**Clinical Medicine Research**
2022; 11(6): 150-158
http://www.sciencepublishinggroup.com/j/cmr
doi: 10.11648/j.cmr.20221106.11
ISSN: 2326-9049 (Print); ISSN: 2326-9057 (Online)

# A Neural Network Classifies Traumatic Brain Injury Outcomes: Glasgow Coma Triples Are Needed

**Hermann Prossinger[1, *], Hubert Hetz[2], Alexandra Acimovic[2], Reinhard Berger[2], Karim Mostafa[3], Alexander Grieb[3], Heinz Steltzer[2, 3]**

[1]Department of Evolutionary Biology, University of Vienna, Vienna, Austria

[2]Department of Anesthesiology and Intensive Medical Care, Trauma Center Vienna, Location Meidling, Vienna, Austria

[3]Faculty of Medicine, Sigmund Freud Private University, Vienna, Austria

## Email address:

hermann.prossinger@univie.ac.at (Hermann Prossinger)

[*]Corresponding author

## To cite this article:

Hermann Prossinger, Hubert Hetz, Alexandra Acimovic, Reinhard Berger, Karim Mostafa, Alexander Grieb, Heinz Steltzer. A Neural Network Classifies Traumatic Brain Injury Outcomes: Glasgow Coma Triples Are Needed. *Clinical Medicine Research*. Vol. 11, No. 6, 2022, pp. 150-158. doi: 10.11648/j.cmr.20221106.11

**Received**: September 14, 2022; **Accepted**: October 12, 2022; **Published**: November 11, 2022

**Abstract:** The Glasgow Coma Score (GCS) is statistically dubious because its calculation assumes that (a) the diagnostic scores used to assess degree of consciousness are numerical and (b) there is an implied metric. The assessed diagnostic scores are, however, categorical and there exists no metric; hence, summing is neither permitted nor medically informative. *Novel methods*: In this paper, we statistically analyze the Glasgow Coma Triples (GCTs) of 162 patients (114 males; 48 females; aged 3–93 years) by using unsupervised machine-learning techniques: first, one-hot encoding; second, a dimension reduction autoencoder; and finally KDE (Kernel Density Estimation). *Results*: We find that this sequence can classify how the resulting segmentation (triage) results in (a) the dead patients clustering separately from the survivors, and (b) the survivors clustering into five groups with different hospital discharge outcomes: from those with GCT={1,1,1} to those with GCT={4,6,5}, albeit with varying trajectories. *Conclusions*: The use of machine learning techniques can uncover the medical progressions of TBI patients that are impossible to discover using conventional GCS analysis. We also find a triage for outcomes, including five clusters for surviving patients. Further research is needed to verify what medically determines these varying trajectories and their ranges in probabilities; using GCS cannot contribute to these extended investigations, however.

**Keywords:** Traumatic Brain Injury, Glasgow Coma Score, Kernel Density Estimation, Dimension Reduction, Feature Extraction, Triage, Unsupervised Machine Learning, Glasgow Coma Triples

## 1. Introduction

The Glasgow Coma Scale is an index derived by summing three response assessments: eye, verbal, and motor responses. [1] The summing [2] is only possible if one assumes that the ordinal number on the response scale can be converted to a cardinal number (a computable number) — an assumption we challenge in this paper (and we are not the first to do so [3]). Each response is assessed by a clinician or other medical staff [4] involved in emergency services (for example, at the scene of an accident prior to transportation to a hospital) using the modern version of the assessment scaling (i.e., the revised Glasgow Coma Scale): 1–4 (eye), 1–6 (verbal), and 1–5 (motor). Conventionally, the three assigned scores are summed, based on the aforementioned assumptions, to provide an index that, it is widely claimed, reveals a victim's (patient's) level of consciousness.

The criteria for assigning the response scores are ubiquitous in the published literature, so we do not list them here. It suffices to say that the lowest score ($GCS = 3$) is the lowest level of consciousness (actually: "consciousness non-existent") and $GCS = 15$ describes the lack of detectable brain injury *influencing* consciousness (so assessed), based on this scoring method.

Many clinicians consider this scale to be both useful and helpful [3, 4], despite a plethora of critiques. [5–7] (Green

[5] describes it as a "sacred cow".) For clinicians, an attractive feature is the ability to assess (in their perception) the severity of damage in the case of TBI (Traumatic Brain Injury): $GCS < 9$ is considered severe. Evidently one (medicine-based) critique is whether three responses suffice to diagnose brain injury severity and/or level of consciousness. Another critique is statistical in nature: whether the assessment is particularly sensitive to lack of inter-rater reliability [6, 7]. Other scaling methods have been proposed [8]. Nonetheless, after reviewing the literature, we perceive a widespread consensus among clinicians that GCS is considered useful, despite its shortcomings. We agree with those that dispute its usefulness (see [5] for a listing). In this paper, we elucidate a further facet of why its usefulness is not clinically helpful, even if its usefulness may exist at the extremes ($GCS = 3$ and $GCS > 13$; [9]).

