



SCUOLA DI DOTTORATO
Università degli Studi di Milano-Bicocca

Department of Medicine and Surgery
PhD program in Public Health
Cycle XXXVI
Curriculum in Biostatistics and Epidemiology

**THE USE OF JOINT AND COX MODELS
TO ASSESS THE ASSOCIATION BETWEEN A
LONGITUDINAL MARKER AND A TIME-TO-EVENT:
A SIMULATION STUDY UNDER DIFFERENT
MISSING MECHANISMS
AND APPLICATIONS IN ICU SETTING**

Candidate: PETROSINO MATTEO
Registration number: 876320

Tutor: Prof.ssa Paola Rebora
Co-tutor: Prof.ssa Stefania Galimberti

ACADEMIC YEAR 2022-2023

Contents

1	Introduction	5
2	Longitudinal data	6
2.1	Methods for longitudinal data analysis	7
2.2	Missing Data in Longitudinal Studies	8
2.3	A two-stage approach estimation for longitudinal responses	11
2.3.1	Stage I	11
2.3.2	Stage II	11
2.4	The Linear Mixed Model (LMM)	12
2.4.1	Estimation of LMM	14
2.5	Mixed Models on real data: two examples	15
2.5.1	Management of arterial partial pressure of carbon dioxide in the first week after traumatic brain injury: results from the CENTER-TBI study	16
2.5.2	High arterial oxygen levels and supplemental oxygen administration in traumatic brain injury: insights from CENTER-TBI and OzENTER-TBI	16
3	Survival Analysis	17
3.1	Characteristics of event time data	17
3.2	Basic theory of survival analysis	18
3.2.1	Time-to-event	18
3.2.2	Censoring	19
3.3	Survival times estimations	19
3.3.1	Non-parametric estimation	19
3.3.2	Parametric estimation	20
3.3.3	Likelihood function for censored data	21
3.4	Failure Times Models	22
3.4.1	Parametric regression models	23
3.5	The Cox Model	23
3.6	Survival analysis with time-dependent covariates	24
3.7	The Extended Cox Model (ECM)	25
4	Joint Models for longitudinal and time-to-event data	27
4.1	Joint Model formulation	27
4.1.1	The Survival sub-model	28
4.1.2	The Longitudinal sub-model	29
4.2	Joint Models estimation	29
4.2.1	Random Effects estimation	30
4.3	Asymptotic Inference	31
4.4	Different Joint Model parametrizations	32
4.5	Connection with the Missing Data Framework	33
4.6	Final considerations on Joint modeling of longitudinal and time-to-event data	36

5	The ORANGE study	37
5.1	Background	37
5.2	The Neurological Pupil Index (NPi)	38
5.3	Study design	38
5.3.1	Outcomes and Aims	40
5.4	Statistical methods	40
5.5	Results on descriptive statistics	41
5.5.1	Baseline characteristics and outcomes	41
5.5.2	NPi evaluations	41
5.5.3	Association between abnormal NPi and 6-months outcomes	43
5.6	Results on long-term outcomes	45
5.6.1	Results on short term mortality (ICU): a comparative analysis using ECM and JM	47
5.7	Discussion	49
5.7.1	Limitations	50
6	Evaluating the robustness of Extended Cox and Joint models with missing longitudinal data	55
6.1	Times-to-event simulation with baseline covariates	56
6.1.1	Survival time simulation	56
6.1.2	The proportional hazards data generating process	56
6.2	Complex survival data generating processes: longitudinal and time-to-event data simulation	57
6.3	Simulation study for longitudinal and time-to-event data in the presence of different missing processes on the longitudinal covariate	58
6.3.1	Simulation of longitudinal and event times data	58
6.3.2	simjm: a R package for simulating Joint longitudinal and time-to- event data	59
6.3.3	Missing processes on simulated data	60
6.3.4	missMethods: a R package to simulate missing data	61
6.4	Simulation design	62
6.4.1	Simulation parameters	62
6.4.2	Generated longitudinal data	63
6.4.3	Methods	66
6.4.4	Statistical indices to evaluate estimation performance	66
6.5	Results	67
6.5.1	Low variability of longitudinal markers	67
6.5.2	High variability of longitudinal markers	68
6.6	Discussion	69
6.6.1	Limitations	70
6.6.2	Future works	70
7	Final remarks	75
	Appendices	83
A	Papers ([23],[24])	84
B	Simulations R code	85

Acknowledgements

I sincerely want to thank Prof. Stefania Galimberti and Prof. Maria Grazia Valsecchi for giving me this wonderful opportunity to pursue for a doctorate in a professionally and humanly rich environment, which under their wise guidance grows day after day. Special thanks go to my tutor, Prof. Paola Rebori, for her support, her ideas and her enthusiastic encouragement. A sincere and deep affection goes to Prof. Giuseppe Citerio, "Il Cit", who in addition to having taught me so much, supported me, put up with me and as he always says "made me work very little". I also thank all my PhD classmates and the other researchers in the group with whom I shared many peaceful moments and Prof. Laura Antolini for her comments on my work. A heartfelt thank you goes to Beatrice for a thousand reasons + 1 and to my family.

I dedicate this achievement to my beloved "Nonna Titti".

Chapter 1

Introduction

The widespread use of information technology has affected many facets of life, including healthcare and the hospital setting. This has made it possible to gather vast amounts of information about patients' clinical histories.

Within this thesis, some fundamental aspects regarding the analysis of longitudinal data will be addressed. Such data are the result of repeated measurements on the same subject of certain clinical indicators or physiological signs, which are commonly referred to as *biomarkers* (within the thesis we will use biomarkers or markers to indicate the covariates measured over time, without distinction). During these three years as a doctoral student, I have worked, from an applied perspective, primarily on longitudinal data from the clinical setting of the Intensive Care Unit (ICU). Because of the medical condition in which ICU patients find themselves, they are subjected to close and continuous monitoring, which generally begins from the day of admission and continues over the following weeks. The vital functions of such individuals, observed over time and quantified through specific markers, show sudden changes: this precisely forces close surveillance and constant evaluation of various clinical indicators of interest over time. This produces, in many cases, a large amount of data that can be analyzed. Therefore, it is essential to have a thorough description of the longitudinal trend of specific markers and the impact they may have on the outcome of interest. So starting with our motivating clinical context, within this thesis we will address, from a modeling perspective, of how to quantify the impact that a longitudinal biomarker may have on the trend of the risk of experiencing a binary event over time, such as death. This question will form the background throughout the paper and special emphasis will be given to the Cox model in its extended version and to the Joint models for longitudinal and time-to-event data. We therefore provide a very brief summary of the structure of the thesis, highlighting the main topics that the reader will find within the individual chapters. The second chapter frames the nature of longitudinal data, common strategies for analyzing them (in particular the *Linear Mixed-Effects models*) and frequent issues such as *missing data*, which are typical in the context of dynamic data. *Survival Analysis* will be briefly introduced in order to build the bridge that will connect the following chapters. In the third and fourth chapter, respectively, two widely used approaches for analyzing the link between a longitudinal profile and the risk of a binary-type event (e.g., death) over time, namely the *Cox model* in its extended version (ECM) and *Joint models* (JM), are detailed. The fifth chapter shows the results from the ORANGE study, an international, observational, prospective cohort study which is the clinical motivating context. The aim of the study was to evaluate the association between the NPi, a neurological pupil index measured over time with an automated electronic device and long-term outcomes, in patients with Acute Brain Injury admitted to Intensive Care units. The data were analysed with ECM, but an empirical comparison with JM fitted on the same data to evaluate ICU mortality is shown. The sixth chapter is devoted to the presentation of the results of the simulation work done to evaluate the robustness of ECM and JM when the longitudinal process is affected by different missing mechanisms. Finally the last chapter contains some final remarks.

Chapter 2

Longitudinal data

In a longitudinal study, each experimental or observational unit is measured at baseline and repeatedly over time on the same subject ([1]). The primary characteristic that sets apart longitudinal studies from other types of research in this context is the occurrence of several measurements across time on the same subject. Direct assessment of changes in outcomes of interest over a clinically relevant time window is made possible by continuous monitoring of participants during the observation phase and subsequent data collection. A longitudinal model is mostly employed to inspect two different types of effects:

- cross-sectional effects, that is what differs between groups at a given point in time (for instance, the mean difference between men and women or between two treatment arms);
- longitudinal effects, also known as time effects or distinct time effects between groups of participants, e.g. include average trajectories of mean arterial pressure following the initiation of a therapy or variations in trajectories between males and females.

Therefore, in the longitudinal context we can distinguish two different sources of variability: one between-subjects, also called inter-variability, and one within-subjects, that is, intra-variability, due to the repetition of measurements.

From a statistical perspective, the key feature of the longitudinal study is the presence of a correlation structure that arises due to repeated measures on the same individuals and it must be accounted for in order to obtain valid inferences ([2]). This circumstance necessitates the employment of ad hoc statistical models and contrasts with the assumption of independence among the residuals, typical, for example, of the linear regression model.

Three sources of variability that describe longitudinal data should be taken into consideration by an adequate model ([3]):

- random variability coming from heterogeneity among individual trajectories;
- serial correlation due to residuals close to each other in time are more similar than residuals further apart;
- measurement error to account for small variability unavoidable even from an immediate replication of the measurement (noise variability).

Longitudinal data frequently fail the homoscedasticity assumption because the variability of the data varies depending on when it was taken, that is, the variability is often heterogeneous across measurement occasions. Furthermore, the number of repeated measurements among patients are frequently *unbalanced*: this means that the data observed on participants in a study are not collected at exactly the same times. Longitudinal data allow researchers to assess multiple disease aspects: changes of outcome(s) over time in relation to associated risk factors, timing of disease onset and individual and group patterns over time ([4]). Assessing longitudinal temporal changes is crucial to learning specific time patterns: for example in a clinical context important impairments could be missed ([5]).

Theoretical and practical concerns, however, make longitudinal data analysis challenging. These consist of data that are possibly missing, correlated, and collected across wide irregularly visit times. Modern statistical techniques address these difficulties, but challenges include knowing when to apply them, checking their assumptions and appropriately interpreting their results. Confusion over these issues may result in analysis that is both improper and misleading.

2.1 Methods for longitudinal data analysis

A longitudinal analysis of within-individual developments is conducted in two conceptually separate stages. Within-individual change is described, in the initial stage, using an adequate summary of the variations in the repeated measurements taken on each subject throughout the observation period. Certainly, the simplest longitudinal design is with only two measurements at the start and at the end of an observational period; for example in a pre post-treatment scenario: here a straightforward approach is to analyze the change score or difference score, that is the differences between the measures at each time point, and a valid approach is to use ANOVA ([6]). These estimates of within-individual change are connected to inter-individual variations in a few important factors (e.g. treatment group, smoking status, gender...) in the second stage. Although conceptually distinct from one another, these two analytical steps can be merged into a single statistical model. In fact, this serves as the foundation for the extremely adaptable class of models known as linear mixed-effects models for longitudinal analysis of continuous responses ([7]).

The mixed-effects model will be better described in the next sections as it is used as a building block for Joint models ([8]), which will be a major theme of this thesis.

Longitudinal data from planned experiments have historically been widely analyzed using response profile analysis: either a univariate repeated-measures ANOVA or a multivariate ANOVA (MANOVA) was used to implement it. These approaches, however, present several limitations. Unrealistic assumptions regarding the variances on each occasion and the correlations between pairs of repeated measures are made by univariate repeated-measures ANOVA, which is commonly the case for longitudinal data. Although MANOVA does not impose any restrictions on variances or correlations, many implementations require a complete case data scenario, which may make its use challenging ([6]).

To characterize the longitudinal trajectories arising from repeated measurements, other types of approximations can be used, such as piecewise linear, curvilinear, or polynomial functions that can also parsimoniously summarize within-individual changes in the response over time. A truly flexible approach to estimating longitudinal profiles, which offers great adaptability due to its mathematical properties, is the use of splines: there are many types, characterized by the choice of different *basis functions* ([9]). These methods, though, do not allow the user to directly regress the longitudinal responses on covariates, but they are used to estimate particular, especially non linear trends, over time.

Another widely used approach to estimate longitudinal profiles is *Generalized Estimating Equations (GEE)*. In contrast to mixed-effects models, returning a conditional model with the possibility to get marginal interpretations ([10]), GEE estimate a marginal model for the mean of the trajectory over time ([11]), making the former to be often preferred to the latter.

2.2 Missing Data in Longitudinal Studies

The issue of missing data presents a significant obstacle for the analysis of longitudinal data. It is fairly common for some participants to miss some of their planned assessments for a number of reasons, even though longitudinal studies are meant to collect data on every subject in the sample at a set of predetermined follow-up times. We can distinguish between *monotone* and *non-monotone* missingness based on the characteristics of the patterns of absent data ([1]). When a participant withdraws from the study before it is supposed to be finished, this is known as *attrition* or *dropout*. *Late entry* occurs when a participant does not provide all of her initial response measurements but shows up later and continues to participate in the study until it is finished. Contrarily, non-monotone missingness, often referred to as *intermittent* missingness, is a more generic class that includes situations in which, for instance, a subject's response is missing at one follow-up time, she returns at the next, but can then be missing once more at later time points.

When designing longitudinal research and analyzing the data from these investigations, the possibility of missing data presents a number of difficulties. The loss of efficiency is the first and most evident result, as changes in the average longitudinal evolutions are estimated with less accuracy than they would have been if all data had been accessible. In order to achieve the same levels of power in identifying significant effects, we will have to recruit more people, which has crucial implications for the design of long-term research. The amount of missing data is closely correlated with the loss of precision, and the technique of analysis used also has an impact. Additionally, because not all individuals have the same amount of measurements at a similar set of times, missingness causes datasets to become unbalanced over time. This complicates analytical techniques that call for balanced data, but does not pose any concern for the linear mixed-effect model we previously discussed. Finally, missing data could induce bias and result in erroneous inferences in some situations especially if handled incorrectly. Before going into more details about how bias issues could arise, it is necessary to provide new terminology that enables us to formally characterize the missing value mechanism and assess how this mechanism might affect subsequent conclusions. In general, we assume that each subject in the study is intended to be measured at $j = 1, \dots, n_i$ times, which means that, for this subject, we are expecting to collect the vector of measurements $y_i = (y_{i1}, \dots, y_{in_i})$. The missing data indicator is introduced to help distinguish between the response measures we actually gathered and those we expected to collect. It is defined as:

$$r_{ij} = \begin{cases} 1 & \text{if } y_{ij} \text{ is observed} \\ 0 & \text{otherwise} \end{cases} \quad (2.1)$$

Therefore, we obtain a partition of the *complete* response vector y_i into two sub-vectors, the observed data sub-vector y_i^o containing those y_{ij} for which $r_{ij} = 1$, and the missing data sub-vector y_i^m containing the remaining components. The vector $r_i = (r_{i1}, \dots, r_{in_i})^T$ and the process generating r_i are referred to as the *missing data process* and the mechanism for missing data determines whether certain techniques for analyzing incomplete longitudinal data are appropriate. The missing data mechanism can be viewed as a probability model outlining the connection between the processes for missing data (r_i) and response data (y_i). A taxonomy of missing data mechanisms, first proposed by Rubin in [12], is based on the conditional density of the missing process r_i given the complete response vector $y_i = (y_i^o, y_i^m)$:

$$p(r_i | y_i^o, y_i^m; \phi_r)$$

where ϕ_r denotes the corresponding parameter vector. The three types of mechanisms are:

- **Missing Completely at Random (MCAR)** which corresponds to the mechanism such that neither the exact values that they would have been observed nor the collection of observed responses have any bearing on the probability that responses are missing. In other words, longitudinal data are MCAR when r_i is unrelated to both y_i^o and y_i^m , i.e.:

$$p(r_i | y_i^o, y_i^m; \phi_r) = p(r_i; \phi_r).$$

Indeed, the observed data y_i^o can be viewed as a random sample of the complete data, which is an important feature of MCAR. As a result, the distribution of the observed data is the same as the distribution of the entire set of data, meaning that the observed data are just a random sample of all the data ([1]). Because of this, MCAR allows us to draw valid conclusions from the data at hand using any appropriate statistical technique while ignoring the process(es) responsible for the missing values.

- **Missing at Random (MAR)** postulates that the probability of missing is independent of the results that should have been attained and instead depends on the collection of observed responses. In other words, longitudinal data are MAR if r_i is conditionally independent of y_i^m given y_i^o . An example of MAR is when a patient's removal from a study is mandated by the study's protocol because their response value exceeds a predetermined threshold. The observed data cannot be regarded as a random sample from the target population in this instance because the missing process is controlled by external factors and is only related to the observed values y_i^o ([1]). Due to the fact that the missing data mechanism depends on y_i^o , the distribution of y_i^o does not match the distribution of y_i . The only distribution that is the same as the distribution of the corresponding observations in the target population is the distribution of each subject's missing values, y_i^m , conditioned on her observed values, y_i^o . Therefore, under a model for the joint distribution $\{y_i^o, y_i^m\}$, missing values can be correctly predicted using the observed data. This characteristic of MAR has a significant impact in that sample moments are not unbiased estimates of the the same moments in the target population. Statistical inferences based on these moments without taking MAR into account, such as scatterplots of sample average longitudinal evolutions, may therefore be incorrect. As long as the model for the measurement process y_i is properly specified, likelihood-based analyses on the observed data under MAR can still produce trustworthy inferences even if we ignore the contribution of r_i . This is demonstrated by the fact that the factorization of the likelihood contribution of the full set of data (y_i^o, y_i^m, r_i) for the i -th subject is as follows:

$$\begin{aligned}
L_i(\phi) &= \int p(y_i, r_i; \phi) dy_i^m = \\
&= \int p(y_i^o, y_i^m; \phi_y) p(r_i | y_i^o, y_i^m; \phi_r) dy_i^m \\
&= \int p(y_i^o, y_i^m; \phi_y) p(r_i | y_i^o; \phi_r) dy_i^m \\
&= p(y_i^o; \phi_y) p(r_i | y_i^o; \phi_r) \\
&= L_i(\phi_y) \times L_i(\phi_r)
\end{aligned}$$

Thus, if ϕ_y and ϕ_r are disjoint in the sense that the parameter space of the full vector $\phi = (\phi_y^T, \phi_r^T)^T$ equals to the product of the parameter spaces of ϕ_y and ϕ_r , respectively, inference for ϕ_y can be based on the marginal observed data density $p(y_i^o; \phi_y)$ ignoring the likelihood of the missing process. This property of likelihood-based inferences under MAR is known as *ignorability*.

- **Missing Not at Random (MNAR)** posits that a fraction of the responses we would have observed affects how likely it is that longitudinal responses would be missing. In particular, the distribution of r_i depends on at least some elements of the subvector y_i^m , even if we condition on y_i^o , i.e.,

$$p(r_i | y_i^m; \phi_r) \text{ or } p(r_i | y_i^o, y_i^m; \phi_r).$$

As a consequence, the missing data are non-ignorable, meaning that the observed data alone are not sufficient to infer about the population ([12]). In such cases, we have to conclude that the response and non-response represent not only different but also unique parts of the true data ([13]). Similarly to MAR, MNAR missingness is often called *non-random missingness*, and in the case of dropout *non-random dropout*. An

example of MNAR longitudinal data arises in pain studies in which patients may ask for rescue medication when their pain levels exceed the threshold they can tolerate. In this situation, we do not have a record of their outcome because of dropout. As was also the case in MAR, under MNAR the observed data do not constitute a random sample from the target population. However, differently from MAR mechanism, the predictive distribution of y_i^m conditional on y_i^o is not the same as in the target population, but depends on both the observed y_i^o and on the missing probability distribution mechanism conditional on the observed responses $p(r_i|y_i)$. For this, the model assumed for the missing process is crucial and must be included in the analysis. Missing data assumptions represent our beliefs about how much the observed data may also apply to the data that are absent. We distinguish between models that rely on information that has not been captured in the observed data (i.e., Missing Not At Random; [12]) and those for which the observed data alone would be sufficient for obtaining valid inference (i.e., Missing Completely At Random or Missing At Random). Unfortunately, based only on the observable data, we can never clearly discriminate between MAR and MNAR mechanisms ([13]). Indeed we should be aware that, as explained in [14], "Every missing not at random model has a missing at random counterpart with equal fit".

In this section we focused on the special issue of missingness in longitudinal data, e.g. on a biomarker resulting from the repeated measurements on the same individuals: the longitudinal profile may present a different amount of missing data due to different causes and the modeling approach used in this context is crucial to get appropriate inferences about it. In particular, we distinguished among two types of missingness in the longitudinal data, that is, *monotone* and *non-monotone* missingness.

Mixed-effects models represent a suitable strategy to analyse data affected by missingness, in particular when the nature of the missing data generating process can be considered to be MAR ([15],[16],[17]).

In chapter 6 we will show the results of a simulation work in the context of longitudinal and time-event data, affected by different missing mechanisms; we will inspect this issue in a survival analysis context, indeed, where Extended Cox Models (ECM) and Joint Models (JM, [8]) will be compared in terms of robustness in estimating an *Hazard Ratio* (HR), that is a measure of the association between the longitudinal trajectory of a biomarker and the risk of experiencing a binary event (e.g. death) throughout time. This is why we will focus the attention on mixed-effects model, which is a widely used method in the context of longitudinal data analysis and, as it will be seen, constitutes a building block for JMs.

2.3 A two-stage approach estimation for longitudinal responses

As previously recognized, the structure of a longitudinal dataset prevents it from being analysed using traditional (linear) models. A parametric linear model for longitudinal data (and also for clustered or hierarchical data), the linear mixed model (LMM), which will be explained below, quantifies the associations between a continuous dependent variable and, possibly, several covariates. In this model, two kind of effects are estimated to establish the relationship between independent (continuous or categorical) and dependent covariate:

- fixed effects, that describe the mean structure model between covariates and response variable in the whole sample;
- random effects, that are related to random cluster-specific (e.g. subjects, hospitals, etc.) variations from the overall mean structure.

It is important to note that the second set of effects likewise controls the correlation between repeated measurements on the same subject. The model gets its name from having both fixed and random effects.

The two-stage approach proposed by Verbeke and Molenberghs ([18]) will be employed to characterize the LMM. In the first stage, a vector of a few estimated subject-specific regression coefficients will be used to try to explain the longitudinal response of interest for each subject. In the second stage, another regression model will link the estimates from the first stage to known covariates like treatment, disease classification, patient demographics, and baseline characteristics. The generic formulation of the linear mixed model will be provided by combining the two stages into a single statistical model.

We assume to follow N individuals, each of whom is measured at n_i different time points ($i = 1, \dots, N$), albeit n_i does not always equal n_j for $i \neq j$, with $i, j = 1, \dots, N$. A dependent random variable denoted by the letter \mathcal{Y} , which is assumed to be continuous (different distributions are allowed in the context of Generalized LMMs - GLMMs), is used to model the response of interest. The fixed effects are modelled by random vectors with both continuous and categorical components. The random effects, denoted by the matrix \mathcal{Z} , are defined as those factors that, in addition to a fixed effect, can have an effect that differs from subject to subject.

2.3.1 Stage I

For $i = 1, \dots, N$ and $j = 1, \dots, n_i$ where N is the number of subjects and n_i is the number of repeated measurements for the i -th subject, let Y_{ij} be the random variable denoting the outcome of interest for the i -th subject measured at time j . If so, the vector of continuous responses for the i -th subject is $Y_i = (Y_i, 1, \dots, Y_i, n_i)$. We begin by assuming that Y_i complies with the linear regression model.

$$Y_i = Z_i \gamma_i + \epsilon_i \quad (2.2)$$

where Z_i is a $(n_i \times q)$ matrix whose columns represent the values of i -th subject's q covariates across time. Equation 2.2 in this case models how the response changes over time for the i -th subject.

In light of this, γ_i is the vector of q subject-specific regression coefficients and the random vector of residuals is assumed to follow a multidimensional Gaussian distribution centered in zero with covariance matrix Σ_i of dimension n_i , that is $\epsilon_i \sim \mathcal{N}_{n_i}(0, \Sigma_i)$. It should be noted that multilevel or clustered data could also be used using eq. 2.2. In such cases, the matrix Z contains some variables required to specifically identify repeated measurements within the same cluster (for example, patients in the same ward).

2.3.2 Stage II

The purpose of the second step is to provide an explanation for the between-subjects variability by modeling the relationship between the $\gamma_i = (\gamma_1, \dots, \gamma_q)$ coefficients produced in

the first stage and a collection of known covariates contained in a $(q \times p)$ matrix indicated with K_i . The model is:

$$\gamma_i = K_i\beta + b_i \quad (2.3)$$

where $b_i \sim \mathcal{N}_q(0, D)$ is a q -dimensional residual vector and D its covariance matrix, β is a p -dimensional vector of unknown regression parameters. The estimate of the regression parameters γ_i which can be obtained by sequentially fitting the two models in the two stages, serves as the fundamental component of the model. The examination (second stage) of the summary statistics generated in the first stage, which this sequential fitting can be understood as, entails at least two issues ([18]):

- the estimated vector of effects $\hat{\gamma}_i$ summarizes the information on the longitudinal response Y_i for the subject i , obtained in the first stage, but involves a loss of information;
- in the second stage, the replacement of γ_i with their estimates $\hat{\gamma}_i$ is another source of variability.

In order to overcome these two issues, the Linear Mixed Model -which will be discussed in the following section- combines the two stages into a single model and implements a simultaneous parameter estimation procedure. Despite these two problems, the two-stage estimation is computationally cheap and could be applied in the real world when the Linear Mixed Model has convergence problems.

2.4 The Linear Mixed Model (LMM)

In order to obtain a single model, we can replace γ_i of the second stage in the first stage, yielding:

$$\begin{aligned} Y_i &= Z_i\gamma_i + \epsilon_i \\ &= Z_i(K_i\beta + b_i) + \epsilon_i \\ &= Z_iK_i\beta + Z_ib_i + \epsilon_i \\ &= X_i\beta + Z_ib_i + \epsilon_i \end{aligned}$$

in which the existence of both random subject-specific effects b_i and fixed effects gives rise to the term "linear mixed model". As a result, the longitudinal outcome for each subject can be viewed as a linear regression model with subject-specific deviations from the mean population and population-specific effects (i.e., effects that are common to the entire group of patients). In conclusion, LMM is defined as follows using Laird and Ware's approach ([7]).

Definition 2.4.1. *A linear mixed-effects model is any model which satisfies the following relationship for each subject $i = 1, \dots, N$*

$$\left\{ \begin{array}{l} Y_i = X_i\beta + Z_ib_i + \epsilon_i \\ b_i \sim \mathcal{N}_q(0, D) \\ \epsilon_i \sim \mathcal{N}_{n_i}(0, \Sigma_i) \\ b_1, \dots, b_q; \epsilon_1, \dots, \epsilon_{n_i} \quad \text{independent} \end{array} \right.$$

where, Y_i is the n_i -dimensional response vector for subject i , X_i and Z_i are $(n_i \times p)$ and $(n_i \times q)$ dimensional matrices, β is a p -dimensional vector, b_i is a q -dimensional vector and ϵ_i is a n_i -dimensional vector of residual components. Finally, D is a $(q \times q)$ covariance matrix while the matrix Σ_i is a $(n_i \times n_i)$ covariance matrix. For the sake of simplicity, we can write the following, for each $i = 1, \dots, N$:

- Y_i is the response vector for subject i ,
- X_i and Z_i are matrices whose elements are the known values of covariates for subject i ,

- β contains the fixed effects,
- b_i contains the random effects,
- D models the associations among the random factors in Z ,
- Σ_i is a subject-specific covariance matrix whose dimension depends on the number of repeated visits done by the i -th subject (n_i) and represents the relationship among the residuals for the i -th subject.

The LMM enables the researchers to establish a-priori a structure for the two matrices or to estimate all entries of the covariance matrices for the random effects D and the residuals Σ_i .

The most typical method is to start with the D matrix and put it up as a matrix with only variance components:

$$D = \begin{pmatrix} \sigma_{Z_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{Z_2}^2 & \cdots & 0 \\ 0 & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{Z_q}^2 \end{pmatrix}$$

where $\sigma_{Z_q}^2$ is the p -th random effect's variance. The researcher in this instance posits the independence of the random effects. Another popular option, even if more computationally expensive, is provided by the use of an unstructured covariance matrix, where all of the half matrix's elements are estimated from the data:

$$D = \begin{pmatrix} \sigma_{Z_1}^2 & \sigma_{1,2}^2 & \cdots & \sigma_{1,q}^2 \\ \sigma_{1,2}^2 & \sigma_{Z_2}^2 & & \sigma_{2,q}^2 \\ \vdots & & \ddots & \vdots \\ \sigma_{1,q}^2 & \cdots & \cdots & \sigma_{Z_q}^2 \end{pmatrix}$$

where $\sigma_{a,b}^2$ is the covariance between the a -th and the b -th random effect, with $a \neq b$ and $a, b = 1, \dots, q$. The most widely used statistical softwares include and allow for the definition of a number of other matrices. Because the random effects are formally employed to explain the random variation from one subject to another, or from one cluster to another in the case of a multilevel model, the model given in definition 2.4.1 is known as a subject-specific model. A marginal model can be derived from def. 2.4.1 to examine the relationship between the fixed components and the result. The random components of the equation are not used directly. By modeling a mean trajectory across the whole sample while accounting for the subject variability, it is feasible to determine the marginal influence of a covariate, such as time, on the result. The following definition provides a general outline of a population-averaged model.

Definition 2.4.2. *A population-averaged model is any model which satisfies the following conditions for each subject $i = 1, \dots, N$:*

$$\begin{cases} Y_i = X_i\beta + \epsilon_i^* \\ \epsilon_i^* \sim \mathcal{N}(0, V_i^*) \end{cases}$$

where, X_i is the design matrix with dimensions $(n_i \times p)$, β is the vector of the fixed effects, ϵ_i represents a vector of marginal residuals errors and V_i^* is a n_i -dimensional matrix. Furthermore $V_i^* = Z_i D Z_i^T + \Sigma_i$ where Σ_i , Z_i and D are as in def. 2.4.1.

In synthesis, def. 2.4.2 becomes for each subject:

$$Y_i \sim N_{n_i}(X_i\beta, V_i),$$

with

$$V_i = Z_i D Z_i^T + \Sigma_i$$

and in standard vector notation including all subjects it becomes:

$$Y = X\beta + \epsilon.$$

Mixed models have the advantage of allowing for the estimation of parameters that define how the mean response varies in the population of interest as well as the prediction of how individual response trajectories vary over time. The combined modeling framework for longitudinal and time-to-event data ([8]), which will be covered in chapter 4, uses these models primarily for this reason. Importantly, mixed models can account for any level of data *imbalance*, thus we are not required to collect the same number of measurements for each subject or to collect them at the same times.

Furthermore, the correlation between each subject's repeated measurements is effectively explained by the random effects. Due to the fact that they share the same random effect b_i , the longitudinal responses of the i -th subject will be marginally correlated. Another way to state it is to say that we assume that the longitudinal responses of a subject are independent conditionally on her random effect, i.e.,

$$p(y_i|b_i; \theta) = \prod_{j=1}^{n_i} p(y_{ij}|b_i; \theta) \quad (2.4)$$

2.4.1 Estimation of LMM

The criteria of maximum likelihood (ML) applied on the marginal density derived in def. 2.4.2 for the i -th subject provides the basis for the estimation of the parameters for the linear mixed model:

$$Y_i \sim \mathcal{N}_{n_i}(X_i\beta, V_i) \\ V_i = V_i(\alpha) = Z_i D Z_i^T + \Sigma_i$$

where:

- α is the column vector of all parameters of the covariance matrix (variance components) found in V_i . In α there are at most $\frac{q(q+1)}{2}$ different elements of D and in Σ_i ; the actual number of different elements depends on the choice of the shape of the variance-covariance matrices as explained in section 2.4.
- $\theta = (\beta^T, \alpha^T)^T$ is the column vector of all parameters in the marginal model for Y_i .

The log-likelihood function, accordingly with the maximum likelihood (ML) approach is:

$$l_{ML}(\theta, y) = \sum_{i=1}^N \log p(y_i; \theta) = \sum_{i=1}^N \log p(y_i; \beta, \alpha) \\ = \sum_{i=1}^N \int p(y_i|b_i; \beta, \alpha) p(b_i; \theta_b) db_i \quad (2.5)$$

with $\theta_b = \text{vech}(D)$, and

$$p(y_i; \beta, \alpha) = (2\pi)^{-\frac{n_i}{2}} |V_i(\alpha)|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y_i - X_i\beta)^T V_i^{-1}(\alpha)(y_i - X_i\beta)\right)$$

with $|V|$ denoting the determinant of the square matrix V .

Assuming α known, the ML estimate for β is given by

$$\hat{\beta}_{ML}(\alpha) = \left(\sum_{i=1}^N X_i^T V_i^{-1} X_i \right)^{-1} \sum_{i=1}^N X_i^T V_i^{-1} y_i$$

with $V_i = V_i(\alpha)$. In this instance, even though V_i is unknown but an estimate of its components α is provided, we can estimate by substituting α with $\hat{\alpha}$ and obtaining $\hat{V}_i = \hat{V}_i(\hat{\alpha})$

It is possible to estimate β_{ML} when V_i is unknown by maximizing $l_{ML}(\theta, y)$ with respect to α and replacing with its ML estimate. The drawback is that it estimates β beforehand from the data rather than β and α concurrently. Because it ignores the loss of degrees of freedom brought on by estimating the fixed-effect parameters in β , the ML estimate of α is thus biased downward. To address this problem the REML -Residuals/Restricted Maximum Likelihood- has been developed ([19]). This approach is a particular form of ML estimation that does not base estimates on a ML fit of all the information, but instead uses a likelihood function calculated from a transformed set of data, so there is no effect related to the nuisance parameters. Then, in the case of variance component estimation as this, the likelihood function is generated from the probability distribution of set of contrasts calculated from the data. In a nutshell, the goal of the REML estimation is to determine an estimate for α not relying on β and viceversa. Additionally, REML can generate unbiased estimates of the parameters shaping variance and covariance matrix.

In other words, the idea in REML estimation of V_i is to eliminate β from the likelihood so that it is defined only in terms of V_i . By maximizing the slightly modified log-likelihood function, REML estimation is carried out:

$$l_{REML} = -\frac{1}{2} \log \left| \sum_{i=1}^N X_i^T V_i^{-1} X_i \right| + l_{ML}(\hat{\beta}(\alpha), \alpha; y)$$

with respect to all parameters (α and β) simultaneously .

2.5 Mixed Models on real data: two examples

In this section some examples of analysis on real data I did during the last three years of doctoral work where Mixed-effects models were employed to answer some clinical questions are briefly detailed. They will be shown two instances of analysis of real data collected for the CENTER-TBI project. CENTER-TBI is a large European project that aims to improve the care for patients with Traumatic Brain Injury (TBI). It forms part of the larger global initiative InTBIR: International Initiative for Traumatic Brain Injury Research with projects currently ongoing in Europe, the US and Canada.

To be consistent with the examples, we recall here that we can distinguish among two different settings of applications of these models. Indeed, mixed-effects models, are suitable to be deployed in situations where:

- the data are in longitudinal form, as a result of repeated measurements on the same subjects of some biomarkers of interest. In this framework, for example, mixed models may be used to investigate averaged longitudinal trajectories of biomarkers as the result of some fixed covariates and some latent traits, accounting for the serial correlation among repeated measurements: therefore the random effects are subject-specific. This was the context of the first work presented.
- the data have a hierarchical structure: that is a situation in which observational data of biomarkers, in longitudinal or aggregate form, refer to different subjects from explicit clusters. To clarify, one can think of longitudinal data, resulting from the observation of one or multiple biomarkers over time on a number of patients who are in different centers or hospitals. In this context, one might want to take into account the reasonable correlation that exists due to the repeated measurements on the same subjects and a residual correlation between data from the same cluster, induced, for example, by center-specific practices or different starting conditions among the places where these centers are located (social, cultural..). For this reason, such models are sometimes appended by the name of *multilevel* models or *hierarchical* models ([20]). This was the context of the second work presented.

2.5.1 Management of arterial partial pressure of carbon dioxide in the first week after traumatic brain injury: results from the CENTER-TBI study

The purpose of the study was to describe the management of arterial partial pressure of carbon dioxide (PaCO_2) in severe traumatic brain-injured (TBI) patients and the optimal target of PaCO_2 in patients with high intracranial pressure (ICP). The primary aim was to describe current practice in PaCO_2 management during the first week of intensive care unit (ICU) after TBI, focusing on the lowest PaCO_2 values. We also assessed PaCO_2 management in patients with and without ICP monitoring and with and without intracranial hypertension. We evaluated the effect of profound hyperventilation (defined as $\text{PaCO}_2 < 30$ mmHg) on long-term outcome. We used the median odds ratio (MOR, ([21])) to estimate the between-centers heterogeneity in targeting a PaCO_2 of 35–45 mmHg. MOR was obtained from a longitudinal logistic mixed effects models on daily lowest PaCO_2 adjusted for the IMPACT core covariates ([22]), ICP monitoring, and daily evidence of elevated ICP (at least one $\text{ICP} > 20$ mmHg during the day); and with a hierarchical random intercept effect's structure (i.e., patients within centers), to estimate the between-centers heterogeneity in targeting a PaCO_2 of 35–45 mmHg. We remind the reader to the Appendix A to get a complete vision of the work ([23]).

2.5.2 High arterial oxygen levels and supplemental oxygen administration in traumatic brain injury: insights from CENTER-TBI and OzENTER-TBI

The aim of the study was to evaluate whether the exposure to high blood oxygen levels and high oxygen supplementation was independently associated with outcomes in TBI patients which underwent mechanical ventilation and were admitted to the ICU. This was a secondary analysis of two multicenter, prospective, observational, cohort studies performed in Europe and Australia. In TBI patients admitted to ICU, we describe the arterial partial pressure of oxygen (PaO_2) and the oxygen inspired fraction (FiO_2). We explored the association between high PaO_2 and FiO_2 levels within the first week with clinical outcomes. Furthermore, in the CENTER-TBI cohort, we investigate whether PaO_2 and FiO_2 levels may have differential relationships with outcome in the presence of varying levels of brain injury severity (as quantified by levels of glial fibrillary acidic protein (GFAP) in blood samples obtained within 24 hours of injury. The role of $\text{PaO}_{2,max}$, $\text{FiO}_{2,max}$, $\text{FiO}_{2,max}$, $\text{FiO}_{2,mean}$ or $\Delta\text{PaO}_{2,mean}$, $\Delta\text{FiO}_{2,mean}$ (one at a time) on 6-month mortality and unfavorable neurological outcome was evaluated through mixed-effect logistic regression models, adjusting for the IMPACT core covariates (age, Glasgow Coma Scale (GCS) motor score and pupillary reactivity) and injury severity score (ISS), with the center as a random effect. We remind the reader to the Appendix A to get a complete vision of the work ([24]).

Chapter 3

Survival Analysis

Many different types of outcomes are collected in follow-up studies, including the multiple longitudinal responses we saw in the previous chapter and the time-to-event(s) of particular interest (such as death, relapse, hospitalization, etc.). The duration between a clinically significant starting point, known as the baseline, and the occurrence of an event of interest can be studied using techniques and models which undergo the umbrella of so called survival analysis. Within this framework, the outcome is typically described as failure time, survival time, or event time.

3.1 Characteristics of event time data

The first factor that needs to be considered in the statistical analysis of failure times, which are typically represented with the letter T (continuous or discrete), is the form of their distribution. Event times must be positive and their distributional patterns are frequently skewed. Therefore, statistical techniques that rely on normality are not immediately applicable and may yield erroneous findings when used to survival data. Using an appropriate transformation of the event times, such as the logarithm or square root, this is frequently a simple problem to overcome.

The most crucial feature that sets the study of survival times apart from other statistical fields is *censoring*, that is the event times of interest are not fully observed on all individuals, which is the distinguishing characteristic of censored data. There are two implications: first, as result of their assumption that we have complete information, common statistical methods like the sample average and standard deviation, the t-test, and linear regression provide biased estimates of the distribution of event times and associated variables. Second, inferences, as opposed to complete data, may be more susceptible to an incorrect definition of the survival time distribution.

The censoring might be of different nature ([25]):

- **Administrative censoring:** The censoring that takes place after the study observation period ends is known as administrative censoring. All participants in the study complete it, and at its conclusion, it is known whether they experienced either failure or survival.
- **Lost-to-follow-up:** This type of censoring occurs when an event may have occurred after the last time a person was under observation, but the specific timing of the event is unknown. In contrast to administrative censoring, which coincides with the end of the analytical period and can be placed precisely in time, lost-to-follow-up is a non-event.

Both these types of censoring are known as *right-censoring*.

A second classification of censoring mechanisms concerns whether the probability of a subject being censored depends on the failure process. More precisely, we can distinguish between:

- **Informative Censoring:** when a participant leaves the study for a reason that is directly connected to the expected failure time, such as when her prognosis deteriorates. A censoring mechanism is informative if, at any time t , the failure rates that apply to participants who are still in the study differ from those that apply to individuals who have dropped out of the study, to put it more technically. Informative censoring resembles the MNAR missing data framework in longitudinal studies that was discussed in chapter 2 in certain ways.
- **Non-informative censoring:** when a participant leaves the research for causes unrelated to her prognosis, although it may be influenced by variables. Using longitudinal research as an example, non-informative censoring is comparable to a MCAR missing data context, introduced in chapter 2.

Different inferential techniques should be used depending on the type of censoring mechanism. Because they are the most common, methods that can handle right censored data have received the greatest attention in the literature when it comes to the first classification. When it comes to the second classification and the informative censoring mechanism, unfortunately, there aren't many options available. The issue is that there isn't enough information in the observed data to model the censoring process, which is a problem that also affects longitudinal studies' MNAR missing data mechanisms. Therefore, the chances for a useful analysis in these situations are fairly slim, unless external information is given. The majority of the literature has concentrated on non-informatively censored data approaches as a result of these issues.

3.2 Basic theory of survival analysis

3.2.1 Time-to-event

Let T^* denote the random variable of failure times under study. The function that is primarily used to describe the distribution of T^* is the survival function.

Definition 3.2.1. *Let T be a continuous random variable defined for $t \in [0, +\infty)$ with cumulative distribution function $F(\cdot)$ and probability function $f(\cdot)$. Its survival function is defined as:*

$$S(t) = 1 - F(t) = \mathbb{P}[T > t] = \int_t^{+\infty} f(s) ds$$

with $S(t = 0) = 1$ always, the survival function must be non-increasing as t increases. The hazard function is another important function in a survival analysis. This is defined as the instantaneous risk for the occurrence of an event in the time period $[t, t + dt]$ assuming survival up to t .

Definition 3.2.2. *Let T be a continuous r.v. defined for $t \in [0, +\infty)$ with cumulative distribution function $F(\cdot)$ and probability function $f(\cdot)$. The hazard function is defined as:*

$$h(t) = \lim_{dt \rightarrow 0} \frac{\mathbb{P}[t \leq T < t + dt | T \geq t]}{dt}, \quad t > 0$$

It turns out that that the two functions are linked through the following differential equations, making them mathematically equivalent:

$$h(t) = \frac{f(t)}{S(t)} = -\frac{dS(t)/dt}{S(t)} = -\frac{d}{dt}[\log S(t)]$$

Therefore we can write:

$$S(t) = \exp\left(-\int_0^t h(s) ds\right) = \exp(-H(t))$$

where $H(\cdot)$ is known as the cumulative risk (or cumulative hazard) function that describes the accumulated risk up until time t . Function $H(t)$ can also be interpreted as the expected number of events to be observed by time t .

3.2.2 Censoring

In the preceding section, the generic formulation for the survival, hazard, and cumulative hazard functions was given in relation to a random variable T that models the failure times. The notion of censoring, which is the primary feature that sets survival analysis apart, must be taken into account: indeed, follow-up studies frequently have censored data, that is, subjects who are lost during the period in which the study is conducted. This condition may have several reasons and must be considered when formulating a model to estimate the time to event. Hence, if in section 3.1 this topic was introduced from a descriptive point of view, below it will be discussed from a mathematical point of view.

The main effects of censoring include the inability to apply common techniques like sample average, t-test, and linear regression, as well as the possibility that inferences could be susceptible to incorrect distribution of event timings.

When censoring occurs, the outcome can be thought of as comprising two dimensions: an event indicator and a time at risk. With a little variation from the previous section, for each subject $i = 1, \dots, N$ let T_i^* (and not T_i) denote the random variable of the failure times under study and let C_i be a non-negative variable which models the censoring times, then only the random variable $T_i = \min(T_i^*, C_i)$ is observed due to censoring.

We also get to see the event indicator $\delta_i = \mathbb{I}(T_i^* \leq C_i)$ in addition to T_i : note that, under non-informative censoring, T_i and C_i are independent for each i , as well as T_1, \dots, T_n being independent and identically distributed (i.i.d.), and C_1, \dots, C_n being assumed to be i.i.d.. In general, the goal of a survival analysis is to estimate the features of the distribution of T using just the data T_i and δ_i for each $i = 1, \dots, N$. There are two methods for estimating the survival function:

- by developing an empirical estimate of the survival function, i.e. a non-parametric estimation;
- by specifying a parametric model for $h(t)$ on a particular density function $f(t)$.

If no censoring occurs, an empirical estimator of the survival function is:

$$\hat{S}(t) = \frac{\sum_{i=1}^N \mathbb{I}(T_i > t)}{N} = \frac{\# \text{ of individuals with } T > t}{\text{total sample size}}$$

Under the assumption of i.i.d. sample, it holds:

$$\hat{S}(t) \sim \text{Binomial}(n, S(t))$$

and for large sample sizes, by the central limit theorem:

$$\hat{S}(t) \sim \mathcal{N}\left(S(t), \frac{S(t)(1 - S(t))}{N}\right)$$

Otherwise, $\hat{S}(t)$ is not a reliable indicator of the real $S(t)$ in the presence of censored observations, necessitating the adoption of additional non-parametric approaches, the most common of which is discussed in sec. 3.3.

3.3 Survival times estimations

3.3.1 Non-parametric estimation

The product limit estimator, generally regarded as the most well-known estimator of $S(t)$ when censoring occurs, was developed by Kaplan and Meier (KM) ([26]). This estimator does not make any assumptions about the failure times' underlying distribution because it is non-parametric. This estimator is based on the observed data and the cumulative distribution function $F(\cdot)$ calculation:

$$\hat{F}(t) = \frac{\# \text{ of individuals who experienced the event up } t}{\text{total sample size}}$$

This function, that assumes the name of *empirical distribution function* is a useful and simply way to summarize and display survival data.

Its complementary survival function $\hat{S}(t)$ also called empirical survivor function is:

$$\hat{S}(t) = 1 - \hat{F}(t)$$

The main issue with using these functions is related to the existence of censoring that is not taken into account. The KM estimator instead can take censoring into consideration based on conditional probability. Suppose:

- $t_0 \leq t_1 \leq \dots \leq t_j \leq \dots \leq t_k \leq t < t_{k+1}$ are different failure times in a sample size of N individuals and $t_{k+1} = +\infty$;
- d_j is the number of subjects who experience the event at time t_j , $j = 0, \dots, k$;
- m_j is the number of censored subjects in the interval $[t_j, t_{j+1})$;
- $r_j = (d_j + m_j) + \dots + (d_k + m_k)$ is the number of subject at risk at a time just prior to t_j (risk set).

The probability of failure at t_j given that you are at risk before t_j is:

$$\mathbb{P}[T^* = t_j | T^* > t_{j-1}] = F(t_j) - F(t_j^-) = \frac{d_j}{r_j}$$

and the Kaplan-Meier estimator of the survival probability beyond t is given by:

$$\begin{aligned} \hat{S}_{\text{KM}}(t_k) &= \mathbb{P}[T^* > t_k] \\ &= \mathbb{P}[T^* > t_k \cap T^* > t_{k-1} \cap \dots \cap T^* > t_1 \cap T^* > t_0] \\ &= \mathbb{P}[T^* > t_1] \prod_{j=2}^k \mathbb{P}[T^* > t_j | T^* > t_{j-1}] \\ &= \prod_{j=1}^k \{1 - \mathbb{P}[T^* = t_j | T^* > t_{j-1}]\} \end{aligned}$$

The KM estimator can be written as:

$$\hat{S}_{\text{KM}}(t) = \prod_{t_j < t} (1 - \lambda_j)$$

where $\lambda_j = \frac{d_j}{r_j}$.

