions are annotated by different experts, i.e. it could be unclear if the lesions appeared bigger due to the contrast/scanner used or due to different annotation protocols.

6.5 Outcome Analysis

Most lesion types and groups showed larger volumes in patients with unfavourable outcome (Figures 5.7 and A.2). Further analysis was conducted through the construction of four generalised linear models. The first model showed a significant association between SV and patient outcome ($p = 0.0466$), and correctly predicted outcome in 66.7% of Verio patients (unknown to the model). The second and third model included lesion volume with respect to lesion group and lesion location, respectively. Both revealed no significant association between any group or location and patient outcome, despite initially significant associations (i.e. before p-value correction) between patient outcome and ILC ($p = 0.0394$, test accuracy $= 66.7\%$), IPO (0.0464, 57.1%) and the ventricles (0.0241, 61.9%). Lastly, the fourth model showed a significant association between ILC and patient outcome, when the left parietal lobe was also included in the model ($p = 0.047$). Petechial haemorrhages showed particularly poor p-values and accuracies in the second and fourth model. However, this is expected as PH are representative of DAI when annotated on SWI. Because of this, the PH impact should be assessed with regards to their distribution across the brain and not their volumes, which, due to the diffuse nature of DAI, is much less substantial. Comparing the test and training error rates of all four different models (see Table A.3 and Figure A.7), suggested that the fourth model was the most accurate, since it showed the lowest error rates (training error rate: 27.6%; test error rate: 33.0%).

These results support the hypothesis that unfavourable outcomes are associated with bigger lesion volumes, i.e. patient outcome can be predicted with respect to lesion volume. Although it seems likely that larger lesions lead to worse outcome, these results must be taken cautiously for multiple reasons.

First, patient selection reduced the sample size in nearly half to select hyper-acute TBI (patients scanned within three days post-injury), such that only 58 out of 125 patients could be considered for the analysis. This reduction compromised outcome comparison across lesion types, e.g. EDH only occurred in patients with favourable outcome, and lesion locations, as almost half of the brain regions in the atlas were neglected due to a lack of patients with lesion volumes in those regions.

Secondly, all models were validated by training them on only Trio scanner data, and fitting them to Verio data. The Trio scanner was used for training simply because the majority of patients in this group were scanned on it (see Table 4.3; Section 4.5) allowing a bigger training cohort. This validation offers some insight on how well the model will predict outcome in unseen patients (Verio patients were excluded during the training phase all together). Outcome prediction after TBI is an inherently difficult task, and validating the models on a held out test data set may be one explanation for the low prediction accuracy. Furthermore, low accuracy might be caused by the vast heterogeneity of lesion patterns observed in TBI patients. Future analysis will likely benefit from larger data sets, and the introduction of other clinical variables such as pupillary reactivity, which has been shown to be relevant in TBI prognosis [98].

6.6 Future Research Directions

This dissertation investigated the biases introduced by scanners and MR sequences and their effect on lesion characteristics, such as volume and connectivity. Furthermore the impact of lesion volume, group and location on patient outcome after TBI was examined. Future research could focus on the following points:

- MR sequence and scanner variation impacted lesion volume and cluster count in most lesion groups. This finding can be further explored with an emphasis on how these biases impact patient outcome, including its modelling and prediction.
- The lesion co-occurrence matrix in Figure 5.2 allows for the identification of highly frequent pairs of lesions. This method is scalable, therefore it can be extended for lesion pattern analysis, by including larger sets of lesions. Eventually, this analysis could be integrated into patient outcome modelling.
- The localization method enabled direct comparisons between lesion types with regards to location. This can be used to investigate the reasons behind highly-frequent lesion pairs, for instance lesion spatial proximity or paired annotation. On top of this, a quality control procedure can be implemented to discard errors in localization. An example of this would be to discard scans that contain regions with high deviations from their expected volumes.
- The outcome models were trained with hyper-acute TBI data. However, the same methods can also be applied to other fixed time windows, thus facilitating a TBI stage comparison. More ambitiously, time can be introduced as a continuous variable, in a shift to lesion progression analysis.
- Besides volume, several other features can be attributed to lesions, groups and locations. A reproduction of the present methodology with increased sample sizes and more lesion features can consolidate or extend the reported findings.

Chapter 7

Conclusion

This dissertation aimed to advance knowledge on TBI, particularly in prognosis, or patient outcome prediction. To do so, manually annotated TBI lesions were analyzed to compute their volumes, connectivity and location. These lesions were annotated on different MR scans (FLAIR and SWI/GE), which were obtained through two different scanners (Trio and Verio). This variation introduced biases in the lesion characteristics, as lesion volumes and number of clusters were found to be bigger on SWI than on FLAIR. This comparison was made for the same lesions, therefore it is likely that it was indeed the change in MR sequence that increased the lesion volume or cluster count. Moreover, the fact that SWI inherently enhances haemorrhagic lesions seems to be a plausible explanation for this finding. Future research could further explore this by comparing this effect (larger volumes/cluster counts on SWI) on haemorrhagic and non-haemorrhagic lesions. The two scanners also showed different average volumes and cluster counts for the same lesion groups. However, in this case, the comparison is being made on different patients, therefore different lesions, since only one patient was scanned on both scanners. Because of this, it is unclear if the differences found were due to the scanner or to the set of patients associated with that scanner. Scanners biases seem to be inevitable in multi-scanner data, and establishing its true extent may be infeasible, since it would require patients to be scanned on two or more scanners in a short time window. Following bias analysis, four different models were examined to understand the impact of lesion volumes on functional outcome after TBI. These models examined lesion volume in four different ways: total lesioned volume in a patient, average lesion group volume, average lesion volume in a brain ROI, and the combination of group and ROI. All models showed relatively similar error rates, from 27.6 to 38.1%. In summary, the addition of location and/or group to the base model, which only included age and GCS, showed a slight improvement in outcome prediction, as the contusion core lesion group was found to be significantly associated with patient outcome, when combined with the parietal lobe. The methodology described above was applied to 58 patients that showed an hyper-acute (within three days since injury), severe case of TBI. The cohort included a decent number of patients (for medical imaging, in particular as they were manually annotated), nonetheless, TBI is a highly heterogeneous disease, thus any attempt to model patient outcome after TBI will benefit from larger sample sizes as statistical power increases. Future experiments will benefit from automated lesion segmentation, which could substantially increase sample sizes by bypassing the laborious and time-consuming task of manually segmenting lesions [109–111].

Appendix A

Supplementary Figures and Tables

Table A.1: Description of main subject features.