Our critique of using this sum of the three response assessments (henceforth called an index, although it isn't one) deals with another difficulty, which we address in this paper. Each score for each response is not a number, but rather a (ranked) categorical variable. The three response scores are the components of a vector, which we henceforth label GCT (Glasgow Coma Triple). Adding categorical variables is not permitted in *any* statistical analyses [3]; we therefore argue that the sum obfuscates rather than reveals medical information about levels of consciousness between "deeply unconscious" and "fully conscious." This shortcoming can be easily recognized if we convert the (categorical) scores from (ordinal) integers to letters from some alphabet. Thus, for example, a diagnosis of $\{1,3,2\}$ would be replaced by a diagnosis $\{A, C, B\}$. In so doing it is impossible to calculate $GCS = 1 + 3 + 2 = 6 \neq A + B + C$, because $A + C + B$ is undefined. We refute the suggestion to replace the letters with *cardinal* integers to calculate the score, because the replacement implies a metric: the "distance" from 2 to 4 on a (cardinal) integer scale is the same as from 1 to 3, while it is not on an ordinal scale (nor is it, arguably, a medically sound procedure). We suspect that no review of GCS's utility has tested this implied metric as supported by statistical analyses (see, in addition, [5]). There have been attempts to overcome this scale choice, which appears to be the result of a habit, as suggested by the authors of the original article [1], but not justified by a rigorous statistical assessment. An attempt to overcome this scale choice is the Extended Early Barthel Index [9, 10]. Some responses are to be assigned "–50", some "+5", and some "+10"; the sum of diagnosed responses is then the Barthel index. Where do these numbers come from? Or, more precisely: what statistical analysis was used to construct such a metric, differing as it does from the Glasgow Coma Scale metric?

Another criticism is the ambiguity due to the summation used to compute the score. There are 14 different triples that have the same $GCS = 7$ (Table 1 and Figure 1). We question whether these triples measure the same consciousness level, let alone brain injury severity — indeed, we will show that they do not.

The claim that GCS assesses consciousness level implies a definition of consciousness which is at variance with the medical condition of a TBI patient. In summary: we investigate the implications by statistically analyzing the *triples* of categorical variables (the assessments/scores), rather than following the conventional procedure of converting the scores to cardinal numbers, assigning a metric to the categorical scores while doing so, and subsequently summing them to a single number/index for each patient at the time he/she is diagnosed. One could ask whether the resulting GCS is itself a categorical or metric variable. However, because we reject (and justify the rejection of) the procedure for calculating GCS, we need not address the issue of whether GCS is (or isn't) a categorical variable.

Categorical variables can be encoded without a metric, but in a higher-dimensional space. One-hot encoding constructs the registrations in this space. Each patient is represented by a set of points in this high-dimensional space (it will be a space of 51 dimensions in this paper), but the nonzero scorings (components) in this space are not independent, so we apply methods of dimension reduction to derive a metric that can then be analyzed by modern unsupervised ML (machine learning) techniques — specifically, by finding a non-parametric probability density function over the space of reduced dimensionality. We discover that the brain injury severity is not uniformly distributed in this space and each cluster (group) of patients has different, medically inferable characterizations. Remarkably, as we show, conventional GCS within each group is not uniformly distributed — one (further) argument we use to request abstaining from the use of GCS as a medical indicator.

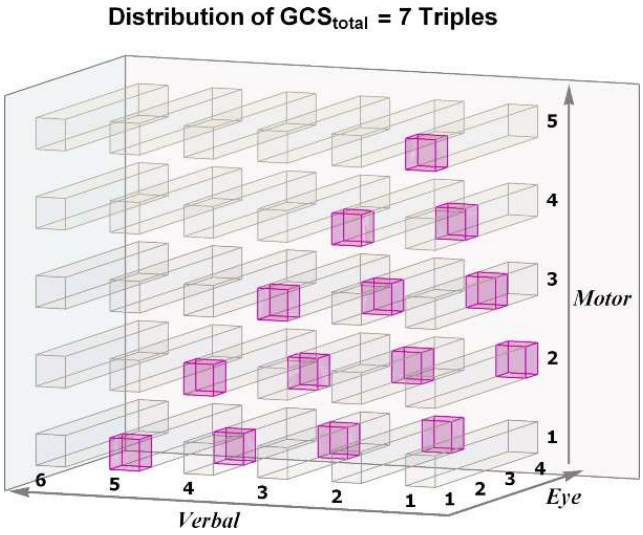


***Figure 1.*** *The distribution of Glasgow Coma Triples with Glasgow Coma Score $GCS = 7$. The Manhattan distance between the lowest GCS (i.e. $GCT = \{1,1,1\}$) and the score $GCS = 7$ is 4, and there are 14 possibilities for patients to have this diagnosis $GCS = 7$.*

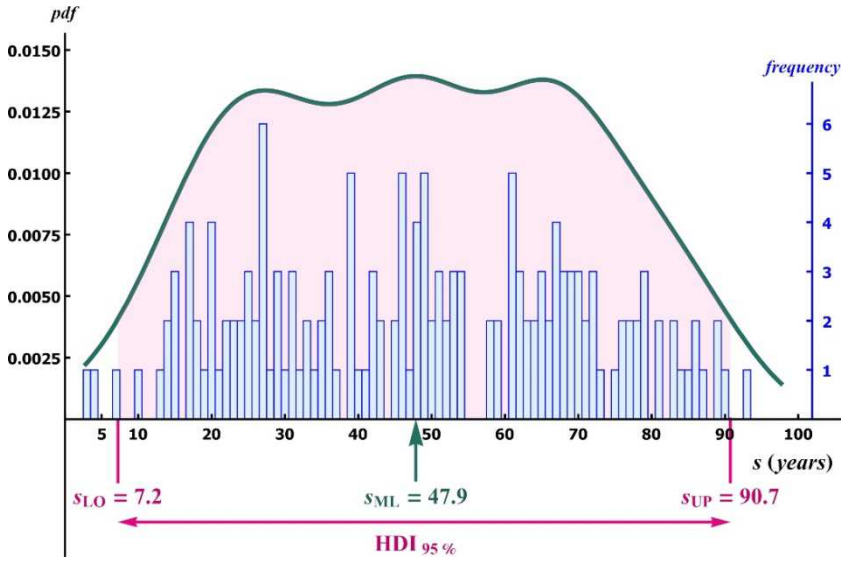


***Figure 2.*** *The (non-parametric) pdf of ages (estimated using KDE) of the TBI patients in this sample. The pdf is almost constant over a large part of its range, indicating that inferences found in this paper are hardly age-dependent. Superimposed on the pdf is the histogram of the ages of the individuals. There are two individuals younger than 5 years and one individual older than 90 years. Point estimators are listed in Table 2. The mode (which cannot be estimated from the raw data) is at 47.9 years, the* $HDI_{95\%}$ *(highest density 95%-confidence interval; [17]) is from 7.2 years to 90.7 years. We note that this uncertainty interval is not symmetric about the mode and is not derived via point estimators.*