It has been proven that the KM estimator, also in presence of censoring, is consistent and asymptotically normal ([27]). Moreover, it has been shown that the KM estimator is also a non-parametric maximum likelihood estimator ([28]). The plot of the KM estimate of the survival function is a step-function in which the estimated survival probabilities are constant between adjacent failure times and decreasing at each failure time.

3.3.2 Parametric estimation

The estimation of the survival experience can also be attained by assuming a parametric form for the distribution of survival times T^* .

We can use distributions such that for $Y \in \mathbb{R}$, by taking into consideration $T^* = e^Y$, so that $Y = \log T^*$ represents the log failure time.

The *exponential* and the *Weibull* distributions are largely used in applications: for this reason we limited the outline about the theory of parametric estimation of survival times to these types of probability distributions, but others are available.

The exponential distribution

The Exponential distribution is used when the hazard function $h(t)$ is *constant*. The instantaneous failure rate is independent of t , so that the conditional chance of failure in a time interval of specified length is the same regardless of how long the individual has been on study; this is referred to as the *memoryless property* of the exponential distribution. It holds:

$$\begin{aligned} h(t) &= \lambda \\ f(t) &= \lambda e^{-\lambda t} \\ S(t) &= \int_0^t f(s) ds = e^{-\lambda t} \\ H(t) &= \int_0^t h(s) ds = \lambda t \end{aligned}$$

The Weibull distribution

Because it allows a power dependence of the hazard on time, the Weibull, a two-parameter (shape and scale) distribution, is an important generalization of the exponential distribution. The hazard function is given by:

$$h(t) = \phi \lambda t^{\lambda-1}; \quad \phi, \lambda > 0$$

therefore, we have that:

$$\begin{aligned} f(t) &= \phi \lambda t^{\lambda-1} e^{-\phi t^\lambda} \\ S(t) &= \int_0^t f(s) ds = e^{-\phi t^\lambda} \\ H(t) &= \int_0^t h(s) ds = \phi \lambda t^\lambda \end{aligned}$$

3.3.3 Likelihood function for censored data

The maximum likelihood technique is frequently used for parameter estimation when the survival function $S(\cdot)$ is believed to have a particular parametric form. Let T^* be a continuous random variable on $[0, +\infty)$ with cumulative distribution function $F(\cdot)$. Assume that $F(\cdot)$ depends on a parameter θ belonging to some sample space. Let C be a censoring random variable with cumulative distribution function $G(\cdot)$. Furthermore, for $i = 1, \dots, N$, assume:

- T_1^*, \dots, T_N^* independent copies of T^* , so that $F(t) = \mathbb{P}[T_i^* \leq t]$
- C_1, \dots, C_N independent copies of C
- $T_i = \min(T_i^*, C_i)$
- $\mathbb{I}(T_i^* \leq C_i)$

Definition 3.3.1. A censoring mechanism is said to be *non-informative* or *random* if

$$\mathbb{P}[T_i^* > t | C_i = t] = \mathbb{P}[T_i^* > t] \quad \forall t > 0, i = 1, \dots, N$$

Now, consider a patient i with complete observation at time t_i :

$$\{T_i = t_i, \delta_i = 1\} = \{T_i^* = t_i, C_i > t_i\}$$

His contribution to the likelihood is given by:

$$\begin{aligned} & \lim_{dt \rightarrow 0} \frac{\mathbb{P}[t_i \leq T^* \leq t_i + dt, C_i > T]}{dt} \\ &= \lim_{dt \rightarrow 0} \frac{\mathbb{P}[t_i \leq T^* \leq t_i + dt] \mathbb{P}[C_i > T]}{dt} \\ &= f(t_i)(1 - G_i(t_i)). \end{aligned}$$

Similarly, if the patient i is censored at t_i , his contribution to the likelihood function is:

$$g(t_i)(1 - F_i(t_i))$$

Therefore, the contribution of the patient i to the likelihood is given by:

$$L_i(\theta) = [f(t_i)(1 - G_i(t_i))]^{\delta_i} [g(t_i)(1 - F_i(t_i))]^{1 - \delta_i}$$

The overall likelihood results to be:

$$\begin{aligned} L(\theta) &= \prod_{i=1}^N L_i(\theta) \\ &= \prod_{i=1}^N f(t_i)^{\delta_i} (1 - G_i(t_i))^{\delta_i} g(t_i)^{1 - \delta_i} (1 - F_i(t_i))^{1 - \delta_i} \\ &= \prod_{i=1}^N f(t_i)^{\delta_i} (1 - F_i(t_i))^{1 - \delta_i} \times k \\ &= \prod_{i=1}^N f(t_i)^{\delta_i} (S_i(t_i))^{1 - \delta_i} \times k \end{aligned}$$

where k is a multiplicative constant that can be ignored because depends on G which does not depend on the parameter θ of interest. Using the relation between $S(\cdot)$ and $h(\cdot)$ (see sec. 3.2) on the log-scale we have that the log-likelihood for the censored data is given by:

$$\log L(\theta) = \sum_{i=1}^N \left(\delta_i \cdot \log h_i(T_i; \theta_i) - \int_0^{T_i} h_i(s; \theta) ds \right).$$

All subjects contribute to the log-likelihood through the cumulative hazard function evaluated at the corresponding observed event time t_i . The subjects who experienced the event additionally contribute an amount equal to the log hazard function evaluated at t_i . Once the log-likelihood has been formulated, iterative optimization procedures (e.g. Newton-Raphson algorithm) can be utilized to give the maximum likelihood estimates $\hat{\theta}$. Inference then proceeds under the classical asymptotic maximum likelihood theory paradigm.

3.4 Failure Times Models

For modeling a population's survival experience, we have seen a variety of survival distributions. However, determining if and how failure time may depend on various explanatory variables is typically of interest. As a result, it becomes important for generalizing models to consider patients' information. We want to model the failure time $T > 0$ taking into account a set of p covariates $\mathbf{X} = (X_{i,1}, \dots, X_{i,p})$ are available for each patient i at baseline ($t = 0$) (qualitative and/or quantitative variables, such as information on treatment, biomarkers, age, and so on). The primary goal is to assess how certain factors affect T , but we also include covariates to take individual variation into account.

3.4.1 Parametric regression models

It is possible to generalize the exponential distribution to create a regression model where the failure rate depends on a collection of covariates X . The hazard function at time t for an individual can be expressed as follows if the failure rate is considered to be constant over time and dependent on the covariates X .

$$h(t|X) = h_0(X)$$

The $h(\cdot)$ function may be modelled through a linear combination $\beta'X$:

$$h(t|X) = h_0 \cdot c(\beta'X)$$

where the vector β is the set of regression parameters that quantifies the effect of each covariate on the hazard, h_0 here is a constant baseline hazard and c is a specified functional form. There are different specific forms for c , and the most used is the form $c(s) = \exp(s)$ for which the hazard function assumes the form:

$$h(t|X) = h_0 \cdot \exp(\beta'X)$$

and the conditional density function of T given $X = x$ becomes:

$$f(t|X) = h_0 \cdot \exp(\beta'X) \cdot \exp[-h_0 \cdot \exp(\beta'X) \cdot t]$$

Analogously, hypothesizing a Weibull distribution for the hazard function, we have that the conditional hazard given x :

$$h(t|X) = \phi \lambda t^{\lambda-1} \cdot \exp(\beta'X)$$

and the conditional density is:

$$f(t|X) = \phi \lambda t^{\lambda-1} \cdot \exp(\beta'X) \cdot \exp[\phi \lambda \cdot \exp(\beta'X) \cdot t^\lambda]$$

Two distinct generalizations are suggested by the forms of the previous regression models. First, Due to the variables' addicting influence on $Y = \log(T)$, both of these models are log-linear, and we can generate a more diverse class of log-linear models known as the *Accelerated Failure Time Models* (AFT). The AFT models, which are commonly employed when it is assumed that the effect of the variables is to accelerate or decelerate the life course of the disease, is not further addressed in this thesis. Second, the hazard function is multiplicatively affected by the covariates, and this relationship suggests the *Relative Risk Model*, often known as the *Cox Model* (see sec. 3.5), which is a more general model that will be discussed in the next sections.

3.5 The Cox Model

Let $h_i(t|X)$ be the hazard function at time t for a subject i whose baseline covariates X were collected. The Relative Risk model ([29]) posits that covariates have a multiplicative effect on the hazard for an event and it is described as:

$$\begin{aligned} h_i(t) &= \lim_{dt \rightarrow 0} \frac{\mathbb{P}[t \leq T^* < t + dt | T^* \geq t, W]}{dt} \\ &= h_0(t) \cdot \exp(\gamma'W_i) \end{aligned} \tag{3.1}$$

where $h_0(t)$ is an arbitrary *unspecified baseline hazard* function and corresponds to the hazard function for a subject for whom $\gamma'W = 0$

The conditional density function of T given X becomes:

$$f(t|W) = h_0(t) \cdot \exp(\gamma'W) \cdot \exp\left(-\exp(\gamma'W) \cdot \int_0^t h_0(s) ds\right)$$

The estimation of all parameters in the model, that are the regression coefficients γ and the parameters in the specification of the $h_0(t)$, proceeds by maximizing the corresponding log-likelihood function. However, Cox ([29]) showed that the estimation of the regression coefficients γ can be based on the Partial Likelihood:

$$L(\gamma) = \prod_{i=1}^n \left[\frac{\exp(\gamma' W_i)}{\sum_{j \in R_i} \exp(\gamma' W_j)} \right]^{\delta_i}$$

where R_i is the risk set at the time just prior to t ; or equivalently expressed as Partial Log-Likelihood:

$$pl(\gamma) = \sum_{i=1}^N \delta_i \cdot \left[\gamma' W_i - \log \left(\sum_{T_j \geq T_i} \exp(\gamma' W_j) \right) \right]$$

that does not require specification of $h_0(\cdot)$, that is, without having to specify the distribution of T_i^* . The relative risk model obtained without a specific baseline function is a *semi-parametric* model because it does not make any assumption for the distribution of the event times, but assumes that the covariates have a multiplicative effect on the hazard rate. Indeed rewriting model (3.1) in the log scale, $\log(h_i(t|W_i)) = \log(h_0(t)) + \gamma_1 w_{i1} + \gamma_2 w_{i2} + \dots + \gamma_p w_{ip}$, we observe that the regression coefficient γ_j , for predictor w_j , denotes the change in the log hazard at any fixed time point t if w_j is increased by one unit while all other predictors are held constant. Analogously, $\exp(\gamma_j)$ denotes the ratio of hazards (HR) for one unit change in w_j at any time t . In general, the HR for a subject i with covariate vector W_i compared to a subject k with covariate vector W_k is:

$$\frac{h_i(t|w_i)}{h_k(t|w_k)} = \exp\{\gamma'(W_i - W_k)\}$$

3.6 Survival analysis with time-dependent covariates

As we have seen, baseline covariates are those that are measured at baseline and which their values remain constant throughout the follow-up period. However, in many clinical studies, the question of whether the covariates (such as biomarkers) change over time are related to the hazard is of interest. Time-dependent covariates are those factors that can change over time and can be either *endogenous* (also known as internal) or *exogenous* (also known as *external*) covariates.

If $\mathcal{Y}_i(t) = \{y_i(s), 0 \leq s < t\}$ denotes the covariate history up to t , we have the following:

Definition 3.6.1. *A covariate is called exogenous if the future path of the covariate up to any time $t > s$ is not affected by the occurrence of an event at time point s ; i.e.:*

$$\mathbb{P}[s \leq T_i < s + ds | T_i^* \geq s, \mathcal{Y}_i(s)] = \mathbb{P}[s \leq T_i^* < s + ds | T_i^* \leq s, \mathcal{Y}_i(t)] \quad (3.2)$$

for all s, t such that $0 < s \leq t$, and $ds \rightarrow 0$. An equivalent and more clear definition is

$$\mathbb{P}[\mathcal{Y}_i(t) | \mathcal{Y}_i(s), T_i^* \geq s] = \mathbb{P}[\mathcal{Y}_i(t) | \mathcal{Y}_i(s), T_i^* = s], \quad s \leq t \quad (3.3)$$

which outlines the concept that $y_i(\cdot)$ is associated with the rate of failures over time, but its future path up to any time $t > s$ is not affected by the occurrence of failure at time s . In particular the value of an exogenous covariate at each time t is known infinitesimally before t , making it a predictable process. Endogenous time-varying covariates, on the other hand, do not meet 3.2 or, equivalently, 3.3. For external covariates, and under condition 3.2 or 3.3, we can directly define the survival function conditional on the covariate path, using its relation to the hazard function, i.e.,

$$S_i(t | \mathcal{Y}_i(t)) = \mathbb{P}[T_i^* > t | \mathcal{Y}_i(t)] = \exp \left\{ - \int_0^t h_i(s | \mathcal{Y}_i(s)) ds \right\} \quad (3.4)$$

Endogenous covariates, on the other hand, typically arise as time-dependent measurements taken on the subjects under study. These include biomarkers and other clinical parameters. The statistical analysis of such covariates is complicated for different reasons. The first crucial feature of endogenous variables is that they frequently depend on the subject's survival in order to exist. As a result, when a subject's death is considered as a failure, the subject's path contains explicit information about the failure time. Indeed, in such a case:

$$S_i(t|\mathcal{Y}_i(t)) = Pr(T_i^* > t|\mathcal{Y}_i(t)) = 1$$

On the other hand, failure of the subject at time s corresponds to the covariate not existing at time s , which directly implies that the endogeneity criterion has been violated. Another feature of endogenous covariates is that they are typically measured with error. Instead of the error caused by the process or equipment that determines the value of this kind of covariate, this measurement error mainly pertains to the biological fluctuation created by the patient herself. Therefore, for such factors, it would be more natural to infer that the reported marker levels are actually a version of the true marker levels that has been tainted by biological variance. Nevertheless, it should be noted that measurement error is not a defining characteristic of endogenous covariates because some external variables (such as air pollution) may also be evaluated inaccurately. Endogenous covariates have one more significant implication: their entire trajectory up to any given time t is not completely seen. Hence, the distinction among endogeneous and exogeneous covariates becomes of practical importance from an inferential point of view, but it pertains more a logical point rather than a practical one, that is, once the information about longitudinal covariate is available it is possible to estimate the extent of association, e.g. through an HR, it has with the risk of an event, regardless of its nature.

3.7 The Extended Cox Model (ECM)

In the relative risk model described in sec. 3.5, we made the assumption that the hazard depends exclusively on covariates, such as baseline age, sex, and randomized treatment, whose values remain constant over time. Nevertheless, in many researches, it may also be interesting to look into whether time-dependent factors are linked to the risk of an event. Environmental elements, biochemical and clinical markers observed during follow-up, and modifications to therapy dose are a few examples of these.

To formally introduce the notation of a time-dependent variable, let $y_i(t)$ denote the covariate vector at time t for subject i , and $\mathcal{Y}_i(t) = \{y_i(s), 0 \leq s < t\}$ denotes the covariate history up to t .

The *counting process* formulation can be used to extend the Cox model presented in sec. 3.5 to handle time-dependent covariates. Here, the precise mathematical formulation of the ECM in terms of counting process is omitted for the sake of brevity, but all the details can be found in [30]. What is important, particularly in view of the next chapters and the simulation work presented in the last chapter, is to understand how the counting process formulation handles time-dependent variables: it is assumed that the longitudinal covariate changes value at the follow-up visits and stay constant in the period between them. Then the observed values of each longitudinal value is directly used in the model: therefore the ECM assumes that this covariate is not affected by random measurement error ([31]). This notion is fundamental and it has to be kept in mind to understand the differences on how ECM and Joint model, that will be discussed in the next chapter, treat the longitudinal covariate. The stepwise extrapolation used by the ECM takes the name of *Last Observation Carried Forward* (LOCF), of which a graphical representation is given in fig. 3.1 as example.

Thereby, the model postulates that the hazard for an event, at any time point t , is associated with the extrapolated value of the covariate at the same time point. The formulation of the model is as follows:

$$h_i(t|\mathcal{Y}_i(t), w_i) = h_0(t) \cdot \exp(\gamma'_i w_i + \alpha y_i(t)) \quad (3.5)$$

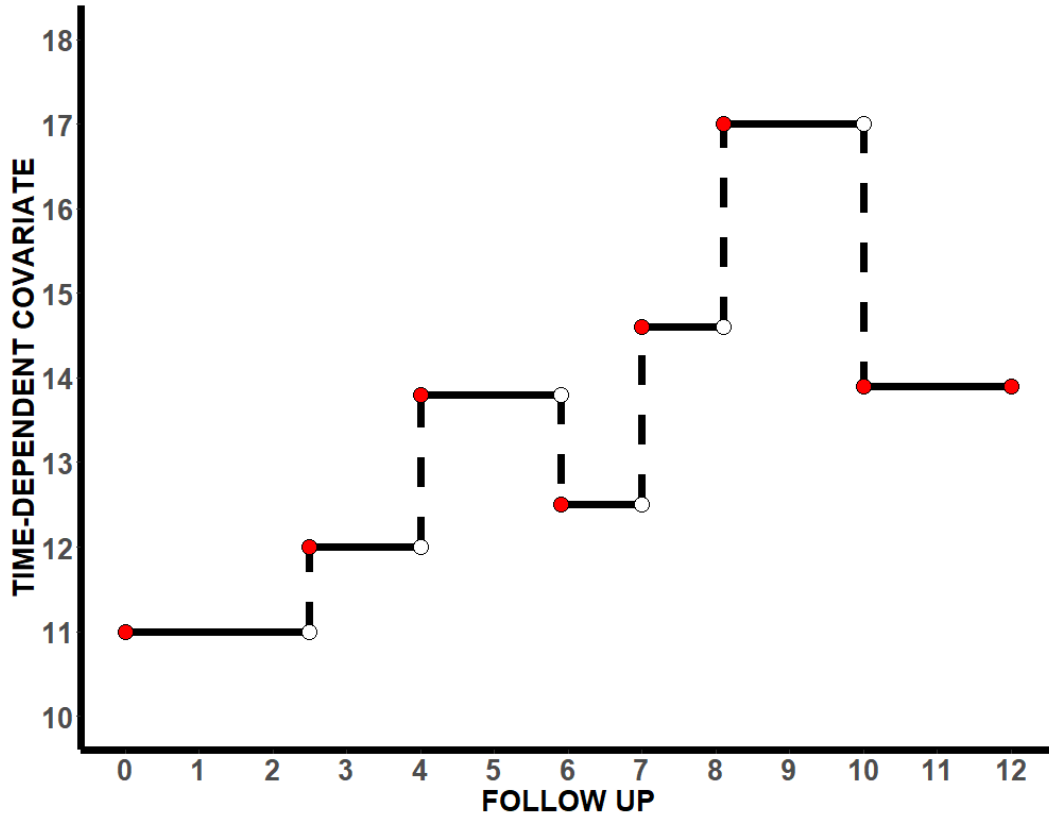


Figure 3.1: Stepwise extrapolation of the time-dependent covariate in the ECM. Red points show the observed biomarker value which remains constant between one measurement occasion and the next.

where w_i denotes a vector of baseline covariates and $y_i(t)$ denotes a time-dependent covariate (possibly multiple covariates).

The regression coefficient α has the exact same meaning as for the vector γ'_i for the covariates w_i , as discussed in 3.5 . For example, if there is only one time-dependent covariate $y_i(t)$, at any given time point t , $\exp(\alpha)$ signifies the relative increase in the risk of an event occurring at that time, due to a one-unit increase in $y_i(t)$ at that same time point.

As a final note we raise the point that the ECM should not be used when the time-dependent factors are endogenous. This is due to the Extended Cox model's assumption that time-dependent variables are deterministic processes that can be predicted, measured accurately, and fully specified throughout their whole course. There are few doubts that this step-function approximation is impractical and unrealistic for several endogenous factors, such as biomarkers. However, we will not explore this theoretical aspect throughout this thesis.

Chapter 4

Joint Models for longitudinal and time-to-event data

This chapter is devoted to the description of the Joint Model (JM) for longitudinal and time-to-event data, which is formulated out of a linear mixed-effects and a relative risk models introduced in the previous chapter. Most of the notation and the flow of the presentation follows from the book by D.Rizopoulos ([8]). This type of modeling belongs to the family of *shared parameters models*.

A Joint Model for longitudinal and time-to-event data allows for the analysis of the association structure between a number of measurable biomarkers gathered over several visits and the amount of time prior to a particular event of interest takes place.

In the survival analysis with time-dependent covariates, seen in chapter 3, the assumption regarding the longitudinal covariate was that it is observed error-free and that its value only varies at each observation time point. Instead of using the observed biomarker values as subject-specific measurements, as if they were the longitudinal response at the observation times, one of the fundamental components of a JM is to use a linear mixed model to model the longitudinal trajectory of a biomarker, assuming that the collected observations of the individual longitudinal profile of the biomarker(s) of interest is actually the result of a true and unobserved profile plus a random measurement error. Hence this a major difference among these two (similar) approaches to analyse longitudinal and time-to-event data. Indeed, the ECM assumes no measurement error on the longitudinal covariates, using in its formulation the observed (extrapolated) value; on the contrary, the JM assumes that the longitudinal profile is affected by measurement error and the time-dependent covariate is instead interpolated from the gathered individual data by using a regression model, in particular a (generalised) mixed-effects model (see chapter 2). Due to this different assumptions, another major point that distinguishes ECM and JM is the fact that the latter allows to make *dynamic predictions*, which is not a feature of the ECM, both on the individual longitudinal trajectories and associated survival experiences, even beyond the follow-up period of the study. This topic will be not addressed within this thesis, but theory and details about it can be found, for example, in [8],[32],[33].

4.1 Joint Model formulation

Let N be the number of subjects and let $\mathcal{D}_N = \{T_i, \delta_i, \mathcal{Y}_i(t); i = 1, \dots, N\}$ denote a sample from the target population, where $T_i = \min(T_i^*, C_i)$ is the observed event time for the i -th subject, with T_i^* being the random variable of the failure times and C_i a non-negative censoring variable. In addition to observing T_i we also get to see the event indicator $\delta_i = \mathbb{I}(T_i^* \leq C_i)$. We focus on the endogenous time-dependent covariate $\mathcal{Y}_i(t) = \{y_i(s), 0 \leq s \leq t\}$ where $\mathcal{Y}_i(t)$ is the vector of n_i observed repeated measurements y_i of a biomarker collected up time t for the i -th subject. We assume that $\mathcal{Y}_i(t)$ is univariate, although the generalisation to the multivariate case is possible in the JM context in the sense that more than one

longitudinal biomarker can be taken into account. The multivariate aspect of the JM will not be covered within this thesis but this extension is pretty straightforward and the details can be found, for example, in [34],[35],[36].

4.1.1 The Survival sub-model

Our goal is to quantify the relationship between the longitudinal marker level and the likelihood of an event. With this purpose, we introduce the term $m_i(t)$, which stands for the true and unobserved longitudinal outcome value at time t . Indeed, $m_i(t)$ differs from $y_i(t)$, the latter being the longitudinal outcome's measurement error-contaminated value at time t . In order to measure the strength of the relationship between $m_i(t)$ and the risk for an event, a relative risk model is used:

$$\begin{aligned} h_i(t|\mathcal{M}_i(t), w_i) &= \lim_{dt \rightarrow 0} \frac{\mathbb{P}[t \leq T_i^* \leq t + dt | T_i^*, \mathcal{M}_i(t), w_i]}{dt} \\ &= h_0(t) \cdot \exp\{\gamma' w_i + \alpha m_i(t)\} \end{aligned} \quad (4.1)$$

where $h_0(\cdot)$ denotes the baseline risk function, $\mathcal{M}_i(t) = \{m_i(s), 0 \leq s < t\}$ denotes the whole true unobserved longitudinal process up to time point t , and w_i is a vector of baseline covariates with a corresponding vector of regression coefficients γ . The α parameter quantifies the effect of the underlying longitudinal response to the risk for an event and $\exp(\alpha)$ should be interpreted as relative increase in the risk for an event at time t resulting from a one unit increase in $m_i(t)$ at the same time point. Note that specifying a model as in 4.1 implies that the risk for an event at time t depends only on the current value of the time-dependent marker $m_i(t)$.

This is not true for the survival function, though. Specifically, we find that using the established relationship between the survival function and the cumulative hazard function (see chapter 3), the corresponding survival function depends on the whole covariate history $\mathcal{M}_i(t)$:

$$\begin{aligned} S_i(t|\mathcal{M}_i(t), w_i) &= \mathbb{P}[T_i^* > t | \mathcal{M}_i(t), w_i] \\ &= \exp\left(-\int_0^t h_0(s) \exp\{\gamma' w_i + \alpha m_i(s)\} ds\right). \end{aligned} \quad (4.2)$$

Unlike the Cox model, in order to complete the specification of 4.1, the baseline risk function $h_0(\cdot)$ has to be specified for JMs. Indeed, as we saw in chapter 3, in order to avoid the effects of specify incorrectly the distribution of survival times, it is common to leave $h_0(\cdot)$ fully undefined in traditional survival analysis. However, it turns out that taking this approach may result in an underestimation of the standard errors of the parameter estimates within the joint modeling framework ([37]), enforcing the necessity to explicitly declare $h_0(\cdot)$ to prevent such issues. One possible, common choice for the baseline hazard function is to relating to a parametric distribution, as Weibull, log-normal, exponential and Gamma distributions, frequently utilized in the context of classic survival analysis. As an alternative, and perhaps better, JM allows to specify a parametric baseline risk function while keeping it flexible. In the literature, a number of methods have been suggested for modeling the baseline risk function flexibly. Two simple options that often work quite satisfactorily in practice are the piecewise-constant and splines approximations ([9]). Under the piecewise-constant approximation, the baseline risk function takes the form:

$$h_0(t) = \sum_{q=1}^Q \psi_q \mathbb{1}(\nu_{q-1} < t \leq \nu_q) \quad (4.3)$$

where $0 = \nu_0 < \nu_1 < \dots < \nu_Q$ denotes a split of the time scale, with ν_Q being larger than the largest observed time, and ψ_q denotes the value of the hazard in the interval $(\nu_{q-1}, \nu_q]$. As the number of knots increases the specification of the baseline hazard becomes more flexible. In the limiting case where each interval $(\nu_{q-1}, \nu_q]$ contains only a single true event

time (assuming no ties), this model is equivalent to leaving $h_0(\cdot)$ completely unspecified and estimating it using non-parametric maximum likelihood. For the regression splines model the log baseline risk function $\log h_0(t)$ is expanded into B-spline basis functions for cubic splines as follows:

$$\log h_0(t) = r_0 + \sum_{d=1}^m r_d B_d(t, q) \quad (4.4)$$

where $r^T = (r_0, r_1, \dots, r_m)$ are the spline coefficients, q denotes the degree of the B-splines basis functions $B(\cdot)$, and $m = l+q-1$, with l denoting the number of interior knots. Similarly to the piecewise-constant model, increasing the number of knots increases the flexibility in approximating $h_0(\cdot)$. However in both approaches, we should keep a balance between bias and variance and avoid overfitting. After the number of knots has been decided, their location is typically based on percentiles of either the observed event times $T_i = \min(T_i^*, C_i)$ or only the true event times $\{T_i : T_i^* \leq C_i, i = 1, \dots, n\}$, to allow for more flexibility in the region of greatest density.

4.1.2 The Longitudinal sub-model

We used $m_i(t)$ to signify the true value of the underlying longitudinal covariate at time point t in the definitions of the survival models (4.1). However, as was previously already indicated, longitudinal data is actually acquired irregularly and prone to error for each participant. Therefore, we have to estimate $m_i(t)$ and effectively reconstruct the entire longitudinal history $\mathcal{M}_i(t)$ for each participant in order to quantify the impact of the longitudinal covariate on the probability for an event. To this end, a mixed-effects model suitable for representing subject-specific temporal evolutions is postulated. Assuming that longitudinal outcomes are normally distributed, we have:

$$\begin{cases} y_i(t) = m_i(t) + \epsilon_i(t) \\ m_i(t) = x_i^T(t)\beta + z_i^T(t)b_i \\ b_i \sim \mathcal{N}(0, D), \quad \epsilon_i(t) \sim \mathcal{N}(0, \sigma^2) \end{cases} \quad (4.5)$$

where the design vectors $x_i(t)$ for the fixed effects β and $z_i(t)$ for the random effects b_i , as well as the error terms $\epsilon_i(t)$, are possibly time-dependent. Moreover, it is assumed that the error terms are mutually independent, independent of the random effects, and normally distributed with mean zero and variance σ^2

By assuming that the observed level of the longitudinal outcome, $y_i(t)$, is equal to the true level, $m_i(t)$, plus a random error term, the mixed model solves the measurement error problem. Additionally, by using subject-specific random effects and the temporal structure in the definitions of $x_i(t)$ and $z_i(t)$, it is possible to reconstruct the whole path determined by each subject's time-dependent process $\mathcal{M}_i(t)$.

4.2 Joint Models estimation

Maximum likelihood (semi-parametric) estimation is the main technique that has been suggested for JMs ([37],[38],[39]). Here, we provide the fundamentals of the maximum likelihood estimation method for JMs.

We will assume that the vector of time-independent random effects jointly underlies both the longitudinal and survival processes in order to establish this joint distribution: for that reason JMs go by the name of *shared parameter models*. This implies that the link between the longitudinal and event outcomes as well as the correlation between the repeated measurements in the longitudinal process are both explained by these random effects (conditional independence). Formally we have:

$$p(T_i, \delta_i, y_i | b_i; \phi) = p(T_i, \delta_i | b_i; \phi) p(y_i | b_i; \phi) \quad (4.6)$$

$$p(y_i | b_i; \phi) = \prod_j p\{y_i(t_{ij}) | b_i; \phi\} \quad (4.7)$$

where $\phi = (\phi_t^T, \phi_y^T, \phi_b^T)^T$ denotes the full parameter vector, with ϕ_t denoting the parameters for the event time outcome, ϕ_y the parameters for the longitudinal outcomes and ϕ_b the unique parameters of the random-effects covariance matrix, and y_i is the $n_i \times 1$ vector of longitudinal responses of the i -th subject. In addition, it is assumed that, given the observed history, the censoring mechanism and the visiting process are independent of the true event times and future longitudinal measurements. These assumptions imply that there is no additional dependence on underlying, latent, subject characteristics associated with the prognosis. Instead, decisions on whether a subject withdraws from the study or visits the clinic for a longitudinal measurement depend on the observed past history (longitudinal measurements and baseline covariates). When one of the two processes depends on the random effects, these presumptions are broken. This is due to the fact that such a dependence implies a dependence on upcoming longitudinal measures. Since the observed data do not provide enough information to suggest differently, evaluating the plausibility of the non-informativeness for the visiting and censoring processes typically requires external information from subject-matter experts.

These assumptions allow for the formulation of the log-likelihood contribution for the i -th subject as follows:

$$\log p(T_i, \delta_i, y_i; \phi) = \log \int p(T_i, \delta_i, y_i, b_i; \phi) db_i \quad (4.8)$$

$$= \log \int p(T_i, \delta_i | b_i; \phi_t, \beta) \left[\prod_j p\{y_i(t_{ij}) | b_i; \phi_y\} \right] p(b_i; \phi_b) db_i \quad (4.9)$$

with the conditional density for the survival part $p(T_i, \delta_i | b_i; \phi_t, \beta)$ having the form:

$$p(T_i, \delta_i | b_i; \phi_t, \beta) = h_i(T_i | \mathcal{M}_i(T_i); \phi_t, \beta)^{\delta_i} S_i(T_i | \mathcal{M}_i(T_i); \phi_t, \beta) \quad (4.10)$$

$$= [h_0(T_i) \exp\{\gamma^T w_i + \alpha m_i(T_i)\}]^{\delta_i} \times \exp\left(-\int_0^{T_i} h_0(s) \exp\{\gamma^T w_i + \alpha m_i(s)\} ds\right) \quad (4.11)$$

where $h_0(\cdot)$ can be any positive function of time, such as the piecewise-constant hazard as in 4.3, or the B-spline approximation given by 4.4 or the hazard function of any known distribution, and the survival function is given by 4.2.

The joint density for the longitudinal responses and the random effects is given by:

$$p(y_i | b_i; \phi) p(b_i; \phi) = \prod_j p\{y_i(t_{ij}) | b_i; \phi_y\} p(b_i; \phi_b) \quad (4.12)$$

$$= (2\pi\sigma^2)^{-n_i/2} \exp\{-\|y_i - X_i\beta - Z_i b_i\|^2 / 2\sigma^2\} \times (2\pi)^{-q_b/2} \det(D)^{-1/2} \exp(-b_i^T D^{-1} b_i / 2)$$

where q_b denotes the dimensionality of the random-effects vector, and $\|x\|$ denotes the Euclidean vector norm. Maximization of the log-likelihood function with respect to θ is a computationally challenging task. This is mainly because both the integral with respect to the random effects, and the integral in the definition of the survival function do not have a closed-form analytical solution, except in rare cases. Standard numerical integration techniques such as Gaussian quadrature and Monte Carlo have been successfully applied in the joint modeling framework ([38],[39]). However, computational complexity increases with the dimension of the random-effects. For this reason, a pseudo-adaptive Gauss-Hermite quadrature rule which is considerably faster than the standard Gauss-Hermite rule was proposed and the mathematical details of this methodology can be found in [40].

4.2.1 Random Effects estimation

Up to this point, our main attention has been on ϕ parameter estimation and inference for the joint model. The relationship between the longitudinal and event time processes was built by using the random effects b_i as a way to characterize heterogeneity in the patient longitudinal evolutions. However, determining patient-specific forecasts for either of the outcomes may be of importance in many circumstances and an estimation of the subject-specific random

effects vector b_i is needed to obtain such predictions. It is natural to estimate the random effects using the Bayesian paradigm because it is assumed that they are random variables. In particular, assuming that $p(b_i; \phi)$ is the prior distribution, and that $p(T_i, \delta_i | b_i; \phi) p(y_i | b_i; \phi)$ is the conditional likelihood part, we can derive the corresponding posterior distribution as:

$$p(b_i | T_i, \delta_i, y_i; \phi) = \frac{p(T_i, \delta_i | b_i; \phi) p(y_i | b_i; \phi) p(b_i; \phi)}{p(T_i, \delta_i, y_i; \phi)} \quad (4.13)$$

$$\propto p(T_i, \delta_i | b_i; \phi) p(y_i | b_i; \phi) p(b_i; \phi) \quad (4.14)$$

This posterior distribution in JMs does not have a closed-form solution and must be computed numerically, in contrast to the linear mixed models framework where it is a multivariate normal distribution. However, as the number of longitudinal measurements n_i increases, this distribution will converge to a normal distribution.

To describe this posterior distribution, standard summary measures are often utilized. For its location the *mean* or the *mode* are typically used, defined here as:

$$\left\{ \begin{array}{l} \bar{b}_i = \int b_i p(b_i | T_i, \delta_i, y_i; \phi) db_i, \\ \hat{b}_i = \operatorname{argmax}_b \{ \log p(b | T_i, \delta_i, y_i; \phi) \} \end{array} \right\} \quad (4.15)$$

respectively, and as a measure of dispersion we may use the posterior variance or the inverse Hessian matrix of the random effects, that is:

$$\left\{ \begin{array}{l} \operatorname{var}(b_i) = \int (b_i - \bar{b}_i)^2 p(b_i | T_i, \delta_i, y_i; \phi) db_i \\ H_i = \left\{ - \frac{\partial^2 \log p(b | T_i, \delta_i, y_i; \phi)}{\partial b^T \partial b} \Big|_{b=\hat{b}_i} \right\}^{-1} \end{array} \right\} \quad (4.16)$$

4.3 Asymptotic Inference

Having fitted the joint model under a maximum likelihood framework, the standard asymptotic likelihood inference tests are appropriate. If, in general, we are interested in testing the null hypothesis:

$$\left\{ \begin{array}{l} H_0 : \phi = \phi_0 \\ H_1 : \phi \neq \phi_0 \end{array} \right\} \quad (4.17)$$

we are allowed to use:

- **Likelihood Ratio Test**

$$LRT = -2 \ln \left[\frac{L_{ML}(\hat{\phi}_0)}{L_{ML}(\phi)} \right] \rightarrow \chi_{df}^2$$

- **Score Test**

$$U = U_\phi(\hat{\phi}_0)' \{I(\hat{\phi}_0)\}^{-1} U_\phi(\hat{\phi}_0) \rightarrow \chi_{df}^2$$

where $U_\phi(\hat{\phi}_0)$ is the score function computed from the JM log-likelihood and $I(\hat{\phi}_0)$ is the observed information matrix (i.e., the negative of the inverse Hessian matrix)

- **Wald Test**

$$W = (\hat{\phi} - \phi_0)' I(\hat{\phi})(\hat{\phi} - \phi_0)$$

Under the null hypothesis, the asymptotic distribution of each of these tests is a χ^2 distribution on p degrees of freedom, with p denoting the number of parameters being tested. For a single parameter ϕ_j the Wald test is equivalent to $(\hat{\phi}_j - \phi_{0j}) / \operatorname{std.err.}(\hat{\phi}_j)$, which under the null hypothesis follows an asymptotic standard normal distribution. These test statistics are approximately low-order Taylor series expansion of each other and they are asymptotically equivalent.

Because standard errors do not account for the variability introduced by estimating the variance components, i.e., the covariance matrix for the random effects, they underestimate the true variability in the Wald test for testing the fixed effects in the classical linear mixed

model ([41]). For this reason, instead of using the conventional χ^2 distribution that the Wald test requires, typically an approximate F distribution with adequate degrees of freedom is employed. This issue might be exaggerated in JMs, since we neglect to estimate the survival process in addition to estimating the variance components. Although there has not been much research in the joint modeling literature to look into the Wald statistic's properties in finite samples, asymptotically, we expect it to follow the alleged χ^2 distribution. Therefore, although likelihood ratio tests need more processing resources, it is often better to choose them.

Only two nested models can be compared using the three standard tests we have seen so far, because the model under the null hypothesis is an exception to the model under the alternative. Information criteria are often applied when comparing non-nested models is of interest. The fundamental goal of these criteria is to compare two models on the basis of their maximal log-likelihood values while penalizing overparameterization. The Bayesian Information Criterion ([42]) and the Akaike's Information Criterion ([43]) are the two most widely used information criteria.

4.4 Different Joint Model parametrizations

The typical joint model, which was described before, makes the assumption that the true level of the longitudinal marker at a given time point t , determines the risk for an event at that time point. The α parameter captures the magnitude of the association between the risk and the present level of the marker. Despite the fact that this parameterization has a straightforward meaning for α , it is unrealistic to anticipate that it will always be the best choice for representing the proper link between the two processes. In this section we present several alternative parameterizations that extend the standard parameterization 4.1 in different ways. These different parameterizations can be seen as special cases of the following general formulation of the association structure between the longitudinal marker and the risk for an event:

$$h_i(t) = h_0(t) \cdot \exp[\gamma' w_{i1} + f\{m_i(t - c), b_i, w_{i2}; \alpha\}] \quad (4.18)$$

where $f(\cdot)$ is a function of the true level of the marker $m_i(\cdot)$, of the random effects b_i and extra covariates w_{i2} .

- *Interaction Effects*: the parameterization 4.1.1 assumes that the effect of the true level of the marker is the same in all subgroups of the target population, but this is evidently a strong assumption that may not be true when the marker behaves differently for different subgroups of subjects. Including interaction terms between the marker and the relevant baseline covariates in the linear predictor of the relative risk model is a natural extension to handle such situations, i.e.:

$$h_i(t) = h_0(t) \cdot \exp[\gamma' w_{i1} + \alpha^T \{w_{i2} \times m_i(t)\}]$$

- *Lagged Effects*: in some circumstances there is the need to consider lagged effects, that is when the association between the relative risk model and the biomarkers is better captured by considering a time past observation. In these cases the following formulation of the relative risk model is used:

$$h_i(t) = h_0(t) \cdot \exp[\gamma' w_{i1} + \alpha m_i\{\max(t - c, 0)\}]$$

which postulates that the risk at time t depends on the true value of the longitudinal marker at time $t - c$, where c specifies the time lag of interest.

- *Time-Dependent Slopes Parameterization*: since for each patient the marker follows a trajectory in time, it is also reasonable to consider parameterizations that allow the risk for an event to also depend on other features of this trajectory. For these situations we can consider a JM in which the risk depends on both the current true value of the

trajectory and the slope of the true trajectory at time t (or just the slope). Therefore we will consider the first derivative of the longitudinal sub-model:

$$m_i^*(t) = \frac{d}{dt}m_i(t) = \frac{d}{dt}\{x_i^T(t)\beta + z_i^T(t)b_i\}$$

or a cumulative effect of the biomarker by considering the integral of the longitudinal sub-model, that is:

$$m_i^*(t) = \int_0^t m_i(s)ds = \int_0^t x_i^T(s)\beta + z_i^T(s)b_i ds$$

where for any particular time point t , α measures the strength of the association between the risk for an event at time point t and the rate of change of the biomarker values and the area under the longitudinal trajectory up to the same time t , with the area under the longitudinal trajectory regarded as a suitable summary of the whole trajectory, respectively.

- *Spline based parametrization*: to allow to capture non-linear effects of the biomarker trajectory, a spline approximation can be used. In this case, a general formulation of the longitudinal sub-model is as follows:

$$m_i(t) = (\beta_0 + b_{i0}) + (\beta_k + b_{ik})^T B(t, 4, 4)$$

where $B(t, df, q)$ denotes a B-spline basis matrix for $q - 1$ degree splines with $df - q + 1$ internal knots placed at the corresponding percentiles of the follow-up times, and β_k and b_{ik} denote the vectors of fixed and random effects corresponding to the B-splines matrix.

4.5 Connection with the Missing Data Framework

In the second chapter of this thesis, we addressed the issue of missing data regarding individual profiles of a time-dependent variable. Indeed, we pointed out how longitudinal data are often affected by this issue and also provided, from a practical point of view, how to deal with this situation. To this end, the mixed-effects model is an appropriate approach for analyzing time evolutions that may have missing data. This approach is valid both in contexts where the missing process is assumed to be MCAR or MAR.

So, when the longitudinal outcome is the main concern, the patient's longitudinal process often ends when a specific event occurs. This is due to the fact that either follow-up measurements are unable to be taken or their distribution changed after the event and is now deemed irrelevant. In particular, it may not be conceptually appropriate to evaluate the values of the longitudinal outcome after the event time in some clinical studies, where, for example, death is the terminating event. Although we designed a mixed model for the observed longitudinal responses, it is important to keep in mind that the JM implicitly includes assumptions for the 'full' longitudinal response vector, including observations that would have been made after the event or censoring ([8]).

To better illustrate this, let us define for each participant the observed and missing parts of the longitudinal response vector in order. The observed part comprises of all longitudinal measures of the i -th subject that have been observed before the event time, i.e. $y_i^o = \{y_i(t_{ij}) : t_{ij} < T_i^*, j = 1, \dots, n_i\}$. The longitudinal measurements that would have been taken up until the end of the research, had the event not occurred, form the missing part, that is the set $y_i^m = \{y_i(t_{ij}) : t_{ij} \geq T_i^*, j = 1, \dots, n_i^+\}$. The dropout mechanism, which is the conditional distribution of the time-to-dropout given the whole vector of longitudinal

responses (y_i^o, y_i^m) , can be derived from these definitions:

$$p(T_i^* | y_i^o, y_i^m; \phi) = \int p(T_i^*, b_i | y_i^o, y_i^m; \phi) db_i = \quad (4.19)$$

$$= \int p(T_i^* | b_i, y_i^o, y_i^m; \phi) p(b_i | y_i^o, y_i^m; \phi) db_i = \quad (4.20)$$

$$= \int p(T_i^* | b_i; \phi) p(b_i | y_i^o, y_i^m; \phi) db_i, \quad (4.21)$$

where the simplification in the last line is due to the conditional independence assumption 4.6.

The posterior distribution of the random effects $p(b_i | y_i^o, y_i^m; \phi)$, reveals that the time-to-dropout is dependent on y_i^m , indicating that joint models are consistent with a MNAR missing data mechanism. A deeper look at 4.19 reveals that the random effects b_i is the main driver of the process in joint models. Indeed, and as already implicitly shown, under the JM, the survival and longitudinal sub-models share the same random effects. Due to this feature, joint models belong to the class of the so called *shared parameter models* ([44], [45]). This missing data mechanism means that patients who exhibit steep rises in their longitudinal profiles may be more (or less) likely to drop out of the study under an ordinary random-effects structure (i.e., random intercepts and random slopes).

The relationship between the association parameter and the kind of the missing data mechanism is an important consideration in this case. Because the dropout process is independent of both missing and observed longitudinal responses, once available covariates are taken into account, a null value for the link parameter α specifically corresponds to a MCAR missing data mechanism ([46]). Additionally, since the parameters in the two sub-models are distinct when $\alpha = 0$, it is possible to factorize the joint probability of the dropout and longitudinal processes as follows, allowing the parameters estimation of the two sub-models independently:

$$\begin{aligned} p(T_i, \delta_i, y_i; \phi) &= p(T_i, \delta_i; \phi_t) p(y_i; \phi_y, \phi_b) \\ &= p(T_i, \delta_i; \phi_t) \int p(y_i | b_i; \phi_y) p(b_i; \phi_b) db_i, \end{aligned} \quad (4.22)$$

The estimated parameters, however, derived from maximizing the log-likelihood of the longitudinal process $l(\phi_y) = \sum_i \log p(y_i; \phi_y)$, remain valid under a MAR missing data mechanism, that is, under the assumption that dropout depends only on the observed responses ([8]). Therefore, even though $\alpha = 0$ corresponds to a MCAR mechanism technically speaking, the parameter estimates we will get will still be valid under MAR. As an aside, we should also point out that, in practice, a subject leaving the study due to censoring also results in the cessation of data collection for the longitudinal process. However, we made the assumption that the censoring mechanism may be dependent on the observed past of longitudinal responses and covariates, but is independent of future longitudinal outcomes when formulating the likelihood function of joint models. In light of this, censoring is equivalent to a MAR missing data mechanism.

The shared parameter models structure has the additional benefit of making it very simple to handle intermittent missingness and attrition (see chapter 2). To demonstrate how this is done, we write the observed data log-likelihood for the longitudinal outcome under the

complete data model y_i^o, y_i^m as follows:

$$l(\phi) = \sum_{i=1}^n \log \int p(T_i, \delta_i, y_i^o, y_i^m; \phi) dy_i^m \quad (4.23)$$

$$= \sum_{i=1}^n \log \int \int p(T_i, \delta_i, y_i^o, y_i^m | b_i; \phi) p(b_i; \phi) dy_i^m db_i \quad (4.24)$$

$$= \sum_{i=1}^n \log \int p(T_i, \delta_i | b_i; \phi) \left\{ \int p(y_i^o, y_i^m | b_i; \phi) dy_i^m \right\} p(b_i; \phi) db_i \quad (4.25)$$

$$= \sum_{i=1}^n \log \int p(T_i, \delta_i | b_i; \phi) p(y_i^o | b_i; \phi) p(b_i; \phi) db_i \quad (4.26)$$

The missing longitudinal responses y_i^m are solely involved in the density of the longitudinal sub-model, according to the first conditional independence assumption 4.6. Additionally, the longitudinal responses depending on the random effects are independent with each other under the second conditional independence assumption 4.7, making it simple to drop the integral with regard to y_i^m . Therefore, even if some subjects have irregularly missing data, it is simple to determine the likelihood of a joint model without needing to integrate with respect to the missing responses. A practical problem in the handling of missing data in longitudinal outcomes is the fact that the observed data alone cannot distinguish between a MAR and a MNAR dropout mechanism ([47]).