Category	Feature	Description
Subject	Age	Continuous
	Sex	Binary: male or female
	Education	Categorical: below GCSE, diploma, GCSE/AS/A levels equivalent and university degree
	GCS score	Nominal: 3 to 15
	Severity	Categorical: mild, moderate and severe
	Injury mechanism	Categorical: fall, road traffic accident, cycling accident, struck by object, violence/assault or other

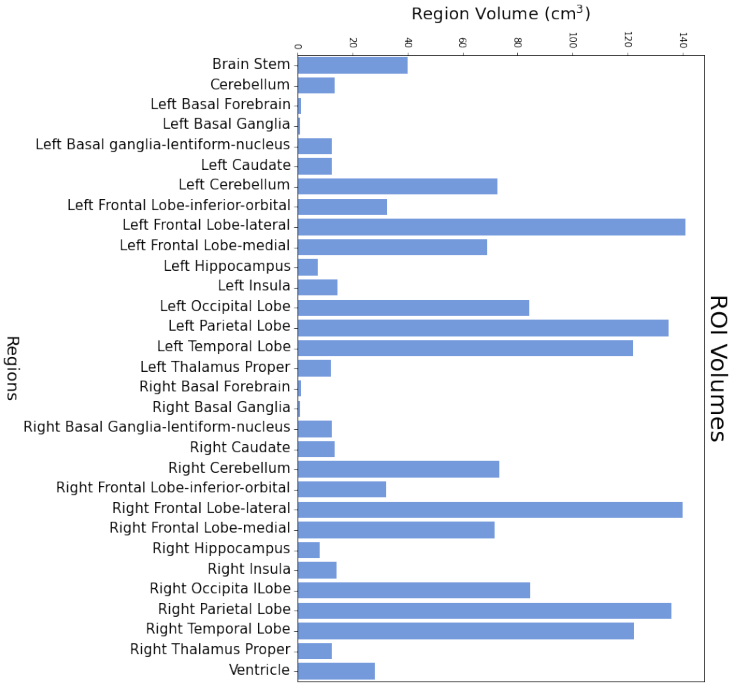

Figure A.1: *ROI volumes pattern across the 31 regions.*

A. SUPPLEMENTARY FIGURES AND TABLES

Table A.2: Outcome groups comparison across lesion groups in average and median group volume and number of subjects presenting each lesion group ($N = 58$).

Lesion group	Favourable outcome		Unfavourable outcome	
	Avg. Volume (Median)	Subjects	Avg. Volume (Median)	Subjects
ILC	9.4 (3.6)	31	21.7 (7.1)	27
IPO	11.7 (6.5)	31	21.8 (11.3)	27
PH	0.6 (0.3)	35	0.8 (0.3)	23
EHL	10.8 (8.9)	32	13.0 (9.2)	26
VCH	0.6 (0.2)	35	2.7 (0.3)	23

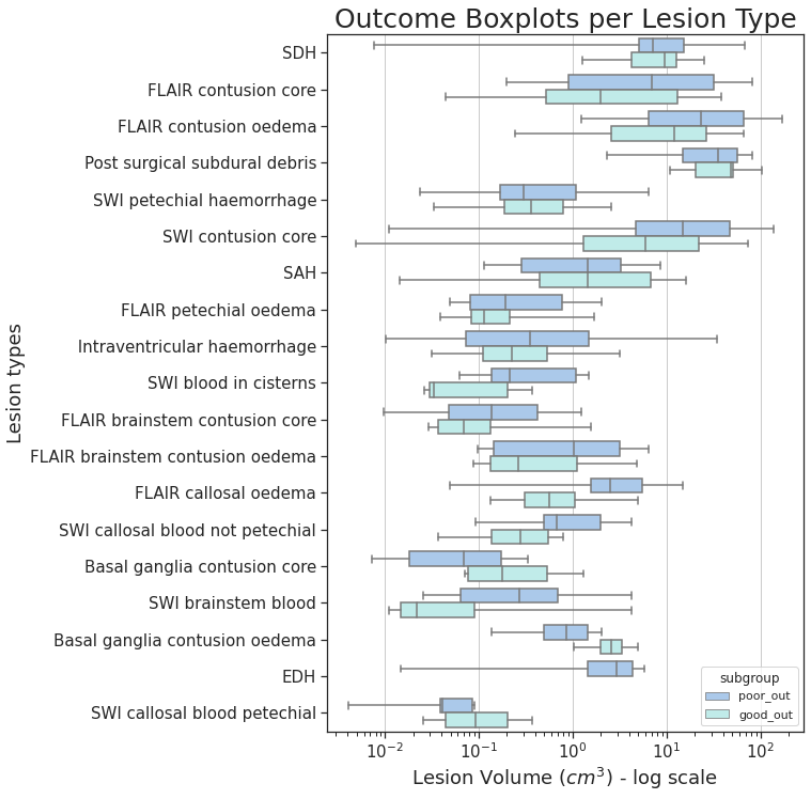

Figure A.2: *Outcome groups boxplots for each lesion type (y-axis), across lesion volume. Volumes were averaged for each lesion type within each group. Patient outcome is color-coded: unfavourable outcome (blue) and favourable outcome (cyan blue).*

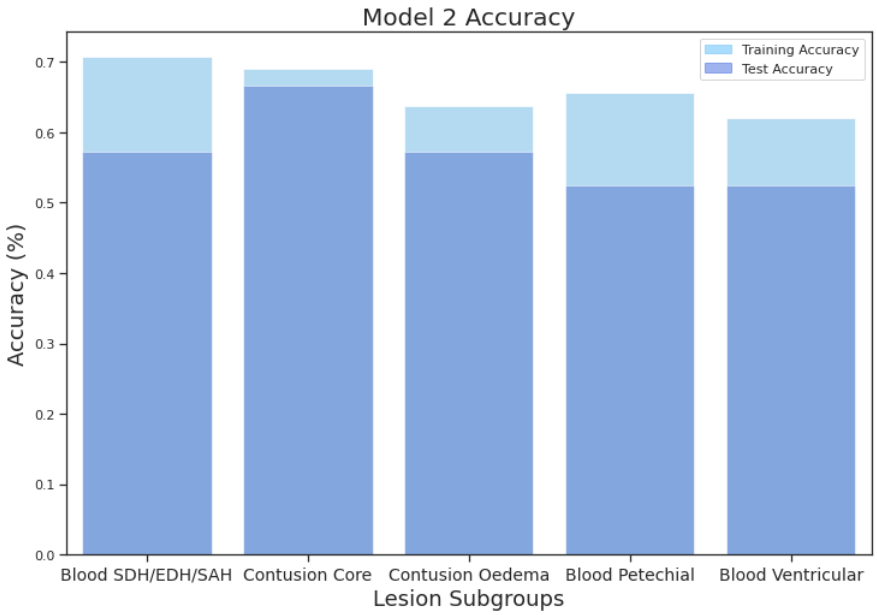

Figure A.3: *Model 2 training and testing accuracies (y-axis) for each lesion group (x-axis).The top bar (light cyan) represents training accuracy while the bottom bar (light blue) comprises the more conservative testing accuracy.*

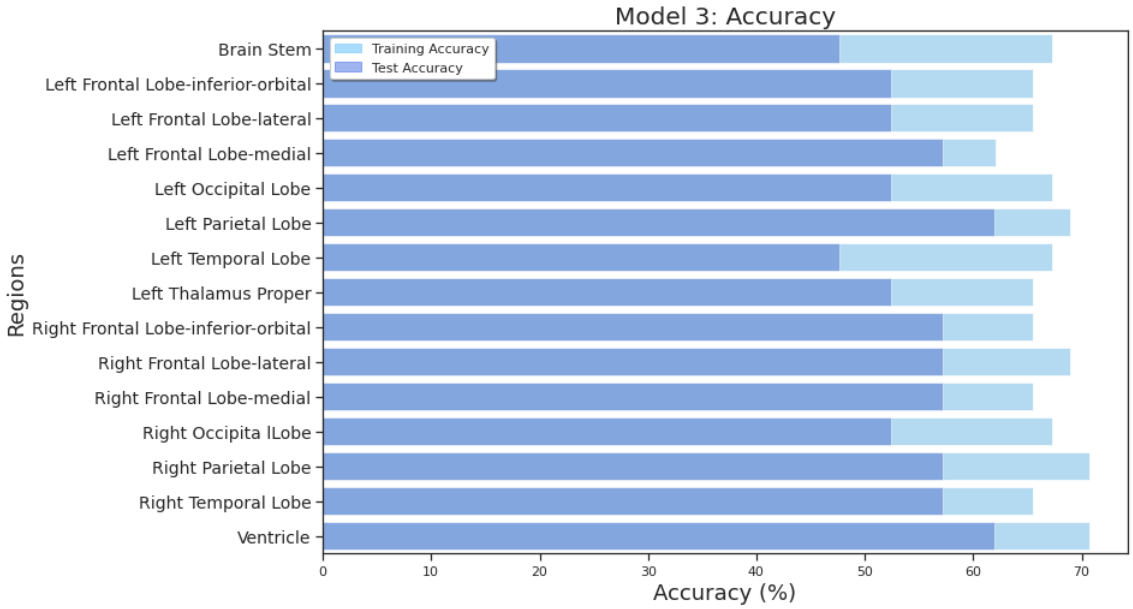

Figure A.4: *Model 3 training and testing accuracies (x-axis) for each region (y-axis).Training accuracy is represented by light cyan colored bars whereas light blue bars show the test accuracy.*