***Table 1.*** *The number of possible triples that yield the same GCS (Glasgow Coma Score), listed by increasing GCS. Clearly, as GCS increases, the number of possible triples initially increases. The score* $GCS = 7$*, below which intervention is oftentimes considered medically critical, describes 14 different triples (see Figure 2); and for* $GCS \leq 7$ *there exist 34 different critical conditions. In this paper, we argue that these 34 different conditions are insufficiently medically differentiated by the five different integers from* $GCS = 3$ *to* $GCS = 7$*.*

| Glasgow Coma Score | Number of possible Triples |
|---|---|
| 3 | 1 |
| 4 | 3 |
| 5 | 6 |
| 6 | 10 |
| 7 | 14 |
| 8 | 17 |
| 9 | 18 |
| 10 | 17 |
| 11 | 14 |
| 12 | 10 |
| 13 | 6 |
| 14 | 3 |
| 15 | 1 |

## 2. Materials & Methods

The data set consists of 162 TBI patients (114 males and 48 females; further parameter characterizations in Table 2) that were brought to our trauma center for emergency treatment. Each patient's GCT was diagnosed three times (not necessarily by the same clinician): (a) after leaving the trauma room and being accepted in the ICU, (b) after discharge from the ICU, and (c) after discharge from the hospital (Table 3). Not all TBI patients were alive at discharge; indeed, some died even before acceptance to the ICU. Whether a TBI victim was alive at each of these assessment times was also registered. At discharge, 128 patients were alive (89 males and 39 females); 31 died in the ICU (22 males and 9 females), and 3 were dead prior to acceptance in the ICU (only males). The age ranges, point estimators, and the $HDI_{95\%}$ uncertainty estimator are listed in Table 2 and illustrated in Figure 2.

To justify our rejection of calculating a sum of the Glasgow Coma *scores*, we dispute the approach of inferring the simplifying assumptions about the distributions of categorical variables (a $\chi^2$-distribution, for instance) and the results of tallying techniques displayed as bar charts. The scores of the categorical variables may or may not be independent. For example, $\mathrm{GCS} = 1 + 1 + 1 = 3$ (in the conventional index calculation) may either indicate that a victim is dead or deeply unconscious. Using statistically detectable dependencies can remove this and other ambiguities (see below). Even more important, however, is presenting groups of TBI severities that cannot be revealed by converting categorical variables to cardinal integers and then calculating the sum of these converted responses.

We will infer a metric (albeit not explicitly, but one derived using a neural network algorithm) to replace the triples of GCT. ML techniques, primarily unsupervised dimension reduction algorithms [11], specifically auto-encoders [12–15], can then map the GCT into a lower-dimensional subspace (in this paper a 2D subspace — namely, a plane).

In the AI (artificial intelligence) world, unsupervised learning is standard; in the medical world of statistical analysis, it seems to be novel — or at least not widespread. We know of one prediction model using ML [16]; it does not, however, use unsupervised learning (rather, it uses other AI techniques for outcome evaluation [12, 13] that differ from the ones presented in this paper). In the next analysis step, the KDE (Kernel Density Estimation) algorithm [14] uses as

input what has been 'learned' and outputted by the auto-encoder algorithm.

*One-hot encoding*: The response registrations in the Glasgow Coma scenario are categorical variables. Prior to statistical analyses, the categorical variable scorings are encoded via the one-hot encoding algorithm (an example is shown in Figure 3), into points in an appropriate multi-dimensional space (details are elucidated in the caption of Figure 3). Each patient's responses are therefore represented by a 51-dimensional feature vector.

*Dimension reduction*: The scorings in the (here) 51-dimensional space are not independent. An unsupervised dimension reduction algorithm (here: a neural network autoencoder with 7 layers) finds non-linear combinations of the feature vector scorings that can be mapped to a space of lower dimension (in the present analyses: 2-dimensional).

*Kernel Density Estimation*: Each patient's feature vector has been reduced to a point in a 2-dimensional (Euclidean) space. We apply a KDE [14] algorithm to the distribution of points in this plane to find a non-parametric *pdf* (probability density function; Figure 4) with an optimum Parzen window (found by the KDE algorithm).

*Segmentation*: The *pdf* shows marked peaks (Figure 4). We search for a threshold likelihood to segment these. We thus obtain five specific regions (Figure 5) with likelihoods higher than their environs, which we call groups in the subsequent analyses. As each patient is a point in the reduced 2-dimensional space, we use Monte Carlo integration (by generating 130000 points randomly distributed according to the *pdf*) to determine the relative probability of patients' being in one of these five groups. Each patient in each group can be characterized by his/her initial, intermediate and final GCT (not: GCS). We look for — and find — medical characteristics the patients in each group have in common. We also succeed in an extended triage; the deceased patients are segmented via the unsupervised algorithms into a further, separate group.