4.6 Final considerations on Joint modeling of longitudinal and time-to-event data

JMs are frequently used in researches where determining the 'real' trajectory of the longitudinal variable is the main focus and, as already explained, where patients leaving the study can result in missing values in longitudinal studies. The linear mixed model can be deployed to investigations when these values are MCAR or MAR to get unbiased estimates. Ignoring the missing process and analyzing the longitudinal evaluations using, e.g., a mixed effects model, can lead to biased inferences regarding group differences in the average longitudinal evolutions: thus, joint modeling of the measurement and missing processes is required to account for informative non-response ([48]). Dropout is considered "informative" and is sometimes referred to as informative censoring when the rate of dropout is correlated with the level of the longitudinal variable, and this kind of missing data is also known as non-ignorable missingness or "not missing at random." Failure to correctly handle the type of missing process may lead to biased estimate ([8]) and JMs can address this issue by using the survival component to inform the longitudinal process of missingness. Shared parameter models, e.g. JMs, account for the dependence between the measurement and missing process by means of latent variables (not observed) as random effect ([1],[44]) Hence, JMs can be applied to correct the longitudinal trajectory for dropout bringing both the survival sub-model (modeling the survival outcome, that is the time-to-event data) and the mixed model (modeling the longitudinal outcome): thus the random effects in the linear mixed model are included (shared) in the survival model, in order to capture the relationship between the longitudinal and the survival outcome on an individual level ([49]). This relationship works in both directions: the survival outcome is used to correct the longitudinal trajectory on dropout due to the event and the longitudinal process is linked to the survival experience, allowing to associate the longitudinal information with the event under study. Also, we underline the concept that the other reason to deploy a JM concerns the need to account and correct for the *random measurement error* in the longitudinal expressions of some personal process going on ([50]), due to the endogenous nature of the dynamic observations measured throughout the time on an individual level. Finally we remind that a third reason to adopt a JM approach, especially instead of the ECM, is the possibility to make dynamic predictions of both individual longitudinal profiles and survival experiences, which were not discussed as being not a major concern of this thesis.

Chapter 5

The ORANGE study

Throughout this chapter, the ORANGE study will be presented. Part of this material is published in *The Lancet Neurology* in a paper for which my contribute was in the statistical analyses, writing and review ([51]). The objective of the study was to investigate the strength of the association between the Neurological Pupil Index (NPi) and outcomes at 6 months. A further evaluation at shorter follow-up time (ICU) was also carried out in this thesis, employing both ECM and JM, previously introduced from a theoretical point of view (chapters 3 and 4). This in order to outline, from a practical point of view, what information can be gathered by looking at these two approaches' outputs.

5.1 Background

The management of ICU patients following brain injury urges the usage of non-invasive procedures, enhancing the odds of their survival during ICU staying and after discharge. Those are needed to provide the neurointensivists a bedside quantitative tool to predict early prognostication in patients with Acute Brain Injury (ABI), without the necessity of adopting procedures that may also aggravate their condition. In this context, the NPi evaluation, which will be detailed in the next section, has emerged as valuable bedside monitoring assessment to get a fast, non-invasive, quantifiable and standardized neurological examination of pupillary function, which may allow rapid diagnosis of intracranial pathology that affects clinical decision making ([52]). The investigation of the value of the NPi in predicting the outcome of high-risk (due to potential worsening of the intra-cerebral lesions) critically ill, severely brain-injured patients ([53]), and its probable association with a corresponding increase of Intracranial blood pressure (ICP) is a topic of great interest in the Neurocritical care field ([54],[55]), being a non-invasive monitoring procedure. Preliminary single-center data recently demonstrated that abnormal NPi is associated with worse outcome in patients with traumatic ([54]), and haemorrhagic ABI ([56],[57]). Emerging data from single-center retrospective cohort studies in patients with severe non-anoxic acute brain injury suggest the potential prognostic value of NPi ([53],[58]). Three-month non-survivors had significantly lower NPi than did survivors at all time points in an observational cohort of sedated, mechanically ventilated VA-ECMO patients during the early phase after ECMO insertion (first 72 hours) [59] and abnormal values of NPi ($\text{NPi} < 3$) was 100% specific for 90-day mortality ([53],[59]). In patients suffering hypoxic-ischemic brain injury after cardiac arrest, where the most likely cause of abnormal NPi is direct irreversible brainstem anoxia, low NPi (defined as ≤ 2) was associated with poor neurological outcomes ([60],[61],[62]). In this setting, one single pathological NPi measurement at least 24 hours after the initial injury was highly predictive of poor outcomes. Conversely, following non-anoxic ABI, brainstem injury may more often be secondary to brainstem compression or herniation, and therefore more susceptible to be reversed by therapeutic interventions, such as osmotherapy, with subsequent improvement in NPi values ([63],[54]). Indeed, about two-third of patients who had an NPi of 0 during their stay had values > 3 , which reinforces the importance of integrating repeated NPi measurements when assessing coma prognostication after ABI. In

patients with traumatic brain injury and aneurysmal subarachnoid haemorrhage, NPi has been found to be associated with the severity of the injury ([53],[56]). However, the added predictive value of NPi to baseline characteristics remains poorly defined and confirmatory studies from large prospective cohorts lack, thus the ORANGE study was planned.

5.2 The Neurological Pupil Index (NPi)

An accurate and precise neurological examination of ICU patients is essential, since pre-existing or acquired neurological disorders may significantly affect their short and long-term outcomes. The pupillary examination is an important part of the neurological assessment, especially in the setting of acutely brain-injured patients, since pupillary abnormalities are associated with poor outcomes ([64]). The standard care of ICU patients with Acute Brain Injury (ABI) involves the pupillary light reactivity (PLR) evaluation as part of daily management, commonly done through a visual assessment of pupil size, shape, symmetry, and pupillary light reflex ([52]). Hence, PLR is a fundamental clinical evaluation done on ICU patients having a strong diagnostic and prognostic value ([65],[66]). Classically, PLR has been performed using a hand-held light source, such as a penlight or flashlight, which provide a non-standardized qualitative measurement. However, these subjective assessments of PLR may lead to imprecision, mainly because of significant inter-observer variability and heterogeneity in the technique used ([67]). On the other hand, Automated Infrared Pupillometry (AIP) ([68],[69]) furnishes a standardized, quantitative, highly reproducible measurement of the PLR and other pupillary variables, including amplitude, latency, constriction and dilation velocity. By employing an electronic pupillometer instead, such as NPi Pupillometer *NPi-200* or *NPi-300* (NeuroOptics, Inc., Irvine, CA), dynamic quantifications can be made over time, capturing subtle changes in pupillary functions ([64],[53]). These measures can be integrated and through an algorithm translated into a numerical index to compute the Neurological Pupil index (NPi) ([63],[70]), which is a composite numerical index (ranging from 0 to 5) of pupillary reactivity and global midbrain function. It is important to note that the NPi is not influenced by sedation–analgesia, at the doses used in neurocritical care practice, and by mild hypothermia ([63],[71],[72]). The deployment of AIP evaluations through an electronic pupillometer in severely brain-injured ICU patients has increased over recent years ([73]). Several single-center retrospective cohort studies have suggested a potential prognostic value of NPi in patients with severe acute brain injuries, but large prospective cohorts are needed to confirm these findings ([53],[54],[74],[58]).

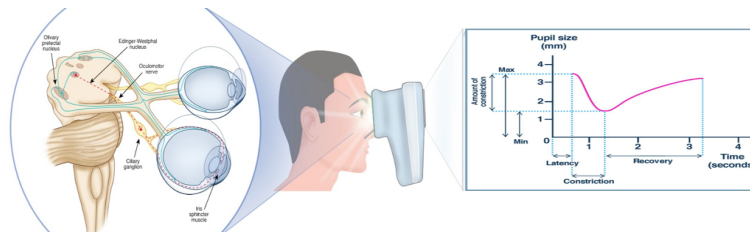


Figure 5.1: The NPi is derived from the integration of multiple measured pupillary variables through an electronic device (pupillometers *NPi-200*, *NPi-300*)

5.3 Study design

As there have never been any large multicenter cohort studies investigating the potential diagnostic and prognostic value of NPi in patients with ABI, the ORANGE study outlines an important turning point. The ORANGE (Outcome pRognostication of Acute brain injury

with the NeuroloGical pupil indEx) study ([64]) is an international, multicenter, prospective, observational trial that enrolled 18 years older patients admitted to the ICU, requiring intubation and mechanical ventilation after Traumatic Brain Injury (TBI), Subarachnoid Haemorrhage (SAH), or Intracranial Haemorrhage (ICH), at 13 hospitals in eight countries in Europe and the United States. Patients were excluded if they had facial trauma that could alter the use of AIP. Patients underwent AIP assessment every 4 hours to compute NPi during the first 7 days after admission to the ICU. Clinical teams were not blinded to the NPi measurements, as it was part of the standard clinical practice. Conversely, the outcome assessors were blinded. After the enrollment, the patients were followed up for 6 months after the injury, for outcome assessment.

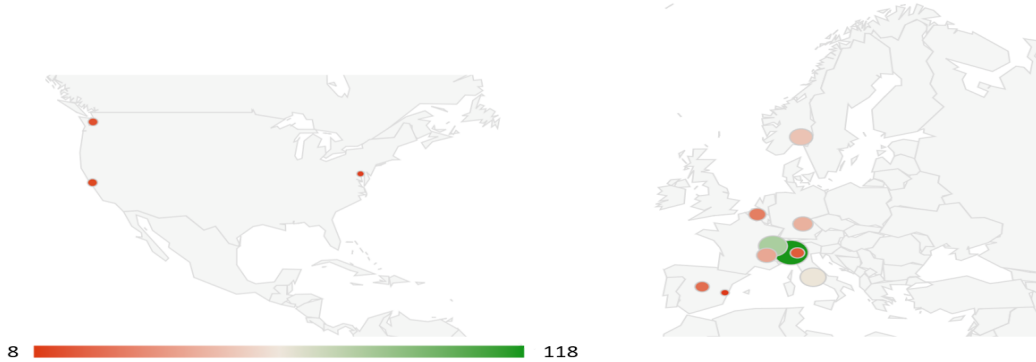


Figure 5.2: The specialized centers involved in the study: Monza, Italy; Lausanne, Switzerland; Roma, Italy; Oslo, Norway; Erlangen, Germany; Grenoble, France; Bruxelles, Belgium; Madrid, Spain; Brescia, Italy; Washington, US; UCSF, US; John Hopkins, US; Valencia, Spain. The size of the dots, locating the centers on the map, is proportional to the number of patients enrolled.

Between November 1, 2020 and May 3, 2022, 1938 patients with ABI were screened, of whom 514 (27%) were enrolled for the study (figure 5.3).

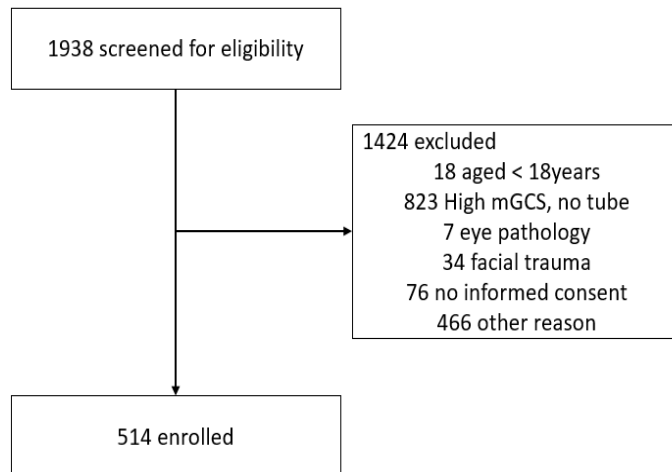


Figure 5.3: Flowchart.

NeurOptics NPi-200 and NPi-300 (after July 2021) pupillometers were used in the participating centers; the two devices have a high level of agreement, as recently reported ([75]), and can be used interchangeably. For all included patients, the data were gathered in an electronic Case Report Form (eCRF) developed in REDCap ([76]) and hosted at the University of Milano-Bicocca. The NPi data were imported automatically into the eCRF, using

smart-card technology (SmartGuard® Reader), thereby avoiding errors related to manual data transfer.

5.3.1 Outcomes and Aims

The ORANGE study aimed to evaluate the association between abnormal pupillary function, assessed by the NPi, and short and long-term outcomes in patients with ABI. In particular the co-primary outcomes of the study were functional neurological outcome (assessed with the extended Glasgow Outcome Scale -GOSE- ([77])) and mortality at 6 months ([64]). The secondary outcome was the mortality in ICU. Scores on the extended Glasgow Outcome Scale (GOSE) were collected by trained personnel who were blinded to the pupillometry results, using a validated questionnaire via telephone-structured interviews with patients or family members. A poor neurological outcome was defined as a GOSE score of four or less (i.e., low and upper disability, vegetative state, and death).

5.4 Statistical methods

Description of baseline characteristics and between-group comparisons were performed through the Wilcoxon or χ^2 test, as appropriate. The NPi values for both the right and left eye were collected. In order to investigate the association between the NPi and the clinical outcomes of interest, the lowest measurement obtained from each assessment on both eyes was used, since deemed as most pathological ([64]). The association between NPi and a poor neurological outcome at 6 months was evaluated through logistic regression, adjusting for age, ABI diagnosis, and motor Glasgow Coma Scale (mGCS) score on admission. We also considered alternative confounders to adjust for illness severity (ie, worst mGCS score, picking the lowest daily mGCS over the week, and pathological radiographic examinations at baseline defined as a Marshall classification ([78]) of three or more for traumatic brain injury, a modified Fisher ([79]) grade three or more for SAH, and a volume of 30 mL or greater for ICH) as supportive analysis. The longitudinal NPi values during the first week were summarised with the following quantities: the relative frequency of NPi less than three, defined as abnormal by the protocol, and the relative frequency of NPi equalling 0. The results were presented as odds ratios (ORs) and corresponding 95% confidence interval (CI). To evaluate the relationship of NPi with 6-month mortality, the extended version of the Cox regression model was used, entering NPi as a time-dependent covariate and age, ABI diagnosis, and mGCS on intensive care unit admission as fixed covariates. To account for individual dynamic variations of NPi over the first week of assessment, NPi was considered in the following three ways: first, categorised in two (NPi < 3 vs NPi \geq 3, as defined by the protocol, and NPi= 0 vs NPi > 0) or three levels (NPi < 3 vs NPi 3-4 vs NPi \geq 4); second, as a continuous variable; and third, considering the actual (NPi[t_i]) and the preceding (NPi[t_{i-1}]) NPi measurement, defining four categories according to the presence of NPi equal to zero or not on both occasions. We used the ECM also to evaluate the association between the NPi and mortality in shorter follow-up time (ICU). The assumption of proportional hazards was assessed by visually inspecting the plots of Schoenfeld residuals and using appropriate statistical tests for all covariates; the linearity of the effect for continuous variables was evaluated using splines. The results are presented as hazard ratios (HRs) and 95% CIs. As comparison we also used the JM on the same outcome. For the sake of brevity, we will only present the results of the ECM using the NPi as a continuous time-dependent covariate and of a JM with piecewise constant baseline hazard and a linear mixed effects specification to model the NPi trajectories over time, both adjusted for age, ABI diagnosis and GCS motor. Among the possible JMs specifications, it was chosen the combination of baseline hazard and random effects structure that resulted in lowest AIC and BIC here. Very similar results were obtained with other parameterizations, though.

5.5 Results on descriptive statistics

5.5.1 Baseline characteristics and outcomes

Among the 514 enrolled patients, 224 (44%) had TBI, 139 (27%) had SAH, and 151 (29%) had ICH. The median age was 61 years (IQR 46–71), and 309 (60%) patients were male. The median Glasgow Coma Scale score on admission was 8 (IQR 5–11). Baseline characteristics are presented in table 5.1

Six-month outcomes were available for 497 (97%) patients: 206 (41%) were alive with a good neurological outcome (GOSE >4), and 291 (59%) had a poor outcome (GOSE ≤4), of whom 160 (32%) died (table 5.2). Before ICU discharge 124 (24%) patients died (table 5.2); out of these ~80% died before or soon after the first week of ICU stay.

5.5.2 NP_i evaluations

40071 pupillometry examinations from both eyes were collected, with a median of 40 measurements per patient (IQR 20–50) during the study period. At baseline and over time, right and left NP_i values were not significantly different. The overall distribution of NP_i values is presented in figure 5.4.A, showing two peaks at 0 and 4.7 and a median NP_i of 4.3 (Inter Quartile Range -IQR- 3.7–4.7). The distribution of the 20194 lowest NP_i values at each timepoint is shown in figure 5.4.B which shows a similar shape to the previous one, with two peaks at 0 and 4.7 and a median NP_i of 4.2 (IQR 3.5–4.6).

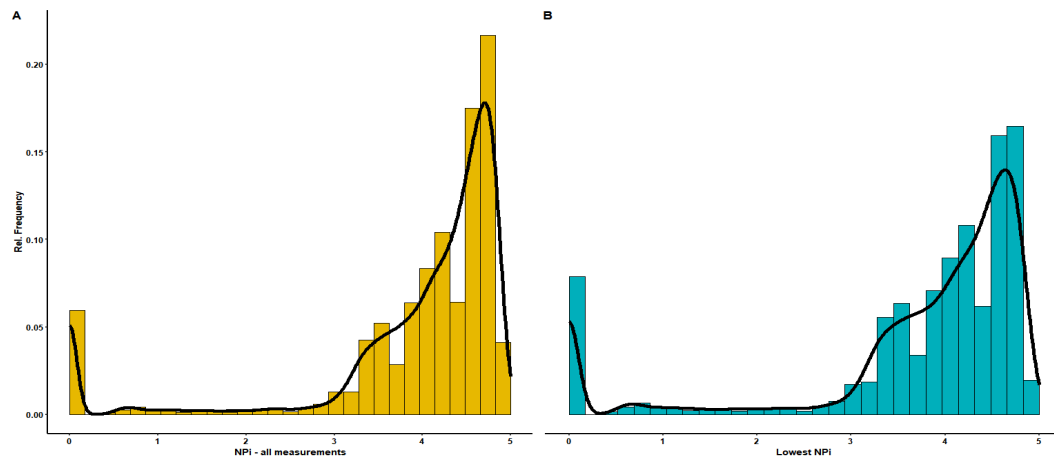


Figure 5.4: (A) Distribution of all ($n=40071$) NP_i measurements on 514 patients. (B) Distribution of lowest NP_i measurements ($n=20194$).

Among the NP_i values, considering the lowest measure at each assessment between the two eyes, abnormal NP_i (NP_i < 3) was observed at least once in 241 (47%) of 514 patients. At least one NP_i equal to zero was recorded in 132 (26%) of 514 patients (two of whom did not have a 6-month outcome), most frequently in the ICH group (table 5.3).

In figure 5.5 are shown some examples of NP_i trajectories measured through the time on the left and the right eye for a few selected patients. In the majority of the cases, the longitudinal profiles are stable and linear with respect to the time, but it can be seen how, in some occasions, this biomarker exhibits quite complex dynamics, with many wiggles and multiple ups and downs close in time.

Table 5.1: Baseline characteristics.

	ALL (n=514)	TBI (n=224)	SAH (n=139)	ICH (n=151)
Age, years	61 (46-71)	54 (34-72)	59 (51-71)	64 (54-71)
Gender				
Male	309 (60)	170 (76)	46 (33)	93 (62)
Female	205 (40)	54 (24)	93 (67)	58 (38)
Glasgow Coma Scale score				
3-5	141 (29)	69 (33)	38 (29)	34 (24)
6-8	131 (27)	70 (33)	21 (16)	40 (28)
9-15	211 (44)	72 (34)	71 (55)	68 (48)
NA	31	13	9	9
Motor Glasgow Coma Scale				
None	120 (23)	63 (28)	31 (22)	26 (17)
Extension	40 (8)	14 (6)	13 (9)	13 (9)
Abnormal flexion	27 (5)	12 (5)	5 (4)	10 (7)
Normal flexion	63 (12)	35 (16)	11 (8)	17 (11)
Localises/obeys	263 (51)	99 (44)	79 (57)	85 (56)
NA	1	1	0	0
Pupil reactivity				
Reactive	414 (82)	186 (85)	117 (85)	111 (75)
One unreactive	32 (6)	11 (5)	9 (7)	12 (8)
Both unreactive	60 (12)	22 (10)	12 (9)	26 (17)
NA	8	5	1	2
Pathological severity by radiographical examinations				
Pathological	334 (65)	117 (52)	121 (87)	96 (64)
Non-pathological	180 (35)	107 (48)	18 (13)	55 (36)
Any cardiovascular disease				
Yes	248 (48)	83 (37)	71 (51)	94 (62)
No	408 (79)	190 (85)	109 (78)	109 (72)
Any liver diseases				
Yes	20 (4)	6 (3)	5 (4)	9 (6)
No	494 (96)	218 (97)	134 (96)	142 (94)
Any neurological diseases				
Yes	72 (14)	23 (10)	21 (15)	28 (29)
No	442 (86)	201 (90)	118 (85)	123 (81)
Any oncological diseases				
Yes	44 (9)	14 (6)	12 (9)	18 (12)
No	470 (91)	210 (94)	127 (91)	133 (88)
Any respiratory diseases				
Yes	31 (6)	9 (4)	5 (4)	17 (11)
No	483 (94)	215 (96)	134 (96)	134 (89)
Any psychiatric disturbances				
Yes	47 (9)	26 (12)	13 (9)	18 (11)
No	467 (91)	198 (88)	126 (91)	133 (89)
Any renal diseases				
Yes	15 (3)	2 (1)	3 (2)	10 (7)
No	499 (97)	222 (99)	136 (98)	141 (93)
Any eye diseases				
Yes	16 (3)	7 (3)	4 (3)	5 (3)
No	498 (97)	217 (97)	135 (97)	146 (97)

Data are n (% of available data) or median (IQR). NA=not available.

Table 5.2: ICU mortality, 6 months mortality and 6 months neurological outcome evaluated with the GOSE, overall and divided by pathology.

		ALL	TBI	SAH	ICH
		(n=514)	(n=224)	(n=139)	(n=151)
ICU Mortality, n(%)	Survivors	124 (24)	46 (20)	32 (23)	46 (30)
	Non-survivors	324 (76)	178 (80)	107 (67)	105 (70)
	NA	0	0	0	0
6 months Mortality, n (%)	Survivors	337 (68)	155 (72)	97 (71)	85 (59)
	Non-survivors	160 (32)	61 (28)	39 (29)	60 (41)
	NA	17	8	3	6
Neurological outcome, n (%)	Poor (GOS-E 1-4)	291 (59)	106 (49)	79 (58)	106 (73)
	Good (GOS-E 5-8)	206 (41)	110 (51)	57 (42)	39 (27)
		17	8	3	6

5.5.3 Association between abnormal NPi and 6-months outcomes

Figure 5.6 illustrates the distribution of NPi values according to different outcomes. The median NPi value was lower in patients with a poor neurological outcome (median 4.0, IQR [3.3–4.5]) compared with those with a good neurological outcome (median 4.3, IQR [3.9–4.6]; $p < 0.0001$). 113 (39%) of 291 patients with a poor neurological outcome had at least one NPi equal to zero, and 17 (8%) of 206 patients with a good neurological outcome had at least one NPi equal to zero ($p < 0.0001$). An abnormal NPi was measured at least once in 179 (62%) of 291 patients with a poor neurological outcome and in 59 (29%) of 206 patients with a good neurological outcome ($p < 0.0001$). The median NPi value was lower in non-survivors (median 3.9, IQR [3.0–4.5]) than in survivors (median 4.3, IQR [3.7–4.6]; $p < 0.0001$); 80 (50%) of 160 non-survivors had at least one NPi equal to zero, and 50 (15%) of 337 survivors had at least one NPi value equal to zero ($p < 0.0001$). None of the 35 patients who had all NPi measurements equal to zero survived (median number of measures 7, IQR [5–10]; median time to death 2.1 days, IQR [1.3–2.5]). An abnormal NPi was measured at least once in 113 (71%) of 160 non-survivors and in 125 (37%) of 337 survivors ($p < 0.0001$).

Table 5.3: NPi values in the first 7 days from admission to ICU.

	All (n=514)	TBI (n=224)	SAH (n=139)	ICH (n=151)
NPi first measure right eye				
Median (IQR)	4.0 (3.4–4.5)	4.0 (3.4–4.5)	4.0 (3.4–4.4)	3.9 (2.5–4.4)
n	514	224	139	151
NPi first measure left eye				
Median (IQR)	4.0 (3.4–4.5)	4.1 (3.4–4.5)	4.1 (3.4–4.5)	3.7 (3.2–4.5)
n	514	224	139	151
NPi right eye				
Median (IQR)	4.4 (3.8–4.7)	4.4 (3.9–4.7)	4.4 (3.8–4.7)	4.3 (3.6–4.6)
n	19976	8466	6112	5398
NPi left eye				
Median (IQR)	4.3 (3.7–4.7)	4.4 (3.9–4.7)	4.4 (3.8–4.7)	4.3 (3.6–4.6)
n	20095	8640	6124	5511
Lowest NPi value				
Median (IQR)	4.2 (3.5–4.6)	4.2 (3.6–4.6)	4.2 (3.6–4.6)	4.1 (3.4–4.5)
n	20194	8499	6163	5532
Patients with at least one NPi=0	132 (26%)	45 (20%)	32 (23%)	55 (36%)
0	382 (74%)	179 (80%)	107 (77%)	96 (64%)
1	22 (4%)	7 (3%)	5 (4%)	10 (7%)
2	12 (2%)	7 (3%)	3 (2%)	2 (1%)
3	8 (2%)	4 (2%)	0	4 (3%)
>3	90 (18%)	27 (12%)	24 (17%)	39 (26%)
Patients with at least one NPi<3	241 (47%)	89 (40%)	70 (50%)	82 (54%)
0	273 (53%)	135 (60%)	69 (50%)	69 (46%)
1	57 (11%)	23 (10%)	17 (12%)	17 (11%)
2	22 (4%)	8 (4%)	6 (4%)	8 (5%)
3	24 (5%)	9 (4%)	6 (4%)	9 (6%)
>3	138 (27%)	49 (22%)	41 (30%)	48 (32%)

Data are n (%), unless otherwise stated. NPi=Neurological Pupil index.

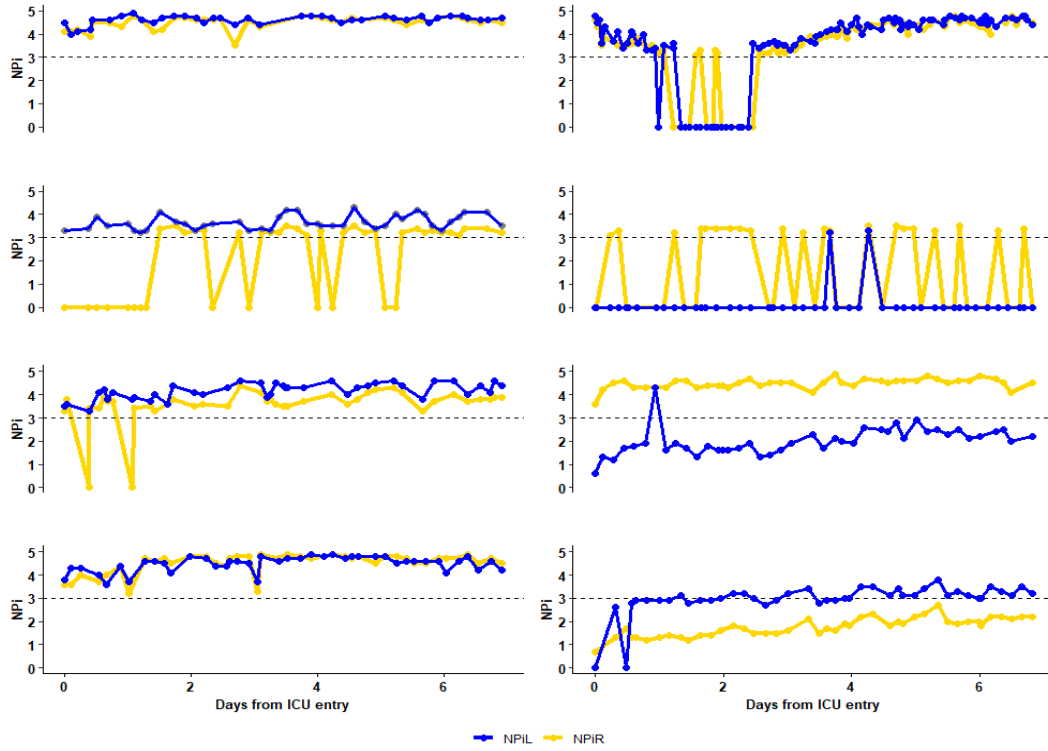


Figure 5.5: Six patients' NP_i trajectories as a result of the measurement on left and right eye over time.

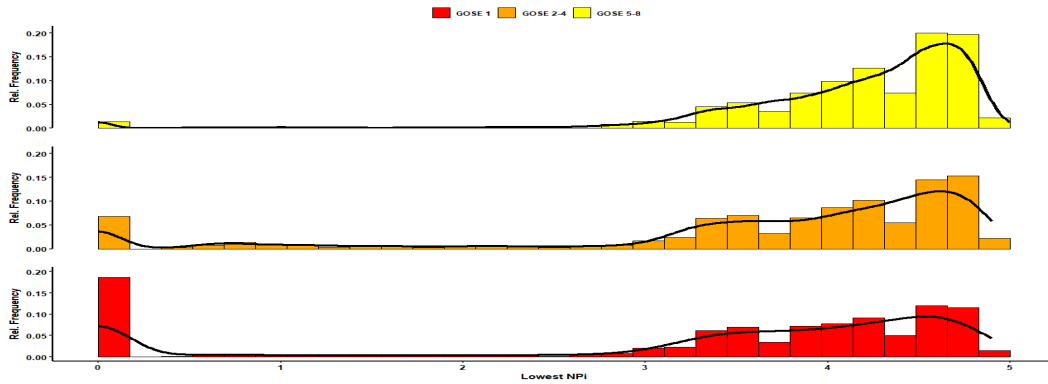


Figure 5.6: Distribution of the 19427 lowest NP_i measurements by 6-month neurological outcome (GOSE 1 [dead], $n=160$ patients, $n=5725$ NP_i measurements; GOSE 2–4 [poor outcome], $n=131$ patients, $n=5997$ NP_i measurements; GOSE 5–8 [good outcome], $n=206$ patients, $n=7705$ NP_i measurements).

5.6 Results on long-term outcomes

The analysis of the relationship of NP_i with the neurological outcome (GOSE ≤ 4) evaluated at 6 months is reported in table 5.4 in terms of odds ratios (OR) with their corresponding 95% CIs. We found that a 10% increase in the frequency of abnormal NP_i was associated with poor neurological outcome (OR 1.42, 95% CI [1.27–1.64]; $p < 0.0001$). Similarly, a 10% increase in the frequency of NP_i equal to zero was associated with poor neurological outcome (OR 1.70, 95% CI [1.37–2.38]; $p < 0.0001$). These results were consistent when TBI, SAH, and ICH were considered separately (results not shown). The sensitivity analyses adjusting

for the worst mGCS score and pathological severity by radiographic examinations further confirmed the results (results not shown).

Table 5.4: Table shows the association between frequencies of $NPi < 3$ and $NPi = 0$ value and 6-month poor neurological outcome ($GOSE \leq 4$; model 1-2), adjusted for age, ABI diagnosis, and motor Glasgow Coma Scale on admission to the intensive care unit. Models 1-2 included data for 497 patients, of whom 291 had a $GOSE \leq 4$.

	NPi and GOSE at 6 months	OR (95 % CI)	p value
Model 1	10% increase in the frequency of $NPi < 3$	1.42 (1.27–1.64)	<0.0001
Model 2	10% increase in the frequency of $NPi = 0$	1.70 (1.37–2.38)	<0.0001

The analysis of the association between the NPi and within 6 months mortality is presented in table 5.5. After adjusting for covariates, abnormal NPi was associated with an increased risk of mortality (HR 5.58, 95% CI [3.92–7.95]; $p < 0.0001$). An association was also found for NPi equal to zero (HR 12.05, 95% CI [7.86–18.48]; $p < 0.0001$). A one-unit decrease in the NPi value was independently associated with a higher risk of mortality (HR 1.80, 95% CI [1.62–1.99]; $p < 0.0001$). In the analysis including individual dynamic NPi variations over time and using two consecutive values of NPi greater than zero as the reference category, the occurrence of two successive NPi values equal to zero was associated with an increased risk of mortality (HR 13.92, 95% CI [8.94–21.67]; $p < 0.0001$). Also, deterioration of an NPi value to zero was associated with an increased risk of mortality (HR 8.37, 95% CI [2.52–27.87]; $p = 0.0007$), and the risk of mortality was not increased when NPi improved from zero to a higher value (HR 1.32, 95% CI [0.32–5.41]; $p = 0.6995$). Finally, NPi less than three (HR 7.10, 95% CI [4.77–10.57]; $p < 0.0001$), and NPi values between three and four (HR 1.70, 95% CI [1.13–2.56]; $p = 0.0186$) were associated with a higher risk of mortality than were higher NPi values (≥ 4). Internal cross-validation showed robustness in the results of the multivariable models that were performed on both outcomes.

Table 5.5: Results of the multivariable time-dependent Cox model on the association between longitudinal NPi value and six-month mortality adjusting for age, ABI diagnosis, and motor Glasgow Coma Scale on admission to the ICU.

	NPi and mortality at 6 months	HR (95 % CI)	p value
Model 1	$NPi(t_i) < 3$	5.58 (3.98–7.95)	< 0.0001
	$NPi(t_i) \geq 3$	1	..
Model 2	$NPi(t_i) = 0$	12.05 (7.86–18.48)	< 0.0001
	$NPi(t_i) > 0$	1	..
Model 3	$NPi(t_i)$	1.80 (1.62–1.99)	<0.0001
Model 4	$NPi(t_{i-1}) = 0, NPi(t_i) = 0$	13.92 (8.94–21.67)	<0.0001
	$NPi(t_{i-1}) = 0, NPi(t_i) > 0$	1.32 (0.32–5.41)	0.6995
	$NPi(t_{i-1}) > 0, NPi(t_i) = 0$	8.37 (2.52–27.87)	0.0005
	$NPi(t_{i-1}) > 0, NPi(t_i) > 0$	1	..
Model 5	$0 \leq NPi(t_i) < 3$	7.10 (4.77–10.57)	<0.0001
	$3 \leq NPi(t_i) < 4$	1.70 (1.13–2.56)	0.0163
	$NPi(t_i) \geq 4$	1	..

5.6.1 Results on short term mortality (ICU): a comparative analysis using ECM and JM

Table 5.6 reports the results of the ECM on mortality within ICU showing an estimated HR of 1.89, very close to the estimated HR by the analogous modeling for 6 months mortality (124 died in ICU vs 160 died within 6 months) (table 5.12). In this analysis we were able to apply also a JM as the outcome was assessed in a shorter time period, i.e. close to the NPi evaluations (first 7 days). In fact for 6 months mortality, although the LOCF approximation in the Cox model is, at least, questionable to be used to assess this long term outcome, it has the advantage of using values of the NPi that are naturally within the definition range [0,5]. This problem, on the other hand, may occur using a JM that, based on available NPi measurements, estimates longitudinal trajectories, providing out-of-range predictions in some cases. In according with ECM, also the JM found that a one-unit decrease in the NPi values was statistically associated with an increase in the risk of death during ICU stay. Nevertheless the magnitudes of the HRs are different. In particular, a one-unit decrease in the NPi value corresponds to an estimated 90% increase in the risk of death during the first week of ICU using the ECM, and a 44% increase using JM (table 5.7). These results showed that these two approaches may give quite different results due to the LOCF approximation of ECM and the estimation of the NPi trajectories with a JM strategy. In fact, as previously noted, NPi trajectories often exhibited highly non-linear patterns over time, with rapid oscillations and abrupt changes and due to this the best regression technique on these longitudinal data remains a challenging task. Indeed, we could have used other mixed models to approximate such longitudinal profiles, for example by deploying regression splines (see chapter 2). With such flexible approximations we could have better captured the shape of the NPi dynamic evolutions over time, but a disadvantage would have been in the interpretation of the results, especially in terms of HR. Thus, we will not delve further into the analysis of these data, but we conclude by saying that all the modeling methodologies used have consistently confirmed the prognostic value of the NPi on the risk of experiencing death: in particular we found that a deterioration of NPi should be sounded as an alarm and, careful and intense monitoring of the longitudinal evolutions by clinicians is desirable and recommended.

Table 5.6: *Extended Cox regression model's hazard ratios with 95% CI of mortality within ICU for NPi over the first week as a continuous variable, age, ABI diagnosis (Subarachnoid hemorrhage, Intracranial hemorrhage, TBI), GCS Motor (no motor response, extension to pain, abnormal flexion, normal, localizes/obey). *Benjamini-Hochberg adjusted.*

	HR (95% CI)	p-value
Age	1.06 (1.04-1.07)	<0.0001
NPi		
NPi	1.89 (1.63-2.11)	<0.0001
Diagnosis		
Subarachnoid haemorrhage	0.79 (0.50-1.26)	0.3306
Intracranial haemorrhage	0.97 (0.63-1.49)	0.8995
Traumatic Brain Injury	1	..
GCS Motor Score		
None	1.24 (0.76-2.02)	0.3945
Extension	2.23 (1.33-3.72)	0.0022
Abnormal flexion/ Normal flexion	1.24 (0.67-2.29)	0.4962
Localizes/obeys	1	..
n. NPi obs. = 20106, n. of deaths = 124 (n. patients 514)		

Table 5.7: JM's hazard ratios with 95% CI of mortality within ICU for NPi over the first week as a continuous variable, age, ABI diagnosis (SAH, ICH, TBI), GCS Motor (no motor response, extension to pain, abnormal flexion, normal, localizes/obey).

Variance Components:		
	Std. Dev.	Corr.
Intercept	1.43	
time (days)	0.13	-0.34
residual	0.63	
Longitudinal Process		
	Coeff. (95% CI)	p-value
Intercept	3.50 (3.37-3.62)	<0.0001
time (days)	0.04 (0.02-0.05)	<0.0001
Event Process		
	HR (95% CI)	p-value
Age	1.05 (1.04-1.07)	<0.0001
NPi		
NPi	1.44 (1.30-1.60)	<0.0001
Diagnosis		
Subarachnoid haemorrhage	0.85 (0.53-1.35)	0.4661
Intracranial haemorrhage	1.19 (0.77-1.84)	0.4218
Traumatic Brain Injury	1	..
GCS motor score		
None	2.12 (1.33-3.39)	0.0016
Extension/ Abnormal flexion	2.85 (1.72-4.74)	0.0001
Normal Flexion	1.37 (0.74-2.55)	0.3192
Localize/obeys	1	..
n. NPi obs. = 20106, n. of deaths = 124 (n. patients 514)		

5.7 Discussion

To the best of our knowledge, the ORANGE study is the most extensive prospective study to investigate the prognostic value of NP_i as a standardised quantitative measurement of pupillary reactivity and global midbrain function in patients with non-anoxic brain injury. Overall, our results strongly suggest that repeatedly abnormal NP_i values, including the most extreme values of zero, in the first week after ABI predict poor outcome. Abnormal NP_i (<3) and NP_i values of zero were more frequently observed in patients with poor (GOSE \leq 4) versus good (GOSE >4) outcomes. An increase in the number of abnormal NP_i measurements over time was associated with a higher probability of poor neurological outcome. Two consecutive NP_i measurements equal to zero, or deterioration of NP_i to a value of zero, were associated with an increased mortality risk. By contrast, the mortality risk was not increased when an NP_i value of zero recovered to a higher value. These findings indicate the importance of the trajectories of NP_i. Our results also offer further valuable insights into the interpretation of NP_i values. Indeed, while the absence of pupillary light reactivity (i.e., NP_i=0) is a well-established indicator of poor outcome, an abnormal NP_i has been previously defined as below 3. This cut-off was recently challenged by a multi-center study in patients with ABI after cardiac arrest, where a value below 2 was found to hold a one-hundred percent specificity of poor prognosis. Here, repeated NP_i measurements enhanced the sensitivity analysis and allowed the identification of an NP_i range, between 3 and 4, that was already associated with an increased risk of mortality. In this setting, NP_i monitoring might identify at-risk patient population that would benefit from careful, intensive observation to manage secondary brain deterioration, and target specific interventions before irreversible damage may occur. In this context, instead of focusing on a single measurement or cut-off, clinicians should view NP_i as a tool for quantifying in a timely fashion the extent of midbrain dysfunction, ranging from very severe (NP_i= 0), to severe (NP_i< 3) and moderate (NP_i 3-4).

The ORANGE study was thereby designed to assess the prognostic value of multiple NP_i measurements over time after adequately adjusting for other known baseline predictors. We used a dataset with high data granularity, a minimum of six NP_i assessments per day, an automated digital system for data downloading, and a rigorous, effective, and blindly evaluated long-term outcome follow-up, all of which added to the robustness of the current study analysis. In patients with non-anoxic brain-injuries, repeated NP_i measurements are crucial for predicting outcome. Indeed, serial NP_i assessments offer a more comprehensive and accurate understanding of the evolution of brain damage over time in this context. Our approach had several advantages. Detecting NP_i changes allows clinicians to monitor patient conditions and identify improvements, deteriorations, or persistent abnormal NP_i values. This information can help to identify patients at high risk of poor outcome (i.e., those with a high percentage of abnormal NP_i measurements over time). Additionally, including reliable non-invasive neurological monitoring in clinical practice is highly beneficial, as it represents a safe alternative to invasive procedures, reducing the associated risks and complications for the patient. Our analysis also showed that NP_i variations between two consecutive measurements have significant prognostic value, particularly in patients with NP_i deterioration without recovery or with persistent NP_i measurements at a value of zero. Due to the dynamic changes of NP_i over time, a single abnormal measure should prompt health-care providers to retest NP_i to minimise measurement errors. Furthermore, as NP_i can improve over time, repeated measurements would enable clinicians to assess the efficacy of therapeutic interventions over time. Our results also offer further valuable insights into the interpretation of NP_i values. Although the absence of pupillary light reactivity (i.e., an NP_i value of zero) is a well established indicator of poor outcome, an abnormal NP_i has been previously defined as below three. Repeated NP_i measurements enhanced the sensitivity analysis and identified an NP_i range between 3 and 4 already associated with an increased mortality risk. In this setting, NP_i monitoring might identify at-risk patient populations that would benefit from careful intensive observation to manage secondary brain deterioration and target specific interventions before irreversible damage can occur. In this context, instead of focusing on a single measurement or cut-off, clinicians could view NP_i

as a tool for quantifying in a timely fashion the extent of midbrain dysfunction, ranging from very severe (NPi= 0), to severe (NPi <3) and moderate (NPi 3–4). Integrating NPi with other available tools for assessing the severity of brain injury could ultimately lead to targeted and effective diagnostic and treatment strategies for patients with varying degrees of ABI. In conclusion, in this prospective international multicenter study, abnormal NPi was strongly associated with long-term mortality and a poor neurological outcome after an ABI, irrespective of age, primary diagnosis, and severity of cerebral damage. Repeated NPi measurements provided relevant prognostic information. Our study also identified novel NPi pathological thresholds (<4) following ABI that could assist clinicians in detecting brain damage and monitoring the response to therapeutic interventions in this setting.

5.7.1 Limitations

Our study has several limitations that should be acknowledged. First, the observational design and the lack of standardised treatment protocols across centers might compromise the robustness of certain results. The study strives to depict real-life situations accurately. The staff conducting the study were not blinded to the NPi evaluation because NPi evaluation is an integral part of the clinical evaluation practice. As a result, the observed NPi changes could have influenced some actions. Moreover, the blinded evaluation of NPi changes and the effect on outcome enhances the reliability of the findings. This approach reinforces the integrity of the assessment by minimising potential biases. Second, our focus was solely on NPi. We did not assess the potential value of other variables obtained from AIP assessment, such as pupillary constriction or dilation velocities. However, this limitation could be seen as an advantage since NPi, contrary to pupillary constriction or dilation velocities, is only minimally influenced by sedatives and analgesics. Third, it is yet to be determined whether the measurement duration over the 7 days following admission to the intensive care unit was the most optimal approach. It remains uncertain whether the findings from our study can be extrapolated to other types of brain injuries. Finally, although our data provided robust associations between NPi and patient prognosis using a large dataset and internal cross-validation, additional confirmation of these findings in diverse settings is needed, including in centers with varying expertise in pupillometry utilisation or different protocols regarding the limitation of life-sustaining therapies. Based on this evidence, we expect that future trial designs will explore the potential of AIP as both a diagnostic tool for decision making and an interventional tool in conjunction with standardised therapy.

Table 5.8: Logistic regression model's odds ratios with 95% CI of GOSE at 6-month follow-up for 10% increment for frequency of NPi < 3, age, ABI diagnosis (SAH, ICH, TBI), GCS Motor (no motor response, extension to pain, abnormal flexion, normal, localizes/obeys). *Benjamini-Hochberg adjusted.

	Odds ratio (95% CI)	p-value
Intercept	0.04 (0.02-0.10)	<0.0001
Age	1.04 (1.03-1.05)	<0.0001
NPi		
10% increment for frequency of NPi < 3	1.42 (1.27-1.64)	* <0.0001
Diagnosis		
Subarachnoid haemorrhage	1.32 (0.80-2.18)	0.2735
Intracranial haemorrhage	2.29 (1.36-3.89)	0.0021
Traumatic Brain Injury	1	..
GCS Motor Score		
None	1.87 (1.07-3.32)	0.0306
Extension	2.84 (1.15-7.56)	0.0284
Abnormal flexion	1.02 (0.39-2.66)	0.9680
Normal flexion	1.88 (0.98-3.71)	0.061
Localizes/obeys	1	..
Number of patients = 497, GOSE ≤ 4 = 291		

Table 5.9: Logistic regression model's odds ratios with 95% CI of GOSE at 6-month follow-up for 10% increment for frequency of NPi = 0, age, ABI diagnosis (SAH, ICH, TBI), GCS Motor (no motor response, extension to pain, abnormal flexion, normal, localizes/obeys). *Benjamini-Hochberg adjusted.

	Odds ratio (95% CI)	p-value
Intercept	0.06 (0.02-0.13)	<0.0001
Age	1.04 (1.03-1.05)	<0.0001
NPi		
10% increment for frequency of NPi = 0	1.70 (1.37-2.38)	* <0.0001
Diagnosis		
Subarachnoid haemorrhage	1.33 (0.81-2.19)	0.2548
Intracranial haemorrhage	2.27 (1.36-3.85)	0.0020
Traumatic Brain Injury	1	..
GCS Motor Score		
None	2.02 (1.16-3.57)	0.0143
Extension	2.75 (1.12-7.22)	0.0320
Abnormal flexion	1.14 (0.44-2.92)	0.7830
Normal flexion	1.83 (0.95-3.57)	0.0728
Localizes/obeys	1	..
Number of patients = 497, GOSE ≤ 4 = 291		

Table 5.10: Extended Cox regression model's hazard ratios with 95% CI of mortality at 6-month follow-up for NPi over the first week categorized in two (using a cut-off of 3), age, ABI diagnosis (SAH, ICH, TBI), GCS Motor (no motor response, extension to pain, abnormal flexion, normal, localizes/obey). *Benjamini-Hochberg adjusted.

	Hazard ratio (95% CI)	p-value
Age	1.05 (1.04-1.07)	<0.0001
NPi		
NPi(t_i) < 3	5.58 (3.92-7.95)	<0.0001
NPi(t_i) ≥ 3	1	..
Diagnosis		
Subarachnoid haemorrhage	0.78 (0.52-1.18)	0.2353
Intracranial haemorrhage	1.20 (0.83-1.75)	0.3287
Traumatic Brain Injury	1	..
GCS Motor Score		
None	1.92 (1.28-2.85)	0.0014
Extension	2.87 (1.70-4.83)	<0.0001
Abnormal flexion	1.26 (0.61-2.61)	0.5248
Normal flexion	1.72 (1.04-2.85)	0.0341
Localizes/obeys	1	..
n. NPi obs. = 19427, n. of deaths= 160 (n. patients 497)		

Table 5.11: Extended Cox regression model's hazard ratios with 95% CI of mortality at 6-month follow-up for NPi over the first week categorized in two (using a cut-off of 0), age, ABI diagnosis (SAH, ICH, TBI), GCS Motor (no motor response, extension to pain, abnormal flexion, normal, localizes/obey). *Benjamini-Hochberg adjusted.