A. SUPPLEMENTARY FIGURES AND TABLES

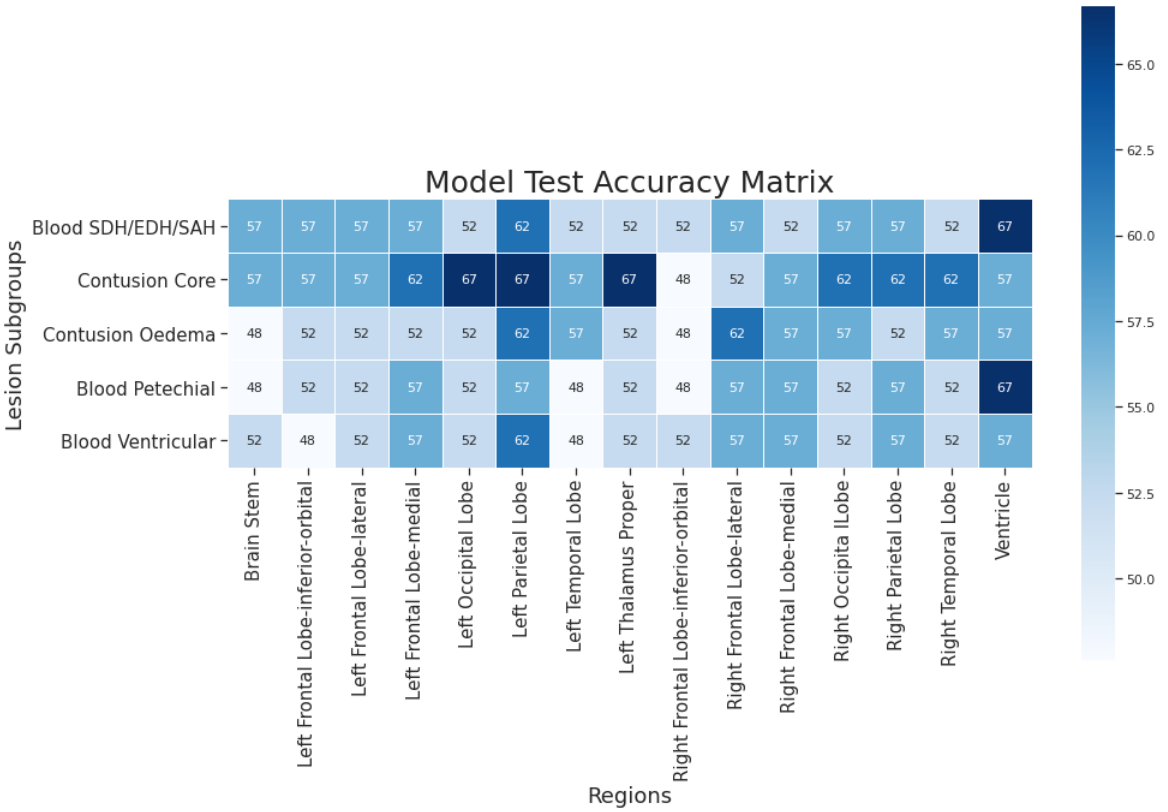

Figure A.5: *Model 4 test accuracy matrix. Each square contains the model 4 test accuracy for a that specific lesion group (y-axis) plus region (x-axis) combination.*

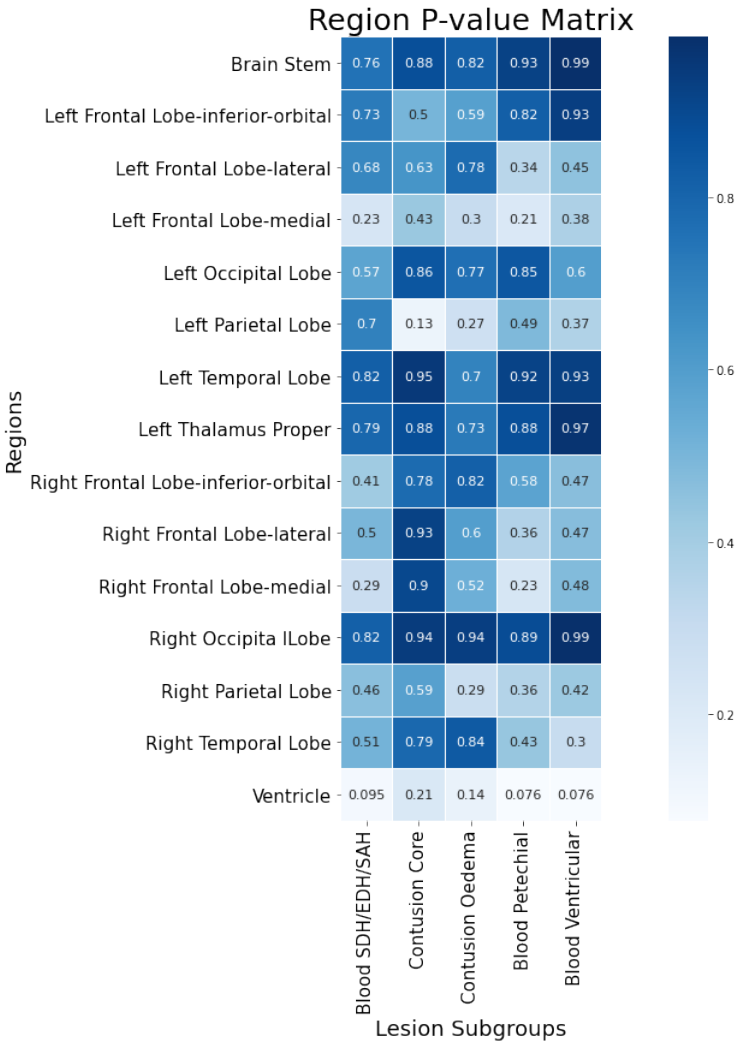

Figure A.6: *Region p-values obtained from model 4.*

Table A.3: Training and test error rates for each model. Values are obtained directly from training and test accuracy.