***Table 2.*** *Age parameters of the 162 TBI patients. The point estimators (means) are not meaningfully interpretable; preferable is the 95% highest density uncertainty interval ($HDI_{95\%}$ [17]) and the mode (Figure 2).*

| **Age range (years)** | **3–93** |
|---|---|
| Point Estimators: | |
| mean (years) | 48.35 |
| $\text{mean}_{\sigma}$ (years) | 46.41 |
| $\text{mean}_{\varphi}$ (years) | 52.65 |
| $HDI_{95\%}$ (years) | 7.2–90.7 |
| Mode (years) | 47.9 |

***Table 3.*** *Classification of the GCTs at the three times for all 128 surviving patients.*

| **Time** | **frequency of GCT = {1, 1, 1}** | **frequency of GCT ≠ {1, 1, 1}** |
|---|---|---|
| Post Trauma Room | 82 | 46 |
| Post ICU | 75 | 53 |
| Post Hospital | 9 | 119 |

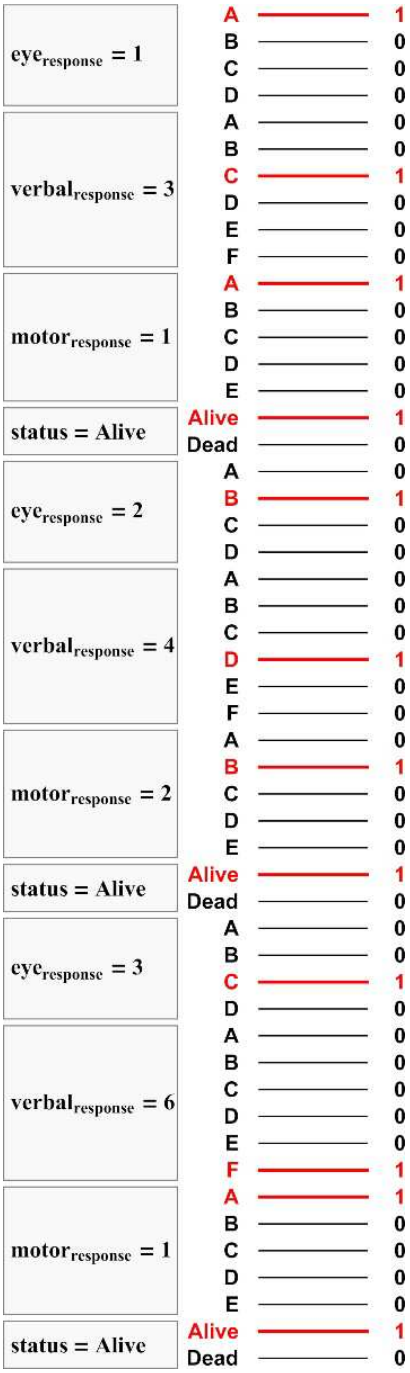


***Figure 3.*** *Application of the algorithm that uses one-hot encoding to convert a listing of categorical variables into a multidimensional feature vector, exemplified for individual No. 83. The three Glasgow Coma triples are* $GCT_1 = \{1,3,1\} = \{A, C, A\}$, $GCT_2 = \{2,4,2\} = \{B, D, B\}$, *and* $GCT_3 = \{3,6,1\} = \{C, F, A\}$*; the respective GCSs are:* $GCS_1 = 5$, $GCS_2 = 8$, *and* $GCS_3 = 10$. *The diagnosed responses are categorical variables that have been illustrated as (capital) letters from the Latin alphabet. The resultant 51-dimensional feature vector is a string of 0s and 1s. For illustration purposes, the components that are nonzero have been highlighted in red (thereby explaining the name for the algorithm: hot wires are insinuated to glow red). We note that there is an inherent interdependence between the components of this feature vector. If, as in this example, the 7th component is "1", then the 5th, 6th, 8th, 9th, and 10th components must be "0". Also: if the 50th component is "1" (as here), then the 51st component must be "0"; furthermore: the 16th and the 33rd component must be "1", while the 17th and the 34th component must be "0".*

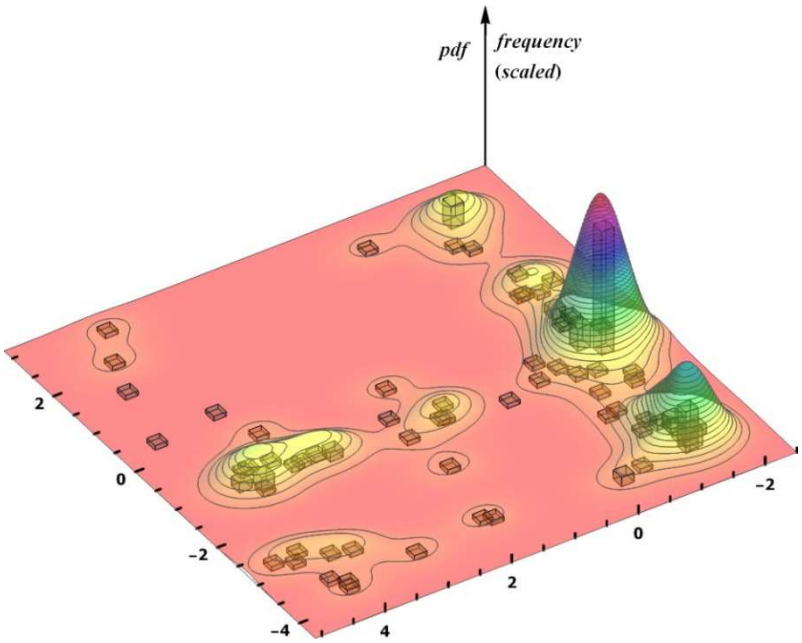


***Figure 4.*** *The pdf (probability density function) of the dimension-reduced GCTs of the surviving patients, superimposed on the scaled bar chart of frequencies of the individual survivors' 2D-scores (shown as stacked cubes). The pdf has been estimated using a KDE algorithm. The bar chart exhibits a 'natural' binning, because it can be identical for several surviving patients. The color-coding of the pdf shows there are five groups of high likelihood (beyond brown) for survivors. The bar chart entries (frequencies) of lowest likelihood (rendered as height in the graph) are singletons. We observe that the very high likelihood 2D-scores are only within the five groups. The scales along the two (horizontal) axes have no direct medical interpretation; they are therefore unlabeled.*