	Hazard ratio (95% CI)	p-value
Age	1.05 (1.04-1.07)	<0.0001
NPi		
NPi(t_i) = 0	12.05 (7.86-18.48)	*<0.0001
NPi(t_i) > 0	1	..
Diagnosis		
Subarachnoid haemorrhage	0.88 (0.59-1.32)	0.5439
Intracranial haemorrhage	1.14 (0.78-1.65)	0.5015
Traumatic Brain Injury	1	..
GCS Motor Score		
None	1.46 (0.95-3.96)	0.0828
Extension	2.28 (1.34-3.90)	0.0025
Abnormal flexion	1.95 (0.96-3.96)	0.0655
Normal flexion	1.52 (0.91-2.51)	0.1061
Localizes/obeys	1	..
n. NPi obs. = 19427, n. of deaths = 160 (n. patients 497)		

Table 5.12: Extended Cox regression model's hazard ratios with 95% CI of mortality at 6-month follow-up for NPi over the first week as a continuous variable, age, ABI diagnosis (Subarachnoid hemorrhage, Intracranial hemorrhage, TBI), GCS Motor (no motor response, extension to pain, abnormal flexion, normal, localizes/obeys). *Benjamini-Hochberg adjusted.

	Hazard ratio (95% CI)	p-value
Age	1.05 (1.04-1.07)	<0.0001
NPi		
NPi(t_i)	1.80 (1.62-1.99)	*<0.0001
Diagnosis		
Subarachnoid haemorrhage	0.79 (0.53-1.19)	0.2640
Intracranial haemorrhage	1.05 (0.72-1.54)	0.7813
Traumatic Brain Injury	1	..
GCS Motor Score		
None	1.30 (0.84-1.99)	0.2345
Extension	2.24 (1.32-3.83)	0.0030
Abnormal flexion	1.50 (0.73-3.05)	0.2680
Normal flexion	1.49 (0.90-2.47)	0.1234
Localizes/obeys	1	..
n. NPi obs. = 19427, n. of deaths = 160 (n. patients 497)		

Table 5.13: Extended Cox regression model's hazard ratios with 95% CI of mortality at 6-month follow-up for NPi over the first week considering the actual [NPi(t_i)] and the preceding [NPi(t_{i-1})] NPi, defining four categories according to a cut-off of 0 on both, age, ABI diagnosis (SAH, ICH, TBI), GCS Motor (no motor response, extension to pain, abnormal flexion, normal, localizes/obeys). *Benjamini-Hochberg adjusted. **Four patients had only one NPi collected.

	Hazard ratio (95% CI)	p-value
Age	1.05 (1.04-1.07)	<0.0001
NPi		
NPi(t_{i-1}) = 0 & NPi(t_i) = 0	13.92 (8.94-21.67)	*<0.0001
NPi(t_{i-1}) = 0 & NPi(t_i) > 0	1.32 (0.32-5.41)	*0.6995
NPi(t_{i-1}) > 0 & NPi(t_i) = 0	8.37 (2.52-27.87)	*0.0007
NPi(t_{i-1}) > 0 & NPi(t_i) > 0	1	..
Diagnosis		
Subarachnoid haemorrhage	0.89 (0.59-1.34)	0.5787
ICH	1.07 (0.73-1.56)	0.7205
Traumatic Brain Injury	1	..
GCS Motor Score		
None	1.30 (0.83-2.02)	0.2480
Extension	2.16 (1.25-3.73)	0.0059
Abnormal flexion	2.01 (0.99-4.09)	0.0533
Normal flexion	1.57 (0.95-2.58)	0.0762
Localizes/obeys	1	..
n. NPi obs. = 18947, n. of deaths = 156 (n. patients 493**)		

Table 5.14: Extended Cox regression model's hazard ratios with 95% CI of mortality at 6-month follow-up for NPi over the first week categorized in three levels (<3, 3-4 and 4-5), age, ABI diagnosis (SAH, ICH, TBI), GCS Motor (no motor response, extension to pain, abnormal flexion, normal, localizes/obeys). *Benjamini-Hochberg adjusted.

	Hazard ratio (95% CI)	p-value
Age	1.06 (1.04-1.07)	<0.0001
NPi		
$0 \leq \text{NPi}(t_i) < 3$	7.10 (4.77-10.57)	*<0.0001
$3 \leq \text{NPi}(t_i) < 4$	1.70 (1.13-2.56)	*0.0186
$\text{NPi}(t_i) \leq 4$	1	..
Diagnosis		
Subarachnoid haemorrhage	0.74 (0.49-1.12)	0.1519
Intracranial haemorrhage	1.11 (0.76-1.62)	0.5855
Traumatic Brain Injury	1	..
GCS Motor Score		
None	1.70 (1.13-2.56)	0.0102
Extension	2.68 (1.59-4.52)	0.0002
Abnormal flexion	1.20 (0.58-2.48)	0.6256
Normal flexion	1.72 (1.05-2.83)	0.0315
Localizes/obeys	1	..
n. NPi obs. = 19427, n. of deaths = 160 (n. patients 497)		

Chapter 6

Evaluating the robustness of Extended Cox and Joint models with missing longitudinal data

Recognising that longitudinal studies often suffer from missing data, including our motivating clinical context, within this thesis work we investigated the ability and the degree of robustness of the Cox model in its extended version (ECM), and of the Joint model (JM), to correctly estimate the strength of the association between a dynamic marker and a time-to-event (event time). In particular we focused on situations in which the dynamic profile of the time-dependent covariate is affected by specific missing processes in different ways. To this aim, we simulated longitudinal and time-to-event data and we assessed the performances of those two approaches considering different functional forms for the individual longitudinal markers (linear and quadratic shape) and numerous missing processes affecting their evolutions over time. The first part of this chapter will cover some aspects, both from a theoretical and a practical point of view, concerning how to simulate time-to-event data. To carry out the simulations, we simulated the event times in a particular circumstance, namely one in which we have a longitudinal covariate. The approach we used will be briefly explained in section 6.2 but, before, the simpler case of simulating event times with only fixed baseline covariates, assuming a parametric or a semi-parametric baseline hazard will be treated. The second part will be devoted to presenting the simulation workflow and the results of the study. We will conclude the chapter by discussing some limitations of this work and possible future extensions.

6.1 Times-to-event simulation with baseline covariates

6.1.1 Survival time simulation

The survival function for individual i is the probability that its “true” event time T_i^* is greater than the current time t :

$$S_i(t) = \mathbb{P}[T_i^* > t]$$

A survival function’s complement to one is the corresponding probability of having experienced the event at or before time t . In other words, the probability of failure is:

$$F_i(t) = \mathbb{P}[T_i^* \leq t] = 1 - S_i(t)$$

It turns out that the definition of the probability of failure in the last equation coincides to the definition of the cumulative distribution function (CDF) for the distribution of event times. The probability integral transform theorem ([80]) states that transforming a continuous random variable by its own CDF leads a new random variable following a uniform distribution on the range $[0,1]$. That is, $F_X(x) \sim U(0,1)$ where $F_X(\cdot)$ denotes the CDF for the continuous random variable X . Therefore, a new random variable obtained by taking the complement to 1 of X transformed by its own CDF ($F_X(x)$) must also follow a uniform distribution on the range $[0,1]$, i.e. $1 - F_X(x) \sim U(0,1)$. Thanks to this probabilistic result one can conclude that the survival probability of individual i at its true event time is distributed according to a uniform random variable taking values on $[0,1]$. Hence,

$$S_i(T_i^*) = U_i \sim U(0,1) \tag{6.1}$$

From 6.1 one can then generate the survival times T_i^S for each subject by using the relation between the survival function $S_i(\cdot)$ and the cumulative hazard function $H_i(\cdot)$, inverting the latter.

6.1.2 The proportional hazards data generating process

It is possible to extend these aforementioned results to generate times-to-event to the case of a proportional hazards (PH) model with baseline fixed covariates which is generally referred to as *cumulative hazard inversion method* ([81]). Under a PH model the survival probability for individual i at its event time T_i^* can be expressed as:

$$S_i(T_i^*) = \exp(-H_0(T_i^*) \exp(\gamma' w_i)) \tag{6.2}$$

where $H_0(t) = \int_0^t h(s)ds$ is the cumulative baseline hazard evaluated at time t , and x_i is a vector of covariates with associated population-level (i.e. fixed effect) parameters γ . This is because the PH assumption also implies proportional cumulative hazards. Hence, using equation 6.1, the survival probability $S_i(T_i^*)$ can be replaced by the uniform random variable U_i . Moreover, since the objective is to simulate a new event time for individual i ($i = 1, \dots, N$), rather than to evaluate the survival probability at the known true event time, one can replace T_i^* with the simulated event time T_i^S . This leads to

$$U_i = \exp(-H_0(T_i^S) \exp(\gamma' x_i))$$

To obtain the formula for the cumulative hazard inversion method, one simply needs to rearrange the latter equation to solve for the event time ([82]):

$$T_i^S = \exp(-\log(U_i) \exp(-\gamma' w_i)).$$

This method allows one to simulate event times under a PH model data generating process with any parametric formulation for the baseline hazard: all that is required is an invertible cumulative baseline hazard function.

For the purpose of exposure, we report the mathematical relations that can be used to simulate times-to-event assuming an exponential baseline hazard:

$$\begin{aligned} T_i &\sim \text{Exp}(\lambda) \\ H_0(t) &= \lambda t \\ H_0^{-1}(t) &= \lambda^{-1}t \\ T_i &= -\frac{\log(U_i)}{\lambda \exp(\gamma' w_i)} \end{aligned}$$

Weibull and Gompertz baseline hazards corresponding equivalences can be found in [81].

6.2 Complex survival data generating processes: longitudinal and time-to-event data simulation

If one can obtain an algebraic closed-form solution for the inverse cumulative baseline hazard $H_0^{-1}(\cdot)$, then a major benefit of the *cumulative hazard inversion* method is that it is simple and computationally efficient. The method can be used to generate event times for a variety of parametric baseline hazards, including standard choices such as the exponential, Weibull or Gompertz distributions. However to simulate event times considering a time-dependent covariate poses many obstacles but, based on the cumulative hazard inversion method, an extension of this method to find closed-form expressions to simulate event times data under Cox-PH models with time-dependent covariates exists ([83]). Generally speaking, there are two possible obstacles that may arise in simulating survival data considering time-varying covariates. First, it's possible that the cumulative baseline hazard $H_0(t)$ cannot be solved in closed form. Secondly, it is possible that the cumulative baseline hazard cannot be inverted. This is due to the fact that the cumulative hazard integral is frequently an intractable integral due to the general shape of the time-dependency in the hazard function, which is dependent on the longitudinal sub-model specification. This is the case if one simulated data under a JM setting because of its complex structure linking the event and longitudinal sub-models. In particular to find a closed form expression for the inverse of its cumulative baseline hazard may be a challenging computational task since it involves many terms and the integration over the random effects.

Within our simulations where we generated longitudinal and time-to-event data under a JM setting, we deployed an extension of the cumulative hazard inversion method ([84]). This extension allows for more adaptable baseline hazard functions and offers a much more broad framework that can be expanded beyond the typical parametric distributions. The method includes numerical quadrature and/or numerical root discovery. Numerical approximations can be employed when an analytical form for the inverted cumulative baseline hazard cannot be obtained and needed when it is necessary to compute complex integrals due to the time-dependent covariate.

Indeed, the root finding is used to numerically solve for T_i^S in circumstances where the cumulative baseline hazard function cannot be inverted analytically, as in our case. A convenient choice in this case is the univariate root finder algorithm ([85]), which guarantees to find a solution to the equation:

$$S_i(T_i^S) - U_i = 0$$

or, equivalently,

$$\exp(-H_i(T_i^S)) - U_i = 0$$

by treating T_i^S as the single unknown. The quadrature is used to numerically evaluate the cumulative hazard function in this setting, where it does not have a simple and tractable form. A standard choice of algorithm is either Gauss-Legendre or Gauss-Kronrod quadrature, whereby the cumulative hazard $H_i(T_i^S) = \int_0^{T_i^S} h_i(s)ds$ can be approximated by

$$H_i(T_i^S) \approx \frac{T_i^S}{2} \sum_{q=1}^Q \nu_q h_i\left(\frac{T_i^S(1+z_q)}{2}\right)$$

where ν_q and z_q are, respectively, the standardised weights and locations ("abscissa") for the $q = 1, \dots, Q$ quadrature nodes. The method involves iterating between numerical quadrature and numerical root finding until the root finding equation is properly solved and an appropriate solution for t is obtained.

6.3 Simulation study for longitudinal and time-to-event data in the presence of different missing processes on the longitudinal covariate

The aim of our simulation study was to evaluate the extent of bias of the estimated Hazard Ratio in the ECM in the JM to assessing the strength of the association between the marker and the time-to-event, when the longitudinal profile of the marker, considering two different functional forms (linear and quadratic), undergoes different missing processes. Therefore, the focus is on the ability of these two data modeling approaches to correctly estimate the α parameter, following notation of chapter 4, when the longitudinal trajectories of the markers are affected by missing data in different ways.

6.3.1 Simulation of longitudinal and event times data

The data were generated under the following univariate JM with a current value association between the PH event sub-model and the longitudinal sub-model, specified through a polynomial trajectory in time with individual random effects and fixed baseline effects.

- The longitudinal sub-model took the form:

$$\begin{aligned}
Y_i(t) &\sim \mathcal{N}(\mu_i(t), \sigma_y^2) \\
y_i(t) &= \mu_i(t) + \epsilon_i(t) \\
\mu_i(t) &= \beta_{0i} + \beta_{1i}t + \beta_{2i}t^2 + \gamma_1x_{1i} + \gamma_2x_{2i} \\
\beta_{0i} &= \beta_0 + b_{0i} \\
\beta_{1i} &= \beta_1 + b_{1i} \\
(b_{0i}, b_{1i})^T &\sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_b^2) \quad , \quad \epsilon_i(t) \sim \mathcal{N}(0, \sigma_y^2) \\
x_{1i} &\sim \mathcal{N}(\nu, \sigma_\nu^2) \quad , \quad x_{2i} \sim \mathcal{B}(p)
\end{aligned} \tag{6.3}$$

Each observed biomarker value $\{y_i(t) : i = 1, \dots, N; t \in [0, 14]\}$ was the result of the "true" (unobserved) biomarker level $\mu_i(t)$ plus a random measurement error $\epsilon_i(t)$, whose variability is controlled by the variance parameter σ_y^2 . The "true" marker level $\mu_i(t)$ was the composite result of an intercept value β_0 and a temporal trend $\beta_1t + \beta_2t^2$, the individual random effects (b_{0i}, b_{1i}) and the fixed baseline effects γ_1 of a continuous covariate x_1 and γ_2 of a dichotomous covariate x_2 . These baseline covariates were independently drawn from a Normal distribution $\mathcal{N}(\nu, \sigma_\nu^2)$ and a Bernoulli distribution $\mathcal{B}(p)$, respectively. The random effects represented the individual random deviations from the population-level fixed trajectory: they entered into the model as a random intercept b_{0i} and a random slope b_{1i} , drawn from a multivariate normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_b^2)$, centered around the $\mathbf{0}$ vector and with 2 by 2 inter-subject variance-covariance matrix $\mathbf{\Sigma}_b^2$, whose diagonal elements represent their variances, and the anti-diagonal entries their correlation.

- For the event sub-model we posited:

$$h_i(t) = \lambda(t^{\lambda-1})\exp(\theta_0 + \theta_1w_{1i} + \theta_2w_{2i} + \alpha\mu_i(t))$$

that is, a PH model with a Weibull baseline hazard, where λ is the shape parameter and θ_0 is the scale parameter on log hazard scale (intercept term), and θ_1 and θ_2 are the true coefficient (log hazard ratio) for the continuous and binary covariate in the event sub-model. The α parameter stands for the association coefficient between the two sub-models.

To simulate longitudinal and time-to-event data under this framework we used the R package *simjm* ([86]). We could have used, in principle, also the R package *simsurv* ([82]) to simulate survival data. However, utilizing the *simsurv* package to generate survival data under a JM setting necessitates writing a substantial amount of computational code in R. To facilitate this process, *simjm* ([86]) was developed: a user-friendly interface is provided by the package, which serves as a wrapper for the *simsurv* package and is designed specifically for simulating longitudinal and times-to-event data under a JM specification.

6.3.2 *simjm*: a R package for simulating Joint longitudinal and time-to-event data

The use of the *simjm* package is briefly demonstrated below through an example. Arguments are provided in the call to the *simjm()* function which, in this instance, produces one of the complete case scenarios considered within the simulations.

```
# Use simjm to generate Joint Longitudinal and Event data
simdat <- simjm(M = 1,
               n = 400,
               max_fuptime = 14,
               family = gaussian(),
               betaEvent_aux = 1,
               max_yobs=10,
               balanced=FALSE,
               fixed_trajectory = "quadratic",
               betaLong_quadratic = 0.1,
               betaLong_linear = 0.25,
               betaLong_intercept = 10,
               betaLong_continuous = 1,
               mean_Z2 = 2,
               sd_Z2 = 3,
               betaLong_binary = -1,
               prob_Z1 = 0.5,
               random_trajectory = "linear",
               b_rho = 0,
               b_sd = c(1.5,1.5),
               betaEvent_assoc = 0.25,
               betaEvent_binary = 0,
               betaEvent_continuous = 0)
```

The number of distinct markers is set to be one. The longitudinal data (marker profiles) for 400 individuals are generated as random draws from a (conditional on covariates and random effects) Gaussian distribution specified in the data generating model, incorporating measurement error $\epsilon(t) \sim \mathcal{N}(0, 1)$; then it is possible to specify the maximum number of measurement occasions, with balanced or unbalanced visit times. The functional form of the marker can be quadratic or linear and accordingly the "true" values are chosen for each of the parameters to generate the population-averaged trajectory. Then the fixed effects of baseline covariates which follow a Gaussian and a Bernoulli distribution, respectively, are set. The structure of the, correlated or uncorrelated, random effects can be linear or quadratic as well: in the former case, in addition to a random intercept, random slopes will also be generated for the linear effect of time on the marker, while, in the latter case also the subject-specific quadratic trend will be composed of a population-level fixed effect and a random one. The values of the standard deviations for each random effect considered in the longitudinal data-generating model must then be supplied. Then the true values of the parameters for the event sub-model are specified: the value of the association parameter α representing the HR and quantifying the impact of the marker on event risk over time and the fixed effects at baseline of the same continuous and discrete covariates used for generating marker profiles, but which may have a different marginal effect on event risk (in

this example equal to zero). The baseline hazard is set to Weibull by default, so it is not necessary to specify it explicitly (as would be necessary if *simSurv* were used directly), and it is possible to provide the scaling and shape parameters that characterize the distribution (omitted here, retaining their default value). The simulated data are returned in a separate data frame for each sub-model, that is, one for the longitudinal sub-model and one for the event sub-model: the first data frame shows the multiple-row per-individual marker data; the second data frame shows the single-row per individual time-to-event data.

The simulation routine just described allows to consider an administrative censoring that coincides with the specified time duration for follow-up. Then, all the subjects that do not die within this time window of observation are followed up the end of the study (`max.fuptime`). However, since survival times are commonly affected by right censoring, in order to be consistent with a more realistic scenario, a random censoring mechanism was manually introduced by considering a right-censoring time random variable C_i^S , uniformly distributed on the range $[0, 10 * \text{max.fuptime}]$, getting around 10% of the patients censored. Then, since the event data of each patient is described through the random couple (T_i^S, δ_i^S) , the time-to-event T_i^S will be set to be equal to $\min(T_i^S, C_i^S)$ and the event status δ_i^S is modified according to:

$$\begin{cases} \delta_i^S = 0, & \text{if } C_i^S \leq T_i^S \\ \delta_i^S = \delta_i^S, & \text{else} \end{cases} \quad (6.4)$$

Recall that, either $\delta_i^S = 0$, if the patient is censored during the follow-up or he/she had no event within it, or $\delta_i^S = 1$ if the patient experiences the event. T_i^S and C_i^S are assumed to be independent conditional on the biomarker trajectory $\mu_i(t)$, as commonly done in survival analysis.

6.3.3 Missing processes on simulated data

In medical studies, longitudinal responses $y(t)$ are prone to be missing. The missingness of some data gives rise to “incomplete” longitudinal data. The nature of the missing data patterns is a major concern in these kind of studies and, generally, some assumption have to be made about the nature of the missingness before conducting analyses ([87]). From a theoretical point of view, a taxonomy of missing processes in the longitudinal context was detailed in chapters 2 and 4. As already said in chapter 4, a practical issue in the handling of missing data in longitudinal outcomes is the fact that the observed data alone cannot precisely distinguish on the nature of missing generating process ([47]) but, in the simulations, we could manage the missing mechanism and in this thesis we focused on the case of Missing Not at Random (MNAR). In particular, we considered different scenarios in which a specific region of the overall distribution of the data is more likely to be affected by missingness, depending on the value of the longitudinal observations themselves. We used a probabilistic approach to generate the missing data and, in particular, recreated the conditions under which higher marker values were more likely to be missing. Thus it was assumed, that in addition to being predictors of the event, higher values of the marker were also more likely to be missing, due to the fact that we set a positive value for the α parameter linking the longitudinal and event processes. To generate the missing data on the simulated longitudinal data, we relied on the *missMethods* R package ([88]) and a quick overview of which function we used will be provided in the next subsection.

Starting from the “complete case” longitudinal data, we branched out into two possible scenarios. In the first case, after the missing process (“amputation”) is applied, the biomarker profiles lose some observations, resulting in *Intermittent* patterns. In such a situation, the marker $y_i(t)$ has some observed values $y_i^o(t)$, while some others are unobserved, i.e. $y_i^m(t)$. In the second case, instead, we took into account a scenario in which as soon as a visit is missed, that is after the first missing value in the biomarker profile, the following values are no longer observed. In such a situation, that we called “*lost assessment*” case, the marker $y_i(t)$ still has some observed values $y_i^o(t)$, and unobserved values $y_i^m(t)$, but the resulting profile is of the type $y_i(t) = (y_i^o(0), \dots, y_i^o(t_j), y_i^m(t_{j+1}), \dots, y_i^m(T_i^S))$. In figure 6.1 we make use of a graphical representation to indicate which were the two types of missing

mechanisms considered. Thereby the distributions of the "complete case" marker profile $y_i(t)$ and after amputation are different. In both cases, the random pair (T_i^S, δ_i^S) is still available for the analysis, i.e. the state and time at the patient's event are observed and unaltered. In other words, the missingness of some values in the individual markers profiles does not imply the drop-out of the individual. We emphasize that we induced longitudinal data amputation starting from the third value to avoid patients with only one observation, especially in the second scenario. This setting was chosen to be logically consistent with the application of JMs with intercept and especially random slope; thus we simply wanted to accommodate the idea that to draw a straight line, estimated for each subject, we at least needed two points in the space of observations for each individual marker.

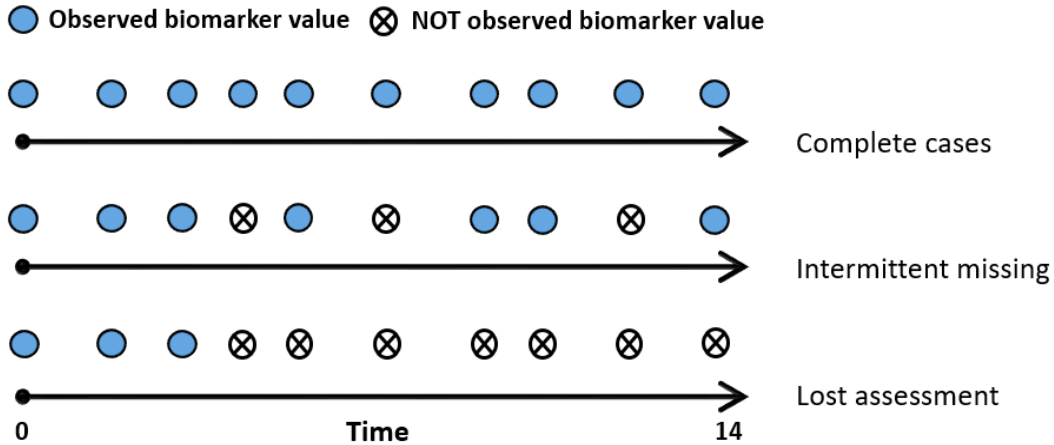


Figure 6.1: The two missing mechanisms on a fictitious longitudinal profile

6.3.4 missMethods: a R package to simulate missing data

In order for "high" marker values to be more prone to be missing than "low" values, we used a procedure called `delete_MNAR_1_to_x()` of the R package `missMethods` ([88]). This routine allowed us to obtain different missing scenarios, by controlling for the following arguments:

- *missing variable*: the variable ending with missing values
- *cutoff*: splitting point dividing the data into two parts
- *odds*: odds of missing in one group of the split data
- *where*: controls where missing values are created ("lower", "upper" or "both")
- *Missing Rate (MR)*: the final percentage of missing data

We now briefly detail the functioning of this procedure. At first, the rows of the longitudinal data returned from the `simjm` procedure (see subsec. 6.3.2) are divided into two groups based on the cutoff value, used to calculate the splitting point for the variable that is going to be amputated. The group A consists of the rows whose values of the unobserved marker values (measurement error-free) are below the calculated cutoff value. The group B consists of the remaining rows, which values are above the cutoff point. Then, each observation is assigned a probability of being missing such that the odds for a value belonging to group B to be unobserved are $x : 1$ compared to an observation in group A. That means that the probability for a value to be missing in group B divided by the probability for a value to be missing in group A equals x (*where="upper"*). Finally, the MR parameter allowed to specify the percentage of missing data of longitudinal markers.

6.4 Simulation design

In order to make explicit and accessible the simulation setup in this section the parameters used to generate the data, following the same notation as in subsec. 6.3.1, and the methods used to analyse them are listed.

6.4.1 Simulation parameters

We drew $n_{sim} = 1000$ simulations for each scenario. We generated longitudinal and times-to-event data for $N = \{150, 400\}$ subjects in the linear case and $N = \{400\}$ in the quadratic longitudinal marker case, with unbalanced (non-regular) measurements occasions over a follow-up period of 14 days, with a maximum number of observations of 10 per patient.

- *Longitudinal sub-model.* We considered two different functional forms for the longitudinal markers: a *linear* temporal trend and a *quadratic* one, with (uncorrelated) random intercept and random slope in both cases.
 - Linear case: the fixed effect parameters of the longitudinal sub-model were chosen to be $\beta_0 = 10$, $\beta_1 = 0.25$, $\beta_2 = 0$.
 - Quadratic case: the fixed effect parameters of the longitudinal sub-model were chosen to be $\beta_0 = 10$, $\beta_1 = -1$, $\beta_2 = 0.10$.

The β parameters, different for the two functional forms considered, were chosen to obtain similar ranges spanned by the markers. The regression coefficient of the continuous and discrete variables (x_1, x_2) in the longitudinal sub-model were chosen to be $\gamma_1 = 0.75$, $\gamma_2 = -1$, respectively, in both cases. The baseline covariates x_1 and x_2 were independently drawn from a Normal $\mathcal{N}(2, 3)$ and a Bernoulli $\mathcal{B}(0.5)$ distribution, respectively. The random intercept b_{0i} and the random slope b_{1i} , for each individual, were generated from a Normal distribution $\mathcal{N}(0, \Sigma_b^2)$, with different variances σ_b^2 , specifying what we called a "low" and a "high" variability case for the longitudinal markers. For the former, we specified Σ_b^2 having $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 0.25$ and for the latter $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 2.25$. In both cases we set correlation of random effects $\rho_{b_0 b_1} = 0$. The value of each random measurement error $\epsilon_i(t)$ was drawn from a Normal distribution $\mathcal{N}(0, 1)$. This can produce a relatively large measurement error, but it can be considered reasonable from an empirical point of view. In fact, overall, no significant discrepancy was found between the actual values of the markers and the corresponding version contaminated by the measurement error random.

- *Event sub-model.* The parameters of the Weibull baseline hazard were kept to their default values, i.e. $\lambda = 1.2$ and $\theta_0 = -7.5$.

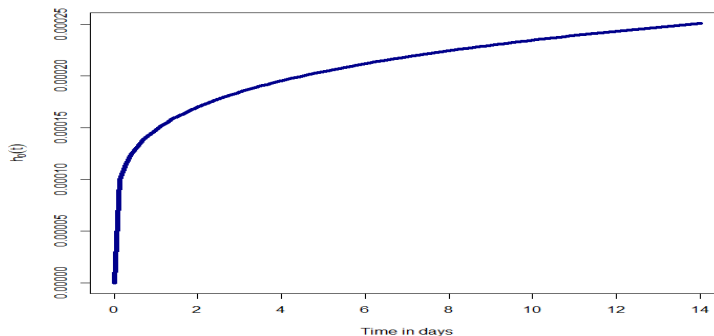


Figure 6.2: Weibull baseline hazard deployed for simulations

The fixed effects for the covariates w_1 and w_2 in the event sub-model were set $\theta_1 = \theta_2 = 0$. Finally, we set the true HR $\alpha = 0.25$, which corresponds to an increase in risk due to the unit increase in the marker of approximately 30%.

- *Missing process.* As previously pointed out (subsec.6.3.4), the data amputation process, basically, relies on the specification of three parameters to indicate the missing rate (MR), a cutoff value and the odds. Within the simulations, we considered two possible MR values, namely 30% and 40%; a single cutoff value equal to the 70th percentile of the overall distribution of the values taken by the markers of all subjects; two different odds values, 4 and 6, which let that marker values above the cutoff have an increasing probability to be missing.

6.4.2 Generated longitudinal data

Complete case

The simulated marker trajectories (N=400) with linear and quadratic functional form and two different random effects distributions are displayed in the spaghetti plots below. Profiles of those who were censored or still alive at the end of the follow-up are shown in light grey, while of those who died in dark grey. The population-averaged trajectory, i.e. the dynamic evolution of the longitudinal marker when the random effects are zero (and no baseline fixed effects), is represented by the black dashed line.

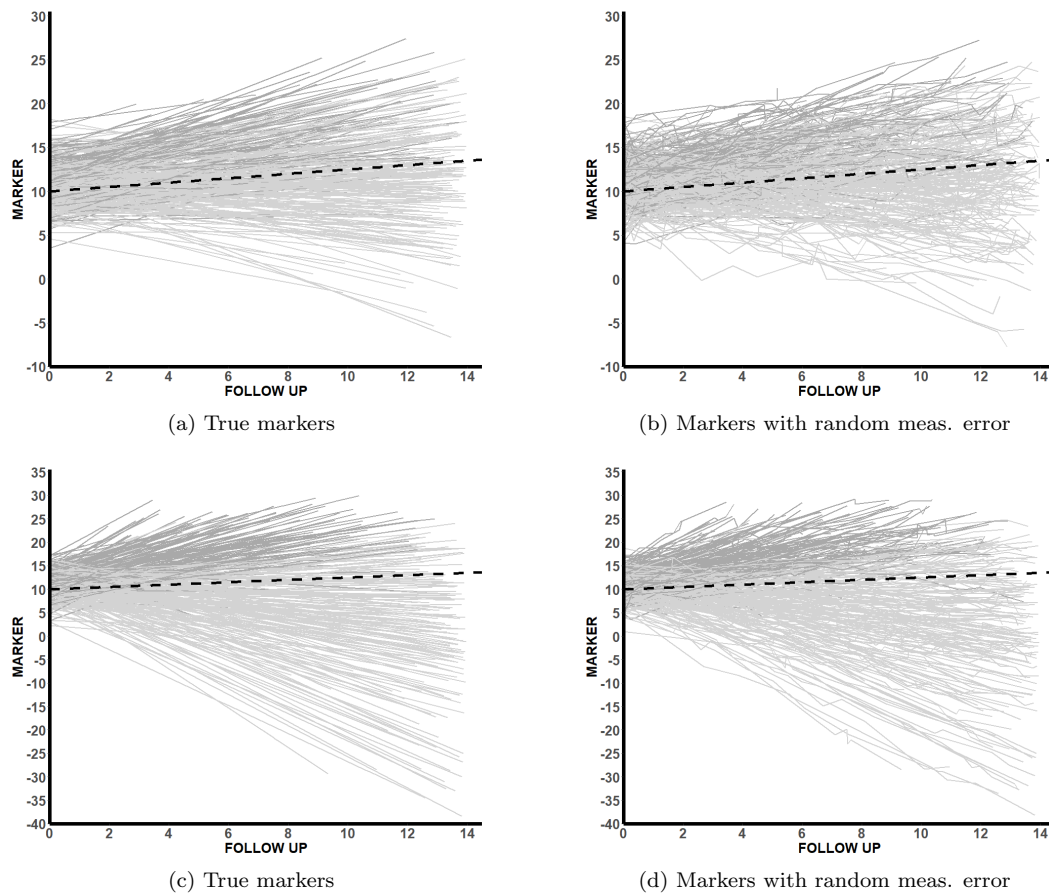


Figure 6.3: True and random measurement error affected markers trajectories with linear temporal trend with $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 0.25$ (first line) and $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 2.25$ (second line)

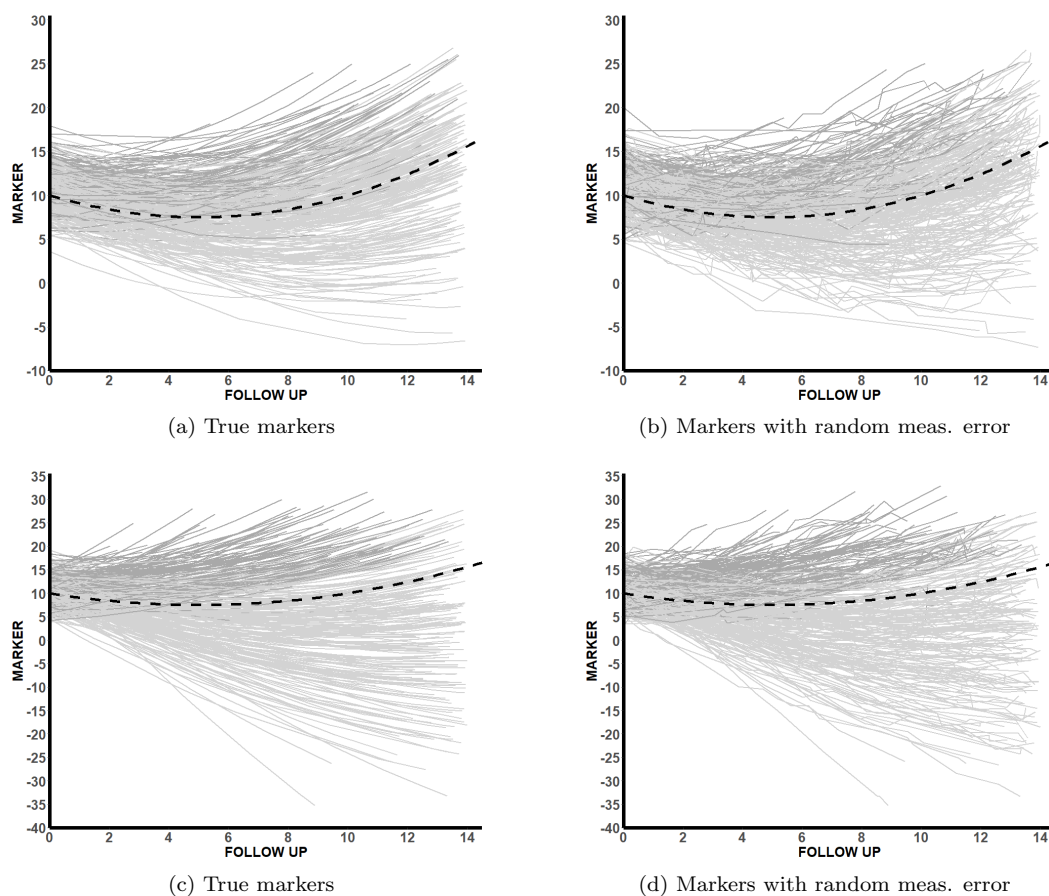


Figure 6.4: True and random measurement error affected markers trajectories with quadratic temporal trend with $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 0.25$ (first line) and $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 2.25$ (second line)

Intermittent missing and Lost Assessment

Within the simulations, we considered a great number of different settings so it would not be feasible to show all the generated markers trajectories after amputation. Nevertheless, we display here a few examples for the purpose of exposition. We graphically report the spaghetti plots referring to the "complete case" shown above (linear and quadratic, low and high variability) after they underwent data amputation. Side by side, the markers in the case of intermittent missing and "lost assessment" are arranged. The MR missing rate was 30% and the missing odds ratio was 4. The population-averaged trajectory of the complete case scenario is represented by the black dashed line.

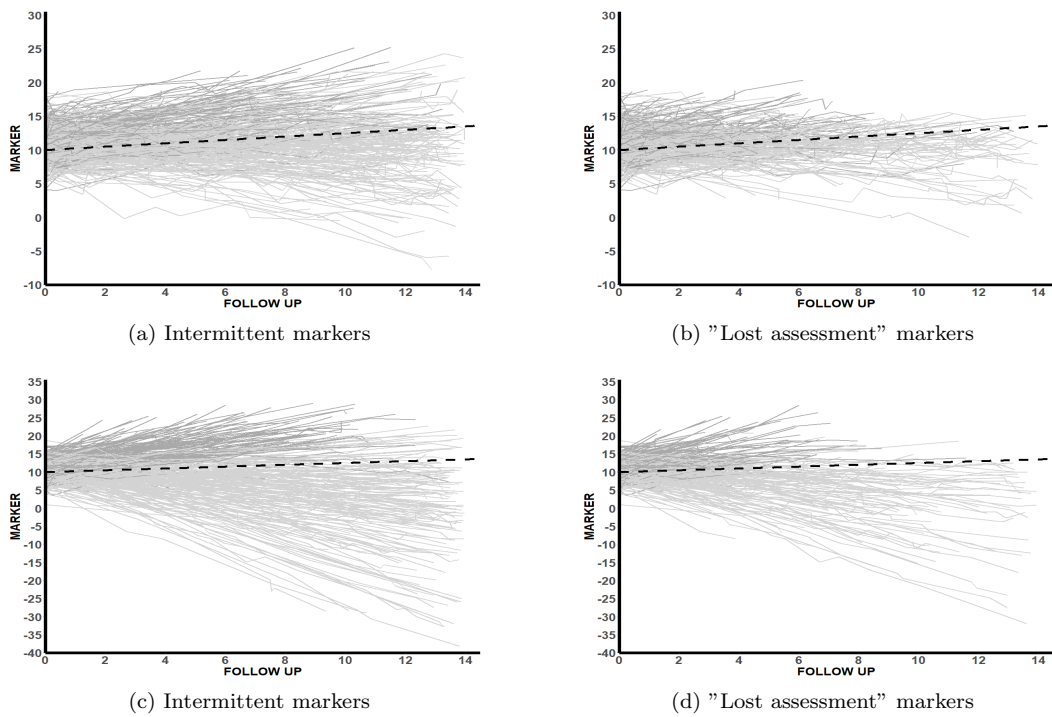


Figure 6.5: Intermittent missing and "lost assessment" for linear marker trajectories with $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 0.25$ (first line) and $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 2.25$ (second line) ($MR=30\%$, $odds=4$)

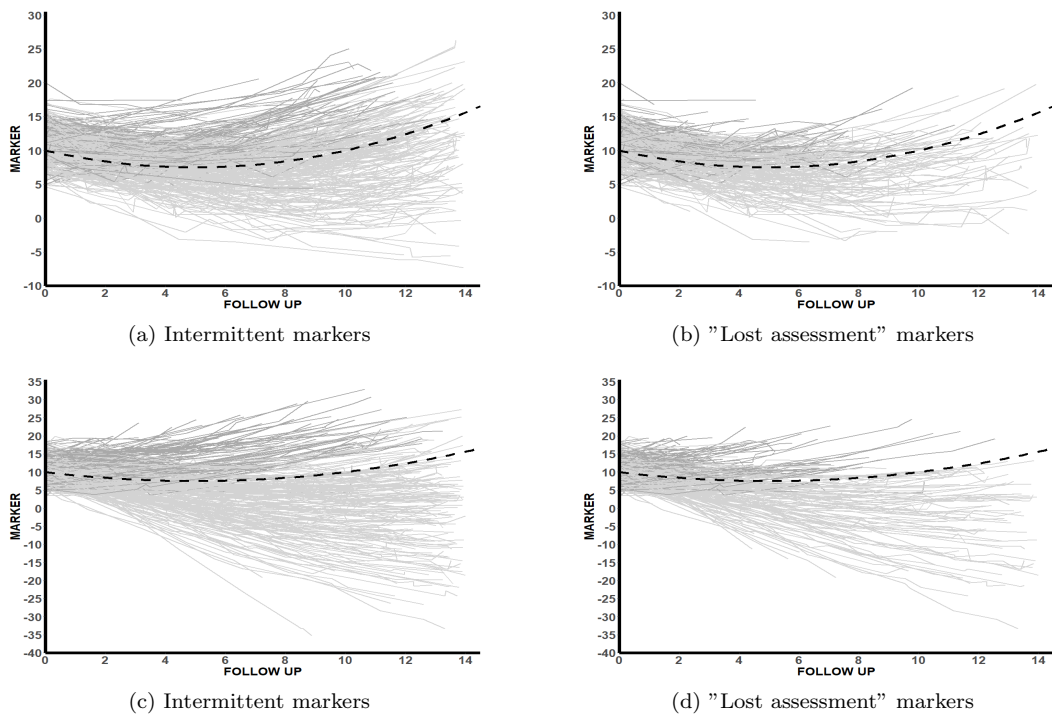


Figure 6.6: Intermittent missing and "lost assessment" for quadratic marker trajectories with $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 0.25$ (first line) and $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 2.25$ (second line) ($MR=30\%$, $odds=4$)

6.4.3 Methods

The modeling approaches used throughout the simulations were: i) the ECM considering the observed updated value of the marker; ii) three different JMs considering a piecewise constant baseline hazard with 7 internal knots, an approximated baseline hazard considering B-splines with 9 internal knots and a Weibull baseline hazard considering the updated value of the marker. For the linear case, the ECM and the three types of JM corresponding to the three different baseline hazards were fit on the simulated data; for the quadratic shape, on the other hand, the ECM and just the JM with baseline hazard approximated by the B-splines were deployed, by virtue of the fact that overall with this specification we had better performances compared to the other options, thus obtaining a discount in terms of computational time.

The ECM was run using the *survival* R package ([89]) and the JMs were fitted using the R package *JM* ([90]). The *JM* package offers two options for numerical integration: the standard Gauss-Hermite rule and the pseudo-adaptive Gauss-Hermite rule. It has been shown that the latter can be more effective than the former in the sense that typically fewer quadrature points are required to obtain an approximation error of the same magnitude and computational burden is reduced ([8]), so that this procedure was used for the JMs fitted on the simulated data. For all other options in JM as well as in ECM, the defaults were used in the analyses. The R code used for simulations can be found in Appendix B.

6.4.4 Statistical indices to evaluate estimation performance

We summarized the results using: mean and median, bias and percentage bias, Mean-Squared Error (MSE), Asymptotic Standard Error (ASE), Empirical Standard Error (ESE) and 95% coverage probabilities (CP) of the simulation estimates of the association parameter α , as recommended in [91]. The ASE was computed as the average of the estimated standard errors, and the ESE as the standard deviation of the estimates of α .

Here we explicit the mathematical expressions of the simulations summary measures in order to be accessible to the reader: α indicates the true parameter value linking the longitudinal and event sub-models in the data generating process and $\hat{\alpha}$ the estimated α parameter obtained after fitting ECM or JMs on each of the simulated datasets.

- Mean($\hat{\alpha}$)

$$E[\hat{\alpha}] = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \hat{\alpha}_i$$

- Bias

$$E[\hat{\alpha}] - \alpha = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \hat{\alpha}_i - \alpha$$

- %Bias

$$\frac{\text{Bias}}{\alpha} \times 100$$

- Asymptotic Standard Error (ASE)

$$\frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \text{Std.Err.}(\alpha_i)$$

- Empirical Standard Error (ESE)

$$\sqrt{\text{Var}(\hat{\alpha})} = \sqrt{\frac{1}{n_{sim} - 1} \sum_{i=1}^{n_{sim}} (\hat{\alpha}_i - E[\hat{\alpha}])^2}$$

- Mean Squared Error (MSE)

$$E[(\hat{\alpha} - \alpha)^2] = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} (\hat{\alpha}_i - \alpha)^2$$

- CP

$$P[\hat{\alpha}_{low} \leq \alpha \leq \hat{\alpha}_{upp}] = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \mathbb{1}(\hat{\alpha}_{low} \leq \alpha \leq \hat{\alpha}_{upp})$$

where $\hat{\alpha}_{low}$ and $\hat{\alpha}_{upp}$ are the 95% CI boundaries of the estimated HRs.

The results of the simulations can be found in tables 6.1, 6.2, 6.3, 6.4, 6.5, 6.6 at the end of this chapter. The tables of results will report by rows the proportion of missing data (MR), the actual proportion of missing data (eff. MR) different between "intermittent" and "lost assessment" missing data patterns, indicated by the letters F and T in the "Lost" column, respectively, and the odds of missing (see subsec. 6.3.4).

6.5 Results

6.5.1 Low variability of longitudinal markers

In the "low" biomarkers volatility case, achieved by considering small magnitudes for the variances of the random effects ($\sigma_{b_0}^2 = \sigma_{b_1}^2 = 0.25$), looking at the simulations results there was no evidence to declare a glaring superiority of one approach in terms of estimation of the HR when the marker profiles followed a linear trajectory. In the quadratic case, however, the two modeling techniques showed different performance, with JM more robust than ECM. Distinguishing by the functional forms of the longitudinal markers, we had the following:

- Linear case (tables 6.1, 6.3): with complete case data, both ECM and JM estimated correctly the HR in all the explored scenarios, for all different levels of missing rate, whether "intermittent" or "Lost to Assessment" data, the two modeling strategies showed good performances in terms of HR estimation (α), with a small amount of bias in general and a coverage of the confidence intervals always near to the nominal value of 95%. Overall, the JMs, showed slightly higher absolute bias in "lost assessment" case, than in the case of intermittent missing data at the same MR levels. With small sample size (N=150) the JM, especially with a piecewise approximated baseline hazard, had a mildly higher amount of bias than ECM but, this already small amount of bias decreased, however, with a larger sample size (N=400). Hence, both ECM and JMs were robust here and this suggested that using the LOCF method or an interpolation of the individual trajectories, as done by JMs, had no substantial difference, from a practical point of view, if the aim of the study was to estimate the average HR on time. The results were fairly easy to discern, taking into account the fact that there is no obvious difference between using a stepwise approximation of the profile or an estimated trajectory obtained from the mixed-effects regression when there is a low variability of the markers.
- Quadratic case (table 6.5): with complete case data, both ECM and JM estimated well the HR. Overall when longitudinal marker data were intermittently missing, the ECM performed worse than JM, with non negligible extents of bias in almost all cases. In fact, in these cases, the ECM vastly underestimated the HR value, with very low coverage probabilities, in contrast to JM which showed excellent performance, even with high values of MR in the data. In the "lost assessment" scenario, the JM did not seem to be affected in terms of estimation ability; instead, ECM "improved" its performances. This latter result was not expected and, at the moment, we can only justify it from an heuristic point of view (see section 6.6).

Regarding the framework of low variability, when the longitudinal trends of the marker resembled a linear shape, we found that ECM and JMs were robust in terms of estimation of the association between the time-dependent covariate and the risk of event over time, both in the case of intermittent missing data and "lost assessment". However, this result was not confirmed in the quadratic case where only the JM showed good performances overall. ECM obtained biased estimates of the HR, particularly in cases of intermittent missing data.