Models	Training Error Rate	Test Error Rate
Model 1	29.3 %	33.3 %
Model 2	31.0 %	33.3 %
Model 3	29.3 %	38.1 %
Model 4	27.6%	33.0%

A. SUPPLEMENTARY FIGURES AND TABLES

Figure A.7: *Model comparison barplot. Test error rate is represented by light cyan bars, while blue bars represent training error rates. An horizontal grid is included to aid visual comparison.*

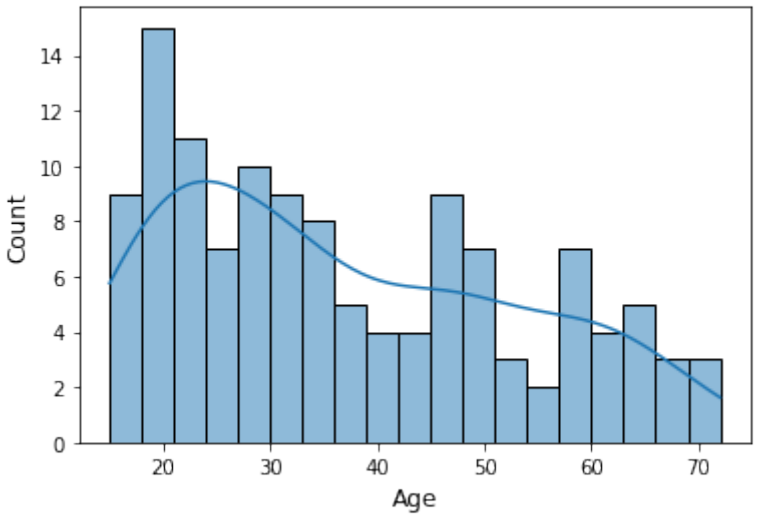

Figure A.8: *Patient age distribution featuring a kernel density estimate.*

Bibliography

[1] Andrew IR Maas, Nino Stocchetti, and Ross Bullock. "Moderate and severe traumatic brain injury in adults". In: *The Lancet Neurology* 7.8 (Aug. 2008), pp. 728–741. ISSN: 1474-4422. DOI: 10.1016/S1474-4422(08)70164-9. URL: https://doi.org/10.1016/S1474-4422(08)70164-9.

[2] Marek Majdan et al. "Epidemiology of traumatic brain injuries in Europe: a cross-sectional analysis". In: *The Lancet Public Health* 1.2 (Dec. 2016), e76–e83. ISSN: 2468-2667. DOI: 10.1016/S2468-2667(16)30017-2. URL: https://doi.org/10.1016/S2468-2667(16)30017-2.

[3] Andrew I. R. Maas et al. "Traumatic brain injury: integrated approaches to improve prevention, clinical care, and research". In: *The Lancet Neurology* 16.12 (Dec. 2017), pp. 987–1048. ISSN: 1474-4422. DOI: 10.1016/S1474-4422(17)30371-X. URL: https://doi.org/10.1016/S1474-4422(17)30371-X.

[4] Andrew Gardner, Elizabeth Adamova, and Ross Zafonte. "Prognosis, Outcome Measures, and Prevention". In: Jan. 2019, pp. 227–239. ISBN: 9780323544566. DOI: 10.1016/B978-0-323-54456-6.00016-5.

[5] Karen Chua et al. "A brief review of traumatic brain injury rehabilitation". In: *Annals of the Academy of Medicine, Singapore* 36 (Feb. 2007), pp. 31–42.

[6] Joseph T. King, Patricia M. Carlier, and Donald W. Marion. "Early Glasgow Outcome Scale Scores Predict Long-Term Functional Outcome in Patients with Severe Traumatic Brain Injury". In: *Journal of Neurotrauma* 22.9 (2005). PMID: 16156710, pp. 947–954. DOI: 10.1089/neu.2005.22.947. eprint: https://doi.org/10.1089/neu.2005.22.947. URL: https://doi.org/10.1089/neu.2005.22.947.

[7] Frank Hillary and Bharat Biswal. "Automated Detection and Quantification of Brain Lesions in Acute Traumatic Brain Injury Using MRI". In: *Brain Imaging Behav* 3 (June 2009), pp. 111–122. DOI: 10.1007/s11682-008-9053-0.

[8] T.L.A. van den Heuvel et al. "Automated detection of cerebral microbleeds in patients with traumatic brain injury". In: *NeuroImage: Clinical* 12 (2016), pp. 241–251. ISSN: 2213-1582. DOI: https://doi.org/10.1016/j.nicl.2016.07.002. URL: http://www.sciencedirect.com/science/article/pii/S221315821630122X.

[9] A. Rao et al. "Contusion segmentation from subjects with Traumatic Brain Injury: A random forest framework". In: *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. 2014, pp. 333–336. DOI: 10.1109/ISBI.2014.6867876.

[10] Erin D. Bigler et al. “Diffuse damage in pediatric traumatic brain injury: A comparison of automated versus operator-controlled quantification methods”. In: *NeuroImage* 50.3 (2010), pp. 1017–1026. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2010.01.003. URL: http://www.sciencedirect.com/science/article/pii/S105381191000008X.

[11] J. M. Wardlaw, V. J. Easton, and P. Statham. “Which CT features help predict outcome after head injury?” eng. In: *Journal of neurology, neurosurgery, and psychiatry* 72.2 (Feb. 2002). PMC1737755[pmcid], pp. 188–151. ISSN: 0022-3050. DOI: 10.1136/jnnp.72.2.188. URL: https://doi.org/10.1136/jnnp.72.2.188.

[12] Wynne E. Morrison et al. “Gender and age effects on outcome after pediatric traumatic brain injury*”. In: *Pediatric Critical Care Medicine* 5.2 (2004). ISSN: 1529-7535. URL: https://journals.lww.com/pccmjournal/Fulltext/2004/02000/Gender_and_age_effects_on_outcome_after_pediatric.8.aspx.

[13] J.T. LINDSAY WILSON, LAURA E.L. PETTIGREW, and GRAHAM M. TEASDALE. “Structured Interviews for the Glasgow Outcome Scale and the Extended Glasgow Outcome Scale: Guidelines for Their Use”. In: *Journal of Neurotrauma* 15.8 (1998). PMID: 9726257, pp. 573–585. DOI: 10.1089/neu.1998.15.573. eprint: https://doi.org/10.1089/neu.1998.15.573. URL: https://doi.org/10.1089/neu.1998.15.573.

[14] Sandra Gauthier et al. “Acute prediction of outcome and cognitive-communication impairments following traumatic brain injury: The influence of age, education and site of lesion”. In: *Journal of Communication Disorders* 73 (2018), pp. 77–90. ISSN: 0021-9924. DOI: https://doi.org/10.1016/j.jcomdis.2018.04.003. URL: https://www.sciencedirect.com/science/article/pii/S0021992417301417.

[15] Nasser M. Aldossary, Mamdouh Ali Kotb, and Ahmed M. Kamal. “Predictive value of early MRI findings on neurocognitive and psychiatric outcomes in patients with severe traumatic brain injury”. In: *Journal of Affective Disorders* 243 (2019), pp. 1–7. ISSN: 0165-0327. DOI: https://doi.org/10.1016/j.jad.2018.09.001. URL: https://www.sciencedirect.com/science/article/pii/S016503271831173X.

[16] Cody A. Chastain et al. “Predicting Outcomes of Traumatic Brain Injury by Imaging Modality and Injury Distribution”. In: *Journal of Neurotrauma* 26.8 (2009). PMID: 19317591, pp. 1183–1196. DOI: 10.1089/neu.2008.0650. eprint: https://doi.org/10.1089/neu.2008.0650. URL: https://doi.org/10.1089/neu.2008.0650.

[17] Marleen Maria van Eijck et al. “Diffuse axonal injury after traumatic brain injury is a prognostic factor for functional outcome: a systematic review and meta-analysis”. In: *Brain Injury* 32.4 (2018). PMID: 29381396, pp. 395–402. DOI: 10.1080/02699052.2018.1429018. eprint: https://doi.org/10.1080/02699052.2018.1429018. URL: https://doi.org/10.1080/02699052.2018.1429018.