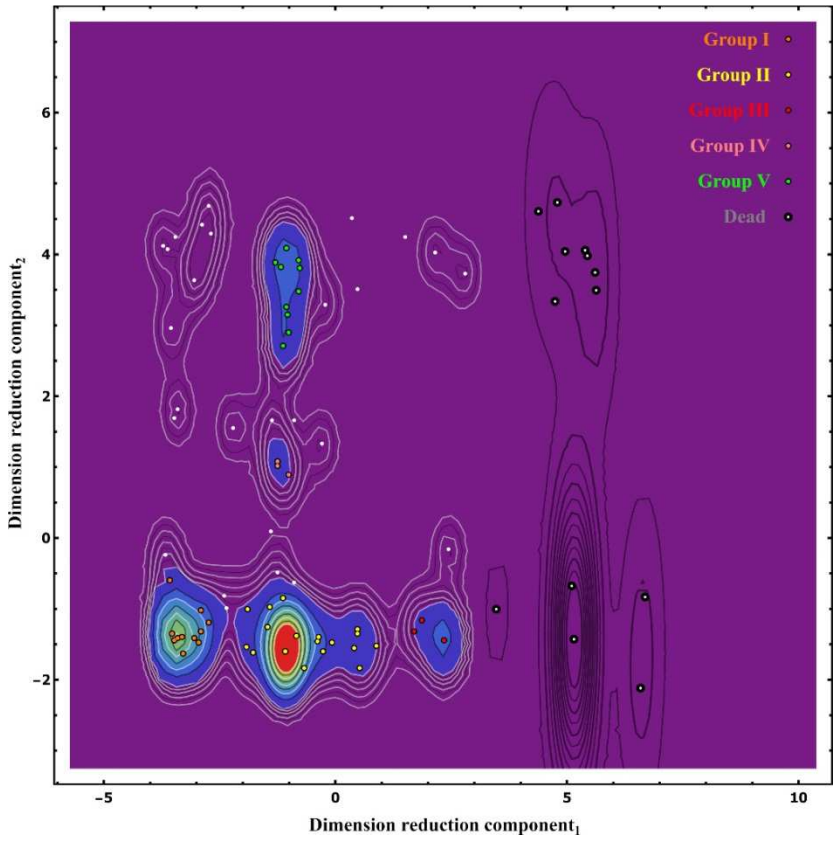


***Figure 5.*** *The projection of the KDE-derived pdf of the dimension-reduced coordinates of the three GCTs of the surviving TBI patients. Superimposed are the coordinates of the 2D-scores and the segmentation results. The white contour lines help reading the graph; they are of no importance for the analyses. The black contour lines and the black dots with white centers are the coordinates of the 34 TBI patients that did not survive. Their 2D-scores are also segmented. However, in this paper we do not analyze these segmentations (we are only investigating the temporal progression of the survivors' GCTs). We observe that the 2D-scores of the TBI patients who died do not overlap the 2D-scores of the 128 TBI survivors — an indicator that the unsupervised dimension reduction algorithm is highly successful and the pdf derived via KDE must be considered reliable.*