6.5.2 High variability of longitudinal markers

In the "high" marker volatility case, achieved by considering larger values for the variances of the random effects ($\sigma_{b_0}^2 = \sigma_{b_1}^2 = 2.25$), overall, the ECM and JM, showed fairly good performances in the linear case of the longitudinal covariate but there was a relative tendency in the coverage of both to decrease, below the 95% value, in both the intermittent and the "lost assessment" data scenarios. The performances deteriorated severely in the quadratic case. Distinguishing by the functional forms of the longitudinal markers, we had the following:

- Linear case (tables 6.2, 6.4): with complete case data, both ECM and JM showed good performances in terms of HR estimation. Overall the JMs, as expected, had better performances than the ECM in the case of intermittent data, for both sample sizes. The robustness was less and less evident as the percentage of missing data increased, although, especially for the JMs, evidently poor performances were only observed in the case of MR of 40%: these results led us to conclude that, in such cases, these models may exhibit less than satisfactory performance and, as expected, things get worse considering increasing missing rates in biomarker values. Again, the ECM surprisingly seemed to perform better with respect to expectations in the case of "lost assessment", although with coverage values still below 95%.
- Quadratic case (table 6.6): in the scenario with complete data both approaches had good performances. However, with intermittent longitudinal data, the coverage of the confidence intervals of both approaches, presented overall lower values than in the situations previously considered. The worst performances were recorded by the ECM, whose coverage probabilities settled at values far away from 95% and underestimated more severely the α parameter, except that for one case. Also with JM emerged a slight tendency to underestimate the α parameter: particularly, with MR values around 30% and intermittent missing data this approach still retained some degree of robustness but, increasing MR, the negative bias extended. In the "lost assessment" case the greatest inefficiencies were noticeable also for JM; with MR of 40% its performance had become inadequate with non-negligible %bias values and very low coverage probabilities. The ECM, on the other hand, again showed somewhat better performance in the "lost assessment" scenario than the corresponding intermittent data cases: overall the α parameter was increasingly overestimated as MR and missing data odds increased, absolute bias got smaller and coverage probabilities increased as well.

Regarding the framework of high variability in the longitudinal markers, it might be concluded that ECM and JM are less robust in terms of HR estimation with respect to the low variability case, both in the case of intermittent missing data and "lost assessment" following the first missing value, when the temporal trends resemble a linear shape. In the quadratic case we found certainly worse performances, both for ECM and for JM. In the case of intermittent data, in particular, both methodologies have rather negative bias and poor coverages. In the "lost assessment" case, as expected, the JM performed even worse, while the ECM tended to overestimate the HR, resulting in apparently better results.

6.6 Discussion

In this simulation work, we set ourselves the objective of evaluating the robustness of the ECM and the JM in estimating the level of association between a longitudinal marker and the risk of event over time, through the HR, when the first one is affected by different missing processes. Two functional forms for the longitudinal covariate, i.e. linear and quadratic time trends were considered, collected on 400 individuals at different occasions. As a starting point we also considered a lower sample size ($N=150$), only for the linear trajectory case. In all the complete case scenarios, both ECM and JMs performed well independently of the sample size and functional forms of the longitudinal trajectories, with very similar results, validating the usage of both the modeling approaches from a practical point of view. We then perturbed the complete case distribution of longitudinal observations by generating missing data in different ways. We designed two possible scenarios considering different ways on how missingness acted on longitudinal data: one in which the longitudinal profile was subjected to intermittent amputations and, a second case, in which the marker was no longer observed after the first missed measurement occasion, resulting in "truncated" versions of the longitudinal marker profiles. In both settings the missing probability was higher for marker values higher than a certain level (i.e. 70th percentile of the overall distribution) and we considered two global missing rates (MR), i.e. 30% and 40%.: then just some of the longitudinal profiles were altered, but many other were unchanged.

Overall, the results of the simulations testified to a good degree of robustness, for both ECM and JMs (independently of the specified baseline hazard shape), whenever the individual trajectories followed a linear trend over time, with low variability in the observed marker values, obtained by small values for the random effects distribution variance. Unlike ECM, for which a consistent negative bias showed up with longitudinal quadratic profiles with intermittent missing, the JM confirmed its good properties, regardless of the MR. Once we shifted our attention to longitudinal markers with higher volatility, we observed some tendency for both techniques to perform worse, even in the linear longitudinal functional form case. This behavior was largely accentuated in the case of quadratic shape of the longitudinal profiles, which led us to think that we witnessed a distinct estimation capability due to some extent to the functional form of the trajectories. It is important to note that in the case of intermittent missing data our findings were expected because an underestimation of α was coherent with "misclassified" trajectories, as somehow shown for the ECM in the presence of measurement error ([8],[92]), and, by analogy, with the effect of non-differential misclassified exposures in epidemiological studies ([93],[94]). In fact, in the case of non-differential misclassification a "bias towards the null" is expected ([93],[94]).

In the case of "lost assessment", however, we obtained results regarding the ECM, that often deviated from what we expected, and we had difficulty providing absolutely convincing justifications for the causes of these unexpected results. Nevertheless we are inclined to believe that there was some degree of interplay between the underestimation due to misclassified longitudinal profiles and an overestimation because of the truncated trajectories. We surmised that the reasons for these results were thus attributable to a kind of "compensatory effect" because, for truncated trajectories, the LOCF approximation caused them to settle to lower values, on average, resulting in less distinct and less scattered profiles between events and non-events. This overall lowering of trajectories, that mainly affected those who died being more likely to have profiles with large values by design, corroborated an expected overestimation of HR. In fact, originally large marker values, because they were subsequently missing, were approximated by smaller values, underestimating the (unobserved) trajectories while the same event process was observed throughout the time (and addressed in the ECM estimation stage). If such a speculation found room already after we tried to interpret the results in the linear trend cases, it was reinforced after conducting the simulations for the quadratic case. Indeed, here we witnessed an even more serious underestimation of HR in the case of intermittent data, but again with better results in the "lost assessment" scenarios. Therefore, it seemed sensible to assume that the truncation of the trajectories, in the quadratic case, was followed by an even more tangible compensatory effect precisely because, for the same time intervals, the biomarker would rise faster than in the linear case,

and the LOCF approximation diverged more from the true value.

We are fully aware that such heuristic speculations are rather partial, but such insights seemed reasonable and give us the opportunity to pave the way for future work. For this reason, we are even more aware that further analysis should be done, to better understand the matrix of these results and provide more robust and accurate explanations, and this goal is deferred to future work.

6.6.1 Limitations

The simulation work has several limitations. Firstly, we considered only a few scenarios, as the simulations were extensively time-consuming (~ 2 h for each data generation and ~ 6 h for each JM fitting): two different sample sizes ($N=150$; $N=400$) for the linear longitudinal sub-model, and only one sample size ($N=400$) when we specified a quadratic shape for the longitudinal profiles.

Among the possible limitations of this work we can certainly mention the approach we used to obtain the missing data. In fact, although through this methodology we managed to obtain the two types of missing data patterns we wanted, the imposition of a single cutoff may be seen as an inelegant solution as well as a limitation. Therefore it might have been sensible to consider other cutoff values, but it would have been very time consuming and we would not expect much differences in the results. Moreover, once the cutoff is set, all marker observations that fall above and below it have the same probability of missing as other observations in their own side without any possibility of further refining this mechanism. However, this approach was chosen for its easy implementation thanks to the dedicated R package (see subsec. 6.3.4) and still remains valid for reproducing a missing data generation process. Also, it is worth noting that, although the problem of missing data is common in longitudinal studies, the missing rate values we have chosen can be considered somewhat extreme, especially in the case of 40%. However, it is true that although true MR is not verifiable in practice with the available data, a 40% MR may simulate a situation where, as in our motivating clinical ICU setting, monitoring and subsequent information gathering may not be very frequent.

Another limitation of the work could be due to the fact that we considered only a single value for the α parameter, linking the longitudinal and the survival sub-models in the JM, but we limited ourselves to this single case both for reasons of time and because we do not expect that if we had changed the true value of the HR we would have obtained widely different results, although such scenarios remain unexplored at the moment.

Finally, we considered only two functional forms for the longitudinal profiles of the time-dependent marker but if we had broadened the functional possibilities, in addition to having to implement the new options, we would probably have complicated the scenarios too much which would have given rise to additional difficulties that we already had with linear and quadratic shapes in explaining some results. Finally, we set the effects of continuous and binary covariates in the event sub-model equal to zero, in order to achieve some benefit in terms of computational burden, leaving unexplored the case in which other baseline covariates affected the risk of event over time. Nevertheless, this was probably not a true limit since we did not expect different conclusions introducing other effects due to baseline covariate in the event sub-model during the data generating process, but at most an increase in the time needed to fit all the models.

6.6.2 Future works

In order to overcome the limitations of our simulations, we can think about future extensions and refinements of this work. First, it might be useful to consider another methodology to generate the missing longitudinal data: in particular, it might be useful to employ a marginal model to build the missing data generation process, using the longitudinal marker as the covariate responsible for assigning missing probabilities. This would give us greater control over the entire process, potentially helping us to better understand the reasons for some unexpected results. We are also interested in extending the simulation scenarios by consid-

ering other sample sizes, different follow-up periods, and other values for the α parameter. In particular it would be useful to see if the results in a mirror situation with negatively trending marker profiles and/or negative HR to evaluate whether our speculation on the interplay of under/overestimation was reasonable. In addition, we limited our simulations to only two functional forms but it would be of interest to examine other forms to better characterize the problem of missing data in longitudinal models.

Indeed, these are the goals we hope to achieve in the near future, as we believe that, despite the obvious limitations of the work, improvements might lead to more robust results with useful implications from a practical and methodological point of view.

Table 6.1: $n_{sim} = 1000$; $N = 150$; (true) linear marker profile $\mu(t) = 10 + 0.25t$; $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 0.25$. For JMs, the corresponding baseline hazard is reported. Column "Lost" (F, T) stands for (intermittent, "lost assessment") missing data scenarios.

Missing Rate (MIR)	Extended Cox Model										JM B-splines					JM Weibull													
	Est. (mean)	Est. (median)	Bias	%Bias	MSE	ESE	Coverage	Est. (mean)	Est. (median)	Bias	%Bias	MSE	ESE	Coverage	Est. (mean)	Est. (median)	Bias	%Bias	MSE	ESE	Coverage								
-	0.251	0.250	0.00	0.37	1.35	0.04	95.4	0.261	0.259	0.01	4.3	1.6	0.04	94.9	0.257	0.255	0.01	3.0	1.54	0.04	94.7	0.256	0.255	0.01	2.4	1.41	0.04	95.6	
0.30	F	0.248	0.246	-0.00	-0.60	1.44	0.04	95.9	0.255	0.253	0.00	2.04	1.52	0.04	96.2	0.252	0.249	0.00	0.68	1.50	0.04	96.3	0.250	0.248	0.00	0.19	1.42	0.04	96.0
	T	0.258	0.256	0.01	3.33	2.52	0.05	95.9	0.268	0.263	0.02	7.06	3.01	0.05	96.1	0.263	0.258	0.01	5.20	2.75	0.05	96.5	0.264	0.258	0.01	5.69	3.25	0.05	96.6
0.40	F	0.247	0.246	-0.00	-1.18	1.46	0.04	96.4	0.253	0.252	0.00	1.34	1.55	0.04	96.0	0.250	0.249	0.00	0.04	1.54	0.04	96.1	0.249	0.247	-0.00	-0.41	1.47	0.04	95.7
	T	0.263	0.258	0.01	5.13	2.88	0.05	95.7	0.270	0.267	0.02	7.89	3.23	0.05	96.0	0.265	0.262	0.01	6.13	2.96	0.05	96.6	0.266	0.261	0.02	6.55	3.01	0.06	96.1
0.40	F	0.245	0.243	-0.00	-2.02	1.56	0.04	97.3	0.254	0.251	0.00	1.54	1.73	0.04	96.4	0.251	0.248	0.00	0.41	1.72	0.04	96.1	0.250	0.248	0.00	0.15	1.68	0.04	95.7
	T	0.262	0.258	0.01	4.70	3.20	0.05	94.7	0.272	0.270	0.02	8.97	3.87	0.06	95.7	0.267	0.263	0.02	6.80	3.48	0.06	96.1	0.269	0.266	0.02	7.60	3.85	0.06	94.8
0.40	F	0.257	0.253	0.01	2.77	2.31	0.05	96.4	0.265	0.264	0.01	6.20	2.53	0.05	96.2	0.262	0.259	0.01	4.94	2.41	0.05	96.4	0.263	0.259	0.01	5.10	2.42	0.05	96.4
	T	0.268	0.263	0.01	5.53	3.29	0.06	95.7	0.275	0.271	0.03	10.1	3.96	0.06	96.4	0.270	0.267	0.02	7.91	3.67	0.06	97.0	0.271	0.267	0.02	8.59	3.80	0.06	96.8

Table 6.2: $n_{sim} = 1000$; $N = 150$; (true) linear marker profile $\mu(t) = 10 + 0.25t$; $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 2.25$. For JMs, the corresponding baseline hazard is reported. Column "Lost" (F, T) stands for (intermittent, "lost assessment") missing data scenarios.

Missing Rate (MIR)	Extended Cox Model										JM B-splines					JM Weibull													
	Est. (mean)	Est. (median)	Bias	%Bias	MSE	ESE	Coverage	Est. (mean)	Est. (median)	Bias	%Bias	MSE	ESE	Coverage	Est. (mean)	Est. (median)	Bias	%Bias	MSE	ESE	Coverage								
-	0.254	0.254	0.00	1.65	0.70	0.03	95.2	0.260	0.259	0.01	4.0	0.88	0.03	92.3	0.256	0.256	0.01	2.4	0.82	0.03	94.4	0.254	0.254	0.00	1.8	0.71	0.03	95.0	
0.30	F	0.240	0.238	-0.01	-4.04	0.65	0.02	92.9	0.253	0.251	0.00	1.28	0.09	0.03	93.3	0.248	0.246	-0.00	-0.73	0.77	0.03	94.4	0.249	0.248	-0.00	-0.47	0.84	0.03	94.3
	T	0.254	0.253	0.00	1.74	1.03	0.03	94.9	0.248	0.245	-0.00	-0.6	1.22	0.03	95.2	0.242	0.240	-0.01	-3.18	1.22	0.03	91.4	0.241	0.239	-0.01	-3.38	1.14	0.03	90.9
0.40	F	0.238	0.236	-0.01	-4.91	0.72	0.03	93.1	0.251	0.249	0.00	0.47	0.88	0.03	94.0	0.246	0.245	-0.00	-1.41	0.86	0.03	92.9	0.244	0.244	-0.01	-2.54	6.51	0.08	93.0
	T	0.258	0.255	0.01	3.08	1.11	0.03	96.3	0.247	0.245	-0.00	-1.03	1.24	0.03	92.6	0.241	0.237	-0.01	-3.77	1.24	0.03	91.3	0.242	0.237	-0.01	-3.3	4.45	0.06	89.7
0.40	F	0.234	0.233	-0.02	-6.3	0.85	0.03	91.6	0.250	0.244	-0.00	-0.01	16.89	0.13	92.2	0.241	0.238	-0.01	-3.7	1.02	0.03	91.7	0.240	0.239	-0.01	-3.85	0.95	0.03	90.6
	T	0.261	0.258	0.02	4.33	1.39	0.04	94.4	0.246	0.242	-0.00	-1.70	1.48	0.04	90.3	0.238	0.235	-0.01	-4.64	1.55	0.04	88.8	0.238	0.236	-0.01	-4.8	1.45	0.04	88.8
0.40	F	0.248	0.246	-0.00	-0.72	0.89	0.03	95.2	0.243	0.241	-0.01	-2.6	1.13	0.03	91.5	0.236	0.235	-0.01	-5.5	1.17	0.03	89.9	0.236	0.234	-0.01	-5.56	1.18	0.03	89.3
	T	0.264	0.261	0.01	5.50	1.55	0.04	95.0	0.245	0.241	-0.04	-1.81	1.51	0.04	90.7	0.238	0.234	-0.01	-4.9	1.51	0.04	88.4	0.237	0.233	-0.01	-5.04	1.47	0.04	87.9

Table 6.3: $n_{sim} = 1000; N = 400$; (true) linear marker profile $\mu(t) = 10 + 0.25t; \sigma_{b_0}^2 = \sigma_{b_1}^2 = 0.25$. For JM's, the corresponding baseline hazard is reported. Column "Lost" (F, T) stands for (intermittent, "lost assessment") missing data scenarios.

Eff. MR (%)	MR odds	Lost (%)	Extended Cox Model										JM B-splines										JM Weibull											
			Est. (mean)	Est. (median)	Bias %	MSE	ASE	ESE	CP	Est. (mean)	Est. (median)	Bias %	MSE	ASE	ESE	CP	Est. (mean)	Est. (median)	Bias %	MSE	ASE	ESE	CP	Est. (mean)	Est. (median)	Bias %	MSE	ASE	ESE	CP				
0	0	-	0.247	0.246	-0.003	-1.35	0.504	0.022	0.022	94.8	0.253	0.252	0.003	1.08	0.525	0.022	0.022	94.9	0.249	0.249	-0.001	-0.34	0.498	0.022	0.022	95.6	0.250	0.249	-0.000	-0.07	0.494	0.022	0.022	95.7
0.30	0.30	F	0.244	0.244	-0.006	-2.24	0.548	0.023	0.023	95.1	0.248	0.247	-0.002	-0.98	0.557	0.023	0.023	94.8	0.244	0.243	-0.006	-2.56	0.571	0.023	0.023	93.3	0.245	0.244	-0.005	-2.08	0.543	0.022	0.023	94.3
0.59	0.60	T	0.254	0.253	0.004	1.71	0.886	0.030	0.029	95.4	0.258	0.257	0.008	3.18	0.995	0.030	0.031	94.9	0.253	0.252	0.003	1.13	0.877	0.030	0.029	95.3	0.256	0.254	0.006	2.23	0.907	0.030	0.029	96.0
0.30	0.30	F	0.243	0.242	-0.07	-2.79	0.552	0.023	0.023	94.5	0.246	0.244	-0.004	-1.67	0.585	0.024	0.023	94.1	0.242	0.241	-0.008	-3.22	0.617	0.023	0.024	93.1	0.243	0.242	-0.007	-2.79	0.580	0.023	0.024	93.1
0.30	0.55	T	0.260	0.260	0.010	3.95	0.974	0.031	0.030	96.4	0.261	0.260	0.011	4.34	1.080	0.030	0.030	94.5	0.256	0.255	0.006	2.21	0.921	0.030	0.030	95.4	0.259	0.258	0.01	3.67	1.225	0.030	0.034	95.0
0.40	0.40	F	0.241	0.240	-0.009	-3.47	0.642	0.025	0.025	94.7	0.245	0.245	-0.005	-1.84	0.628	0.024	0.024	93.7	0.242	0.241	-0.008	-3.34	0.656	0.024	0.025	93.5	0.243	0.242	-0.007	-2.98	0.628	0.024	0.024	93.9
0.68	0.69	T	0.257	0.257	0.008	3.00	1.03	0.032	0.031	95.8	0.263	0.263	0.013	5.37	1.308	0.033	0.035	94.8	0.257	0.256	0.001	2.94	1.096	0.033	0.032	96.9	0.260	0.259	0.010	4.09	1.164	0.032	0.033	95.8
0.40	0.40	F	0.232	0.230	-0.002	0.73	0.804	0.029	0.028	95.0	0.234	0.233	0.004	1.72	0.780	0.027	0.027	95.7	0.230	0.229	0.000	0.12	0.724	0.027	0.026	95.8	0.233	0.233	0.003	1.24	0.752	0.027	0.027	96.0
0.40	0.63	T	0.262	0.261	0.012	4.65	1.210	0.033	0.034	94.6	0.267	0.267	0.017	6.87	1.475	0.033	0.038	94.0	0.261	0.260	0.011	4.44	1.220	0.033	0.034	95.5	0.264	0.264	0.014	5.67	1.317	0.033	0.035	95.3

Table 6.4: $n_{sim} = 1000; N = 400$; (true) linear marker profile $\mu(t) = 10 + 0.25t; \sigma_{b_0}^2 = \sigma_{b_1}^2 = 2.25$. For JM's, the corresponding baseline hazard is reported. Column "Lost" (F, T) stands for (intermittent, "lost assessment") missing data scenarios.

Eff. MR (%)	MR odds	Lost (%)	Extended Cox Model										JM B-splines										JM Weibull											
			Est. (mean)	Est. (median)	Bias %	MSE	ASE	ESE	CP	Est. (mean)	Est. (median)	Bias %	MSE	ASE	ESE	CP	Est. (mean)	Est. (median)	Bias %	MSE	ASE	ESE	CP	Est. (mean)	Est. (median)	Bias %	MSE	ASE	ESE	CP	Est. (mean)	Est. (median)	Bias %	MSE
0	0	-	0.252	0.251	0.002	0.72	0.222	0.015	0.015	96.5	0.256	0.255	0.006	2.38	0.298	0.015	0.017	92.9	0.252	0.251	0.002	0.95	0.274	0.016	0.016	94.8	0.252	0.251	0.002	0.84	0.256	0.015	0.016	95.0
0.30	0.30	F	0.239	0.238	-0.011	-4.45	0.330	0.015	0.018	88.6	0.249	0.248	-0.001	-0.58	0.294	0.016	0.016	93.6	0.244	0.243	-0.006	-2.23	0.333	0.016	0.018	92.0	0.244	0.243	-0.006	-2.23	0.314	0.016	0.017	92.1
0.60	0.60	T	0.250	0.251	0.001	0.42	0.355	0.018	0.018	95.3	0.242	0.242	-0.008	-3.36	0.468	0.019	0.021	90.2	0.236	0.236	-0.014	-5.48	0.586	0.019	0.024	85.3	0.237	0.236	-0.013	-5.29	0.543	0.018	0.023	85.8
0.30	0.30	F	0.235	0.235	-0.015	-5.84	0.422	0.015	0.020	82.5	0.245	0.244	-0.005	-2.08	0.322	0.016	0.018	90.0	0.241	0.240	-0.009	-3.63	0.373	0.017	0.019	90.1	0.241	0.239	-0.009	-3.80	0.372	0.016	0.019	89.6
0.30	0.55	T	0.254	0.254	0.004	1.73	0.413	0.019	0.020	94.9	0.241	0.239	-0.009	-3.69	0.511	0.019	0.022	88.9	0.235	0.234	-0.015	-5.89	0.642	0.019	0.025	83.5	0.236	0.234	-0.014	-5.75	0.605	0.018	0.024	82.8
0.40	0.40	F	0.233	0.231	-0.017	-6.96	0.54	0.016	0.023	79.9	0.238	0.236	-0.012	-4.96	0.473	0.017	0.021	84.7	0.233	0.231	-0.017	-6.92	0.623	0.017	0.024	76.3	0.233	0.231	-0.017	-6.79	0.592	0.016	0.024	77.2
0.69	0.69	T	0.259	0.257	0.008	3.43	0.522	0.020	0.022	93.2	0.259	0.257	0.011	-4.41	0.595	0.020	0.024	87.4	0.253	0.251	-0.017	-6.71	0.765	0.020	0.027	81.9	0.254	0.252	-0.016	-6.55	0.704	0.019	0.026	82.9
0.40	0.40	F	0.244	0.243	-0.00	-2.37	0.342	0.018	0.018	93.8	0.239	0.238	-0.021	-8.35	0.801	0.017	0.028	72.3	0.234	0.232	-0.026	-10.49	1.056	0.017	0.032	62.3	0.234	0.233	-0.026	-10.47	1.057	0.017	0.031	59.6
0.40	0.64	T	0.261	0.260	0.011	4.43	0.601	0.021	0.024	93.7	0.238	0.236	-0.012	-4.78	0.653	0.020	0.025	86.3	0.232	0.230	-0.018	-7.27	0.843	0.020	0.028	79.6	0.232	0.231	-0.018	-7.11	0.773	0.019	0.027	78.6

Table 6.5: $n_{sim} = 1000$; $N = 400$; (true) quadratic marker profile $\mu(t) = 10 - t - 0.10t^2$; $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 0.25$. Column "Lost" (F,T) stands for (intermittent, "lost assessment") missing data scenarios.

				Extended Cox Model								JM B-splines							
MR (%)	Eff. MR (%)	odds	Lost	Est. (mean)	Est. (median)	Bias	Bias %	MSE	ASE	ESE	CP	Est. (mean)	Est. (median)	Bias	%Bias	MSE	ASE	ESE	CP
0	0	-	-	0.251	0.250	0.001	0.31	0.454	0.022	0.021	95.1	0.252	0.251	0.002	0.84	0.468	0.022	0.022	95.9
0.30	0.30	4	F	0.220	0.219	-0.030	-11.94	1.319	0.021	0.021	69.6	0.248	0.247	-0.002	-0.66	0.540	0.023	0.023	94.0
	0.60		T	0.244	0.243	-0.006	-2.27	0.945	0.030	0.030	94.5	0.258	0.256	0.008	3.13	0.998	0.031	0.031	95.0
0.30	0.30	6	F	0.218	0.218	-0.032	-12.66	1.440	0.021	0.021	67.4	0.246	0.245	-0.04	-1.59	0.721	0.023	0.027	93.7
	0.55		T	0.249	0.249	-0.001	-0.48	0.951	0.031	0.031	95.1	0.258	0.255	0.008	3.04	1.044	0.031	0.031	95.2
0.40	0.40	4	F	0.218	0.217	-0.032	-12.94	1.491	0.022	0.021	68.2	0.246	0.245	-0.004	-1.67	0.614	0.025	0.024	93.5
	0.69		T	0.250	0.249	0.015	0.060	1.032	0.032	0.032	94.9	0.261	0.259	0.011	4.22	1.242	0.034	0.034	94.9
0.40	0.40	6	F	0.241	0.240	-0.010	-3.89	0.844	0.027	0.028	91.2	0.252	0.252	0.002	0.82	0.734	0.027	0.027	95.6
	0.63		T	0.253	0.253	0.002	1.12	1.317	0.032	0.034	93.7	0.264	0.262	0.014	5.54	1.418	0.034	0.035	94.4

Table 6.6: $n_{sim} = 1000$; $N = 400$; (true) quadratic marker profile $\mu(t) = 10 - t - 0.10t^2$; $\sigma_{b_0}^2 = \sigma_{b_1}^2 = 2.25$. Column "Lost" (F,T) stands for (intermittent, "lost assessment") missing data scenarios.

				Extended Cox Model								JM B-splines							
MR (%)	Eff. MR (%)	odds	Lost	Est. (mean)	Est. (median)	Bias	Bias %	MSE	ASE	ESE	CP	Est. (mean)	Est. (median)	Bias	%Bias	MSE	ASE	ESE	CP
0	0	-	-	0.244	0.244	-0.006	-2.25	0.204	0.014	0.013	94.0	0.252	0.252	0.002	0.94	0.225	0.015	0.015	96.0
0.30	0.30	4	F	0.228	0.227	-0.022	-8.91	0.650	0.014	0.012	63.1	0.246	0.245	-0.004	-1.65	0.272	0.016	0.016	93.9
	0.60		T	0.260	0.259	0.010	3.80	0.545	0.019	0.021	91.1	0.234	0.233	-0.016	-6.43	0.630	0.019	0.019	82.2
0.30	0.30	6	F	0.225	0.224	-0.025	-10.06	0.794	0.014	0.013	54.7	0.242	0.241	-0.008	-3.27	0.332	0.016	0.016	90.8
	0.55		T	0.264	0.263	0.014	5.43	0.663	0.020	0.022	90.0	0.232	0.232	-0.018	-7.18	0.701	0.019	0.019	80.5
0.40	0.40	4	F	0.224	0.224	-0.026	-10.37	0.859	0.014	0.014	56.0	0.233	0.231	-0.017	-6.91	0.567	0.017	0.016	79.0
	0.68		T	0.268	0.267	0.018	7.25	0.844	0.021	0.023	87.9	0.229	0.227	-0.021	-8.56	0.905	0.020	0.021	74.4
0.40	0.40	6	F	0.245	0.244	-0.005	-1.97	0.383	0.017	0.019	89.9	0.222	0.221	-0.028	-11.26	1.099	0.017	0.018	58.4
	0.63		T	0.269	0.268	0.019	7.43	0.890	0.022	0.024	87.1	0.227	0.226	-0.023	-9.07	0.964	0.020	0.021	71.5

Chapter 7

Final remarks

The extensive adoption of information technology has impacted numerous aspects of life, encompassing healthcare and the hospital environment. This has allowed for the collection of enormous volumes of data regarding the clinical histories of individuals. Repeated measurements on the same subject of a certain marker of interest provide more information than a single assessment, thus obtaining a broader overview of the evolution of the phenomenon being studied. During my doctoral period, I took part in several projects in the ICU context in which data from different markers, measured over time, were available ([23],[24]). These markers are constantly monitored but, often, their relationship to clinical outcome is not fully known. Therefore the increasing availability of data offers a considerable volume of information, and one of the goals is often to assess the association between temporal changes in longitudinal profiles and a clinical end-point. The goal is then to quantify in a dynamic context the intensity of this association, taking into account the correlation between individual observations on the same subject. In addition, during the observation period, some measurements may not be taken for various reasons, and often patients are not observed for the same period of time, often due to the occurrence of a certain clinical end-point, interrupting the measurement of the marker of interest, resulting in what are called unbalanced datasets. These longitudinal profiles are thus affected by missing data over time or truncations due to the interruption of measurements, and these issues, while they may complicate analyses, can be informative when examined through the use of appropriate techniques. All these aspects were covered within the thesis work.

The second chapter then briefly reviewed the main state-of-the-art methods usually employed to model longitudinal profile trajectories. The mixed-effects model was presented as the most relevant approach in this regard, since, unlike a classical regression model in which observations are assumed to be independent, or a Generalized Estimating Equations approach, which estimates a marginal model for the entire population under study, this methodology allows for a conditional estimation of trajectories, taking into account the correlation between observations on the same subject through the introduction of latent variables, named random effects. Indeed, this modeling approach falls into the category of conditional models and thus has both a marginal, overall, interpretation and a further subject-specific interpretation, which is very useful, for example, in the context of precision medicine. Moreover, the estimates obtained are valid even in a context affected by missing data, as long as the nature of the missing process is properly assessed. We discussed in detail from a theoretical point of view, the taxonomy of missing data types because, also motivated by the clinical context in which I have worked, this condition is often present also within prospective clinical studies. To assess the association between a biomarker and an end-point, one of the approaches that is most widely used in practice is the Cox model, which, in its extended version (ECM), allows for one or more time-dependent variables to be taken into account. While this approach has several advantages over parametric regression models that require proper specification of the baseline hazard, it also has several limitations, especially in the way individual-specific observations that update over time are treated. In particular, in addition to disregarding the likely correlation between individual measure-

ments on the same subject, longitudinal trajectories of markers are assumed to be free of measurement error, which is not always reasonable for endogenous variables, as biomarkers, and are assumed to be constant between one visit and the next. This approximation, which goes by the name of LOCF (Last Observation Carried Forward), may be trivial if not entirely questionable, especially if the times considered are not so tight. Among the modeling approaches that allow to overcome this limitation we can find the Joint Models (JM), the origin of which dates back to the last century, but for which until a few years ago there was a lack of software solutions capable of estimating all the parameters in reasonable times and therefore being implemented in applications. JMs are composed of two sub-models: an event sub-model, for which usually a proportional hazard model is specified (for which the shape of the baseline hazard must also be specified), and a sub-model for the longitudinal part that uses the mixed-effects models mentioned above. These two sub-models are joined through the use of the subject-specific random effects and the strength of the association between dynamically observed marker and event risk over time is quantified through a link parameter. The subject-specific random effects in the longitudinal sub-model are thus shared with the event sub-model and, for this reason, this model belongs to the category of "shared parameters models". Thus one of the main differences between the two approaches lies in the fact that the ECM extrapolates longitudinal trajectories using observed values, whereas the JMs, use these to interpolate individual trajectories. This technique has several advantages, especially in a context where data are affected by missingness, where the ECM would struggle in certain situations to decently approximate the temporal trajectories of biomarkers. However the benefits are not always free and interpolation of trajectories presents several challenges, starting from the difficulty in finding the correct model. If in fact the LOCF approximation may be inappropriate, it is equally true that the dynamic predictions obtained from JMs, may suffer from several problems, such as falling outside the range of a certain biomarker: therefore the use of one rather than the other approach has to be evaluated on a context-dependent basis. Using data from the ORANGE study, a comparative analysis was presented within the thesis that used the two approaches to evaluate the association between mortality and the Neurological Pupil Index (NPi), a time-dependent variable, in patients with ABI admitted to ICU. We have seen why in that case ECM might have been a more reasonable approach to assess the association with death at 6 months, while to assess the association with mortality over a shorter period, the use of the JM was appropriate, although it had some limitations. The ORANGE (Outcome pRognostication of Acute brain lesion with the NeuroloGical pupil indEx) study ([64]) was an international, multicenter, prospective, observational study that enrolled adult patients admitted to the Intensive Care Unit (ICU), requiring intubation and mechanical ventilation after traumatic brain injury (TBI), subarachnoid hemorrhage (SAH) or intracranial hemorrhage (ICH), in 13 hospitals in Europe and the United States. Because pupillary abnormalities are associated to poor outcomes, particularly in the case of patients with acute brain injury (ABI), the study required that patients underwent assessment of pupillary function about every 4 hours during the first 7 days after ICU admission, using an electronic device. Pupillary examination is an essential component of neurological assessment, and patient observations were quantified by an algorithm coded within the electronic pupillometer that compositely took into account several parameters to measure pupillary reactivity and global midbrain function, returning the NPi, a numerical value ranging from 0 to 5. The objectives of the study were thus to evaluate the prognostic value of NPi in predicting neurological functional outcome, measured by means of the GOSE - Glasgow Extended Coma Scale and dichotomized as unfavorable $\text{GOSE} \leq 4$ vs. favorable ($\text{GOSE} \geq 5$), and mortality, at 6 months. Logistic regressions, adjusted for age, ABI diagnosis, and motor Glasgow Coma Scale at ICU admission, were used to assess the association between NPi and functional neurological outcome, showing that increasing abnormal values ($\text{NPi} < 3$, $\text{NPi} = 0$) were statistically associated with an unfavorable outcome. Different Extended Cox models adjusted for the same variables, with NPi was included as a time-dependent variable, were used to evaluate the association between NPi and 6-month mortality. From the results, presented in terms of Hazard Ratios (HR), it was seen that an increase in the number of abnormal NPi measurements over time was associated with an increased risk of death. It was also shown that two consecutive

NPi measurements of zero or deterioration of NPi to a value of zero were associated with a consistent increase in mortality risk. In contrast, mortality risk did not increase when an NPi value of zero rose to a higher value. Our results also offered other valuable insights into the interpretation of NPi values. Indeed, while the absence of pupillary reactivity (i.e., $\text{NPi}=0$) is an established indicator of unfavorable outcome, an abnormal NPi was previously defined as less than 3. In this case, repeated NPi measurements improved the sensitivity analysis and identified a range of NPi, between 3 and 4, that was already associated with increased mortality risk. For that, in the context examined within the study, NPi monitoring could identify an at-risk patient population that would benefit from careful and intensive observation to manage secondary brain deterioration and target specific interventions before irreversible damage occurs. Instead of focusing on a single measurement or cut-off, clinicians should consider NPi as a tool for early quantification of the extent of midbrain dysfunction, ranging from very severe ($\text{NPi}=0$), to severe ($\text{NPi}<3$) and moderate ($\text{NPi} 3-4$). These results indicate the importance of NPi trajectories by providing evidence for routine use of this tool. This is of considerable importance as including reliable, non-invasive neurological monitoring in clinical practice is extremely beneficial, as it provides a safe alternative to invasive procedures, reducing the associated risks and complications for the patient. The study results were published in “The Lancet Neurology” ([51]).

We also performed a simulation study whose objective was to evaluate the capabilities and robustness of ECM and JM in correctly estimating the association, in terms of HR, between a dynamic marker and a time-to-event, when the profile of the longitudinal covariate was affected by specific missing processes, in different ways. For this purpose, we generated both individual profiles of longitudinal markers, assuming two different possible functional forms (linear and quadratic) and considering two different degrees of variability, obtained through two different variances of the random effect distributions during the data generation process, and the time-to-event of the subjects, starting from a JM specification. Then, once we obtained the “complete case” datasets, we were going to modify the trajectories of the individual markers, simulating missing processes with a probabilistic approach, considering two missing rates (MR), i.e., 30% and 40%: thus only some longitudinal profiles were modified, but many others remained unchanged. We drew two possible scenarios: in the first case, the biomarker profiles lost some observations, resulting in *intermittent* patterns. In this situation, the marker had some observed values, while others were unobserved. In the second case, we considered a scenario in which subsequent values are not collected after the first missing value in the biomarker profile, i.e., once a visit is missed. In such a situation, which we called the “*lost assessment*” case, the marker still had some observed values and unobserved values but the resulting profile was “truncated.” Therefore, the distributions of the “complete case” marker profile and the resulting distributions following data amputation were different and because the missingness depended on the biomarker value itself, we designed a context equivalent to a Missing Not at Random (MNAR) scenario. In all cases, the time-to-event data remained available for analysis, i.e., the simulated event status and time per patient was observed during follow-up without being altered. In other words, missing some values in the profiles of individual markers did not imply that the individual dropped out of the study. Then on the simulated data we fit the extended Cox model and, for the linear case, three JMs, different from each other according to the specified baseline hazard, while for the quadratic case we used only the JM that seemed more robust in the linear case, specifying an approximate baseline hazard with B-splines, to limit the computational time, because the simulations were extremely time-consuming.

Results were presented by distinguishing the functional form of the longitudinal sub-model and the variability of biomarkers by statistical summaries typically used to assess outcomes in a simulation study ([91]). On the complete case data both approaches showed good HR estimation abilities in all scenarios considered. In the linear case with intermittent data, the two approaches showed a good degree of robustness in the case of low biomarker variability, but the performance worsened when variability increased. In particular, the ECM showed the worst performance already at a 30% MR, while the JMs lost their good properties most noticeably only in the case of 40% MR. In the quadratic form case, on the other hand, the ECM underestimated the HR, both with low and high biomarker variability,

where the performance got even worse. In the "lost assessment" case, on the other hand, we obtained better results for the ECM, compared to the corresponding intermittent data scenarios, while the JM was found to underestimate HR, predominantly in cases of higher MR, as expected. The explanation of these results was addressed in the discussion in chapter six by assuming a sort of "compensatory effect" of different sources of bias. In conclusion, the use of Cox and Joint models proved to be appropriate when the data were largely unaffected by missingness, although the misclassification of trajectories by the ECM due to random measurement error results in a deterioration of its ability to correctly estimate HR, while the performance of the JM suffers a significant deterioration only in cases where the data had a large missing rate. We therefore concluded that further work will be needed to better understand the causes of some unexpected results in scenarios where many longitudinal trajectories were truncated early.

From an applied point of view, in the context of the ICU, among the various objectives for the next works there is that of deploying JMs for the identification of clusters of trajectories of specific biomarkers to evaluate their prognostic validity in patients suffering from ABI, using Joint Latent Class Mixed Models ([95]).

References

- [1] J. G. Ibrahim and G. Molenberghs. “Missing data methods in longitudinal studies: a review”. In: *Test (Madrid, Spain)* 18.1 (May 2009), pp. 1–43.
- [2] G. M. Fitzmaurice, Nan M Laird, and James H Ware. *Applied longitudinal analysis*. John Wiley & Sons, 2012.
- [3] P.J Diggle. et al. *Analysis of Longitudinal Data*. Oxford University Press, New York., 1994.
- [4] T. Garcia and K. Marder. “Statistical Approaches to Longitudinal Data Analysis in Neurodegenerative Diseases: Huntington’s Disease as a Model”. In: *Current Neurology and Neuroscience Reports* 17 (Feb. 2017).
- [5] X. Tan et al. “A time-varying effect model for intensive longitudinal data”. In: *Psychol Methods* 17 (Mar. 2012), pp. 61–77.
- [6] G. M. Fitzmaurice and C. Ravichandran. “A Primer in Longitudinal Data Analysis”. In: *Circulation* 118 (2008), pp. 2005–2010.
- [7] N. M. Laird and J. H. Ware. “Random-effects models for longitudinal data.” In: *Biometrics* 38 (1982), pp. 963–74.
- [8] D. Rizopoulos. *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R (1st ed.)*. Chapman and Hall/CRC., 2012.
- [9] A. Perperoglou et al. “A review of spline function procedures in R”. In: *BMC Medical Research Methodology* 19 (Mar. 2019).
- [10] Y. Lee and J. Nelder. “Conditional and Marginal Models: Another View”. In: *Statistical Science* 19 (May 2004).
- [11] J. J. Locascio and A. Atri. “An Overview of Longitudinal Data Analysis Methods for Neurological Research”. In: *Dementia and Geriatric Cognitive Disorders EXTRA* 1 (2011), pp. 330–357.
- [12] D. B. Rubin. “Inference and Missing Data”. In: *Biometrika* 63 (1976), pp. 581–590.
- [13] R. M. Schouten and G. Vink. “The Dance of the Mechanisms: How Observed Information Influences the Validity of Missingness Assumptions”. In: *Sociological Methods & Research* 50.3 (2021), pp. 1243–1258.
- [14] G. Molenberghs et al. “Every missingness not at random model has a missingness at random counterpart with equal fit”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 70 (2008).
- [15] C. H. Mallinckrodt et al. “Assessing and interpreting treatment effects in longitudinal clinical trials with missing data”. In: *Biological Psychiatry* 53.8 (2003), pp. 754–760.
- [16] J. Twisk et al. “Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis”. In: *Journal of Clinical Epidemiology* 66.9 (2013), pp. 1022–1028.
- [17] J. Schafer and J. Graham. “Missing Data: Our View of the State of the Art”. In: *Psychological methods* 7 (July 2002), pp. 147–77.
- [18] G. Verbeke and G. Molenberghs. *Linear Mixed Models for Longitudinal Data*. Springer Series in Statistics., 2000.

- [19] D. A. Harville. “Bayesian inference for variance components using only error contrasts”. In: *Biometrika* 61 (1974), pp. 383–385.
- [20] A. Gelman and J. Hill. *Data analysis using regression and multilevel/hierarchical models*. Vol. Analytical methods for social research. New York: Cambridge University Press, 2007.
- [21] P. C. Austin and J. Merlo. “Intermediate and advanced topics in multilevel logistic regression analysis”. In: *Statistics in Medicine* 36.20 (2017), pp. 3257–3277.
- [22] A.I.R. Maas et al. “Prognosis and Clinical Trial Design in Traumatic Brain Injury: The IMPACT Study”. In: *Journal of Neurotrauma* 24.2 (2007), pp. 232–238.
- [23] G. Citerio et al. “Management of arterial partial pressure of carbon dioxide in the first week after traumatic brain injury: results from the CENTER-TBI study”. In: *Intensive Care Medicine* 47.9 (Sept. 2021).
- [24] E. Rezoagli et al. “High arterial oxygen levels and supplemental oxygen administration in traumatic brain injury: insights from CENTER-TBI and OzENTER-TBI”. In: *Intensive Care Medicine* 48.9 (2022).
- [25] C. R. Lesko et al. “When to Censor?” In: *American Journal of Epidemiology* 187.3 (Aug. 2017), pp. 623–632.
- [26] E. L. Kaplan and P. Meier. “Nonparametric Estimation from Incomplete Observations”. In: *Journal of the American Statistical Association* 53 (1958), pp. 457–481.
- [27] N. E. Breslow and J. J. Crowley. “A Large Sample Study of the Life Table and Product Limit Estimates Under Random Censorship”. In: *Annals of Statistics* 2 (1974), pp. 437–453.
- [28] A.C. Wong and N. Reid. *“Analysis of Survival Data” by D.R. Cox and D. Oakes: Solutions to Selected Exercises*. Technical report (University of Toronto. Department of Statistics). University of Toronto, Department of Statistics, 1990.
- [29] D.R.Cox. “Regression Models and Life-Tables”. In: *Journal of the royal statistical society series b-methodological* 34 (1972), pp. 187–220.
- [30] P. K. Andersen and R. D. Gill. “Cox’s Regression Model for Counting Processes: A Large Sample Study”. In: *The Annals of Statistics* 10.4 (1982), pp. 1100–1120.
- [31] R. L. Prentice. “Covariate measurement errors and parameter estimation in a failure time regression model”. In: *Biometrika* 69.2 (Aug. 1982), pp. 331–342.
- [32] D. Rizopoulos. “Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data”. In: *Biometrics* 67.3 (2011), pp. 819–829.
- [33] P. Blanche et al. “Quantifying and Comparing Dynamic Predictive Accuracy of Joint Models for Longitudinal Marker and Time-to-Event in Presence of Censoring and Competing Risks”. In: *Biometrics* 71.1 (2015), pp. 102–113.
- [34] J. Long and J. Mills. “Joint modeling of multivariate longitudinal data and survival data in several observational studies of Huntington’s disease”. In: *BMC Medical Research Methodology* 18 (2018).
- [35] G. Hickey et al. “Joint modelling of time-to-event and multivariate longitudinal outcomes: Recent developments and issues”. In: *BMC Medical Research Methodology* 16 (2016).
- [36] C. Proust-Lima et al. “Joint modelling of multivariate longitudinal outcomes and a time-to-event: A nonlinear latent class approach”. In: *Computational Statistics Data Analysis* 53.4 (2009), pp. 1142–1154.
- [37] F. Hsieh, Y.-K. Tseng, and J.-L. Wang. “Joint Modeling of Survival and Longitudinal Data: Likelihood Approach Revisited”. In: *Biometrics* 62.4 (2006), pp. 1037–1043.
- [38] M.S. Wulfsohn and A.A. Tsiatis. “A Joint Model for Survival and Longitudinal Data Measured with Error”. In: *Biometrics* 53.1 (1997), pp. 330–339. (Visited on 06/19/2023).

- [39] R. Henderson, P.J. Diggle, and A. Dobson. “Joint modelling of longitudinal measurements and event time data.” English. In: *Biostatistics* 1.4 (Dec. 2000), pp. 465–480. ISSN: 1465-4644.
- [40] D. Rizopoulos. “Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive Gaussian quadrature rule”. In: *Computational Statistics Data Analysis* 56.3 (2012), pp. 491–501.
- [41] A.P. Dempster, D.B. Rubin, and Robert K. Tsutakawa. “Estimation in Covariance Components Models”. In: *Journal of the American Statistical Association* 76 (1981), pp. 341–353.
- [42] G. Schwarz. “Estimating the Dimension of a Model”. In: *The Annals of Statistics* 6.2 (1978), pp. 461–464.
- [43] H. Akaike. “A new look at the statistical model identification”. In: *IEEE Transactions on Automatic Control* 19 (1974), pp. 716–723.
- [44] M. C. Wu and R. J. Carroll. “Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process”. In: *Biometrics* 44 (1988), pp. 175–188.
- [45] E. Vonesh, T. Greene, and M. Schluchter. “Shared parameter models for the joint analysis of longitudinal data and event times”. In: *Statistics in medicine* 25 (Jan. 2006), pp. 143–63.
- [46] S. Viviani, D. Rizopoulos, and M. Alfó. “Local sensitivity to non-ignorability in joint models”. In: *Statistical Modelling* 14.3 (2014), pp. 205–228.
- [47] G. Papageorgiou and D. Rizopoulos. “An alternative characterization of MAR in shared parameter models for incomplete longitudinal data and its utilization for sensitivity analysis”. In: *Statistical Modelling* 21.1-2 (2021), pp. 95–114.
- [48] R. Tsonaka, G. Verbeke, and E. Lesaffre. “A Semi-Parametric Shared Parameter Model to Handle Nonmonotone Nonignorable Missingness”. In: *Biometrics* 65.1 (2009), pp. 81–87.
- [49] N. C Chesnaye et al. “An introduction to joint models—applications in nephrology”. In: *Clinical Kidney Journal* 13.2 (2020), pp. 143–149.
- [50] J. Ibrahim, H. Chu, and L. Chen. “Basic Concepts and Methods for Joint Models of Longitudinal and Survival Data”. In: *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 28 (June 2010), pp. 2796–801.
- [51] M. Oddo et al. “The Neurological Pupil index for outcome prognostication in people with acute brain injury (ORANGE): a prospective, observational, multicentre cohort study”. In: *LANCET NEUROLOGY* 22 (2023), pp. 925–933.
- [52] B. Lussier, D. Olson, and V. Aiyagari. “Automated Pupillometry in Neurocritical Care: Research and Practice”. In: *Current Neurology and Neuroscience Reports* 19 (Aug. 2019).
- [53] F. Romagnosi et al. “Neurological Pupil Index for the Early Prediction of Outcome in Severe Acute Brain Injury Patients”. In: *Brain Sciences* 12.5 (May 2022), p. 609.
- [54] F.P. Jahns et al. “Quantitative pupillometry for the monitoring of intracranial hypertension in patients with severe traumatic brain injury”. In: *Critical Care* 23 (Dec. 2019).
- [55] J. Pansell et al. “Can Quantitative Pupillometry be used to Screen for Elevated Intracranial Pressure? A Retrospective Cohort Study”. In: *Neurocritical Care* 37 (May 2022).
- [56] S. G. Aoun et al. “Detection of delayed cerebral ischemia using objective pupillometry in patients with aneurysmal subarachnoid hemorrhage”. In: *Journal of neurosurgery* 132.1 (2020), pp. 27–32.
- [57] M. Osm et al. “Correlation of Objective Pupillometry to Midline Shift in Acute Stroke Patients”. In: *Journal of Stroke and Cerebrovascular Diseases* 28 (Apr. 2019).