[18] Hourmazd Haghbayan et al. “The Prognostic Value of MRI in Moderate and Severe Traumatic Brain Injury: A Systematic Review and Meta-Analysis”. In: *Critical Care Medicine* 45.12 (2017). ISSN: 0090-3493. URL: https://journals.lww.com/ccmjournal/Fulltext/2017/12000/The_Prognostic_Value_of_MRI_in_Moderate_and_Severe.37.aspx.

[19] John R. Williams et al. “Prognostic Value of Hemorrhagic Brainstem Injury on Early Computed Tomography: A TRACK-TBI Study”. In: *Neurocritical Care* (July 2021). ISSN: 1556-0961. DOI: 10.1007/s12028-021-01263-8. URL: https://doi.org/10.1007/s12028-021-01263-8.

[20] Mark Faul and Victor Coronado. “Chapter 1 - Epidemiology of traumatic brain injury”. In: *Traumatic Brain Injury, Part I*. Ed. by Jordan Grafman and Andres M. Salazar. Vol. 127. Handbook of Clinical Neurology. Elsevier, 2015, pp. 3–13. DOI: https://doi.org/10.1016/B978-0-444-52892-6.00001-5. URL: https://www.sciencedirect.com/science/article/pii/B9780444528926000015.

[21] A.J. Gardner and R. Zafonte. “Neuroepidemiology of traumatic brain injury”. In: *Neuroepidemiology*. Elsevier, 2016, pp. 207–223. DOI: 10.1016/b978-0-12-802973-2.00012-4. URL: https://doi.org/10.1016/b978-0-12-802973-2.00012-4.

[22] David F. Meaney and Douglas H. Smith. “Biomechanics of Concussion”. In: *Clinics in Sports Medicine* 30.1 (2011). Concussion in Sports, pp. 19–31. ISSN: 0278-5919. DOI: https://doi.org/10.1016/j.csm.2010.08.009. URL: https://www.sciencedirect.com/science/article/pii/S027859191000061X.

[23] J. Kraus and P. Nourjah. “The epidemiology of mild, uncomplicated brain injury.” In: *The Journal of trauma* 28 12 (1988), pp. 1637–43.

[24] David Cifu et al. “VA/DoD Clinical Practice Guideline for Management of Concussion/Mild Traumatic Brain Injury”. In: *Journal of Rehabilitation Research and Development - J REHABIL RES DEV* 46 (Jan. 2009). DOI: 10.1682/JRRD.2009.06.0076.

[25] Mark Sherer and Angelle Sander. *Handbook on the Neuropsychology of Traumatic Brain Injury*. Jan. 2014, pp. 1–420. ISBN: 978-1-4939-0783-0. DOI: 10.1007/978-1-4939-0784-7.

[26] Michael P. Alexander. “Mild traumatic brain injury”. In: *Neurology* 45.7 (1995), pp. 1253–1260. ISSN: 0028-3878. DOI: 10.1212/WNL.45.7.1253. eprint: https://n.neurology.org/content/45/7/1253.full.pdf. URL: https://n.neurology.org/content/45/7/1253.

[27] Victor G. Coronado et al. “Trends in Traumatic Brain Injury in the U.S. and the public health response: 1995–2009”. In: *Journal of Safety Research* 43.4 (2012). CDC Injury Center at 20 Years: Celebrating the past, protecting the future, pp. 299–307. ISSN: 0022-4375. DOI: https://doi.org/10.1016/j.jsr.2012.08.011. URL: http://www.sciencedirect.com/science/article/pii/S0022437512000564.

[28] Mark Faul et al. *Traumatic brain injury in the United States; emergency department visits, hospitalizations, and deaths, 2002-2006*. Report. Dept. of Health et al., 2010. URL: https://stacks.cdc.gov/view/cdc/5571.

[29] Linda Carroll et al. “Prognosis for mild traumatic brain injury: Results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury”. In: *Journal of rehabilitation medicine : official journal of the UEMS European Board of Physical and Rehabilitation Medicine* 43 (Mar. 2004), pp. 84–105. DOI: 10.1080/16501960410023859.

[30] Max Wintermark et al. “Imaging Evidence and Recommendations for Traumatic Brain Injury: Conventional Neuroimaging Techniques”. In: *Journal of the American College of Radiology* 12.2 (2015), e1–e14. ISSN: 1546-1440. DOI: https://doi.org/10.1016/j.jacr.2014.10.014. URL: http://www.sciencedirect.com/science/article/pii/S1546144014006760.

[31] Ann-Christine Duhaime et al. “Common Data Elements in Radiologic Imaging of Traumatic Brain Injury”. In: *Archives of Physical Medicine and Rehabilitation* 91.11 (2010), pp. 1661–1666. ISSN: 0003-9993. DOI: https://doi.org/10.1016/j.apmr.2010.07.238. URL: http://www.sciencedirect.com/science/article/pii/S0003999310007331.

[32] Karl Janich, Ha Nguyen, and Ninh Doan. “Management of Adult Traumatic Brain Injury: A Review”. In: *Journal of Trauma & Treatment* 5 (Jan. 2016). DOI: 10.4172/2167-1222.1000320.

[33] J. Jacob Kazam and Apostolos John Tsiouris. “Brain Magnetic Resonance Imaging for Traumatic Brain Injury: Why, When, and How?” In: *Topics in Magnetic Resonance Imaging* 24.5 (2015). ISSN: 1536-1004. URL: https://journals.lww.com/topicsinmri/Fulltext/2015/10000/Brain_Magnetic_Resonance_Imaging_for_Traumatic.2.aspx.

[34] Barbara Bendlin et al. “Longitudinal changes in patients with traumatic brain injury assessed with diffusion-tensor and volumetric imaging”. In: *NeuroImage* 42 (June 2008), pp. 503–14. DOI: 10.1016/j.neuroimage.2008.04.254.

[35] Erin Bigler et al. “Temporal lobe morphology in normal aging and traumatic brain injury”. In: *AJNR. American journal of neuroradiology* 23 (Mar. 2002), pp. 255–66.

[36] B.A. Cohen et al. “Proton MR Spectroscopy and MRI-Volumetry in Mild Traumatic Brain Injury”. In: *American Journal of Neuroradiology* 28.5 (2007), pp. 907–913. ISSN: 0195-6108. eprint: http://www.ajnr.org/content/28/5/907.full.pdf. URL: http://www.ajnr.org/content/28/5/907.

[37] Brian Levine et al. “The Toronto traumatic brain injury study - Injury severity and quantified MRI”. In: *Neurology* 70 (Apr. 2008), pp. 771–8. DOI: 10.1212/01.wnl.0000304108.32283.aa.

[38] Daniel T. Ginat and Steven P. Meyers. “Intracranial Lesions with High Signal Intensity on T1-weighted MR Images: Differential Diagnosis”. In: *RadioGraphics* 32.2 (2012). PMID: 22411945, pp. 499–516. DOI: 10.1148/rg.322105761. eprint: https://doi.org/10.1148/rg.322105761. URL: https://doi.org/10.1148/rg.322105761.

[39] W G Bradley. “MR appearance of hemorrhage in the brain.” In: *Radiology* 189.1 (1993). PMID: 8372185, pp. 15–26. DOI: 10.1148/radiology.189.1.8372185. eprint: https://doi.org/10.1148/radiology.189.1.8372185. URL: https://doi.org/10.1148/radiology.189.1.8372185.

[40] Chelsea S. Kidwell et al. “Comparison of MRI and CT for Detection of Acute Intracerebral Hemorrhage”. In: *JAMA* 292.15 (Oct. 2004), pp. 1823–1830. ISSN: 0098-7484. DOI: 10.1001/jama.292.15.1823. eprint: https://jamanetwork.com/journals/jama/articlepdf/199622/joc40605.pdf. URL: https://doi.org/10.1001/jama.292.15.1823.