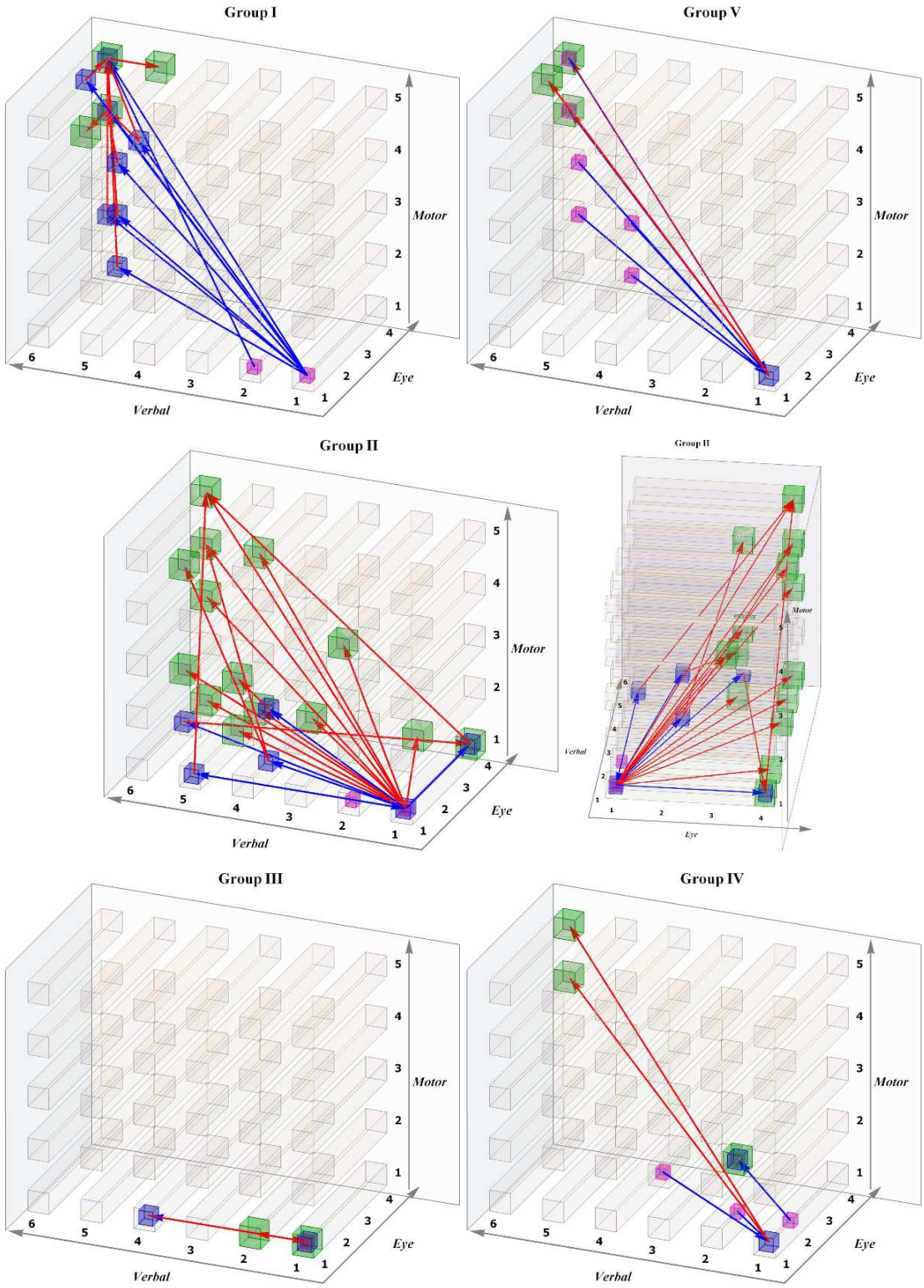


***Figure 6.*** *The temporal evolution of GCTs of 128 TBI patients that were still alive post hospital displayed by groups identified by KDE segmentation (Figure 4). The initial GCT (post TR) is displayed as a magenta cube, the intermediate GCT (post ICU) as a blue cube, and the final outcome (post hospital) as a green cube. The arrows indicate the direction of temporal evolution. In Group I and Group V, all patients regain full or almost full consciousness; in Group I, patients enter the ICU essentially unconscious, while in Group V, patients enter the ICU with reasonably high levels of consciousness, these levels drop by the time of discharge from the ICU, and the reasonably high levels of consciousness are regained before discharge from the hospital. Group II is displayed in two different orientations to ease readability: in this group, the final eye response is high (3–4), but verbal and motor response vary over the whole range (verbal: 1–6; motor: 1–5). In Group III, all patients are unconscious prior to entering the ICU and do not regain consciousness at discharge from the hospital. In Group IV, most patients regain a high level of consciousness, except for one type of outcome: high visual response and intermediate verbal response and no motor response. In all graphs, the multiplicity of each cube is not rendered; therefore, the graphs do not show the frequencies of the outcomes. The probabilities of outcomes are listed in Table 4. We note that these differentiating descriptions cannot be detected using GCSs.*

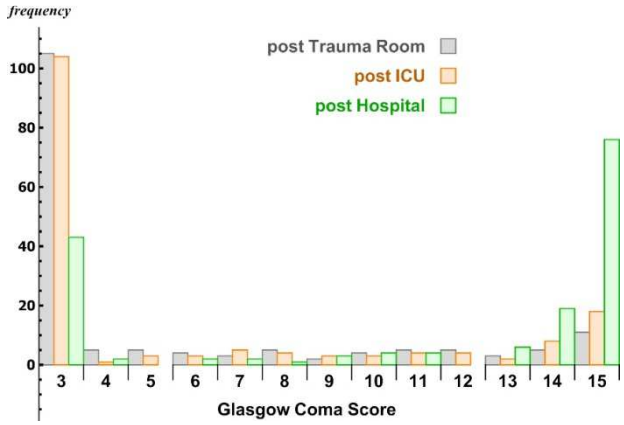


***Figure 7.*** *The distribution of GCSs at the three times we analyze in this data set. The distribution includes all 162 TBI survivors and non-survivors. Upon admission to the ICU, most patients had* $GCT = \{1,1,1\}$*, and at hospital discharge, most had* $GCT = \{4,6,5\}$ *(Table 3). This paper deals with analyzing as to whether GCS, in contrast to GCT, can be considered medically sound because most clinicians argue that it informs adequately; we argue that GCS does not. Bizarrely, none of the patients had* $GCS = 5$ *or* $GCS = 12$ *at hospital discharge.*

***Table 4.*** *Tabulation of the GCTs relative probabilities and frequencies of occurrence for the 128 patients that were alive at discharge from the hospital. We note that very few survivors remain at the state* $GCT = \{1,1,1\}$ *(equivalent to* $GCS = 3$ *— unconscious survivors) at hospital discharge; the probability is 5.8% for a patient being alive at this low level of consciousness. The increase in the number of triples* $GCT \neq \{1,1,1\}$ *necessitates investigating the temporal evolution of GCT of the patients that remained alive at discharge from the hospital (Figure 6), not how the GCSs with* $GCS > 3$ *are distributed.*

*a*

| **Group** | **I** | | **II** | |
|---|---|---|---|---|
| **Probability (%)** | **27.2** | | **52.6** | |
| initial | {1,1,1} | 25 | {1,1,1} | 48 |
| | {1,2,1} | 1 | {1,2,1} | 1 |
| | {3,6,4} | 1 | {3,3,3} | 1 |
| | {4,5,5} | 1 | {3,5,1} | 1 |
| | {4,6,4} | 2 | {3,5,2} | 1 |
| | {4,6,5} | 22 | {3,6,2} | 1 |
| | | | {3,6,4} | 1 |
| | | | {4,1,1} | 1 |
| final | | | {4,2,1} | 1 |
| | | | {4,4,1} | 1 |
| | | | {4,5,1} | 1 |
| | | | {4,5,4} | 1 |
| | | | {4,6,1} | 2 |
| | | | {4,6,3} | 1 |
| | | | {4,6,4} | 6 |
| | | | {4,6,5} | 30 |