- [58] F. Romagnosi, F. Bongiovanni, and M. Oddo. “Eyeing up the injured brain: automated pupillometry and optic nerve sheath diameter”. In: *Current Opinion in Critical Care* 26 (Feb. 2020), p. 1.
- [59] J.P. I Miroz et al. “Neurological Pupil index for Early Prognostication After Venocorporeal Extracorporeal Membrane Oxygenation”. In: (2020).
- [60] L. Peluso et al. “Neurological Pupil Index and its association with other prognostic tools after cardiac arrest: A post hoc analysis”. In: *Resuscitation* 179 (July 2022).
- [61] M. Oddo et al. “Quantitative versus standard pupillary light reflex for early prognostication in comatose cardiac arrest patients: an international prospective multicenter double-blinded study”. In: *Intensive Care Medicine* 44 (Nov. 2018).
- [62] M. Menozzi et al. “Early Neurological Pupil Index Assessment to Predict Outcome in Cardiac Arrest Patients Undergoing Extracorporeal Membrane Oxygenation”. In: *ASAIO Journal* Publish Ahead of Print (Sept. 2021).
- [63] J. Chen et al. “Infrared pupillometry, the Neurological Pupil index and unilateral pupillary dilation after traumatic brain injury: implications for treatment paradigms”. In: *SpringerPlus* 3 (Sept. 2014), p. 548.
- [64] M. Oddo et al. “Outcome Prognostication of Acute Brain Injury using the Neurological Pupil Index (ORANGE) study: protocol for a prospective, observational, multicentre, international cohort study”. In: *BMJ Open* 11.5 (2021).
- [65] M. Hoffmann et al. “Pupil evaluation in addition to the Glasgow Coma Scale (GCS) components in traumatic brain injury”. In: *The British journal of surgery* 99 Suppl 1 (Jan. 2012), pp. 122–30.
- [66] A. Marmarou et al. “Prognostic Value of The Glasgow Coma Scale And Pupil Reactivity in Traumatic Brain Injury Assessed Pre-Hospital And on Enrollment: An IMPACT Analysis”. In: *Journal of Neurotrauma* 24.2 (2007), pp. 270–280.
- [67] D. Olson et al. “Interrater Reliability of Pupillary Assessments”. In: *Neurocritical care* 24 (Sept. 2015).
- [68] C. Sandroni, G. Citerio, and F. Taccone. “Automated pupillometry in intensive care”. In: *Intensive Care Medicine* 48 (June 2022).
- [69] P. Opic et al. “Automated, Quantitative Pupillometry in the Critically Ill: A Systematic Review of the Literature”. In: *Neurology* 97 (May 2021).
- [70] M. Ghauri et al. “Evaluating the Reliability of Neurological Pupillary Index as a Prognostic Measurement of Neurological Function in Critical Care Patients”. In: *Cureus* 14 (Sept. 2022).
- [71] J. Chen et al. “Pupillary reactivity as an early indicator of increased intracranial pressure: The introduction of the Neurological Pupil Index”. In: *Surgical neurology international* 2 (June 2011), p. 82.
- [72] I. Shoyombo et al. “Understanding the Relationship Between the Neurologic Pupil Index and Constriction Velocity Values”. In: *Scientific Reports* 8 (May 2018).
- [73] S. S. Phillips et al. “A systematic review assessing the current state of automated pupillometry in the NeuroICU”. In: *Neurocritical care* 31 (2019), pp. 142–161.
- [74] L. Peluso et al. “Early neurological pupil index to predict outcome after cardiac arrest”. In: *Intensive Care Medicine* 48 (Apr. 2022).
- [75] S. Stutzman et al. “Inter-device reliability of the NPi-200 and NPi-300 pupillometers”. In: *Journal of Clinical Neuroscience* 100 (June 2022), pp. 180–183.
- [76] P. A. Harris et al. “Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support”. In: *Journal of Biomedical Informatics* 42.2 (2009), pp. 377–381.
- [77] 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), pp. 573–585.

- [78] L. F. Marshall et al. “The National Traumatic Coma Data Bank: Part 1: Design, purpose, goals, and results”. In: *Journal of Neurosurgery* 59.2 (1983), pp. 276–284.
- [79] J. Claassen et al. “Effect of Cisternal and Ventricular Blood on Risk of Delayed Cerebral Ischemia After Subarachnoid Hemorrhage:: The Fisher Scale Revisited”. In: *Stroke: Journal of the American Heart Association* 32 (2001), pp. 2012–2020.
- [80] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods (Springer Texts in Statistics)*. Springer-Verlag, 2005.
- [81] R. Bender, T. Augustin, and M. Blettner. “Generating survival times to simulate Cox proportional hazards models”. In: *Statistics in Medicine* 24 (2005).
- [82] S. L. Brilleman et al. “Simulating Survival Data Using the `simsurv` R Package”. In: *Journal of Statistical Software* 97.3 (2020), pp. 1–27.
- [83] P. C. Austin. “Generating survival times to simulate Cox proportional hazards models with time-varying covariates”. In: *Statistics in Medicine* 31.29 (2012), pp. 3946–3958.
- [84] M. Crowther and P. Lambert. “Simulating biologically plausible complex survival data”. In: *Statistics in medicine* 32 (Oct. 2013). DOI: 10.1002/sim.5823.
- [85] R. P. Brent. “An Algorithm with Guaranteed Convergence for Finding a Zero of a Function”. In: *Comput. J.* 14 (1971), pp. 422–425.
- [86] S.L. Brilleman. *simsurv: Simulate Survival Data*. R package version 0.2.0. 2018. URL: <https://cran.r-project.org/package=simsurv>.
- [87] R. M. Schouten and G. Vink. “The Dance of the Mechanisms: How Observed Information Influences the Validity of Missingness Assumptions”. In: *Sociological Methods & Research* 50.3 (2021), pp. 1243–1258.
- [88] T. Rockel. *missMethods: Methods for Missing Data*. R package version 0.4.0. 2022. URL: <https://cran.r-project.org/package=missMethods>.
- [89] T. M. Therneau. *A Package for Survival Analysis in R*. R package version 3.5-7. 2023. URL: <https://CRAN.R-project.org/package=survival>.
- [90] D.s Rizopoulos. “JM: An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data”. In: *Journal of Statistical Software* 35.9 (2010), pp. 1–33.
- [91] A. Burton et al. “The design of simulation studies in medical statistics”. In: *Statistics in Medicine* 25.24 (2006), pp. 4279–4292.
- [92] M. Arisido et al. “Joint model robustness compared with the time-varying covariate Cox model to evaluate the association between a longitudinal marker and a time-to-event endpoint”. In: *BMC medical research methodology* 19 (Dec. 2019), p. 222.
- [93] S. Wacholder et al. “Non-differential misclassification and bias towards the null: a clarification”. In: *Occupational and environmental medicine* 52 (Sept. 1995), pp. 557–8.
- [94] G. B Hamra. “Invited Commentary: Is Bias Towards the Null From Nondifferential Misclassification Wishful Thinking?” In: *American Journal of Epidemiology* 191.8 (May 2022), pp. 1496–1497.
- [95] Cécile Proust-Lima et al. “Joint latent class models for longitudinal and time-to-event data: A review”. In: *Statistical Methods in Medical Research* 23 (2014), pp. 74–90.


Appendix A

Papers ([23],[24])

ORIGINAL



Management of arterial partial pressure of carbon dioxide in the first week after traumatic brain injury: results from the CENTER-TBI study

Giuseppe Citerio^{1,2*} , Chiara Robba^{3,4}, Paola Rebori^{1,5}, Matteo Petrosino⁵, Eleonora Rossi⁶, Letterio Malgeri⁷, Nino Stocchetti^{8,9}, Stefania Galimberti^{1,5}, and David K. Menon¹⁰ on behalf of the Center-TBI participants and investigators

© 2021 The Author(s)

Abstract

Purpose: To describe the management of arterial partial pressure of carbon dioxide (PaCO₂) in severe traumatic brain-injured (TBI) patients, and the optimal target of PaCO₂ in patients with high intracranial pressure (ICP).

Methods: Secondary analysis of CENTER-TBI, a multicentre, prospective, observational, cohort study. The primary aim was to describe current practice in PaCO₂ management during the first week of intensive care unit (ICU) after TBI, focusing on the lowest PaCO₂ values. We also assessed PaCO₂ management in patients with and without ICP monitoring (ICP_m), and with and without intracranial hypertension. We evaluated the effect of profound hyperventilation (defined as PaCO₂ < 30 mmHg) on long-term outcome.

Results: We included 1100 patients, with a total of 11,791 measurements of PaCO₂ (5931 lowest and 5860 highest daily values). The mean (± SD) PaCO₂ was 38.9 (± 5.2) mmHg, and the mean minimum PaCO₂ was 35.2 (± 5.3) mmHg. Mean daily minimum PaCO₂ values were significantly lower in the ICP_m group (34.5 vs 36.7 mmHg, *p* < 0.001). Daily PaCO₂ nadir was lower in patients with intracranial hypertension (33.8 vs 35.7 mmHg, *p* < 0.001). Considerable heterogeneity was observed between centers. Management in a centre using profound hyperventilation (HV) more frequently was not associated with increased 6 months mortality (OR = 1.06, 95% CI = 0.77–1.45, *p* value = 0.7166), or unfavourable neurological outcome (OR 1.12, 95% CI = 0.90–1.38, *p* value = 0.3138).

Conclusions: Ventilation is manipulated differently among centers and in response to intracranial dynamics. PaCO₂ tends to be lower in patients with ICP monitoring, especially if ICP is increased. Being in a centre which more frequently uses profound hyperventilation does not affect patient outcomes.

Keywords: Carbon dioxide, Hyperventilation, Traumatic brain injury, Intracranial pressure, Outcome

*Correspondence: giuseppe.citerio@unimib.it

¹ School of Medicine and Surgery, University of Milano - Bicocca, Monza, Italy

Full author information is available at the end of the article

Giuseppe Citerio and Chiara Robba equally contributed as first authors to this work. Stefania Galimberti and David K. Menon equally contributed as last authors to this work.

CENTER-TBI ICU participants and investigators are listed as non-author contributors in the Acknowledgement section.

Introduction

Changes in the arterial partial pressure of carbon dioxide (PaCO₂), by modifying the extravascular pH, modulate cerebrovascular tone, and hence cerebral blood flow (CBF) and cerebral blood volume (CBV) [1, 2]. Hypercapnia results in perivascular acidosis, which causes cerebral vasodilation, and consequently, an increase in intracranial volume. In patients with poor intracranial compliance, this could raise intracranial pressure (ICP). On the other

hand, hyperventilation (HV) induced alkalosis reduces vascular calibre, and hence CBV, and can represent an effective measure to control intracranial hypertension, when ICP remains elevated despite first-line therapies [3–6]. However, hypocapnic cerebral vasoconstriction can also reduce CBF [7], thus posing the risk of secondary ischaemic insults [8]. In a survey across European trauma centers, the most frequently reported PaCO₂ target was 36–40 mmHg in the absence of intracranial hypertension, which was reduced to 30–35 mmHg when ICP was >20 mmHg [9]. The most recent evidence-based guidelines on TBI management provide no definitive recommendations regarding target PaCO₂ levels due to the low quality of evidence available on this issue [10, 11].

Consequently, although many patients with severe TBI undergo several days of mechanical ventilation, there is little evidence-based guidance on PaCO₂ targets, and clinical practice remains highly variable. A recent consensus on mechanical ventilation in patients with acute brain injury suggested aiming for a physiologic range of PaCO₂ between 35 and 45 mmHg [12], and to only use hyperventilation (with an undefined PaCO₂ target) as a short-term therapeutic option in patients with evidence of brain herniation. However, the document was unable to provide a recommendation on the use of hyperventilation in patients who showed significant ICP elevation, but no evidence of herniation. A management algorithm for patients with intracranial hypertension, based on expert consensus, suggested the use of HV (PaCO₂ 32–35 mmHg) for controlling ICP only as a second-tier treatment, did not support lower PaCO₂ levels and recommended against routine hyperventilation to PaCO₂ below 30 mmHg [13].

The objectives of this study were to assess, in a real-world context, PaCO₂ management and the lowest target of PaCO₂ in a large cohort of mechanically ventilated TBI patients and practice variability between centres to evaluate the association between the use of profound HV and 6-month clinical outcomes.

Methods

Study design and patients

The Collaborative European NeuroTrauma Effectiveness in Research in Traumatic Brain Injury (CENTER-TBI study, registered at clinicaltrials.gov NCT02210221) is a longitudinal, prospective collection of data from TBI patients across 65 centers in Europe. The study was conducted between December 19th, 2014, and December 17th, 2017 and details regarding the design and the results of the screening and enrolment process have been previously described [14–16].

The CENTER-TBI study was approved by the Medical Ethics Committees in all participating centers, and

Take-home message

The manipulation of arterial carbon dioxide levels (PaCO₂) is easy, and hyperventilation (HV) has been a common ICP-lowering strategy for over half a century. However, hyperventilation-induced vasoconstriction is a double-edged sword. It reduces cerebral blood volume and intracranial volume, and therefore, lowers ICP

We observed huge variability among centers in PaCO₂ values and use of HV. Although causal inferences cannot be drawn from these observational data, our results suggest that, in patients with severe intracranial hypertension, HV is not associated with worse long-term clinical outcome

informed consent was obtained according to local regulations (<https://www.center-tbi.eu/project/ethical-approval>). This project on PaCO₂ management was preregistered on the CENTER-TBI proposal platform and approved by the CENTER-TBI proposal review committee before starting the analysis (ESM Document 1). This report complies with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines (ESM Table S1).

We included all patients in the CENTER-TBI Core study who had a TBI necessitating ICU admission, required tracheal intubation and mechanical ventilation, had at least two PaCO₂ measurements in the first 7 days and had been admitted to a study centre that enrolled at least ten patients.

Data collection and definitions

Detailed information on data collection is available on the study website (<https://www.center-tbi.eu/data/dictionary>). For the first week in ICU, the daily lowest and highest PaCO₂ values from arterial blood gases and, if an ICP device was inserted, the hourly ICP measures were used for analysis.

HV was defined as moderate for PaCO₂ ranging between 30 and 35 mmHg and profound for PaCO₂ <30 mmHg [10, 13]. Therapy intensity level (TIL) was calculated according to the most recent TIL scale [17]. Patients with invasive ICP monitoring during the first week of ICU stay were classified as ICP_m, while those who did not receive ICP monitoring during ICU stay as no-ICP_m. Intracranial hypertension was defined as ICP >20 mmHg.

Objectives

The aims of this study are:

1. to describe the PaCO₂ values in the first week from ICU admission in mechanically ventilated TBI patients, and to evaluate practice variability across centers, particularly focusing on the lowest targets of PaCO₂;

2. to assess at a center level the PaCO₂ management in patients with/without ICP monitoring and with/without intracranial hypertension;
3. to evaluate the association between patient outcomes and center propensity to use profound HV.

Outcomes

Mortality and functional outcome (measured as the Extended Glasgow Outcome Score, GOSE) were assessed at 6 months. All responses were obtained by study personnel from patients or from a proxy (where impaired cognitive capacity prevented patient interview), during a face-to-face visit, by telephone interview, or by postal questionnaire around 6 months after injury [18]. All evaluators had received training in the use of the GOSE. An unfavourable outcome was defined as GOSE ≤ 4, which includes both poor functional outcome and mortality.

Statistical methods

Patient characteristics were described by means (± standard deviation, SD), medians (I–III quartiles, Q₁–Q₃) and counts or proportions, as appropriate. The comparison of baseline features according to ICP monitoring was performed using Mann–Whitney *U* test, *t* test and Chi-square test as appropriate. We used the median odds ratio (MOR) to estimate the between-centre heterogeneity in targeting a PaCO₂ of 35–45 mmHg. MOR was obtained from a longitudinal logistic mixed-effect model on daily lowest PaCO₂ adjusted for the IMPACT core covariates [19], ICP monitoring, and daily evidence of elevated ICP (at least one ICP > 20 mmHg during the day); and with a hierarchical random intercept effect's structure (i.e., patients within centers). The same model architecture was used to quantify between-centres heterogeneity in the use of profound HV.

We resorted to an instrumental variable approach to evaluate the association between HV and 6-month outcomes, trying to minimize the potential measured and unmeasured confounding acting in this complex observational study [20]. This was done by considering the propensity of centres to apply profound HV, measured as the proportion of daily lowest PaCO₂ < 30 mmHg, as an instrument in the logistic regression model with a random intercept for centers. This model was adjusted for some subject-specific covariates that included IMPACT core covariates at baseline, ICP monitoring and dose of intracranial hypertension, calculated as the area under the ICP profile above 20 mmHg, named AUC ICP > 20 [21]. The assumptions underlying the IV approach were assessed (ESM-Statistical methods).

Tests were performed with a two-sided significance level of 5%. All analyses were conducted using R statistical software (version 4.03).

Results

Of the 4509 patients included in the CENTER-TBI dataset, 2138 patients with TBI from 51 centers in Europe were admitted to ICU. Among these, 1176 required mechanical ventilation and had at least two PaCO₂ measurements within the first 7 days from ICU admission. Excluding the centres that enrolled less than ten patients, 1100 patients from 36 centers were available for the analysis (ESM Fig. 1). During the first week of ICU admission, a total of 11,791 measurements of PaCO₂ were available (5931 lowest and 5860 highest daily values).

Patient characteristics

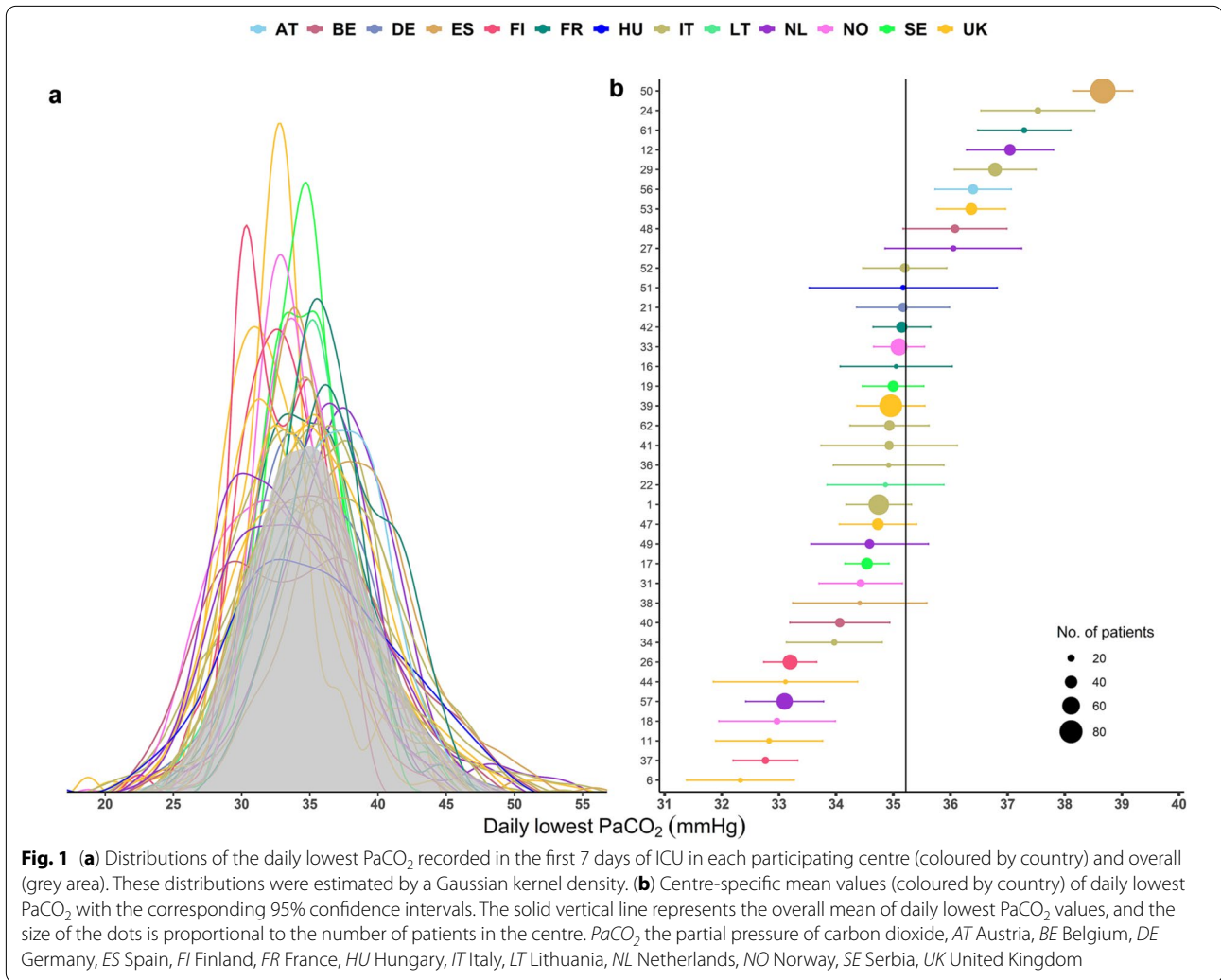
Patient characteristics at hospital admission in the overall population and stratified according to the presence (*n* = 751) or not (*n* = 349) of ICP monitoring, are summarized in Table 1. The median age was 48 years (Q1–Q3 = 29–64), and most patients were male (74%). 64.7% of patients presented with a severe TBI (Glasgow Coma Scale, GCS ≤ 8) and 12.5% of cases were complicated by thoracic trauma. In 727 (97%) ICP_m patients, ICP was inserted by the second day of ICU admission.

In the overall population, the mean PaCO₂ at ICU admission was 39.1 (± 6) mmHg, and the no-ICP_m group had higher PaCO₂ mean values compared to the ICP_m patients (39.9 ± 6.8 vs 38.7 ± 5.6 mmHg, *p* < 0.002).

Lowest PaCO₂ targets according to centers

Daily minimum PaCO₂ distribution during the first week for the whole population, and separated by the centre, are presented in Fig. 1a. The overall mean lowest PaCO₂ was 35.2 ± 5.4 mmHg with substantial heterogeneity between centres, whose means ranged from 32.3 (± 3.7) to 38.7 mmHg (± 5.9). This result seems to be related more to different management strategies at the centre level, rather than reflecting national policies (Fig. 1b). For example, among the UK centers (in yellow), two centers had a mean PaCO₂ value of 32.3 and 36.4 mmHg.

Only 144 (13%) patients had all PaCO₂ measurements between 35 and 45 mmHg, while 588 (53%) patients had at least half of the total PaCO₂ measurements in this range. Using MOR to quantify between-centre differences in targeting the suggested PaCO₂ range of 35–45 mmHg, we found that, after correction for patient and trauma characteristics, there was a 1.72-fold difference in the odds of having a PaCO₂ range of 35–45 mmHg between centres with the highest and lowest rates. After excluding 390 patients with intracranial hypertension, the percentage of patients with all and at least half of the total PaCO₂ measurements between 35 and 45 mmHg raised to 19% (111/593) and 64% (380/593), while MOR decreased to 1.4.



Lowest PaCO₂ targets in the presence or not of ICP monitoring

Mean minimum PaCO₂ values were significantly lower in ICP_m patients compared to no-ICP_m (34.7 ± 4.9 mmHg vs 36.8 ± 5.7 mmHg, $p < 0.001$). Large variability was observed among centers in the management of PaCO₂ targets in both subgroups (Fig. 2 and ESM Fig. 2). Some centres showed no differences in target PaCO₂ when ICP_m was used (i.e. data points near the line of identity in Fig. 2a), but most hospitals tended to adopt lower PaCO₂ targets when ICP was monitored (i.e. data points that deviate substantially from the line of identity in Fig. 2a). For example, three centers showed a reduction greater than 4 mmHg in the mean daily lowest PaCO₂ when ICP monitoring was available (from 38–38.4 mmHg to 33.1–34.2 mmHg).

Lowest PaCO₂ in the presence of intracranial hypertension

In the subgroup of patients with ICP monitoring, we also explored the attitude of centres in response to episodes of intracranial hypertension ($n = 3646$). Some centres showed no differences in target PaCO₂ when ICP was elevated (i.e. data points near the line of identity in Fig. 2b), but most hospitals tended to adopt lower PaCO₂ targets when ICP was monitored (i.e. data points that deviate substantially from the line of identity in Fig. 2b). The mean minimum PaCO₂ was significantly lower in 398 patients with at least one episode of intracranial hypertension compared to the 240 who did not experience increased ICP (34.1 vs 35.6 mmHg, $p < 0.001$). Within the group of patients with ICP monitoring in place, significant inter-centre differences were observed in the mean lowest PaCO₂, both in the absence and presence of intracranial hypertension (ESM Fig. 3).

Table 1 Baseline demographic and clinical characteristics, including trauma characteristics, clinical presentation, and neuroimaging features at ICU admission in the overall population and stratified according to the presence or not of ICP monitoring

Characteristic	Overall (n = 1100)	no-ICP _m (n = 349)	ICP _m (n = 751)	P value
Age (years), median (Q1–Q3)	48 (29–64)	53 (31–69)	46 (28–61)	<0.001
Sex, n (%)				
Female	284 (25.8)	89 (25.5)	195 (26)	0.929
Thoracic trauma, n (%)				
Yes	138 (12.5)	42(12)	96 (12.8)	0.802
ISS, median (Q1–Q3)	34 (25–48)	34 (25–43)	34 (25–48)	0.011
Hypotension, n (%)				
Yes	178 (17.4)	60 (17.7)	118 (17.3)	0.936
Not available	78	10	68	
Hypoxia, n (%)				
Yes	182 (17.9)	53 (15.6)	129 (19)	0.217
Not available	82	10	72	
Severity TBI, n (%)				
GCS ≤ 8	367 (35.3)	147 (44.3)	220 (31)	<0.001
GCS > 8	674 (64.7)	185 (55.7)	489 (69)	
Not available	59	17	42	
Pupillary reactivity, n (%)				
Both reactive	799 (75.8)	280 (82.8)	519 (72.5)	0.001
One reactive	89 (8.4)	22 (6.5)	67 (9.4)	
Both unreactive	166 (15.7)	36 (10.7)	130 (18.2)	
Not available	47	11	35	
GCS motor, n (%)				
None	460 (42.7)	129 (37.7)	331 (45)	<0.001
Extension	51 (4.7)	9 (2.6)	42 (5.7)	
Abnormal flexion	60 (5.6)	10 (2.9)	50 (6.8)	
Normal flexion	89 (8.3)	30 (8.8)	59 (8)	
Localizes/obeys	418 (38.8)	164 (48)	254 (34.5)	
Not available	22	7	15	
Marshall CT classification, n (%)				
1	63 (6.5)	48 (15.6)	15 (2.3)	0.0005
2	416 (42.9)	167 (54.2)	249 (37.7)	
3	98 (10.1)	17 (5.5)	81 (12.3)	
4	19 (2)	3 (1)	16 (2.4)	
5	6 (0.6)	2 (0.6)	4 (0.6)	
6	367 (37.9)	71 (23.1)	296 (44.8)	
Not available	131	41	90	
Overall PaCO ₂ (mmHg), mean (SD)	39.10 (6)	39.93 (6.8)	38.72 (5.6)	0.002
Lowest PaCO ₂ (mmHg), mean (SD)	34.66 (5.98)	35.92 (6.67)	34.09 (5.56)	<0.001
Highest PaCO ₂ (mmHg), mean (SD)	43.68 (8.1)	44.07 (8.6)	43,5 (7.86)	0.287

Hypotension was defined as a documented systolic blood pressure < 90 mmHg; hypoxia was defined as a documented partial pressure of oxygen (PaO₂) < 8 kPa (60 mmHg), oxygen saturation (SaO₂) < 90%, or both; PaCO₂ data refer to values at ICU admission

PaCO₂ the partial pressure of carbon dioxide, SD standard deviation, Q1–Q3 I and III quartiles, ISS injury severity score, TBI traumatic brain injury, GCS Glasgow Coma Scale, ICP_m intracranial pressure monitored, ICU intensive care unit

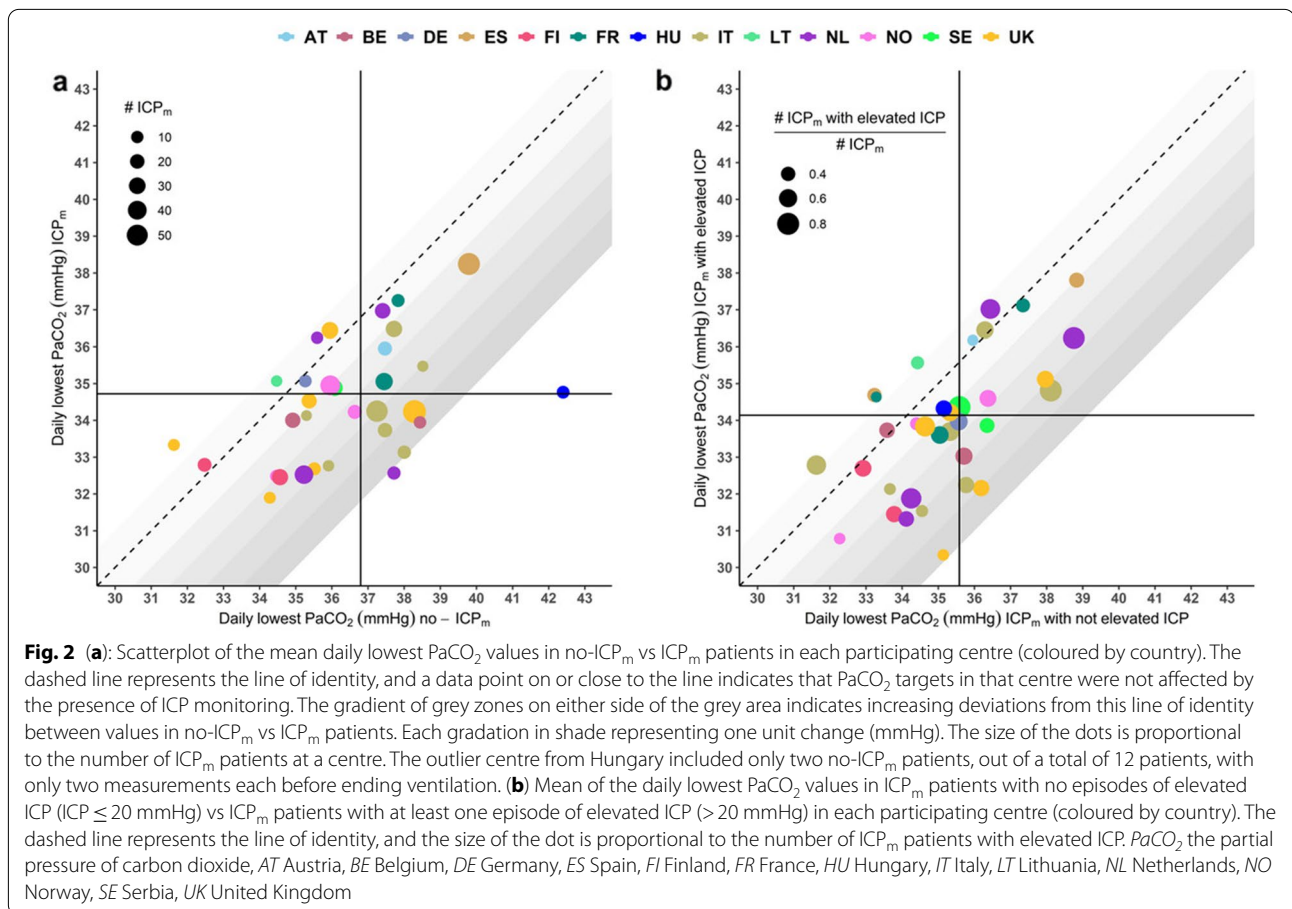
Profound hyperventilation

An episode of profound HV (PaCO₂ < 30 mmHg) was recorded on 727 occasions during the first week of ICU admission in 397 (36%) patients (57% had one, 22% two and 10% three occurrences). Results from the longitudinal mixed-effects model show notable heterogeneity between centres on the use of HV, even after adjusting for patient and trauma characteristics, with a MOR of 2.04 (Fig. 3, ESM Table 1). We found a significant positive association between the occurrence of increased ICP and the use of HV. Among ICP_m patients, even

after correction for covariates, the odds of HV in a day with elevated ICP was nearly three times that in a day with controlled ICP (OR=4.34 95% CI = 4.25-4.44, *p* value < 0.0001 vs OR=1.47 95% CI = 0.97-2.22, *p* value = 0.03167). Finally, HV was less applied from day 1 to 7 (OR of HV per day=0.83; 95% CI=0.82–0.84, *p* value < 0.0001).

Neuromonitoring

Indirect CBF monitoring, using jugular bulb venous oxygen saturation or brain tissue oxygenation, was not



used frequently. No differences were found in their use in patients receiving profoundly HV (jugular bulb venous oxygen saturation, S_{jv}O₂: 2.4% vs profound HV 3.5%, *p* value=0.380; brain tissue oxygenation, PbtO₂: 14.2% vs profound HV 13.9%, *p* value=0.937). However, the use of profound HV was associated with significantly higher use of more aggressive treatment, expressed as mean TIL (9.7 vs 6.3 *p* value<0.001). In particular, patients who received profound hyperventilation were more likely to have decompressive surgery (8.6 vs 4.8, *p* value<0.001) and hyperosmolar therapy (low dose 12.7 vs 5.5, *p* value<0.001; high dose 16.8 vs 5.7, *p* value<0.001).

6 months mortality and neurological outcome

Overall, of the 1100 patient cohort, 165 died before ICU discharge (15%). Of the 970 patients for whom 6-month outcomes were available, 246 (25.4%) died, and 529 (54.5%) experienced unfavourable functional outcomes (GOSE ≤ 4). The 6 months mortality rate was 29% in patients who had at least one episode of profound HV and 23% in those who did not (*p* value=0.045), while the rates of unfavourable GOSE were 64% vs 49% in the two groups, respectively (*p* value<0.001). The percentage

of patients who received profound HV in the first seven days from admission ranged from 1 to 30% between hospitals. In the IV analysis, the propensity to apply profound HV (defined by the use of PaCO₂<30 mmHg) did not significantly increase mortality or unfavourable functional outcome, after adjusting for the dose of intracranial hypertension. Patients in hospitals that used 10% more profound HV had 1.06 higher odds of mortality compared to hospitals where profound HV was applied less often (95% CI=0.77–1.45, *p* value=0.7166) and the OR for the same comparison was 1.12 (95% CI=0.90–1.38, *p* value=0.3138) for an unfavourable functional outcome (Table 2).

Discussion

The current literature is inconclusive regarding the optimal ventilatory strategy to adopt in patients with TBI and, though there is increasing caution surrounding the use of HV, the translation of expert consensus recommendations into clinical practice remains uncertain. This study examined the PaCO₂ management during mechanical ventilation at a centre level in prospectively collected

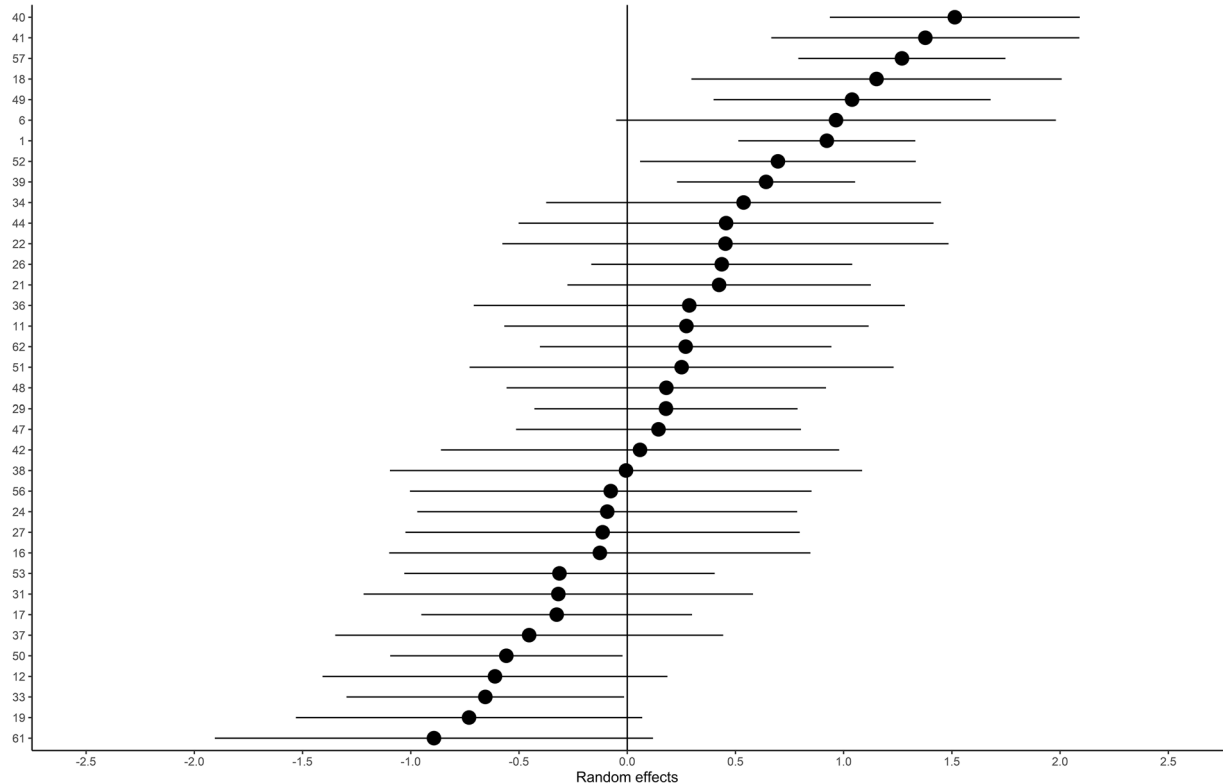


Fig. 3 Caterpillar plot of between-centre variation in using profound HV. The figure shows the predicted random intercepts for each centre, on the log-odds scale, along with their 95% prediction intervals. Higher values indicate a higher propensity to use profound HV. A longitudinal random effect logistic model was used to correct for random variation and adjusted for the core IMPACT covariates and elevated ICP. The MOR summarises the between-centre variation: a MOR = 1 indicates no variation, while the larger the MOR is, the larger the variation present. The median odds ratio (MOR = 2.04) refers to the odds of using profound HV between two randomly selected centres for patients with the same covariates and (comparable) random effects

observational data from a large multicentre cohort of TBI patients, focusing on the use of HV.

Our main findings are:

- there is substantial practice variation among countries and centers regarding PaCO₂ levels and the lowest PaCO₂ adopted in TBI patients;
- patients who received ICP monitoring were managed at lower PaCO₂ compared to patients in whom such monitoring was not used;
- patients who did receive ICP monitoring and experienced episodes of increased ICP were managed at lower PaCO₂ levels than those who did not have ICP elevations; profound HV was commonly used in such patients;
- we observed no association between the risk of mortality or unfavourable functional outcome and more frequent use of profound hyperventilation (PaCO₂ < 30 mmHg).

Appropriate management of PaCO₂ is a critical requirement in mechanically ventilated patients with TBI, since carbon dioxide is one of the major determinants of cerebral vascular physiology, and therefore cerebral blood flow and volume. The effect of the interplay between carbon dioxide and perfusion pressure on the cerebral circulation results in a sophisticated modulation of cerebrovascular resistance and tone, with hypercapnia causing cerebral vasodilation, and hypocapnia, vasoconstriction.

The only randomized controlled trial [22] addressing the benefit of prophylactic hyperventilation was conducted thirty years ago, and randomised TBI patients into three categories: control ($n=41$), hyperventilation ($n=36$), and HV + tromethamine (an H⁺ acceptor used to treat metabolic acidosis; $n=36$). This setting is different from the current context, as the putatively normoventilated controls had PaCO₂ values in the hypocapnic range (35 mmHg), and the HV utilized was

Table 2 Results of the logistic mixed-effect model on 6-month outcomes by the instrumental variable approach with complete data (n = 919)

Outcome	6-month GOSE OR (95% CI) p value	6-month mortality OR (95% CI) p value
Centre HV tendency (per 10% change)*	1.12 (0.9–1.38) 0.3138	1.06 (0.77–1.45) 0.7166
Age	1.04 (1.03–1.05) <0.0001	1.05 (1.04–1.06) <0.0001
GCS Motor Score		
None	2.08 (1.46–2.95) <0.0001	2.28 (1.44–3.62) 0.0004
Extension	5.47 (2.39–12.51) <0.0001	1.82 (0.74–4.48) 0.1886
Abnormal flexion	3.29 (1.63–6.65) 0.0009	1.69 (0.65–4.37) 0.2794
Normal flexion	1.45 (0.82–2.56) 0.1980	1.2 (0.55–2.64) 0.6421
Localizes/obeys	1	1
Pupilar reactivity		
Both reacting	1	1
One reacting	1.98 (1.14–3.43) 0.0146	2.18 (1.16–4.11) 0.0154
Both unreacting	3.29 (2.05–5.27) <0.0001	6.04 (3.69–9.87) <0.0001
ICP monitoring		
No	1	1
Yes	1.79 (1.27–2.51) 0.0008	1.00 (0.65–1.54) 0.9948
AUC ICP > 20 (per one SD change) ^o	3.72 (1.94–7.15) <0.0001	5.15 (2.86–9.25) <0.0001

OR Odds ratio, CI confidence intervals, SD standard deviation

* Centre HV propensity is calculated as the percentage of daily lowest PaCO₂ < 30 mmHg out of all available measures

^oStandardized AUC ICP > 20 is the dose of intracranial hypertension calculated as the area under the ICP profile above 20 mmHg

profound (PaCO₂ 25 mmHg). These discordances with current practice, the limited number of patients, and the low incidence of episodes of intracranial hypertension make the results difficult to interpret.

A recent consensus still recommends targeting a normal range of PaCO₂ values in the absence of increased ICP [12]. However, in the case of increased ICP, no agreement was achieved regarding the role of HV, providing evidence of the current uncertainty in this area [12]. Although induced hypocapnia is considered an efficient second line measure to reduce ICP, clinicians remain worried about potential cerebral ischemic complications of hyperventilation [8, 23]. Coles et al. used positron emission tomography in a cohort of 30 patients to show that the acute application of HV resulted in a reduction of cerebral blood flow and an increase in oxygen extraction fraction and the ischemic brain volume [23]. These results have left an indelible imprint on the way HV is perceived by intensivists, but they do not represent a randomized trial. Other authors suggest that mild HV may reduce ICP without leading to pathological changes of brain metabolism and oxygenation measured through cerebral microdialysis and PbtO₂ [24] or energy failure. Moreover, Diringer et al. demonstrated that HV reduces global cerebral blood flow while increased oxygen extraction fraction leaving cerebral metabolic rate for oxygen

unchanged, concluding that it is unlikely that HV causes neurological injury [25, 26].

Although some concerns still exist, PaCO₂ reduction is still widely used in the clinical setting for ICP control. The most common PaCO₂ target declared by clinicians in the absence of intracranial hypertension (35–40 mmHg) is higher than in the case of raised ICP (30–35 mmHg) [9]. Similarly, in a retrospective study of 151 patients with TBI, the PaCO₂ target adopted in clinically stable ICP was 36 ± 5.7 mmHg, whereas in the case of increased ICP it was 34 ± 5.4 mmHg [27]. Besides, a recent consensus on ICP treatment suggested considering HV to PaCO₂ of 30–32 mmHg when ICP is elevated in patients not responding to Tier 1 and 2 treatment [13].

Our data document a divergence between suggestions from literature and practice: nearly half of the daily lowest PaCO₂ measurements in the first week were < 35 mmHg. Moreover, in presence of ICP monitoring, clinicians use a lower target of PaCO₂. However, we also saw wide variability in PaCO₂ levels between centres, both in terms of the overall values, and the lowest levels of PaCO₂ observed. These differences were seen not just across the whole study cohort, but also in subgroups of patients with and without ICP monitoring, and those with and without episodes of intracranial hypertension in the first week. HV in presence of high ICP was frequently used,

particularly in the first few days after admission, and was often combined with other ICP-lowering therapies such as osmotic agents and decompressive craniectomy. Interestingly, centres that used HV more frequently were not more likely to routinely apply more advanced neuromonitoring techniques for early detection of impaired cerebral blood flow and cerebral oxygen availability.

There is no strong evidence regarding the possible benefits or harms of profound HV on patient outcomes. However, a single retrospective analysis of 251 brain-injured patients [28] reported that, when compared to controls, patients who underwent prolonged HV (PaCO₂: 25–30 mmHg; mean duration = 10, min–max = 5–41 h) experienced lower mortality (9.8 vs. 32.8%) but a higher rate of poor functional outcome.

We found that being treated in a centre where profound hypocapnia is more frequently used compared to centres where it is rarely used was not significantly associated with a higher rate of mortality or poor functional outcome.

In summary, our results suggest that moderate HV is widely used in severely brain-injured patients, especially when ICP is monitored, and in case of elevated ICP.

Limitations

Although our results may provide useful context with an important clinical message for physicians, we believe they should be interpreted with caution for several reasons. First, 6 months GOSE and mortality are influenced by several other factors, such as systemic and ICU complications, as well as post-ICU events. Therefore, based on observational data, it is speculative to draw a direct causal relationship between PaCO₂ and outcome: further randomized controlled studies are needed to assess the effect of PaCO₂ more precisely and in particular HV, on the outcome. Second, this is an analysis of data from a large study, which primarily addressed the epidemiology, clinical care and outcome of TBI. However, as respiratory management was not a primary focus of the study, more specific data on ventilatory management of these patients are missing, and hence unavailable to strengthen our analysis. Data on the incidence and timing of pulmonary complications such as acute respiratory distress respiratory syndrome and ventilator-associated pneumonia, the use of ventilatory strategies used to manipulate PaCO₂, and the ventilator settings used in our study population are unavailable. Third, the outcome was evaluated at 6 months, which can be considered as an early measurement of outcome after TBI, and further long-term evaluations would have been desirable. Fourth, we did not specifically take into consideration the temperature management of the patients, which can importantly affect PaCO₂ values. However, the measurements of PaCO₂ are

automatically corrected for temperature from the arterial blood gases machines, and we aimed to assess the targets of PaCO₂ achieved, regardless of the effects of different factors on its final value.

Finally, in our dataset only the daily lowest and highest PaCO₂ values were collected, thus missing possible changes in PaCO₂ and pulmonary function parameters that may occur suddenly and repeatedly during the day. However, our analysis includes data on daily PaCO₂, thus providing a longitudinal view of PaCO₂ management over time.