[41] R. Ashikaga, Y. Araki, and O. Ishida. “MRI of head injury using FLAIR”. In: *Neuroradiology* 39.4 (Apr. 1997), pp. 239–242. ISSN: 1432-1920. DOI: 10.1007/s002340050401. URL: https://doi.org/10.1007/s002340050401.

[42] M. Wintermark et al. “Imaging Evidence and Recommendations for Traumatic Brain Injury: Advanced Neuro- and Neurovascular Imaging Techniques”. In: *American Journal of Neuroradiology* 36.2 (2015), E1–E11. ISSN: 0195-6108. DOI: 10.3174/ajnr.A4181. eprint: http://www.ajnr.org/content/36/2/E1.full.pdf. URL: http://www.ajnr.org/content/36/2/E1.

[43] Yukinori AKIYAMA et al. “Susceptibility-Weighted Magnetic Resonance Imaging for the Detection of Cerebral Microhemorrhage in Patients With Traumatic Brain Injury”. In: *Neurologia medico-chirurgica* 49.3 (2009), pp. 97–99. DOI: 10.2176/nmc.49.97.

[44] Bejoy Thomas et al. “Clinical applications of susceptibility weighted MR imaging of the brain - A pictorial review”. In: *Neuroradiology* 50 (Mar. 2008), pp. 105–16. DOI: 10.1007/s00234-007-0316-z.

[45] Sahil Sood, Ajay Gupta, and Apostolos Tsiouris. “Advanced Magnetic Resonance Techniques in Neuroimaging: Diffusion, Spectroscopy, and Perfusion”. In: *Seminars in roentgenology* 45 (Apr. 2010), pp. 137–46. DOI: 10.1053/j.ro.2009.09.008.

[46] Nicholas Galloway et al. “Diffusion-Weighted Imaging Improves Outcome Prediction in Pediatric Traumatic Brain Injury”. In: *Journal of neurotrauma* 25 (Oct. 2008), pp. 1153–62. DOI: 10.1089/neu.2007.0494.

[47] Amon Y. Liu et al. “Traumatic Brain Injury: Diffusion-Weighted MR Imaging Findings”. In: *American Journal of Neuroradiology* 20.9 (1999), pp. 1636–1641. ISSN: 0195-6108. eprint: http://www.ajnr.org/content/20/9/1636.full.pdf. URL: http://www.ajnr.org/content/20/9/1636.

[48] Laszlo L. Mechtler, Kalyan K. Shastri, and Kevin E. Crutchfield. “Advanced Neuroimaging of Mild Traumatic Brain Injury”. In: *Neurologic Clinics* 32.1 (2014). Neuroimaging, pp. 31–58. ISSN: 0733-8619. DOI: https://doi.org/10.1016/j.ncl.2013.08.002. URL: http://www.sciencedirect.com/science/article/pii/S0733861913000960.

[49] Cyrus Eierud et al. “Neuroimaging after mild traumatic brain injury: Review and meta-analysis”. In: *NeuroImage: Clinical* 4 (Jan. 2014). DOI: 10.1016/j.nicl.2013.12.009.

[50] Bram Geurts et al. “The reliability of magnetic resonance imaging in traumatic brain injury lesion detection”. In: *Brain injury : [BI]* 26 (June 2012), pp. 1439–50. DOI: 10.3109/02699052.2012.694563.

[51] M. Herman. “Statistical Methods for the Analysis of Biomedical Data”. In: *Technometrics* 45 (2003), pp. 111–111.

[52] Younis Skaik. “The bread and butter of statistical analysis "t-test": Uses and misuses”. eng. In: *Pakistan journal of medical sciences* 31.6 (2015). PJMS-31-1558[PII], pp. 1558–1559. ISSN: 1682-024X. URL: https://pubmed.ncbi.nlm.nih.gov/26870136.

[53] Patrick E. McKnight and Julius Najab. “Mann-Whitney U Test”. In: *The Corsini Encyclopedia of Psychology*. American Cancer Society, 2010, pp. 1–1. ISBN: 9780470479216. DOI: https://doi.org/10.1002/9780470479216.corpsy0524. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470479216.corpsy0524. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470479216.corpsy0524.

[54] A. Hart. “Mann-Whitney test is not just a test of medians: differences in spread can be important”. eng. In: *BMJ (Clinical research ed.)* 323.7309 (Aug. 2001). PMC1120984[pmcid], pp. 391–393. ISSN: 0959-8138. DOI: 10.1136/bmj.323.7309.391. URL: https://doi.org/10.1136/bmj.323.7309.391.

[55] David J. Sheskin. *Handbook of Parametric and Nonparametric Statistical Procedures*. 4th ed. Chapman & Hall/CRC, 2007. ISBN: 1584888148.

[56] R. F. Woolson. “Wilcoxon Signed-Rank Test”. In: *Wiley Encyclopedia of Clinical Trials*. American Cancer Society, 2008, pp. 1–3. ISBN: 9780471462422. DOI: https://doi.org/10.1002/9780471462422.eoct979. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780471462422.eoct979. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780471462422.eoct979.

[57] CLINT W. COAKLEY. “Practical Nonparametric Statistics (3rd ed.)” English. In: *Journal of the American Statistical Association* 95 (Mar. 2000). Book review, p. 332. ISSN: 01621459. URL: https://link.gale.com/apps/doc/A60903613/AONE?u=anon~6d66f9d8&sid=googleScholar&xid=3163fdaa.

[58] E. C. Alexopoulos. “Introduction to multivariate regression analysis”. eng. In: *Hippokratia* 14.Suppl 1 (Dec. 2010). PMC3049417[pmcid], pp. 23–28. ISSN: 1790-8019. URL: https://pubmed.ncbi.nlm.nih.gov/21487487.

[59] Abdulhamit Subasi. “Chapter 3 - Machine learning techniques”. In: *Practical Machine Learning for Data Analysis Using Python*. Ed. by Abdulhamit Subasi. Academic Press, 2020, pp. 91–202. ISBN: 978-0-12-821379-7. DOI: https://doi.org/10.1016/B978-0-12-821379-7.00003-5. URL: https://www.sciencedirect.com/science/article/pii/B9780128213797000035.

[60] Sandro Sperandei. “Understanding logistic regression analysis”. eng. In: *Biochemia medica* 24.1 (Feb. 2014). biochem-24-1-12-4[PII], pp. 12–18. ISSN: 1330-0962. URL: https://pubmed.ncbi.nlm.nih.gov/24627710.

[61] Frank E. Harrell. “Binary Logistic Regression”. In: *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Cham: Springer International Publishing, 2015, pp. 219–274. ISBN: 978-3-319-19425-7. DOI: 10.1007/978-3-319-19425-7_10. URL: https://doi.org/10.1007/978-3-319-19425-7_10.

[62] Tom Fawcett. “An introduction to ROC analysis”. In: *Pattern Recognition Letters* 27.8 (2006). ROC Analysis in Pattern Recognition, pp. 861–874. ISSN: 0167-8655. DOI: https://doi.org/10.1016/j.patrec.2005.10.010. URL: https://www.sciencedirect.com/science/article/pii/S016786550500303X.

[63] Chakkrit Tantithamthavorn et al. “An Empirical Comparison of Model Validation Techniques for Defect Prediction Models”. In: *IEEE Transactions on Software Engineering* 43.1 (2017), pp. 1–18. DOI: 10.1109/TSE.2016.2584050.