*b*

| **Group** | **III** | | **IV** | | **V** | |
|---|---|---|---|---|---|---|
| **Probability (%)** | **5.8** | | **1.7** | | **12.6** | |
| initial | {1,1,1} | 7 | {1,3,2} | 1 | {2,4,2} | 2 |
| | | | {2,1,1} | 1 | {2,4,3} | 1 |
| | | | {2,2,1} | 1 | {2,5,3} | 1 |
| | | | | | {2,5,4} | 2 |
| | | | | | {4,6,4} | 3 |
| | | | | | {4,6,5} | 4 |
| final | {1,1,1} | 6 | {2,2,2} | 1 | {3,6,5} | 1 |
| | {1,2,1} | 1 | {4,6,4} | 1 | {4,6,4} | 3 |
| | | | {4,6,5} | 1 | {4,6,5} | 9 |

# 3. Results

*Clustering into five groups*: The segmentation results in five groups (Figure 5) of survivors, with probabilities $1.2\% < P_{\text{group}} < 36.3\%$ (Table 4). These probabilities have *not* been calculated as point estimators derived from patient occurrence frequencies, but rather by Monte Carlo integration of the group regions defined via the *pdf* (Figure 4). Group membership accounts for 68.9% of all cases observed of alive patients. Two groups account for almost 55.5% of these cases. All triples outside these five groups, accounting for 28 (of 128) scorings, are singletons. Each of their likelihoods is below *pdf* threshold (likelihood threshold). We are surprised that so many patients can be described by so few groups of Glasgow Coma Triples.

*Properties of the five groups*: The groups vary considerably in size and in membership frequencies (Table 4). Patients in Group I and Group II have initial triples $\{1,1,1\}$ and very rarely (only 2 of 75 patients) $\{1,2,1\}$. Arguably, $\text{GCT} = \{1,2,1\}$ can be considered, in the context of the analyses presented here, as no different from $\text{GCT} = \{1,1,1\}$, given the uncertainty of the boundary in the assessment by a clinician of the patient's response at such low levels of consciousness. The patients in Group III, with a probability $P_{\text{Group III}} = 4.0\%$ are the ones with the (medically) bleakest outlook (Table 4 and Figure 6); none regained any semblance to higher states of consciousness at hospital discharge. Group IV, with the lowest probability ( $P_{\text{Group IV}} = 1.2\%$ ), is a group with no straightforwardly identifiable pattern: some patients regained high levels of consciousness, while some did not. We argue that the smallness of the group may include statistical fluctuations that mask any pattern. In Group V, patients' initial responses are characterized by the highest number of different GCTs (namely six) of all groups: from $\text{GCT} = \{2,4,2\}$ to $\text{GCT} = \{4,6,5\}$ (we note in passing that the initial GCSs are $8, 9, 10, 11, 14, 15$ — they are neither consecutive nor are they unique to this group). Occurrence of patients in Group V has the 3rd highest probability ($P_{\text{Group V}} = 8.7\%$), and the patients' outcome triple components are high ($\{3,6,5\}$, $\{4,6,4\}$, and $\{4,6,5\}$). These triples also occur in other groups, so summing the outcome triples in Group V to form GCSs is neither useful for uniquely characterizing TBI patients, nor helpful for segmentation purposes, nor (medically) meaningful in order to predict patient outcomes (Figure 6).

The method we use in this paper also indicates an additional triage: the patients who did not survive is manifested in the KDE *pdf*. Interestingly, the GCSs in each of these groups do not allow for a distinction, while the GCTs do. Relying on GCSs therefore masks information that the GCTs contain.

# 4. Discussion

By showing how the survivors' GCTs segment into five groups, while the GCSs do not (Figure 7), we can document

that computing indices is neither statistically meaningful (because of allowing arithmetic operations for categorical variables), nor medically useful (the consciousness statuses of different patients with the same GCS have very different progressions over time). Indeed, a histogram of the occurrences of the GCSs (Figure 7) shows no pattern that reveals the temporal evolution of the patients during their transition from pre-ICU to hospital discharge.

The segmentation result also shows which patients did not survive prior to hospital discharge (Figure 5) — a segmentation result that had not been explicitly included in the programming code before we began the analysis. The segmentation method thus also supplies a triage: segmenting the five groups of survivors from the deceased.

Clinicians are — and rightfully so — interested in whether the patients in each survivor group have further identifiable diagnostic characteristics. We too are interested in such dependencies on other variables. Clusters of GCSs, we have shown, are not meaningful and only *after* segmentation of GCTs can meaningful associations be found. More rigorously stated: in this paper we have shown that GCSs are not only insufficiently informative, but also unable to reveal any differentiation of the clinical picture of TBI patients.

The segmentation of the survivors into five groups also describes a progress within the ICU and an outcome comparable to GOSE. Rather than relying on GOSE diagnosis criteria, the suite of algorithms we used predicts five types of ICU progression (for the survivors) and three types of outcome (Groups I, IV, and V have the same post-ICU outcome).

# 5. Conclusion

We argue that the novel approach of using unsupervised ML can uncover medical characterizations of TBI patients that are impossible to discover using only the temporal sequence of three GCSs. We would like to add that, by using unsupervised ML and the presented sequence of further algorithms, we avoid fallacies involving assignments of cardinal numbers to categorical variables and relying on the conventional GCS metric that has not been statistically derived.