Conclusions

In a large cohort of mechanically ventilated TBI patients, we found substantial between-centre variations in PaCO₂, but with a large proportion of patients being managed at PaCO₂ levels below those suggested by expert consensus statements. On average, patients who had ICP monitors in place had significantly lower PaCO₂ levels than those that did not, and amongst ICP monitored patients, PaCO₂ levels were lower in patients who had episodes of intracranial hypertension—suggesting that HV is still used for ICP management. Profound hyperventilation (PaCO₂ < 30 mmHg) was not uncommon. However, a centre that had a greater propensity to use profound HV did not worsen 6-month mortality or functional outcome. Notwithstanding this, we believe that the available evidence still makes the case for caution in the use of HV, with careful consideration of risks and benefits on a case-by-case basis. Our data provide no basis for dismissing continuing concerns regarding prophylactic or profound hyperventilation. We need randomized controlled trials and high-level evidence guidelines to support rational choices regarding optimal ventilation management and PaCO₂ targets in patients with TBI.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s00134-021-06470-7>.

Abbreviations

AUC CO₂i: Area below the value of 30 mmHg as a benchmark and the interpolation of the PaCO₂ profile in time; AUC ICP > 20: Area under ICP profile above 20 mmHg; CBF: Cerebral blood flow; CENTER-TBI: Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury; CI: Confidence interval; CO₂: Carbon dioxide; CT: Computed tomography; ESM: Electronic supplementary material; GCS: Glasgow coma scale; GOSE: Glasgow outcome scale extended; HP: Hypocapnia; HR: Hazard rate; HV: Hyperventilation; ICP: Intracranial pressure; ICP_m: ICP monitored; No-ICP_m: No-ICP monitored; ICU: Intensive care unit; ISS: Injury severity score; LOS: Length of stay; MOR: Median odds ratio; OR: Odds ratio; PaCO₂: Partial pressure of carbon dioxide; PbtO₂: Brain tissue oxygenation; SaO₂: Oxygen saturation; S_{jv}O₂: Jugular bulb venous oxygen saturation; SD: Standard deviation; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TBI: Traumatic brain injury; TIL: Therapy intensity level.

Author details

¹ School of Medicine and Surgery, University of Milano - Bicocca, Monza, Italy.

² Neurointensive Care Unit, Ospedale San Gerardo, Azienda Socio-Sanitaria

Territoriale Di Monza, Monza, Italy.³ Anesthesia and Intensive Care, Policlinico San Martino, IRCCS on Oncology and Neuroscience, Genoa, Italy.⁴ Department of Surgical Science and Integrated Diagnostic, University of Genoa, Genoa, Italy.⁵ Bicocca Bioinformatics Biostatistics and Bioimaging Center B4, School of Medicine and Surgery, University of Milano - Bicocca, Milan, Italy.⁶ Department of Clinical-Surgical, Diagnostic and Paediatric Sciences, Unit of Anaesthesia and Intensive Care, University of Pavia, Pavia, Italy.⁷ Anesthesia and Intensive Care, School of Medicine, Messina, Italy.⁸ Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy.⁹ Department of Physiopathology and Transplantation, Milan University, Milan, Italy.¹⁰ Neurocritical Care Unit, Addenbrooke's Hospital, Cambridge, UK.

Acknowledgements

We acknowledge the CENTER-TBI ICU Participants and Investigators listed here as non-authors contributors: Cecilia Åkerlund¹, Krisztina Amrein², Nada Andelic³, Lasse Andreassen⁴, Audny Anke⁵, Anna Antoni⁶, Gérard Audibert⁷, Philippe Azouvi⁸, Maria Luisa Azzolini⁹, Ronald Bartels¹⁰, Pál Barzó¹¹, Romuald Beauvais¹², Ronny Beer¹³, Bo-Michael Bellander¹⁴, Antonio Belli¹⁵, Habib Benali¹⁶, Maurizio Berardino¹⁷, Luigi Beretta¹⁸, Morten Blaabjerg¹⁹, Peter Bragge²⁰, Alexandra Brazinova²¹, Vibeke Brinck²², Joanne Brooker²³, Camilla Brorsson²⁴, Andras Buki²⁵, Monika Bullinger²⁶, Manuel Cabelreira²⁷, Alessio Caccioppola²⁸, Emiliana Calappi²⁹, Maria Rosa Calvi³⁰, Peter Cameron³¹, Guillermo Carbayo Lozano³², Marco Carbonara³³, Simona Cavallo³⁴, Giorgio Chevallard³⁵, Arturo Chierogato³⁶, Giuseppe Citerio³⁷, Hans Clusmann³⁸, Mark Coburn³⁹, Jonathan Coles⁴⁰, Jamie D. Cooper⁴¹, Marta Correia⁴², Amra Ovi⁴³, Nicola Curry⁴⁴, Endre Czeiter⁴⁵, Marek Czosnyka⁴⁶, Claire Dahyot-Fizelier⁴⁷, Paul Dark⁴⁸, Helen Dawes⁴⁹, Véronique De Keyser⁵⁰, Vincent Degos⁵¹, Francesco Della Corte⁵², Hugo den Boogert⁵³, Bart Depreitere⁵⁴, Đula Đilvesi⁵⁵, Abhishek Dixit⁵⁶, Emma Donoghue⁵⁷, Jens Dreier⁵⁸, Guy-Loup Dulière⁵⁹, Ari Ercole⁶⁰, Patrick Esser⁶¹, Erzsébet Ezer⁶², Martin Fabricius⁶³, Valery L. Feigin⁶⁴, Kelly Foks⁶⁵, Shirin Frisvold⁶⁶, Alex Furmanov⁶⁷, Pablo Gagliardo⁶⁸, Damien Galanaud⁶⁹, Dashiell Gantner⁷⁰, Guoyi Gao⁷¹, Pradeep George⁷², Alexandre Ghuysen⁷³, Lelde Giga⁷⁴, Ben Glocker⁷⁵, Jagoš Golubovic⁷⁶, Pedro A. Gomez⁷⁷, Johannes Gratz⁷⁸, Benjamin Gravesteyn⁷⁹, Francesca Grossi⁸⁰, Russell L. Gruen⁸¹, Deepak Gupta⁸², Juanita A. Haagsma⁸³, Iain Haitsma⁸⁴, Raimund Helbok⁸⁵, Eirik Helseth⁸⁶, Lindsay Horton⁸⁷, Jiliske Huijben⁸⁸, Peter J. Hutchinson⁸⁹, Bram Jacobs⁹⁰, Stefan Jankowski⁹¹, Mike Jarrett⁹², Ji-yao Jiang⁹³, Faye Johnson⁹⁴, Kelly Jones⁹⁵, Mladen Karan⁹⁶, Angelos G. Koliás⁹⁷, Erwin Kompanje⁹⁸, Daniel Kondziella⁹⁹, Evgenios Kornaropoulos¹⁰⁰, Lars-Owe Koskinen¹⁰¹, Noémi Kovács¹⁰², Ana Kowark¹⁰³, Alfonso Lagares¹⁰⁴, Linda Lanyon¹⁰⁵, Steven Laureys¹⁰⁶, Fiona Lecky¹⁰⁷, Didier Ledoux¹⁰⁸, Rolf Lefering¹⁰⁹, Valerie Legrand¹¹⁰, Aurelie Lejeune¹¹¹, Leon Levi¹¹², Roger Lightfoot¹¹³, Hester Lingsma¹¹⁴, Andrew I.R. Maas¹¹⁵, Ana M. Castaño-León¹¹⁶, Marc Mægele¹¹⁷, Marek Majdan¹¹⁸, Alex Manara¹¹⁹, Geoffrey Manley¹²⁰, Costanza Martino¹²¹, Hugues Maréchal¹²², Julia Mattern¹²³, Catherine McMahon¹²⁴, Béla Meleghe¹²⁵, David Menon¹²⁶, Tomas Menovsky¹²⁷, Ana Mikolicić¹²⁸, Benoit Misset¹²⁹, Visakh Muralleedharan¹³⁰, Lynnette Murray¹³¹, Auncuta Negrú¹³², David Nelson¹³³, Virginia Newcombe¹³⁴, Daan Nieboer¹³⁵, József Nyírádi¹³⁶, Otesile Olubukola¹³⁷, Matej Oresic¹³⁸, Fabrizio Ortolano¹³⁹, Aarno Palotie¹⁴⁰, Paul M. Parizel¹⁴¹, Jean-François Payen¹⁴², Natascha Perera¹⁴³, Vincent Perlbarg¹⁴⁴, Paolo Persona¹⁴⁵, Wilco Peul¹⁴⁶, Anna Piippo-Karjalainen¹⁴⁷, Matti Pirinen¹⁴⁸, Dana Pisica¹⁴⁹, Horia Ples¹⁵⁰, Suzanne Polinder¹⁵¹, Inigo Pomposo¹⁵², Jussi P. Posti¹⁵³, Louis Puybasset¹⁵⁴, Andreea Radoi¹⁵⁵, Arminas Ragauskas¹⁵⁶, Rahul Raj¹⁵⁷, Malinka Rambadagalla¹⁵⁸, Isabel Retel Helmrich¹⁵⁹, Jonathan Rhodes¹⁶⁰, Sylvia Richardson¹⁶¹, Sophie Richter¹⁶², Samuli Ripatti¹⁶³, Saulius Rocka¹⁶⁴, Cecilie Roe¹⁶⁵, Olav Roise¹⁶⁶, Jonathan Rosand¹⁶⁷, Jeffrey V. Rosenfeld¹⁶⁸, Christina Rosenlund¹⁶⁹, Guy Rosenthal¹⁷⁰, Rolf Rossaint¹⁷¹, Sandra Rossi¹⁷², Daniel Rueckert¹⁷³, Martin Rusnák¹⁷⁴, Juan Sahuquillo¹⁷⁵, Oliver Sakowitz¹⁷⁶, Renan Sanchez-Porras¹⁷⁷, Janos Sandor¹⁷⁸, Nadine Schäfer¹⁷⁹, Silke Schmidt¹⁸⁰, Herbert Schoechl¹⁸¹, Guus Schoonman¹⁸², Rico Frederik Schou¹⁸³, Elisabeth Schwendenwein¹⁸⁴, Charlie Sewalt¹⁸⁵, Toril Skandsen¹⁸⁶, Peter Smielewski¹⁸⁷, Abayomi Sorinola¹⁸⁸, Emmanuel Stamatakis¹⁸⁹, Simon Stanworth¹⁹⁰, Robert Stevens¹⁹¹, William Stewart¹⁹², Ewout W. Steyerberg¹⁹³, Nino Stocchetti¹⁹⁴, Nina Sundström¹⁹⁵, Riikka Takala¹⁹⁶, Viktória Tamás¹⁹⁷, Tomas Tamosiutis¹⁹⁸, Mark Steven Taylor¹⁹⁹, Braden Te Ao²⁰⁰, Olli Tenovuori²⁰¹, Alice Theadom²⁰², Matt Thomas²⁰³, Dick Tibboel²⁰⁴, Marjolijn Timmers²⁰⁵, Christos Toliás²⁰⁶, Tony Trapani²⁰⁷, Cristina María Tudora²⁰⁸, Andreas Unterberg²⁰⁹, Peter Vajkoczy²¹⁰, Shirley Vallance²¹¹, Egils Valeinis²¹², Zoltán Vámos²¹³, Mathieu van der Jagt²¹⁴, Gregory Van der Steen²¹⁵, Joukje van der Naalt²¹⁶, Jeroen T.J.M. van Dijk²¹⁷, Thomas A. van Essen²¹⁸, Wim Van Hecke²¹⁹, Carolinevan Heugten²²⁰,

Dominique Van Praag¹³⁹, Ernest van Veen⁶⁴, Thijs Vande Vyvere¹³⁷, Roel P. J. van Wijk¹⁰¹, Alessia Vargiolu³², Emmanuel Vega⁸³, Kimberley Velt⁶⁴, Jan Verheyden¹³⁷, Paul M. Vespa¹⁴⁰, Anne Vik^{123, 141}, Rimantas Vilcinis¹³², Victor Volovic⁶⁷, Nicole von Steinbüchel³⁸, Daphne Voormolen⁶⁴, Petar Vulekovic⁴⁶, Kevin K.W. Wang¹⁴², Eveline Wiegers⁶⁴, Guy Williams⁴⁷, Lindsay Wilson⁶⁹, Stefan Winzeck⁴⁷, Stefan Wolf¹⁴³, Zhihui Yang¹¹³, Peter Ylén¹⁴⁴, Alexander Youns⁹⁰, Frederick A. Zeiler^{47, 145}, Veronika Zelinkova²⁰, Agate Ziverte⁶⁰, Tommaso Zoerle²⁷

¹Department of Physiology and Pharmacology, Section of Perioperative Medicine and Intensive Care, Karolinska Institutet, Stockholm, Sweden

²János Szentágothai Research Centre, University of Pécs, Pécs, Hungary

³Division of Surgery and Clinical Neuroscience, Department of Physical Medicine and Rehabilitation, Oslo University Hospital and University of Oslo, Oslo, Norway

⁴Department of Neurosurgery, University Hospital Northern Norway, Tromsø, Norway

⁵Department of Physical Medicine and Rehabilitation, University Hospital Northern Norway, Tromsø, Norway

⁶Trauma Surgery, Medical University Vienna, Vienna, Austria

⁷Department of Anesthesiology & Intensive Care, University Hospital Nancy, Nancy, France

⁸Raymond Poincaré hospital, Assistance Publique – Hôpitaux de Paris, Paris, France

⁹Department of Anesthesiology & Intensive Care, S Raffaele University Hospital, Milan, Italy

¹⁰Department of Neurosurgery, Radboud University Medical Center, Nijmegen, The Netherlands

¹¹Department of Neurosurgery, University of Szeged, Szeged, Hungary

¹²International Projects Management, ARTTIC, München, Germany

¹³Department of Neurology, Neurological Intensive Care Unit, Medical University of Innsbruck, Innsbruck, Austria

¹⁴Department of Neurosurgery & Anesthesia & intensive care medicine, Karolinska University Hospital, Stockholm, Sweden

¹⁵NIHR Surgical Reconstruction and Microbiology Research Centre, Birmingham, UK

¹⁶Anesthésie-Réanimation, Assistance Publique – Hôpitaux de Paris, Paris, France

¹⁷Department of Anesthesia & ICU, AOU Città della Salute e della Scienza di Torino—Orthopedic and Trauma Center, Torino, Italy

¹⁸Department of Neurology, Odense University Hospital, Odense, Denmark

¹⁹BehaviourWorks Australia, Monash Sustainability Institute, Monash University, Victoria, Australia

²⁰Department of Public Health, Faculty of Health Sciences and Social Work, Trnava University, Trnava, Slovakia

²¹Quesgen Systems Inc., Burlingame, California, USA

²²Australian & New Zealand Intensive Care Research Centre, Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

²³Department of Surgery and Perioperative Science, Umeå University, Umeå, Sweden

²⁴Department of Neurosurgery, Medical School, University of Pécs, Hungary and Neurotrauma Research Group, János Szentágothai Research Centre, University of Pécs, Hungary

²⁵Department of Medical Psychology, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany

²⁶Brain Physics Lab, Division of Neurosurgery, Dept of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK

²⁷Neuro ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy

²⁸ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia

²⁹Department of Neurosurgery, Hospital of Cruces, Bilbao, Spain

³⁰NeuroIntensive Care, Niguarda Hospital, Milan, Italy

³¹School of Medicine and Surgery, Università Milano Bicocca, Milano, Italy

³²NeuroIntensive Care, ASST di Monza, Monza, Italy

³³Department of Neurosurgery, Medical Faculty RWTH Aachen University, Aachen, Germany

³⁴Department of Anesthesiology and Intensive Care Medicine, University Hospital Bonn, Bonn, Germany

³⁵Department of Anesthesia & Neurointensive Care, Cambridge University Hospital NHS Foundation Trust, Cambridge, UK

- ³⁶School of Public Health & PM, Monash University and The Alfred Hospital, Melbourne, Victoria, Australia
- ³⁷Radiology/MRI department, MRC Cognition and Brain Sciences Unit, Cambridge, UK
- ³⁸Institute of Medical Psychology and Medical Sociology, Universitätsmedizin Göttingen, Göttingen, Germany
- ³⁹Oxford University Hospitals NHS Trust, Oxford, UK
- ⁴⁰Intensive Care Unit, CHU Poitiers, Poitiers, France
- ⁴¹University of Manchester NIHR Biomedical Research Centre, Critical Care Directorate, Salford Royal Hospital NHS Foundation Trust, Salford, UK
- ⁴²Movement Science Group, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, UK
- ⁴³Department of Neurosurgery, Antwerp University Hospital and University of Antwerp, Edegem, Belgium
- ⁴⁴Department of Anesthesia & Intensive Care, Maggiore Della Carità Hospital, Novara, Italy
- ⁴⁵Department of Neurosurgery, University Hospitals Leuven, Leuven, Belgium
- ⁴⁶Department of Neurosurgery, Clinical centre of Vojvodina, Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia
- ⁴⁷Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK
- ⁴⁸Center for Stroke Research Berlin, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany
- ⁴⁹Intensive Care Unit, CHR Citadelle, Liège, Belgium
- ⁵⁰Department of Anaesthesiology and Intensive Therapy, University of Pécs, Pécs, Hungary
- ⁵¹Departments of Neurology, Clinical Neurophysiology and Neuroanesthesiology, Region Hovedstaden Rigshospitalet, Copenhagen, Denmark
- ⁵²National Institute for Stroke and Applied Neurosciences, Faculty of Health and Environmental Studies, Auckland University of Technology, Auckland, New Zealand
- ⁵³Department of Neurology, Erasmus MC, Rotterdam, the Netherlands
- ⁵⁴Department of Anesthesiology and Intensive care, University Hospital Northern Norway, Tromsø, Norway
- ⁵⁵Department of Neurosurgery, Hadassah-hebrew University Medical center, Jerusalem, Israel
- ⁵⁶Fundación Instituto Valenciano de Neurorrehabilitación (FIVAN), Valencia, Spain
- ⁵⁷Department of Neurosurgery, Shanghai Renji hospital, Shanghai Jiaotong University/school of medicine, Shanghai, China
- ⁵⁸Karolinska Institutet, INCF International Neuroinformatics Coordinating Facility, Stockholm, Sweden
- ⁵⁹Emergency Department, CHU, Liège, Belgium
- ⁶⁰Neurosurgery clinic, Pauls Stradins Clinical University Hospital, Riga, Latvia
- ⁶¹Department of Computing, Imperial College London, London, UK
- ⁶²Department of Neurosurgery, Hospital Universitario 12 de Octubre, Madrid, Spain
- ⁶³Department of Anesthesia, Critical Care and Pain Medicine, Medical University of Vienna, Austria
- ⁶⁴Department of Public Health, Erasmus Medical Center-University Medical Center, Rotterdam, The Netherlands
- ⁶⁵College of Health and Medicine, Australian National University, Canberra, Australia
- ⁶⁶Department of Neurosurgery, Neurosciences Centre & JPN Apex trauma centre, All India Institute of Medical Sciences, New Delhi-110029, India
- ⁶⁷Department of Neurosurgery, Erasmus MC, Rotterdam, the Netherlands
- ⁶⁸Department of Neurosurgery, Oslo University Hospital, Oslo, Norway
- ⁶⁹Division of Psychology, University of Stirling, Stirling, UK
- ⁷⁰Division of Neurosurgery, Department of Clinical Neurosciences, Addenbrooke's Hospital & University of Cambridge, Cambridge, UK
- ⁷¹Department of Neurology, University of Groningen, University Medical Center Groningen, Groningen, Netherlands
- ⁷²Neurointensive Care, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
- ⁷³Salford Royal Hospital NHS Foundation Trust Acute Research Delivery Team, Salford, UK
- ⁷⁴Department of Intensive Care and Department of Ethics and Philosophy of Medicine, Erasmus Medical Center, Rotterdam, The Netherlands
- ⁷⁵Department of Clinical Neuroscience, Neurosurgery, Umeå University, Umeå, Sweden
- ⁷⁶Hungarian Brain Research Program—Grant No. KTIA_13_NAP-A-II/8, University of Pécs, Pécs, Hungary
- ⁷⁷Department of Anaesthesiology, University Hospital of Aachen, Aachen, Germany
- ⁷⁸Cyclotron Research Center, University of Liège, Liège, Belgium
- ⁷⁹Centre for Urgent and Emergency Care Research (CURE), Health Services Research Section, School of Health and Related Research (SchARR), University of Sheffield, Sheffield, UK
- ⁸⁰Emergency Department, Salford Royal Hospital, Salford UK
- ⁸¹Institute of Research in Operative Medicine (IFOM), Witten/Herdecke University, Cologne, Germany
- ⁸²VP Global Project Management CNS, ICON, Paris, France
- ⁸³Department of Anesthesiology-Intensive Care, Lille University Hospital, Lille, France
- ⁸⁴Department of Neurosurgery, Rambam Medical Center, Haifa, Israel
- ⁸⁵Department of Anesthesiology & Intensive Care, University Hospitals Southampton NHS Trust, Southampton, UK
- ⁸⁶Cologne-Merheim Medical Center (CMMC), Department of Traumatology, Orthopedic Surgery and Sportmedicine, Witten/Herdecke University, Cologne, Germany
- ⁸⁷Intensive Care Unit, Southmead Hospital, Bristol, Bristol, UK
- ⁸⁸Department of Neurological Surgery, University of California, San Francisco, California, USA
- ⁸⁹Department of Anesthesia & Intensive Care, M. Bufalini Hospital, Cesena, Italy
- ⁹⁰Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany
- ⁹¹Department of Neurosurgery, The Walton centre NHS Foundation Trust, Liverpool, UK
- ⁹²Department of Medical Genetics, University of Pécs, Pécs, Hungary
- ⁹³Department of Neurosurgery, Emergency County Hospital Timisoara, Timisoara, Romania
- ⁹⁴School of Medical Sciences, Örebro University, Örebro, Sweden
- ⁹⁵Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland
- ⁹⁶Analytic and Translational Genetics Unit, Department of Medicine; Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry; Department of Neurology, Massachusetts General Hospital, Boston, MA, USA
- ⁹⁷Program in Medical and Population Genetics; The Stanley Center for Psychiatric Research, The Broad Institute of MIT and Harvard, Cambridge, MA, USA
- ⁹⁸Department of Radiology, University of Antwerp, Edegem, Belgium
- ⁹⁹Department of Anesthesiology & Intensive Care, University Hospital of Grenoble, Grenoble, France
- ¹⁰⁰Department of Anesthesia & Intensive Care, Azienda Ospedaliera Università di Padova, Padova, Italy
- ¹⁰¹Dept. of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands and Dept. of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands
- ¹⁰²Department of Neurosurgery, Helsinki University Central Hospital
- ¹⁰³Division of Clinical Neurosciences, Department of Neurosurgery and Turku Brain Injury Centre, Turku University Hospital and University of Turku, Turku, Finland
- ¹⁰⁴Department of Anesthesiology and Critical Care, Pitié-Salpêtrière Teaching Hospital, Assistance Publique, Hôpitaux de Paris and University Pierre et Marie Curie, Paris, France
- ¹⁰⁵Neurotraumatology and Neurosurgery Research Unit (UNINN), Vall d'Hebron Research Institute, Barcelona, Spain
- ¹⁰⁶Department of Neurosurgery, Kaunas University of technology and Vilnius University, Vilnius, Lithuania
- ¹⁰⁷Department of Neurosurgery, Rezekne Hospital, Latvia
- ¹⁰⁸Department of Anaesthesia, Critical Care & Pain Medicine NHS Lothian & University of Edinburgh, Edinburgh, UK
- ¹⁰⁹Director, MRC Biostatistics Unit, Cambridge Institute of Public Health, Cambridge, UK
- ¹¹⁰Department of Physical Medicine and Rehabilitation, Oslo University Hospital/University of Oslo, Oslo, Norway
- ¹¹¹Division of Orthopedics, Oslo University Hospital, Oslo, Norway
- ¹¹²Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
- ¹¹³Broad Institute, Cambridge MA Harvard Medical School, Boston MA, Mas-

sachusetts General Hospital, Boston MA, USA

¹¹⁴National Trauma Research Institute, The Alfred Hospital, Monash University, Melbourne, Victoria, Australia

¹¹⁵Department of Neurosurgery, Odense University Hospital, Odense, Denmark

¹¹⁶International Neurotrauma Research Organisation, Vienna, Austria

¹¹⁷Klinik für Neurochirurgie, Klinikum Ludwigsburg, Ludwigsburg, Germany

¹¹⁸Division of Biostatistics and Epidemiology, Department of Preventive Medicine, University of Debrecen, Debrecen, Hungary

¹¹⁹Department Health and Prevention, University Greifswald, Greifswald, Germany

¹²⁰Department of Anaesthesiology and Intensive Care, AUVA Trauma Hospital, Salzburg, Austria

¹²¹Department of Neurology, Elisabeth-TweeSteden Ziekenhuis, Tilburg, the Netherlands

¹²²Department of Neuroanesthesia and Neurointensive Care, Odense University Hospital, Odense, Denmark

¹²³Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, NTNU, Trondheim, Norway

¹²⁴Department of Physical Medicine and Rehabilitation, St.Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

¹²⁵Department of Neurosurgery, University of Pécs, Pécs, Hungary

¹²⁶Division of Neuroscience Critical Care, John Hopkins University School of Medicine, Baltimore, USA

¹²⁷Department of Neuropathology, Queen Elizabeth University Hospital and University of Glasgow, Glasgow, UK

¹²⁸Dept. of Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

¹²⁹Department of Pathophysiology and Transplantation, Milan University, and Neuroscience ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milano, Italy

¹³⁰Department of Radiation Sciences, Biomedical Engineering, Umeå University, Umeå, Sweden

¹³¹Perioperative Services, Intensive Care Medicine and Pain Management, Turku University Hospital and University of Turku, Turku, Finland

¹³²Department of Neurosurgery, Kaunas University of Health Sciences, Kaunas, Lithuania

¹³³Intensive Care and Department of Pediatric Surgery, Erasmus Medical Center, Sophia Children's Hospital, Rotterdam, The Netherlands

¹³⁴Department of Neurosurgery, Kings college London, London, UK

¹³⁵Neurologie, Neurochirurgie und Psychiatrie, Charité – Universitätsmedizin Berlin, Berlin, Germany

¹³⁶Department of Intensive Care Adults, Erasmus MC – University Medical Center Rotterdam, Rotterdam, the Netherlands

¹³⁷icoMetrix NV, Leuven, Belgium

¹³⁸Movement Science Group, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, UK

¹³⁹Psychology Department, Antwerp University Hospital, Edegem, Belgium

¹⁴⁰Director of Neurocritical Care, University of California, Los Angeles, USA

¹⁴¹Department of Neurosurgery, St.Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

¹⁴²Department of Emergency Medicine, University of Florida, Gainesville, Florida, USA

¹⁴³Department of Neurosurgery, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

¹⁴⁴VTT Technical Research Centre, Tampere, Finland

¹⁴⁵Section of Neurosurgery, Department of Surgery, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

Author contributions

GC conceived and supervised the project, participated in the data analysis, revised the first version of the manuscript the manuscript, and the supplementary tables. CR participated in the data analysis, drafted the manuscript, the supplementary tables and collected the COIs. SG, MP, and PR analysed the data, drafted the manuscript, and the supplementary material. LM, ER, DKM, and NS were an active part of the manuscript drafting and revision. GC, CR, SG, MP, PR have verified the underlying data. DKM was one of the two coordinators of the CENTER-TBI study, and GC and NS were Work Package leaders. GC, CR, SG and DKM discussed the findings with all the authors. All co-authors

gave substantial feedback on the manuscript and approved the final version of it.

Funding

Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement. The Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI study, registered at clinicaltrials.gov NCT02210221) was funded by the FW7 program of the European Union (602150). Additional funding was obtained from the Hannelore Kohl Stiftung (Germany), from OneMind (USA) and Integra LifeSciences Corporation (USA). The funder had no role in the design of the study, the collection, analysis, and interpretation of data, or in writing the manuscript.

Declarations

Conflict of interest

GC reports grants, personal fees as Speakers' Bureau Member and Advisory Board Member from Integra and Neuroptics. DKM reports grants from the European Union and UK National Institute for Health Research, during the conduct of the study; grants, personal fees, and non-financial support from GlaxoSmithKline; personal fees from Neurotrauma Sciences, Lantmaanen AB, Pressura, and Pfizer, outside of the submitted work. The other authors declare that they have no competing interests.

Ethics approval and consent to participate

The Medical Ethics Committees of all participating centers approved the CENTER-TBI study, and informed consent was obtained according to local regulations. (<https://www.center-tbi.eu/project/ethical-approval>).

Open Access

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 28 May 2021 Accepted: 26 June 2021

Published online: 24 July 2021

References


1. Hoiland RL, Fisher JA, Ainslie PN (2019) Regulation of the Cerebral Circulation by Arterial Carbon Dioxide. In: Compr. Physiol. Wiley. <https://doi.org/10.1002/cphy.c180021>
2. Gouvea Bogossian E, Peluso L, Creteur J, Taccone FS (2021) Hyperventilation in Adult TBI Patients: How to Approach It? *Neurol Front*. <https://doi.org/10.3389/fneur.2020.580859>
3. Lundberg N, Kjallquist A, Bien C (1949) Reduction of increased intracranial pressure by hyperventilation. A therapeutic aid in neurological surgery, *Acta Psychiatr. Scand. Suppl.* 34:1–64. <http://www.ncbi.nlm.nih.gov/pubmed/14418913>.
4. Godoy DA, Seifi A, Garza D, Lubillo-Montenegro S, Murillo-Cabezas F (2017) Hyperventilation Therapy for Control of Posttraumatic Intracranial Hypertension. *Front Neurol*. <https://doi.org/10.3389/fneur.2017.00250>
5. Stocchetti N, Maas AIR, Chierogato A, van der Plas AA (2005) Hyperventilation in Head Injury. *Chest* 127:1812–1827. <https://doi.org/10.1378/chest.127.5.1812>

6. Stocchetti N, Carbonara M, Citerio G, Ercole A, Skrifvars MB, Smielewski P, Zoerle T, Menon DK (2017) Severe traumatic brain injury: targeted management in the intensive care unit. *Lancet Neurol* 16:452–464. [https://doi.org/10.1016/S1474-4422\(17\)30118-7](https://doi.org/10.1016/S1474-4422(17)30118-7)
7. Coles JP, Minhas PS, Fryer TD, Smielewski P, Aigbirihio F, Donovan T, Downey SPMJ, Williams G, Chatfield D, Matthews JC, Gupta AK, Carpenter TA, Clark JC, Pickard JD, Menon DK (2002) Effect of hyperventilation on cerebral blood flow in traumatic head injury: of clinical relevance and monitoring correlates*. *Crit Care Med* 30:1950–1959. <https://doi.org/10.1097/00003246-200209000-00002>
8. Curley G, Kavanagh BP, Laffey JG (2010) Hypocapnia and the injured brain: More harm than benefit. *Crit Care Med* 38:1348–1359. <https://doi.org/10.1097/CCM.0b013e3181d8cf2b>
9. Cnossen MC, Huijben JA, van der Jagt M, Volovici V, van Essen T, Polinder S, Nelson D, Ercole A, Stocchetti N, Citerio G, Peul WC, Maas AIR, Menon D, Steyerberg EW, Lingsma HF (2017) Variation in monitoring and treatment policies for intracranial hypertension in traumatic brain injury: a survey in 66 neurotrauma centers participating in the CENTER-TBI study. *Crit Care* 21:233. <https://doi.org/10.1186/s13054-017-1816-9>
10. N Carney, AM Totten, C O'Reilly, JS Ullman, GWJ Hawryluk, MJ Bell, SL Bratton, R Chesnut, OA Harris, N Kisssoon, AM Rubiano, L Shutter, RC Tasker, MS Vavilala, J Wilberger, DW Wright, J Ghajar (2017) Guidelines for the Management of Severe Traumatic Brain Injury, Fourth Edition. *Neurosurgery*. <https://doi.org/10.1227/NEU.0000000000001432>
11. Roberts I, Schierhout G (1997) Hyperventilation therapy for acute traumatic brain injury. *Cochrane Database Syst Rev*. <https://doi.org/10.1002/14651858.CD000566>
12. Robba C, Poole D, McNett M, Asehnoun K, Bösel J, Bruder N, Chierigato A, Cinotti R, Duranteau J, Einav S, Ercole A, Ferguson N, Guerin C, Siempos II, Kurtz P, Juffermans NP, Mancoff J, Mascia L, McCredie V, Nin N, Oddo M, Pelosi P, Rabinstein AA, Neto AS, Seder DB, Skrifvars MB, Suarez JI, Taccone FS, van der Jagt M, Citerio G, Stevens RD (2020) Mechanical ventilation in patients with acute brain injury: recommendations of the European Society of Intensive Care Medicine consensus. *Intensive Care Med*. <https://doi.org/10.1007/s00134-020-06283-0>
13. Hawryluk GWJ, Aguilera S, Buki A, Bulger E, Citerio G, Cooper DJ, Arrastia RD, Diringner M, Figaji A, Gao G, Geocadin R, Ghajar J, Harris O, Hoffer A, Hutchinson P, Joseph M, Kitagawa R, Manley G, Mayer S, Menon DK, Meyfroidt G, Michael DB, Oddo M, Okonkwo D, Patel M, Robertson C, Rosenfeld JV, Rubiano AM, Sahuquillo J, Servadei F, Shutter L, Stein D, Stocchetti N, Taccone FS, Timmons S, Tsai E, Ullman JS, Vespa P, Videtta W, Wright DW, Zammit C, Chesnut RM (2019) A management algorithm for patients with intracranial pressure monitoring: the Seattle International Severe Traumatic Brain Injury Consensus Conference (SIBICC). *Intensive Care Med*. <https://doi.org/10.1007/s00134-019-05805-9>
14. Maas AIR, Menon DK, Steyerberg EW, Citerio G, Lecky F, Manley GT, Hill S, Legrand V, Sorgner A (2015) Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI). *Neurosurgery*. <https://doi.org/10.1227/NEU.0000000000000575>
15. Steyerberg EW, Wieggers E, Sewalt C, et al (2019) Case-mix, care pathways, and outcomes in patients with traumatic brain injury in CENTER-TBI: a European prospective, multicentre, longitudinal, cohort study. *Lancet Neurol*. [https://doi.org/10.1016/S1474-4422\(19\)30232-7](https://doi.org/10.1016/S1474-4422(19)30232-7)
16. Huijben JA, Wieggers EJA, Lingsma HF, Citerio G, Maas AIR, Menon DK, Ercole A, Nelson D, van der Jagt M, Steyerberg EW, Helbok R, Lecky F, Peul W, Birg T, Zoerle T, Carbonara M, Stocchetti N (2020) Changing care pathways and between-center practice variations in intensive care for traumatic brain injury across Europe: a CENTER-TBI analysis. *Intensive Care Med* 46:995–1004. <https://doi.org/10.1007/s00134-020-05965-z>
17. Huijben JA, Dixit A, Stocchetti N, Maas AIR, Lingsma HF, van der Jagt M, Nelson D, Citerio G, Wilson L, Menon DK, Ercole A (2021) Use and impact of high intensity treatments in patients with traumatic brain injury across Europe: a CENTER-TBI analysis. *Crit Care* 25:78. <https://doi.org/10.1186/s13054-020-03370-y>
18. Wilson JTL, Pettigrew LEL, Teasdale GM (1998) Structured Interviews for the Glasgow Outcome Scale and the Extended Glasgow Outcome Scale: Guidelines for Their Use. *J Neurotrauma*. <https://doi.org/10.1089/neu.1998.15.573>
19. Maas AIR, Marmarou A, Murray GD, Teasdale SGM, Steyerberg EW (2007) Prognosis and Clinical Trial Design in Traumatic Brain Injury: The IMPACT Study. *J Neurotrauma* 24:232–238. <https://doi.org/10.1089/neu.2006.0024>
20. Cnossen M, van Essen TA, Ceyisakar IE, Polinder S, Andriessen T, van der Naalt J, Haitsma I, Horn J, Franschman G, Vos P, Peul W, Menon DK, Maas A, Steyerberg E, Lingsma H (2018) Adjusting for confounding by indication in observational studies: a case study in traumatic brain injury. *Clin Epidemiol* 10:841–852. <https://doi.org/10.2147/CLEPS154500>
21. Vik A, Nag T, Fredriksli OA, Skandsen T, Moen KG, Schirmer-Mikalsen K, Manley GT (2008) Relationship of “dose” of intracranial hypertension to outcome in severe traumatic brain injury. *J Neurosurg* 109:678–684. <https://doi.org/10.3171/JNS/2008/109/10/0678>
22. Muizelaar JP, Marmarou A, Ward JD, Kontos HA, Choi SC, Becker DP, Grueber H, Young HF (1991) Adverse effects of prolonged hyperventilation in patients with severe head injury: a randomized clinical trial. *J Neurosurg*. <https://doi.org/10.3171/jns.1991.75.5.0731>
23. Coles JP, Fryer TD, Coleman MR, Smielewski P, Gupta AK, Minhas PS, Aigbirihio F, Chatfield DA, Williams GB, Boniface S, Carpenter TA, Clark JC, Pickard JD, Menon DK (2007) Hyperventilation following head injury: Effect on ischemic burden and cerebral oxidative metabolism*. *Crit Care Med*. <https://doi.org/10.1097/01.CCM.0000254066.37187.88>
24. Brandi G, Stocchetti N, Pagnamenta A, Stretti F, Steiger P, Klinzing S (2019) Cerebral metabolism is not affected by moderate hyperventilation in patients with traumatic brain injury. *Crit Care*. <https://doi.org/10.1186/s13054-018-2304-6>
25. Diringner MN, Videen TO, Yundt K, Zazulia AR, Aiyagari V, Dacey RG, Grubb RL, Powers WJ (2002) Regional cerebrovascular and metabolic effects of hyperventilation after severe traumatic brain injury. *J Neurosurg* 96:103–108. <https://doi.org/10.3171/jns.2002.96.1.0103>
26. Diringner MN, Yundt K, Videen TO, Adams RE, Zazulia AR, Deibert E, Aiyagari V, Dacey RG, Grubb RL, Powers WJ (2000) No reduction in cerebral metabolism as a result of early moderate hyperventilation following severe traumatic brain injury. *J Neurosurg* 92:7–13. <https://doi.org/10.3171/jns.2000.92.1.0007>
27. Neumann J-O, Chambers IR, Citerio G, Enblad P, Gregson BA, Howells T, Mattern J, Nilsson P, Piper I, Ragauskas A, Sahuquillo J, Yau YH, Kiening K (2008) The use of hyperventilation therapy after traumatic brain injury in Europe: an analysis of the brain database. *Intensive Care Med* 34:1676. <https://doi.org/10.1007/s00134-008-1123-7>
28. Gordon E (1971) Controlled respiration in the management of patients with traumatic brain injuries. *Acta Anaesthesiol Scand*. <https://doi.org/10.1111/j.1399-6576.1971.tb05461.x>

ORIGINAL



High arterial oxygen levels and supplemental oxygen administration in traumatic brain injury: insights from CENTER-TBI and OzENTER-TBI

Emanuele Rezoagli^{1,2}, Matteo Petrosino³, Paola Reborà³, David K. Menon⁴, Stefania Mondello⁵, D. James Cooper^{6,7}, Andrew I. R. Maas⁸, Eveline J. A. Wiegers⁹, Stefania Galimberti³ and Giuseppe Citerio^{1,10*}  on behalf of CENTER-TBI, OzENTER-TBI Participants and Investigators

© 2022 The Author(s)

Abstract

Purpose: The effect of high arterial oxygen levels and supplemental oxygen administration on outcomes in traumatic brain injury (TBI) is debated, and data from large cohorts of TBI patients are limited. We investigated whether exposure to high blood oxygen levels and high oxygen supplementation is independently associated with outcomes in TBI patients admitted to the intensive care unit (ICU) and undergoing mechanical ventilation.

Methods: This is a secondary analysis of two multicenter, prospective, observational, cohort studies performed in Europe and Australia. In TBI patients admitted to ICU, we describe the arterial partial pressure of oxygen (PaO₂) and the oxygen inspired fraction (FiO₂). We explored the association between high PaO₂ and FiO₂ levels within the first week with clinical outcomes. Furthermore, in the CENTER-TBI cohort, we investigate whether PaO₂ and FiO₂ levels may have differential relationships with outcome in the presence of varying levels of brain injury severity (as quantified by levels of glial fibrillary acidic protein (GFAP) in blood samples obtained within 24 h of injury).

Results: The analysis included 1084 patients (11,577 measurements) in the CENTER-TBI cohort, of whom 55% had an unfavorable outcome, and 26% died at a 6-month follow-up. Median PaO₂ ranged from 93 to 166 mmHg. Exposure to higher PaO₂ and FiO₂ in the first seven days after ICU admission was independently associated with a higher mortality rate. A trend of a higher mortality rate was partially confirmed in the OzENTER-TBI cohort ($n = 159$). GFAP was independently associated with mortality and functional neurologic outcome at follow-up, but it did not modulate the outcome impact of high PaO₂ levels, which remained independently associated with 6-month mortality.

*Correspondence: giuseppe.citerio@unimib.it

¹ School of Medicine and Surgery, University of Milano - Bicocca, Monza, Italy

Full author information is available at the end of the article

CENTER-TBI ICU and OzENTER-TBI Participants and Investigators are listed in the Acknowledgements section.

Conclusions: In two large prospective multicenter cohorts of critically ill patients with TBI, levels of PaO₂ and FiO₂ varied widely across centers during the first seven days after ICU admission. Exposure to high arterial blood oxygen or high supplemental oxygen was independently associated with 6-month mortality in the CENTER-TBI cohort, and the severity of brain injury did not modulate this relationship. Due to the limited sample size, the findings were not wholly validated in the external OzENTER-TBI cohort. We cannot exclude the possibility that the worse outcomes associated with higher PaO₂ were due to use of higher FiO₂ in patients with more severe injury or physiological compromise. Further, these findings may not apply to patients in whom FiO₂ and PaO₂ are titrated to brain tissue oxygen monitoring (PbtO₂) levels. However, at minimum, these findings support the need for caution with oxygen therapy in TBI, particularly since titration of supplemental oxygen is immediately applicable at the bedside.

Keywords: PaO₂, FiO₂, Traumatic brain injury, GOSE, Mortality, GFAP

Introduction

In patients with traumatic brain injury (TBI), hypoxemia is a major predictor of hospital and 6-month mortality [1]. Oxygen supplementation aims to reverse tissue hypoxia and, thus, improve cell viability, organ function, and survival in critically ill patients [2]. However, this may lead to administering more oxygen than needed to patients admitted to the intensive care unit (ICU) [3].

While hyperbaric oxygen is known to be neurotoxic [4], it is not clear whether high normobaric oxygen levels may play a detrimental role in the brain [5]. Hyperoxia, i.e., high inspiratory oxygen fraction, may be associated with excitotoxicity in severe TBI [6]. Furthermore, hyperoxemia, i.e., high blood oxygen partial pressure levels, may potentially worsen organ injury and impact the case fatality rate of critically ill patients with TBI [7, 8]. Therefore, not only too low but even extreme hyperoxemia might cause injury in TBI patients, as David et al. showed [9]. Data on more than 36,000 mixed ICU patients mechanically ventilated with early arterial partial pressure of oxygen (PaO₂) suggested an independent U-shape association with hospital mortality [10]. A recent meta-analysis of 32 studies in acute brain-damaged patients highlighted that hyperoxemia, differently defined across studies, was associated with an increased risk of poor neurological outcomes [11]. Patients with a poor neurological outcome also had a significantly higher maximum PaO₂ and mean PaO₂. These associations were present, especially in patients with subarachnoid hemorrhage and ischemic stroke, but not in traumatic brain injured.

Currently, there is no evidence to support the role of hyperoxemia or hyperoxia in a large real-world dataset of critically ill patients admitted to ICU with severe TBI [12–14].

Therefore, we described variability across centers in the blood oxygen levels (i.e., PaO₂) and oxygen supplementation distributions (i.e., inspiratory oxygen fraction, FiO₂)

Take-home message

In two large prospective multicenter cohorts of traumatic brain injured patients, arterial and supplemental oxygen levels varied widely across centers during the first seven days after admission to the intensive care unit.

Exposure to high arterial blood oxygen or high supplemental oxygen—a therapeutic gas immediately titratable at the bedside—was independently associated with 6-month mortality, regardless of brain injury severity.

and investigated whether high PaO₂ and FiO₂ levels are associated with worse 6-month outcomes. We validated our findings in the multicenter Australian OzENTER-TBI database [15]. Finally, we explored whether PaO₂ and FiO₂ levels may contribute differently to outcomes in the presence of increasing levels of glial fibrillary acidic protein (GFAP), a biomarker of brain injury severity.

The aims of this study are to:

1. Describe the values and the differences in PaO₂ and FiO₂ in the first week from ICU admission in mechanically ventilated TBI patients across centers in CENTER-TBI;
2. assess whether high levels of PaO₂ or FiO₂ are independently associated with 6-month mortality and unfavorable neurologic outcome in CENTER-TBI;
3. evaluate whether the impact of high levels of oxygen exposure or high levels of supplemental oxygen on 6-month outcome could be worsened by increasing brain injury severity, as assessed by acute (first 24 h) serum levels of GFAP in the CENTER-TBI cohort.

All these objectives (except the last one) were subsequently validated in an external cohort of patients with traumatic brain injury from OzENTER-TBI. Hypotheses of the current analyses were that exposure to high oxygen and FiO₂ levels in TBI patients mechanically ventilated and admitted to ICU may promote brain injury and have

a negative impact on both functional neurological disability and survival.

Methods

Study design and patients

The Collaborative European NeuroTrauma Effectiveness in Research in Traumatic Brain Injury (CENTER-TBI study, registered at clinicaltrials.gov NCT02210221) is a longitudinal, prospective data collection from TBI patients across 65 centers in Europe between December 2014 and December 2017. The design and the results of the screening and enrollment process have been previously described [12, 13]. The Australia–Europe NeuroTrauma Effectiveness Research in Traumatic Brain Injury OzENTER-TBI Study was conducted in two designated adult major trauma centers in Victoria, Australia, between February 2015 and March 2017 [15]. The Medical Ethics Committees approved both studies in all participating centers, and informed consent was obtained according to local regulations (<https://www.center-tbi.eu/project/ethical-approval>). Therefore, the studies have been performed per the ethical standards of the Declaration of Helsinki and its later amendments.

In the OzENTER-TBI Study, patients or families were allowed to opt out of data collection. OzENTER-TBI was used as an external validation cohort.

Before starting the analysis, this project on PaO₂ management was preregistered on the CENTER-TBI proposal platform and approved by the CENTER-TBI proposal review committee.

We included all patients in the CENTER-TBI Core study who had:

- a TBI necessitating ICU admission,
- tracheal intubation and mechanical ventilation,
- at least two PaO₂ measurements in the first seven days.

These inclusion criteria were also applied to select patients from the OzENTER-TBI study for the validation cohort.

This report complies with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines.

Data collection and definitions

Detailed information on data collection is available on the study website (<https://www.center-tbi.eu/data/dictionary>). The daily lowest and highest PaO₂ and FiO₂ values from arterial blood gases—that were collected as per the case report form—were evaluated in this study. Specifically, we investigated the role of variables representing different aspects of arterial oxygen levels and

supplemental oxygen administration during the first week of ICU admission, including:

- The highest PaO₂ (PaO_{2max}) and FiO₂ (FiO_{2max}) exposures.
- The mean of the highest daily PaO₂ (PaO_{2mean}) and FiO₂ (FiO_{2mean}).
- The mean of the swings of PaO₂ (Δ PaO_{2mean}) and of FiO₂ (Δ FiO_{2mean}). The swings were calculated daily as the difference between the highest and the lowest PaO₂ and FiO₂. They represent the average day-to-day variability of PaO₂ and FiO₂.

Mortality and functional neurological outcome measured as the 8-point Extended Glasgow Outcome Score (GOSE) were assessed six months post-injury. An unfavorable outcome was defined as GOSE \leq 4 (i.e., low and upper severe disability, vegetative state, or dead), including both poor functional outcome and mortality. All responses were obtained by trained study personnel—blinded to the PaO₂ and FiO₂ data—from patients or from a proxy (where impaired cognitive capacity prevented patient interview), during a face-to-face visit, by telephone interview, or by postal questionnaire around six months after injury [16].