[64] William S. Noble. “How does multiple testing correction work?” In: *Nature Biotechnology* 27.12 (Dec. 2009), pp. 1135–1137. ISSN: 1546-1696. DOI: 10.1038/nbt1209-1135. URL: https://doi.org/10.1038/nbt1209-1135.

[65] Sangseok Lee and Dong Kyu Lee. “What is the proper way to apply the multiple comparison test?” eng. In: *Korean journal of anesthesiology* 71.5 (Oct. 2018). kja.d.18.00242[PII], pp. 353–360. ISSN: 2005-6419. DOI: 10.4097/kja.d.18.00242. URL: https://doi.org/10.4097/kja.d.18.00242.

[66] Jarred Garfinkle et al. “Location and Size of Preterm Cerebellar Hemorrhage and Childhood Development”. In: *Annals of Neurology* 88.6 (2020), pp. 1095–1108. DOI: https://doi.org/10.1002/ana.25899. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ana.25899. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ana.25899.

[67] Hidetoshi Matsukawa et al. “Intraventricular hemorrhage on computed tomography and corpus callosum injury on magnetic resonance imaging in patients with isolated blunt traumatic brain injury: Clinical article”. In: *Journal of Neurosurgery JNS* 117.2 (2012), pp. 334–339. DOI: 10.3171/2012.5.JNS112318. URL: https://thejns.org/view/journals/j-neurosurg/117/2/article-p334.xml.

[68] Kent G. Moen et al. “Traumatic Axonal Injury: The Prognostic Value of Lesion Load in Corpus Callosum, Brain Stem, and Thalamus in Different Magnetic Resonance Imaging Sequences”. In: *Journal of Neurotrauma* 31.17 (2014). PMID: 24773587, pp. 1486–1496. DOI: 10.1089/neu.2013.3258. eprint: https://doi.org/10.1089/neu.2013.3258. URL: https://doi.org/10.1089/neu.2013.3258.

[69] Marta Cicuendez et al. “Prognostic value of corpus callosum injuries in severe head trauma”. In: *Acta Neurochirurgica* 159.1 (Jan. 2017), pp. 25–32. ISSN: 0942-0940. DOI: 10.1007/s00701-016-3000-4. URL: https://doi.org/10.1007/s00701-016-3000-4.

[70] Matthew F. Glasser et al. “A multi-modal parcellation of human cerebral cortex”. In: *Nature* 536.7615 (Aug. 2016), pp. 171–178. ISSN: 1476-4687. DOI: 10.1038/nature18933. URL: https://doi.org/10.1038/nature18933.

[71] Simon B. Eickhoff, B. T. Thomas Yeo, and Sarah Genon. “Imaging-based parcellations of the human brain”. In: *Nature Reviews Neuroscience* 19.11 (Nov. 2018), pp. 672–686. ISSN: 1471-0048. DOI: 10.1038/s41583-018-0071-7. URL: https://doi.org/10.1038/s41583-018-0071-7.

[72] Emily S. Finn et al. “Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity”. In: *Nature Neuroscience* 18.11 (Nov. 2015), pp. 1664–1671. ISSN: 1546-1726. DOI: 10.1038/nn.4135. URL: https://doi.org/10.1038/nn.4135.

[73] Christos Davatzikos. “Computational neuroanatomy using brain deformations: From brain parcellation to multivariate pattern analysis and machine learning”. In: *Medical Image Analysis* 33 (2016). 20th anniversary of the Medical Image Analysis journal (MedIA), pp. 149–154. ISSN: 1361-8415. DOI: https://doi.org/10.1016/j.media.2016.06.026. URL: https://www.sciencedirect.com/science/article/pii/S1361841516301037.

[74] Karla L. Miller et al. “Multimodal population brain imaging in the UK Biobank prospective epidemiological study”. In: *Nature Neuroscience* 19.11 (Nov. 2016), pp. 1523–1536. ISSN: 1546-1726. DOI: 10.1038/nn.4393. URL: https://doi.org/10.1038/nn.4393.

[75] Deepthi P. Varikuti et al. “Evaluation of non-negative matrix factorization of grey matter in age prediction”. In: *NeuroImage* 173 (2018), pp. 394–410. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2018.03.007. URL: https://www.sciencedirect.com/science/article/pii/S1053811918301927.

[76] Mark Jenkinson et al. “FSL”. In: *NeuroImage* 62.2 (2012). 20 YEARS OF fMRI, pp. 782–790. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2011.09.015. URL: https://www.sciencedirect.com/science/article/pii/S1053811911010603.

[77] Karl J Friston and John T Ashburner. *Statistical parametric mapping the analysis of functional brain images*. 1st ed. Elsevier, 2007.

[78] Bruce Fischl. “FreeSurfer”. In: *NeuroImage* 62.2 (2012). 20 YEARS OF fMRI, pp. 774–781. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2012.01.021. URL: https://www.sciencedirect.com/science/article/pii/S1053811912000389.

[79] Christian Ledig et al. “Robust whole-brain segmentation: Application to traumatic brain injury”. In: *Medical Image Analysis* 21.1 (2015), pp. 40–58. ISSN: 1361-8415. DOI: https://doi.org/10.1016/j.media.2014.12.003. URL: https://www.sciencedirect.com/science/article/pii/S136184151400187X.

[80] David Alexander Dickie et al. “Whole Brain Magnetic Resonance Image Atlases: A Systematic Review of Existing Atlases and Caveats for Use in Population Imaging”. In: *Frontiers in Neuroinformatics* 11 (2017), p. 1. ISSN: 1662-5196. DOI: 10.3389/fninf.2017.00001. URL: https://www.frontiersin.org/article/10.3389/fninf.2017.00001.

[81] Jean A. Frazier et al. “Structural Brain Magnetic Resonance Imaging of Limbic and Thalamic Volumes in Pediatric Bipolar Disorder”. In: *American Journal of Psychiatry* 162.7 (2005), pp. 1256–1265. DOI: 10.1176/appi.ajp.162.7.1256.

[82] Ahmed M. Radwan et al. “Virtual brain grafting: Enabling whole brain parcellation in the presence of large lesions”. In: *NeuroImage* 229 (2021), p. 117731. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2021.117731. URL: https://www.sciencedirect.com/science/article/pii/S1053811921000082.

[83] B.B. Avants et al. “Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain”. In: *Medical Image Analysis* 12.1 (2008). Special Issue on The Third International Workshop on Biomedical Image Registration – WBIR 2006, pp. 26–41. ISSN: 1361-8415. DOI: https://doi.org/10.1016/j.media.2007.06.004. URL: https://www.sciencedirect.com/science/article/pii/S1361841507000606.

[84] Juan Eugenio Iglesias and Mert R. Sabuncu. “Multi-atlas segmentation of biomedical images: A survey”. eng. In: *Medical image analysis* 24.1 (Aug. 2015). S1361-8415(15)00099-7[PII], pp. 205–219. ISSN: 1361-8423. DOI: 10.1016/j.media.2015.06.012. URL: https://doi.org/10.1016/j.media.2015.06.012.

[85] Stephanie K. Daniels and Anne L. Foundas. “Lesion Localization in Acute Stroke”. In: *Journal of Neuroimaging* 9.2 (1999), pp. 91–98. DOI: https://doi.org/10.1111/jon19999291. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jon19999291. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/jon19999291.

[86] Kazumi Kasahara et al. “Voxel- and atlas-based analysis of diffusion tensor imaging may reveal focal axonal injuries in mild traumatic brain injury — comparison with diffuse axonal injury”. In: *Magnetic Resonance Imaging* 30.4 (2012), pp. 496–505. ISSN: 0730-725X. DOI: https://doi.org/10.1016/j.mri.2011.12.018. URL: https://www.sciencedirect.com/science/article/pii/S0730725X11004802.