This sequence of analysis steps (one-hot-encoding → dimension reduction → KDE of non-parametric *pdf* → segmentation) is unsupervised. Nowhere in the calculations are parameters chosen by us, the authors, except for our choice of threshold (Figure 4) after the completion of all algorithm executions. Other algorithms for segmentation exist; we chose segmentation via non-parametric *pdf* thresholds.

We can use unsupervised ML to analyze the categorical response assessments that clinicians perform to estimate (among other things) the level of consciousness of TBI patients. We show that the outcomes of ML segmentation are much more informative about TBI patients' consciousness landscape than GCSs would lead a clinician to suspect. We can show that the 128 TBI post-hospital survivors predominantly fall into five groups. These groups differ in how their consciousness levels progress in time during their stay in the ICU.

Understandably, a clinician would like to have 'beforehand' information as to what the outcome likelihoods for patients delivered to the trauma room will be. This is very challenging. As we show in this article, using the Glasgow Coma Score is not helpful; it is, in fact, misleading. On the other hand, our segmentation using unsupervised ML can supply a triage (not only of deaths versus survival, but also different survival profiles) and also point out that, at discharge from the ICU, patients with the same GCS have vastly different GCTs. By implication, there must be other parameters that differentiate the outcomes. Further research is required to identify these.

# Author Contributions

HS supervised the project. HH, AA, and RB collected the data. KM and AG checked the data and filled in gaps wherever possible. HP designed the analyses, programmed various algorithms, and analyzed the data. HP wrote the manuscript; HS and HH made corrections and modifications.

# Conflict of Interest

The authors declare that they have no financial interest in the outcomes of the presented research.

# Acknowledgements

We thank the hospital medical staff for their untiring dedication in their attempt to increase the survival chances of the TBI patients under their care. We also thank support staff for carefully collecting data.

# References

[1] Teasdale G, Jennett B. (1974) Assessment of coma and impaired consciousness: a practical scale. *Lancet* 304, 81–84.

[2] Teasdale G, Jennett B, Murray L, Murray, G. (1983) Glasgow Coma Scale: to sum or not to sum. *Lancet*. 2, 678.

[3] Moore L, Lavoie A, Camden S, Le Sage N, Sampalis JS, Bergeron E, Abdous, B. (2006) Statistical validation of the Glasgow coma score. *J Trauma*. 60, 1238–1243.

[4] Iankova, A. (2006) The Glasgow Coma Scale clinical application in an emergency. *Emergency Nurse* 14 (8), 30–35. Doi: 10.7748/en2006.12.14.8.30.c4221.

[5] Green, S. M. (2011) Cheerio, Laddie! Bidding Farewell to the Glasgow Coma Scale *Annals of Emergency Medicine* 58 (5), 427–430. Doi: 10.1016/j.annemergmed.2011.06.009.

[6] Zuercher M, Ummenhofer W, Baltussen A, Walder, B. (2009) The use of Glasgow Coma Scale in injury assessment: a critical review. *Brain Inj* 23, 371–384.

[7] Gill M, Windemuth R, Steele R, Steven, G. (2005) A comparison of the Glasgow coma score to simplified alternative scores for the prediction of traumatic brain injury outcomes. *Ann Emerg Med* 45, 37–42.

[8] Gill M, Steele R, Windemuth R, Green, SM. (2016) A comparison of five simplified scales to the prehospital Glasgow Coma Scale for the prediction of traumatic brain injury outcomes. *Acad Emerg Med* 13, 968–973.

[9] Choi JH, Jakob M, Stapf C, Marshall RS, Hartmann A, Mast H. (2008) Multimodal early rehabilitation and predictors of outcome in survivors of severe traumatic brain injury. *J Trauma* 65, 1028–1035. doi: 10.1097/TA.0b013e31815eba9b.

[10] Rollnik JD, Bertram M, Bucka C, Hartwich Jöbges M, Ketter G, Leineweber B, Mertl-Rötzer M, Nowak DA, Platz T, Scheidtmann K, Thomas R, von Rosen F, Wallesch CW, Woldag H, Peschel P, Mehrholz J, Pohl M. (2016) Criterion validity and sensitivity to change of the Early Re Rehabilitation Index (ERI): results from a German multi-center study. *BMC Res Notes* 9, 356. doi: 10.1186/s13104-016-2154-8.

[11] Kelleher, J. D. (2019) *Deep Learning*. MIT Press, Cambridge, MA, USA.

[12] James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning with Applications in R*. New York: Springer Science+Business (2015). 426 p. doi: 10.1007/978-1-4614-7138-7.

[13] Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction.* New York: Springer (2017). 745 p.

[14] Murphy KP. *Machine Learning. A Probabilistic Perspective.* Cambridge, MA: The MIT Press (2012). 1071 p.

[15] Mohri, M., Rozdamizadeh, A., Talwalkar, A. (2018) Foundations of Machine Learning (Adaptive Computation and Machine Learning Series), Cambridge, MA: The MIT Press. 504 p.

[16] Fleuren, L. M., Klausch, T. L. T., Zwager, C. L., Schoonmade, L. J., Guo, T., Roggeveen, L. F., Swart, E. L., Girbes, A. R. J., Thoral, P., Ercole, A., Hoogendoorn, M. & Elbers, P. W. G. (2020) Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Car Med* 46, 383–400. Doi.org/10.1007/s001.34-019-05872-y.

[17] Kruschke, J. K. (2015) *Doing Bayesian Data Analysis*. *A Tutorial with R, JAGS, and Stan*. Elsevier, Amsterdam.