[87] John Ashburner and Karl J. Friston. “Unified segmentation”. In: *NeuroImage* 26.3 (2005), pp. 839–851. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2005.02.018. URL: https://www.sciencedirect.com/science/article/pii/S1053811905001102.

[88] Christian Ledig et al. “Robust whole-brain segmentation: Application to traumatic brain injury”. In: *Medical Image Analysis* 21.1 (2015), pp. 40–58. ISSN: 1361-8415. DOI: https://doi.org/10.1016/j.media.2014.12.003. URL: https://www.sciencedirect.com/science/article/pii/S136184151400187X.

[89] Neil Graham et al. “Diffuse axonal injury predicts neurodegeneration after moderate–severe traumatic brain injury”. In: *Brain* 143 (Oct. 2020). DOI: 10.1093/brain/awaa316.

[90] Torsten Rohlfing, Daniel B. Russakoff, and Calvin R. Maurer. “Expectation Maximization Strategies for Multi-atlas Multi-label Segmentation”. In: *Information Processing in Medical Imaging*. Ed. by Chris Taylor and J. Alison Noble. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 210–221. ISBN: 978-3-540-45087-0.

[91] Torsten Rohlfing et al. “Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains”. In: *NeuroImage* 21.4 (2004), pp. 1428–1442. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2003.11.010. URL: https://www.sciencedirect.com/science/article/pii/S1053811903007316.

[92] Rolf A. Heckemann et al. “Automatic anatomical brain MRI segmentation combining label propagation and decision fusion”. In: *NeuroImage* 33.1 (2006), pp. 115–126. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2006.05.061. URL: https://www.sciencedirect.com/science/article/pii/S1053811906006458.

[93] Yiming Xiao et al. “Investigation of morphometric variability of subthalamic nucleus, red nucleus, and substantia nigra in advanced Parkinson’s disease patients using automatic segmentation and PCA-based analysis”. eng. In: *Human brain mapping* 35.9 (Sept. 2014). PMC5127820[pmcid], pp. 4330–4344. ISSN: 1097-0193. DOI: 10.1002/hbm.22478. URL: https://doi.org/10.1002/hbm.22478.

[94] Anneke Meyer. “Multi-atlas Based Segmentation of Corpus Callosum on MRIs of Multiple Sclerosis Patients”. In: Sept. 2014, pp. 729–735. ISBN: 978-3-319-11751-5. DOI: 10.1007/978-3-319-11752-2_61.

[95] S. Keihaninejad et al. “Automatic volumetry can reveal visually undetected disease features on brain MR images in temporal lobe epilepsy”. In: *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 2010, pp. 105–108. DOI: 10.1109/ISBI.2010.5490402.

[96] Alexis F. Turgeon et al. “Determination of Neurologic Prognosis and Clinical Decision Making in Adult Patients With Severe Traumatic Brain Injury: A Survey of Canadian Intensivists, Neurosurgeons, and Neurologists*”. In: *Critical Care Medicine* 41.4 (2013). ISSN: 0090-3493. URL: https://journals.lww.com/ccmjournal/Fulltext/2013/04000/Determination_of_Neurologic_Prognosis_and_Clinical.17.aspx.

[97] Pablo Perel et al. “Prognosis following head injury: a survey of doctors from developing and developed countries”. In: *Journal of Evaluation in Clinical Practice* 13.3 (2007), pp. 464–465. DOI: https://doi.org/10.1111/j.1365-2753.2006.00713.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2753.2006.00713.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2753.2006.00713.x.

[98] Hester F. Lingsma et al. "Early prognosis in traumatic brain injury: from prophecies to predictions". In: *The Lancet Neurology* 9.5 (May 2010), pp. 543–554. ISSN: 1474-4422. DOI: 10.1016/S1474-4422(10)70065-X. URL: https://doi.org/10.1016/S1474-4422(10)70065-X.

[99] Ana Mikolić et al. "Prediction of Global Functional Outcome and Post-Concussive Symptoms after Mild Traumatic Brain Injury: External Validation of Prognostic Models in the Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI) Study". In: *Journal of Neurotrauma* (Oct. 2020). DOI: 10.1089/neu.2020.7074. URL: https://doi.org/10.1089/neu.2020.7074.

[100] Jason R. Taylor et al. "The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample". In: *NeuroImage* 144 (2017). Data Sharing Part II, pp. 262–269. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2015.09.018. URL: https://www.sciencedirect.com/science/article/pii/S1053811915008150.

[101] M Schell et al. "Automated brain extraction of multi-sequence MRI using artificial neural networks". In: European Congress of Radiology-ECR 2019. 2019.

[102] Nicholas J Tustison et al. "N4ITK: improved N3 bias correction". In: *IEEE transactions on medical imaging* 29.6 (2010), pp. 1310–1320.

[103] Brian B Avants, Nick Tustison, Gang Song, et al. "Advanced normalization tools (ANTS)". In: *Insight j* 2.365 (2009), pp. 1–35.

[104] Robert K. Kosior et al. "Atlas-Based Topographical Scoring for Magnetic Resonance Imaging of Acute Stroke". In: *Stroke* 41.3 (2010), pp. 455–460. DOI: 10.1161/STROKEAHA.109.567289. eprint: https://www.ahajournals.org/doi/pdf/10.1161/STROKEAHA.109.567289. URL: https://www.ahajournals.org/doi/abs/10.1161/STROKEAHA.109.567289.

[105] Alexandre Abraham et al. "Machine learning for neuroimaging with scikit-learn". In: *Frontiers in Neuroinformatics* 8 (2014), p. 14. ISSN: 1662-5196. DOI: 10.3389/fninf.2014.00014. URL: https://www.frontiersin.org/article/10.3389/fninf.2014.00014.

[106] Gershon Spitz et al. "Detecting Lesions after Traumatic Brain Injury Using Susceptibility Weighted Imaging: A Comparison with Fluid-Attenuated Inversion Recovery and Correlation with Clinical Outcome". In: *Journal of neurotrauma* 30 (Aug. 2013). DOI: 10.1089/neu.2013.3021.

[107] Miriam H. Beauchamp et al. "Detecting Traumatic Brain Lesions in Children: CT versus MRI versus Susceptibility Weighted Imaging (SWI)". In: *Journal of Neurotrauma* 28.6 (2011). PMID: 21501069, pp. 915–927. DOI: 10.1089/neu.2010.1712. eprint: https://doi.org/10.1089/neu.2010.1712. URL: https://doi.org/10.1089/neu.2010.1712.

[108] F.Gaillard. "Susceptibility weighted imaging. Reference article." In: *Radiopaedia* (). DOI: 10.53347/rID-13858.

[109] Frank Hillary and Bharat Biswal. "Automated Detection and Quantification of Brain Lesions in Acute Traumatic Brain Injury Using MRI". In: *Brain Imaging Behav* 3 (June 2009), pp. 111–122. DOI: 10.1007/s11682-008-9053-0.

[110] Salma Alqazzaz et al. "Automated brain tumor segmentation on multi-modal MR image using SegNet". In: *Computational Visual Media* 5 (Apr. 2019). DOI: 10.1007/s41095-019-0139-y.

[111] Konstantinos Kamnitsas et al. “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation”. In: *Medical Image Analysis* 36 (2017), pp. 61–78. ISSN: 1361-8415. DOI: https://doi.org/10.1016/j.media.2016.10.004. URL: http://www.sciencedirect.com/science/article/pii/S1361841516301839.