In CENTER-TBI, the severity of brain injury, traditionally evaluated with clinical and neuroradiologic elements, was also gauged by serum brain injury biomarkers. For this study, a decision was made to use GFAP, a glial cytoskeletal protein, as a proxy measure of brain injury severity. GFAP was the brain injury biomarker with the highest discriminative performance on computed tomography (CT) brain injury [17], and it is strongly associated with mortality and long-term outcomes after injury [18, 19]. GFAP within 24 h after trauma was quantified by an ultrasensitive immunoassay using digital array technology (Single Molecule Arrays, SiMoA)-based assay (Quanterix Corp., Lexington, MA).

Statistical methods

Patient characteristics were described by medians (interquartile range, IQR) or means (standard deviations, SD) as appropriate and counts or proportions. The role of PaO_{2max}, FiO_{2max}, PaO_{2mean}, FiO_{2mean} or Δ PaO_{2mean}, Δ FiO_{2mean} (one at a time) on 6-month mortality and unfavorable neurological outcome was evaluated through mixed-effect logistic regression models, adjusting for the IMPACT core covariates (age, Glasgow Coma Scale (GCS) motor score and pupillary reactivity) and injury severity score (ISS), with the center as a random effect. The assumption of linearity of the effect for continuous variables was evaluated using

splines, and the results of the models were reported as odds ratios (OR) along with the corresponding 95% confidence intervals (CI). To simplify the clinical interpretation of the OR of the exposure variables, PaO₂ and FiO₂ increases were referred to 10 mmHg and 0.1 each, respectively. Then, we enriched the models, including GFAP, which was log-transformed to satisfy the linearity assumption. We also investigated a potential interaction between GFAP and the six variables representing the oxygen status (one at a time) through a flexible approach based on restricted cubic splines and tensor-product splines. The final models were selected using standard statistical performance measures such as Akaike Information Criteria (AIC) and likelihood ratio tests for non-nested and nested models. Finally, we used data from the OzENTER-TBI cohort to validate our findings through the same modeling approach used for CENTER-TBI. However, here we omitted the random term for centers, while including the only two centers in the study as a dummy variable. Analyses were done on complete cases and using the MICE algorithm for multiple imputations of missing data (ten imputed datasets). Tests were performed two-sided with a significance alpha level of 5%. To protect from the risk of alpha inflation in testing the effect of arterial oxygen levels and supplemental oxygen administration on outcomes, we also adjusted the *p* values in the models according to the approach of Benjamini–Hochberg. All analyses were conducted using R statistical software (version 4.03).

Results

Of the 4509 patients included in the CENTER-TBI dataset, 2138 subjects were admitted to ICU and, among these, 1084 (median age was 49 [29–65], and 75% male) from 51 centers fulfilled the inclusion criteria (Supplemental Fig. 1). Half of the population experienced thoracic trauma, which in 41.5% of the cases was major.

All 198 patients included in the OzENTER-TBI dataset were admitted to ICU and, among these, 159 fulfilled the inclusion criteria (Supplemental Figure 1). In OzENTER-TBI, the median age was 39 [24–65], and 77% of the population was male. Almost 55% of the population experienced thoracic trauma, which in 46.5% of the cases was severe or critical. A comprehensive description of the population of the CENTER-TBI and OzENTER-TBI study is reported in Table 1. Patient characteristics stratified by 6-month mortality are described in Supplemental Table 1 (CENTER-TBI) and Supplemental Table 2 (OzENTER-TBI). We focused on the highest PaO₂ and FiO₂ daily levels in the current analysis in both cohorts.

CENTER-TBI

Arterial oxygen levels and supplemental oxygen administration

During the first week of ICU admission, a total of 11,577 measurements of PaO₂ were available (5747 lowest and 5830 highest daily values), for an overall median of PaO₂ and FiO₂ of 112 mmHg (IQR 86–144) and 0.4 (IQR 0.3–0.5), respectively. A total of 526 (48.5%) patients had complete daily measurements of high PaO₂ during the first week (median of 6 measures, IQR 4–7). The remaining patients had, respectively, 6 (136, 12.5%), 5 (72, 6.6%), 4 (89, 8.2%), 3 (94, 8.7%) and 2 (167, 15.4%) daily measurements of PaO₂. The median highest PaO₂ level during the first seven days since ICU admission was 134 mmHg (IQR 113–167). The median of highest FiO₂ levels during the first seven days since ICU admission was 0.45 (IQR 0.40–0.5) (Supplemental Fig. 2). Mean PaO_{2max}, PaO_{2mean} and ΔPaO_{2mean} were 231, 156 and 57 mmHg, respectively. PaO_{2max} showed a strong correlation with ΔPaO_{2mean} ($T_{\text{Kendall}}=0.51$, 95% CI [0.48–0.53]) and with PaO_{2mean} ($T_{\text{Kendall}}=0.66$, 95% CI [0.64–0.68]). Mean FiO_{2max}, FiO_{2mean} and ΔFiO_{2mean} were 0.59, 0.45 and 0.05 mmHg, respectively (Table 1). The highest PaO₂ levels varied widely across centers, with the center-specific median ranging from 88 to 170 mmHg and the highest PaO₂ levels within center ranging from 162 to 612 mmHg. Similarly, the highest median FiO₂ levels during the first seven days since ICU admission varied widely across centers ranging from 0.21 to 0.96. Center variability in PaO₂ (panel A) and FiO₂ levels (panel B) across centers is represented in Fig. 1. Of note, overall median PaO₂ levels in patients with brain tissue oxygen monitoring (PbtO₂) were similar compared to the patient population with no PbtO₂ monitoring (133 versus 137 mmHg, data not shown) (Supplemental Fig. 3).

Arterial oxygen levels and outcomes in TBI patients

Data on mortality and neurological functional score GOSE at 6 months were available in 967 (89.2%) TBI patients. Five hundred and twenty-eight patients (54.6%) had an unfavorable GOSE at a 6-month follow-up, and 252 died within that period (26.1%). After adjusting, we estimated the OR for a 10 mmHg increase in PaO₂. We found that both PaO_{2max} (OR 1.02, 95% CI 1–1.04) and ΔPaO_{2mean} (OR 1.07, 95% CI 1.03–1.12) were independently associated with an unfavorable functional neurologic outcome as expressed by a GOSE score ≤ 4 at 6-month follow-up (Model 1, Table 2 and Supplemental Table 3 for the estimates in the complete regression model). Furthermore, we observed that all the exposure variables to high PaO₂ were positively associated with an increased risk of mortality (PaO_{2max}, OR 1.03, 95% CI 1.01–1.05; PaO_{2mean}, OR 1.08, 95% CI 1.04–1.13;

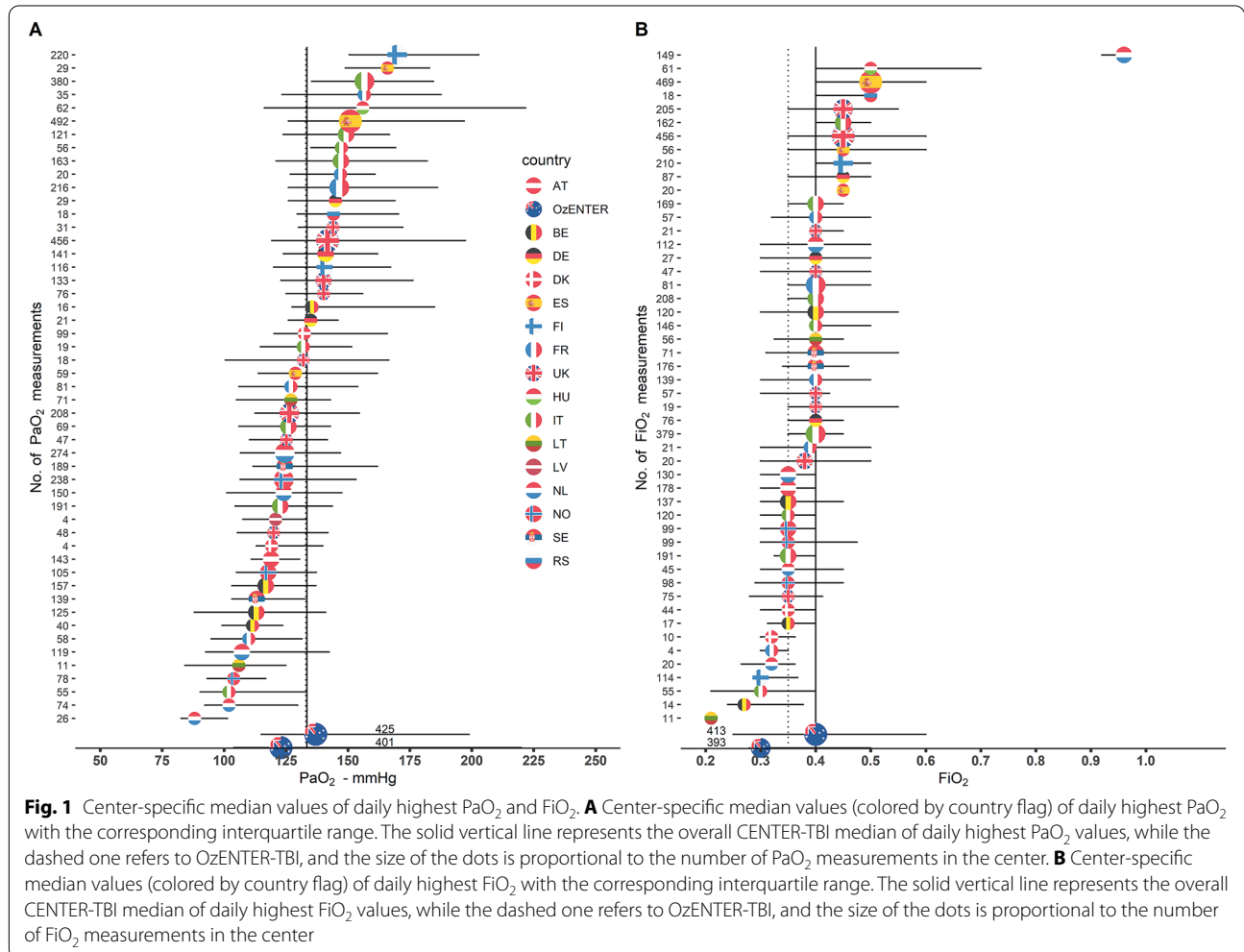
Table 1 Characteristics of the study cohorts from CENTER-TBI and OzENTER-TBI

Variable	Level	CENTER-TBI (N = 1084)	OzENTER-TBI (N = 159)
Demographic characteristics			
Age, median [IQR]		49 [29–65]	39 [24–65]
Sex, n (%)	Female	270 (25)	37 (23)
	Male	814 (75)	122 (77)
Clinical presentation			
Hypotension, n (%)	No	843 (77.9)	116 (73)
	Yes	239 (22.1)	43 (27)
	NA (n)	2	0
Hypoxia, n (%)	No	1030 (95)	157 (98.7)
	Yes	54 (5)	2 (1.3)
Injury Severity Score, median [IQR]		34 [25–45]	29 [25–38]
	NA (n)	3	0
pH, median [IQR]	Lowest	7.34 [7.29–7.39]	7.33 [7.29–7.37]
	NA (n)	20	0
	Highest	7.43 [7.39–7.47]	7.41 [7.38–7.45]
	NA (n)	6	0
Neurological presentation			
Pupillary reactivity, n (%)	Both reactive	790 (72.9)	119 (74.8)
	One reactive	87 (8)	11 (7)
	Both unreactive	157 (14.5)	25 (15.7)
	NA	50 (4.6)	4 (2.5)
GCS Motor Score, n (%)	Localizes/obeys	419 (38.7)	33 (20.7)
	None/extension	493 (45.5)	117 (73.6)
	Any flexion	151 (13.9)	8 (5)
	NA	21 (1.9)	1 (0.7)
GCS score, n (%)	GCS > 8	370 (34.1)	58 (36.5)
	GCS ≤ 8	657 (60.6)	97 (61)
	NA	57 (5.3)	4 (2.5)
ICP at ICU admission, median [IQR]		8 [4–14]	11 [7–15]
	NA (n)	521	108
Mean ICP, median [IQR]		11 [6–15]	11 [8–15]
	NA (n)	521	108
Brain injury severity			
Marshall CT Classification, median [IQR]		3 [2–6]	2 [2–6]
	NA (n)	105	21
GFAP, median [IQR]	ng/mL	20.5 [7–50.8]	/
	NA (n)	198	159
Oxygenation			
Day 1 PaO ₂ overall, mean (SD)	mmHg	207.17 (99.91)	328.18 (144.46)
PaO ₂ mean, mean (SD)	mmHg	155.79 (46.93)	197.79 (73.79)
PaO ₂ max, mean (SD)	mmHg	230.92 (102.95)	356.01 (134.47)
ΔPaO ₂ mean, mean (SD)	mmHg	57 (36.7)	98.20 (59.95)
Day 1—PaO ₂ /FiO ₂ , mean (SD)	mmHg	412.48 (197.08)	453.59 (207.1)
Day 1 FiO ₂ overall, mean (SD)		0.54 (0.21)	0.76 (0.26)
FiO ₂ mean, mean (SD)		0.45 (0.15)	0.48 (0.15)
FiO ₂ max, mean (SD)		0.59 (0.22)	0.82 (0.23)
ΔFiO ₂ mean, mean (SD)		0.05 (0.08)	0.15 (0.11)
Functional neurologic outcome			
GOSE 6-month follow-up, n (%)			
GOSE ≤ 4		528 (48.7)	53 (33.3)
GOSE > 4		439 (40.5)	95 (59.7)
NA		117 (10.8)	11 (7)

Table 1 (continued)

Hypotension was defined as a documented systolic blood pressure < 90 mmHg; hypoxia was defined as a documented partial pressure of oxygen (PaO_2) < 8 kPa (60 mmHg), oxygen saturation (SaO_2) < 90%, or both

CT computed tomography, GCS Glasgow Coma Scale, GFAP gliofibrillar acid protein, GOSE Glasgow Outcome Scale Extended, ICP intracranial pressure, ICU intensive care unit, IQR interquartile range, NA not available, SD standard deviation



$\Delta\text{PaO}_{2\text{mean}}$, OR 1.14, 95% CI 1.08–1.2; all estimates for 10 mmHg) (Model 1, Table 2 and Supplemental Table 4). A detailed description of all confounders estimates for both outcomes is described in Supplemental Tables 3 and 4. The estimated probability of mortality from the regression model by arterial oxygen levels is depicted in Fig. 2 (Panel A, B, C).

We also explored the role of exposure to high blood oxygen levels on the neurologic outcome by further adjusting the model for GFAP levels. GFAP was positively associated with a lower GOSE score and a higher mortality rate. Among the variables representing higher blood oxygenation, the $\Delta\text{PaO}_{2\text{mean}}$ confirmed its positive

association with a lower GOSE, while all the three high oxygenation variables remained positively associated with a higher mortality rate (Model 2, Table 2). A detailed description of all confounders estimates is reported in Supplemental Tables 5 and 6. We explored the interaction between exposure to high $\text{PaO}_{2\text{max}}$ and GFAP levels on GOSE and mortality. We did not find any interaction between the studied variables, as shown in Fig. 3 (panels A and B), where the surfaces that represent the smoothed interactions (on log scale) are mainly flattened on zero.

Table 2 Multivariable models on GOSE and mortality at 6-month follow-up in CENTER-TBI (Models 1, 2 and 3)

CENTER-TBI		6-month GOSE N = 912 patients, 489 GOSE ≤ 4			6-month mortality N = 912 patients, 225 died		
Model 1		OR ^a	95% CI	p value	OR ^a	95% CI	p value
PaO _{2max} (for 10 mmHg increase)		1.02	1–1.04	0.014	1.03	1.01–1.05	0.002
PaO _{2mean} (for 10 mmHg increase)		1.03	1–1.07	0.059	1.08	1.04–1.13	<0.001
ΔPaO _{2mean} (for 10 mmHg increase) ^b		1.07	1.03–1.12	0.001	1.14	1.08–1.20	<0.001
CENTER-TBI		6-month GOSE N = 764 patients, 407 GOSE ≤ 4			6-month mortality N = 764 patients, 175 died		
Model 2		OR ^a	95% CI	p value	OR ^a	95% CI	p value
Logarithm GFAP		1.51	1.33–1.71	<0.001	1.51	1.29–1.77	<0.001
PaO _{2max} (for 10 mmHg increase)		1.02	1–1.03	0.064	1.03	1.01–1.05	0.008
Logarithm GFAP		1.52	1.34–1.72	<0.001	1.52	1.3–1.78	<0.001
PaO _{2mean} (for 10 mmHg increase)		1.03	0.99–1.07	0.092	1.09	1.04–1.14	0.001
Logarithm GFAP		1.52	1.34–1.72	<0.001	1.53	1.3–1.81	<0.001
ΔPaO _{2mean} (for 10 mmHg increase)		1.05	1–1.11	0.031	1.14	1.08–1.21	<0.001
CENTER-TBI		6-month GOSE N = 877 patients, 470 GOSE ≤ 4			6-month mortality N = 877 patients, 212 died		
Model 3		OR ^c	95% CI	p value	OR ^c	95% CI	p value
FiO _{2max} (for 0.1 increase)		1.03	0.96–1.1	0.453	1.18	1.08–1.29	<0.001
FiO _{2mean} (for 0.1 increase)		1.02	0.92–1.14	0.694	1.31	1.13–1.51	<0.001
ΔFiO _{2mean} (for 0.1 increase)		1.03	0.84–1.27	0.761	1.46	1.13–1.88	0.004

Model 1. Adjusted odds ratio with 95% confidence intervals of exposure to high blood oxygen levels within 7 days of ICU admission on GOSE and mortality at 6-month follow-up in CENTER-TBI. Mixed-effect logistic regression models adjusted for age, pupillary reactivity (both reactive, one reactive, both unreactive), GCS motor (any flexion, none/extension, localizes/obey), Injury Severity Score, and, once at a time, PaO_{2max}, PaO_{2mean}, and ΔPaO_{2mean} for CENTER-TBI with center as a random effect. **Model 2.** Model 1 plus the degree of brain injury quantified as GFAP levels. **Model 3.** Adjusted odds ratio with 95% CI of GOSE and mortality at 6-month follow-up in TBI patients exposed to high supplemental oxygen administration within 7 days of ICU admission in CENTER-TBI. Mixed-effect logistic regression models adjusted for age, pupillary reactivity (reactive, one reactive, both unreactive), GCS motor (any flexion, none/extension, localizes/obey) and, once at a time, FiO_{2max}, FiO_{2mean}, and ΔFiO_{2mean} for CENTER-TBI with center as a random effect. Full models with all covariates estimates are reported in the Supplemental material

^a OR is for 10 mmHg increase in PaO₂ covariate

^b 1 patient did not have low PaO₂

^c OR regards 0.1 increments in FiO₂ covariate

Supplemental oxygen administration and outcome

After adjustment for confounders, FiO_{2max}, FiO_{2mean} and ΔFiO_{2mean} had no significant association with neurological outcomes. However, they showed a positive independent association with mortality at 6 months (Model 3, Table 2, and Supplemental Tables 7 and 8). The estimated mortality probability by administering supplemental oxygen is depicted in Fig. 2 (Panels D, E, and F).

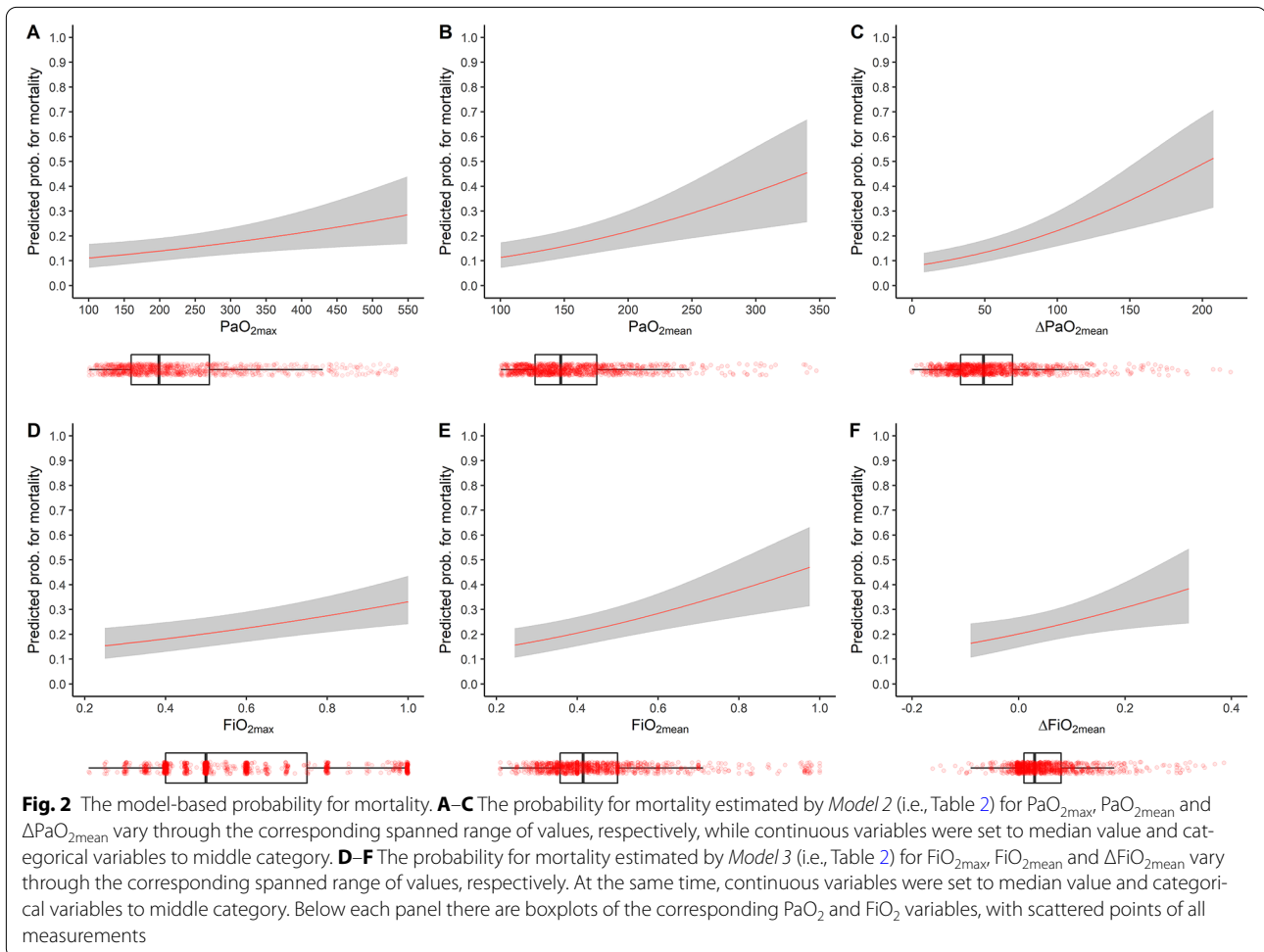
Results concerning PaO₂ and FiO₂ were confirmed when the Benjamini–Hochberg method was applied to control the false discovery rate (results not shown). The sensitivity analyses accounting for missing data also corroborated the findings from the models on complete cases for both PaO₂ and FiO₂ data (Supplemental Table 9). From the descriptive analysis reported in Supplemental Table 10, patients with and without missing data have similar characteristics. As 5 patients died within 48 h with PaO₂ levels beyond 450 mmHg and

PaCO₂ > 60 mmHg and may have undergone an apnea breath test, we performed a sensitivity analysis excluding these patients for all the explored outcomes in the original analysis. No differences were observed as reported in Supplemental Table 11.

OzENTER-TBI

Arterial oxygen levels and supplemental oxygen administration

During the first week of ICU admission, a total of 1651 measurements of PaO₂ were available (825 lowest and 826 highest daily values) for an overall median value of PaO₂ and FiO₂ of 133 (IQR 109–212) and 0.3 (IQR 0.25–0.4), respectively. During the first week, 43.4% had complete daily measurements of PaO₂ (median 6, IQR 3–7). The median of the highest PaO₂ level during the first 7 days since ICU admission was 133 (IQR 109–212) (Supplemental Fig. 2). The highest median FiO₂ levels during the first 7 days since ICU admission was 0.35 (IQR



0.25–0.5) (Supplemental Fig. 2). Mean $\text{PaO}_{2\text{max}}$, $\text{PaO}_{2\text{mean}}$ and $\Delta\text{PaO}_{2\text{mean}}$ were 356, 197 and 98 mmHg, respectively (Table 1). $\text{PaO}_{2\text{max}}$ showed a strong correlation with $\Delta\text{PaO}_{2\text{mean}}$ ($T_{\text{Kendall}}=0.63$, $p < 0.001$) and with $\text{PaO}_{2\text{mean}}$ ($T_{\text{Kendall}}=0.71$, $p < 0.001$). Mean $\text{FiO}_{2\text{max}}$, $\text{FiO}_{2\text{mean}}$ and $\Delta\text{FiO}_{2\text{mean}}$ were 0.82, 0.48 and 0.15 mmHg, respectively. Center variability in PaO_2 (panel A) and FiO_2 levels (panel B) across the 2 centers was represented in Fig. 1.

Arterial oxygen levels and outcomes in TBI patients

Data on mortality and neurological functional score GOSE at 6 months were available for 148 (93.1%) TBI patients. Ninety-five patients (64.2%) had an unfavorable GOSE at 6-month follow-up, and 40 died within that period (27%). After adjusting for multiple confounders, including IMPACT core baseline covariates, ISS and the 2 different centers (i.e., site code), we observed that none of the oxygen exposure variables was independently associated with GOSE (Model 1, Table 3 and Supplemental Table 12). After adjustment

for the same confounders, we observed that $\Delta\text{PaO}_{2\text{mean}}$ (OR 1.08, 95% CI 1–1.18) trended toward a higher mortality rate (Model 1, Table 3 and Supplemental Table 13). A detailed description of all confounders estimates for both outcomes was described in Supplemental Tables 12 and 13.

Supplemental oxygen administration and outcome

After adjustment for confounders, $\text{FiO}_{2\text{max}}$, $\text{FiO}_{2\text{mean}}$ and $\Delta\text{FiO}_{2\text{mean}}$ confirmed the data of CENTER-TBI with no significant association with neurological outcome. However, increases in $\text{FiO}_{2\text{mean}}$ trended toward a higher mortality rate (Model 2, Table 3). A detailed description of all confounders estimates for both outcomes was described in Supplemental Tables 14 and 15.

Discussion

In this study, we investigated whether exposure to high blood oxygen levels and high oxygen supplementation is

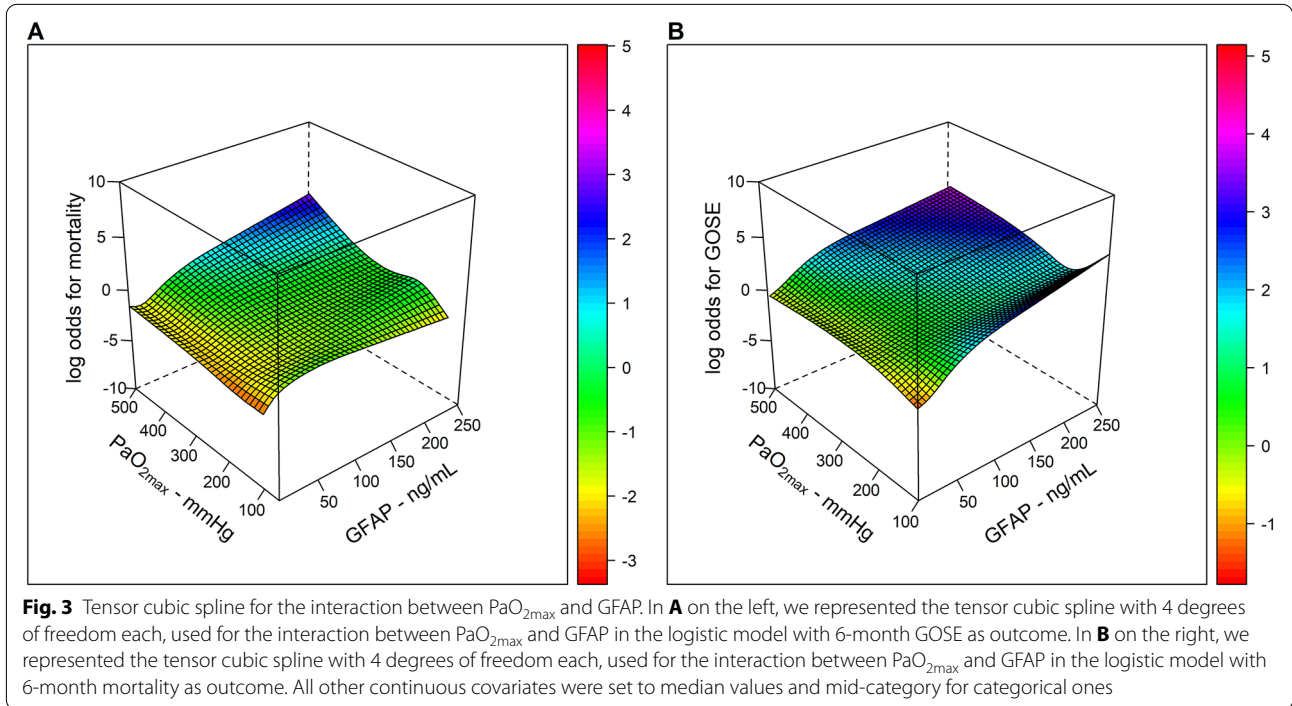


Table 3 Multivariable models on GOSE and mortality at 6-month follow-up in OzENTER-TBI (Model 1 and 2)

OzENTER-TBI	6-month GOSE N = 141 patients, 92 GOSE ≤ 4			6-month mortality N = 141 patients, 39 died		
Model 1	OR ^a	95% CI	p value	OR ^a	95% CI	p value
PaO _{2max} (for 10 mmHg increase)	1.01	0.98–1.04	0.433	1	0.97–1.04	0.898
PaO _{2mean} (for 10 mmHg increase)	1.01	0.96–1.07	0.656	1.05	0.99–1.11	0.118
ΔPaO _{2mean} (for 10 mmHg increase)	1.03	0.96–1.12	0.376	1.08	1–1.18	0.054
OzENTER-TBI	6-month GOSE N = 141 patients, 92 GOSE ≤ 4			6-month mortality N = 141 patients, 39 died		
Model 2	OR ^b	95% CI	p value	OR [*]	95% CI	p value
FiO _{2max} (for 0.1 increase)	1.06	0.89–1.26	0.492	1	0.83–1.23	0.963
FiO _{2mean} (for 0.1 increase)	1.02	0.77–1.34	0.911	1.32	0.98–1.8	0.069
ΔFiO _{2mean} (for 0.1 increase)	1.15	0.79–1.69	0.483	1	0.68–1.48	0.981

Model 1. Adjusted odds ratio with 95% confidence intervals effect of exposure to high blood oxygen levels within 7 days of ICU admission on GOSE and mortality at 6-month follow-up. Validation on OzENTER-TBI. Standard logistic regression models adjusted for age, pupillary reactivity (both reactive, one reactive, both unreactive), GCS Motor (any flexion, none/extension, localizes/obey), Injury Severity Score, and, once at a time, PaO_{2max}, PaO_{2mean} and ΔPaO_{2mean} for OzENTER-TBI with a dummy variable for center. *Model 2.* Adjusted odds ratio with 95% CI of GOSE and mortality at 6-month follow-up in TBI patients exposed to high supplemental oxygen administration within 7 days of ICU admission in OzENTER-TBI. Standard logistic regression models adjusted for age, pupillary reactivity (both reactive, one reactive, both unreactive), GCS Motor (any flexion, none/extension, localizes/obey) and, once at a time, FiO_{2max}, FiO_{2mean} and ΔFiO_{2mean} for OzENTER-TBI with a dummy variable for center. Full models with all covariates estimates are reported in the Supplemental material

^a OR is for 10 mmHg increase in PaO₂ covariate

^b OR regards 0.1 increments in FiO₂ covariate

independently associated with outcomes in TBI patients admitted to ICU and undergoing mechanical ventilation.

The main findings can be summarized as follows:

1. TBI patients were largely exposed, with wide variability between centers, to high levels of PaO₂ during the first week of ICU admission.
2. Exposure to high PaO₂ within seven days after ICU admission was an independent predictor of 6-month

mortality in the CENTER-TBI cohort, even regardless of the severity of brain injury as defined by higher serum concentration of GFAP.

3. A higher average daily variability in PaO₂ (Δ PaO_{2mean}) predicts an unfavorable GOSE at 6 months in CENTER-TBI. These findings were not validated in the OzENTER-TBI cohort, where only Δ PaO_{2mean} trended to a higher mortality rate.
4. Exposure to high levels of supplemental oxygen has an independent positive association with mortality in the CENTER-TBI cohort. In contrast, the association between higher FiO_{2mean} and worse mortality in the OzENTER-TBI cohort showed similar directional trends but did not achieve statistical significance.

The first insight of this study is that more than 50% of TBI patients are exposed to hyperoxemia, defined as PaO₂ levels above 120 mmHg [20, 21], during the first week after ICU admission. Despite hyperoxemia being quite often defined as the presence of a PaO₂ > 120 [20, 22, 23], there is no agreement in the literature about a univocal threshold to define it [7, 8, 24–27]. Understanding if there is a maximum dose of oxygen that may be harmful for the brain tissue and whether a prolonged time of exposure to high oxygen levels may impair brain function and have an impact on mortality is debated. The lack of a clear definition of hyperoxemia and a limited time of oxygen exposure may lead to underestimate an association with outcome in TBI patients [27–30], despite some reports of a higher mortality in TBI patients exposed to higher levels of oxygen [7–9, 24].

This clinical investigation highlights a relevant finding that might have a direct potential clinical implication.

We reported that increasing exposure to high blood oxygen levels within the first 7 days after ICU admission independently correlates with long-term mortality in patients with TBI. This association was observed by exploring either the highest PaO₂ levels (interpreted for each 10-mmHg increase) or the daily highest PaO₂ variability. This may suggest that clinicians should pay attention not just to the absolute values of PaO₂ but also to the daily swings of blood oxygenation. We logically hypothesized that PaO₂ levels are driven by inappropriately high inspiratory levels of oxygen administered to TBI patients. When we explored the role of supplemental oxygen use (i.e., FiO₂), similarly to the association reported between blood oxygenation and mortality, we showed that the highest the levels of FiO₂ or the most elevated average daily swings of FiO₂ within the first 7 days, the higher the mortality rate. These findings highlight a direct potential clinical implication for the management of oxygen administration in critically ill patients mechanically ventilated and admitted to the ICU with TBI. The amount of

oxygen delivered to TBI patients can be easily titrated by ICU physicians by setting FiO₂ levels on the ventilator. In the presence of an isolated TBI, therefore not involving the lung parenchyma that may lead to impaired oxygenation, high oxygen supplementation may be easily avoided on the ventilator by setting FiO₂ levels to target a physiological range of blood oxygenation.

Furthermore, avoiding major changes in daily FiO₂—if not needed to avoid hypoxemia—should prevent a major blood oxygenation variability and limit exposure to high oxygen levels and its detrimental effects. Our findings are in line with the recent guidelines of the European Society of Intensive Care Medicine (ESICM) on the management of mechanical ventilation in patients with an acute brain injury which, with a low level of evidence, recommend targeting normoxia (80–120 mmHg) regardless of the presence of intracranial pressure (ICP) elevation while it remains unknown whether a certain threshold of high PaO₂ should be considered safe in TBI patients [20]. The pathophysiological mechanisms behind the role of oxygen toxicity induced by hyperoxia (i.e., high FiO₂) [31, 32] and hyperoxemia (i.e., high PaO₂) [33, 34] in humans are widely recognized [5, 35]. On the one hand, hyperoxia has been shown to induce direct pulmonary toxicity by alveolar-capillary leak and fibrogenesis in healthy volunteers [36] and to have cytotoxic properties [37–39]. On the other hand, hyperoxemia increases peripheral vascular resistances [40–43], and determines the production of reactive oxygen species [44, 45] with the release of proinflammatory mediators [46]. In a cohort of severe TBI patients studied with advanced multimodality monitoring, hyperoxia had variable effects on lactate and lactate/pyruvate ratio. Microdialysis did not demonstrate a constant increase in the cerebral metabolic rate of oxygen in at-risk tissue [47]. Similar results have been shown in TBI patients exposed to high FiO₂. Hyperoxia marginally reduced lactate levels in brain tissue after TBI. However, the estimated redox status of the cells did not change and cerebral O₂ extraction seemed to be reduced. These data indicate that glucose oxidation was not improved by hyperoxia in cerebral and adipose tissue and might even be impaired [48].

In recent years, the role of oxygen on outcome has been explored in ICU patients to evaluate whether oxygen's inflammatory and cytotoxic effects on organ viability might translate into a worse survival. Two randomized controlled trials (RCTs) in critically ill (Oxygen-ICU) [49] and in septic patients (HYPER-2S) [50] showed that targeting higher levels of PaO₂ or hyperoxia could cause a higher mortality rate. A large meta-analysis including critically ill patients confirmed that a

strategy targeting more elevated levels of PaO₂ increased mortality [51].

In contrast, so far, 4 big RCTs (LOCO₂ trial [52], ICU-ROX trial [53], HOT-ICU trial [54] and O₂-ICU trial [55]) suggested no significant differences in terms of primary study outcome (i.e., mortality [52, 54]; ventilator-free days [53]; and non-respiratory Sequential Organ Failure Assessment (SOFA) score [55]) between patients managed with lower versus higher oxygen targets. However, these trials showed differences in their study design in terms of targeted physiologic variables of oxygenation (i.e., PaO₂, SpO₂ and SaO₂), targets of oxygenation, safety threshold for oxygen conservative therapy [52] and study outcomes. These trials were in broad populations of critically ill patients, and do not specifically address patients with TBI. Indeed, the one trial that specifically reported on patients with brain injury provided data suggesting that patients with neurological disease not due to hypoxic–ischemic encephalopathy may have had worse outcomes with conservative oxygen therapy [53]. In the meantime, the UK-ROX trial (ISRCTN13384956) and the Mega-ROX trial (ACTRN12620000391976)—two large RCTs aimed at exploring the role of oxygen targets on mortality in critically ill patients—are currently ongoing and will shed further light on the role of oxygen targets on outcome in ICU.

We also investigated whether these negative associations of hyperoxia with outcome were modulated by injury severity, as measured by GFAP levels [17, 56]. GFAP is a biomarker representing glial injury [56] and correlates well with the severity of brain injury evaluated by brain computed tomography [17]. Furthermore, GFAP is associated with outcomes in TBI patients [57]. However, we could not demonstrate an interaction between injury severity (as measured by GFAP levels) and the association between oxygen exposure variables and outcome. This corroborates the idea that oxygen exposure may somehow influence the outcome in TBI patients regardless of the severity of brain injury. Therefore, preventing exposure to high oxygen levels in TBI patients might be suggested even in milder TBI.

However, another potential explanation for the lack of interaction between oxygen levels and GFAP may be the temporal misalignment of GFAP and oxygen levels assessment. TBI is not an acute event but an evolving process. Hence, acute GFAP and sub-acute oxygen level measures may capture distinct complementary aspects providing independent prognostic information which can enable a more effective risk-stratification of patients with TBI. Moreover, it is conceivable that high blood oxygen levels could have a differential effect based on the injury pattern/type rather than the severity of structural

brain damage after TBI owing to distinct pathogenetic and pathobiological pathways. In support of such a possibility, robust experimental evidence has indicated specific therapeutic responses according to different injury models as also tracked by circulating GFAP [58, 59].

Strengths

Strengths of this work include the prospective nature of the two multicenter cohorts of patients, with the OzENTER-TBI validation cohort confirming a trend similar to the findings reported in the sizeable CENTER-TBI cohort. Data comes from a large real-world dataset of patients with TBI representing a global population of TBI patients. Evaluating the effect of exposure to oxygen on the outcome is not episodic but integrated over the first week after ICU admission increases the association's credibility. Furthermore, the exposure variables (i.e., PaO₂ and FiO₂) are not evaluated using a pre-set cut-off. Still, their association with the outcome is explored by including them as continuous data, strengthening the findings in the multivariable models. The use of GFAP, which allowed to investigate whether oxygen exposure could play a different contribution to the outcome because of a different degree of brain injury severity, make the results generalizable to most of the spectrum of TBI. Moreover, although we acknowledge that various models were performed, the strong associations we found on mortality were supported even when we accounted for multiple comparisons.

Limitations

Several limitations deserve mention. First, considering the observational nature of the data, it is speculative to draw a direct causal relationship between high arterial oxygen levels and supplemental oxygen administration and their relationship with outcome. Therefore, our results should be taken with caution. Further randomized controlled studies are necessary to assess the effect of high arterial oxygen levels and supplemental oxygen administration on the TBI patients' outcomes. Second, 6-month GOSE and mortality are influenced by several other factors, such as systemic and ICU complications and post-ICU events. To overcome this limitation, we used an analytic model considering the effect of other available confounding factors, particularly patient clinical condition and neuroimaging features.

Besides, in these two cohorts, only a minority of patients had a brain tissue oxygen monitor. As documented by a phase-2 RCT, monitoring brain tissue (PbtO₂) oxygenation could reduce brain tissue hypoxia with a trend toward more favorable outcomes compared to treatment driven by intracranial pressure monitoring only [60]. A recent consensus suggested the possibility, in

the presence of low PbtO₂ values, of elevating the PaO₂ up to 150 mmHg or higher in more severe cases, fine-tuned to the patient's PbtO₂ values [61]. Some phase III randomized trials are ongoing to demonstrate the benefit of exposing hypoxic brain patients to higher oxygen levels. Therefore, our findings are not focused on a population with brain tissue hypoxia but to the overall TBI population, with/without brain hypoxia. However, we did not observe a difference in the distribution of PaO₂ levels between TBI with or without PbtO₂ monitoring. We cannot exclude the possibility that the worse outcomes associated with higher PaO₂ were due to use of higher FiO₂ in patients with more severe injury or physiological compromise. Further, these findings may not apply to patients in whom FiO₂ and PaO₂ are titrated to PbtO₂ levels.

Moreover, the two cohorts were prospectively collected with the primary aim of assessing the epidemiology and clinical practice in the management of TBI patients. As respiratory targets are not included in the primary outcome, more frequent daily data on gas exchange and more specific data on the ventilator management of these patients are missing and would have strengthened our analysis. Further, we do not have detailed data about the presence of hyperoxemia in patients undergoing an apnea breath test. However, only five patients who died within 48 h had PaO₂ levels beyond 450 mmHg with a PaCO₂ > 60 mmHg in the CENTER-TBI dataset, which may suggest an apnea breath test. Sensitivity analyses excluding these patients confirmed the independent association with outcome of both PaO₂ and FiO₂ variables. Finally, our dataset is limited to the first week after TBI. However, our analysis includes data that provides a longitudinal view of PaO₂ management over time.

Conclusions

In two large prospective multicenter cohorts of critically ill patients with TBI arterial oxygen levels and supplemental oxygen, administration varied widely across centers during the first 7 days after ICU admission. Exposure to high arterial blood oxygen and high supplemental oxygen were independently associated with 6-month mortality in the CENTER-TBI cohort. This was not driven by the severity of brain injury quantified by serum levels of GFAP within 24 h. The findings were not externally validated in the OzENTER-TBI cohort likely due to the limited sample size, although the effects were in the same direction of the ones from CENTER-TBI. Titration of supplemental oxygen in the presence of TBI is a practice immediately applicable at bedside. Randomized controlled trials and high-level evidence guidelines are

warranted to help clinicians optimize oxygen exposure management in this cohort of patients.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s00134-022-06884-x>.

Author details

¹ School of Medicine and Surgery, University of Milano - Bicocca, Monza, Italy. ² Department of Emergency and Intensive Care, San Gerardo University Hospital, Extracorporeal Membrane Oxygenation (ECMO) Center, Azienda Socio-Sanitaria Territoriale (ASST) di Monza, Monza, Italy. ³ Department of Medicine and Surgery, Bicocca Bioinformatics Biostatistics and Bioimaging B4 Center, University of Milano - Bicocca, Monza, Italy. ⁴ Division of Anaesthesia, Addenbrooke's Hospital, University of Cambridge, Hills Road, Box 93, Cambridge CB2 0QQ, UK. ⁵ Department of Biomedical and Dental Sciences and Morphofunctional Imaging, University of Messina, Messina, Italy. ⁶ Intensive Care Department, Alfred Hospital, Melbourne, Australia. ⁷ School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia. ⁸ Antwerp University Hospital and University of Antwerp, Edegem, Belgium. ⁹ Department of Public Health, Erasmus MC University Medical Center, Rotterdam, The Netherlands. ¹⁰ NeuroIntensive Care Unit, Neuroscience Department, Hospital San Gerardo, ASST Monza, Monza, Italy.

Acknowledgements

The CENTER-TBI ICU WP6 participants and ICU ONLY investigators: Cecilia Ackerslund: Department of Physiology and Pharmacology, Section of Perioperative Medicine and Intensive Care, Karolinska Institutet, Stockholm, Sweden. Krisztina Amrein: János Szentágotthai Research Centre, University of Pécs, Pécs, Hungary. Nada Andelic: Division of Surgery and Clinical Neuroscience, Department of Physical Medicine and Rehabilitation, Oslo University Hospital and University of Oslo, Oslo, Norway. Lasse Andreassen: Department of Neurosurgery, University Hospital Northern Norway, Tromsø, Norway. Audny Anke: Department of Physical Medicine and Rehabilitation, University Hospital Northern Norway, Tromsø, Norway. Gérard Audibert: Department of Anesthesiology & Intensive Care, University Hospital Nancy, Nancy, France. Philippe Azouvi: Raymond Poincaré hospital, Assistance Publique – Hôpitaux de Paris, Paris, France. Maria Luisa Azzolini: Department of Anesthesiology & Intensive Care, S Raffaele University Hospital, Milan, Italy. Ronald Bartels: Department of Neurosurgery, Radboud University Medical Center, Nijmegen, The Netherlands. Ronny Beer: Department of Neurology, Neurological Intensive Care Unit, Medical University of Innsbruck, Innsbruck, Austria. Bo-Michael Bellander: Department of Neurosurgery & Anesthesia & intensive care medicine, Karolinska University Hospital, Stockholm, Sweden. Habib Benali: Anesthésie-Réanimation, Assistance Publique – Hôpitaux de Paris, Paris, France. Maurizio Berardino: Department of Anesthesia & ICU, AOU Città della Salute e della Scienza di Torino—Orthopedic and Trauma Center, Torino, Italy. Luigi Beretta: Department of Anesthesiology & Intensive Care, S Raffaele University Hospital, Milan, Italy. Erta Beqiri: NeuroIntensive Care, Niguarda Hospital, Milan, Italy. Morten Blaabjerg: Department of Neurology, Odense University Hospital, Odense, Denmark. Stine Borgen Lund: Department of Public Health and Nursing, Faculty of Medicine and health Sciences, Norwegian University of Science and Technology, NTNU, Trondheim, Norway. Camilla Brorsson: Department of Surgery and Perioperative Science, Umeå University, Umeå, Sweden. Andras Buki: Department of Neurosurgery, Medical School, University of Pécs, Hungary and Neurotrauma Research Group, János Szentágotthai Research Centre, University of Pécs, Hungary. Manuel Cabelreira: Brain Physics Lab, Division of Neurosurgery, Dept of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Alessio Caccioppola: Neuro ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy. Emiliana Calappi: Neuro ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy. Maria Rosa Calvi: Department of Anesthesiology & Intensive Care, S Raffaele University Hospital, Milan, Italy. Peter Cameron: ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. Guillermo Carbayo Lozano: Department of Neurosurgery, Hospital of Cruces, Bilbao, Spain. Marco Carbonara: Neuro ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy. Ana M. Castaño-León: Department of Neurosurgery, Hospital Universitario 12 de Octubre, Madrid,

Spain. Simona Cavallo: Department of Anesthesia & ICU, AOU Città della Salute e della Scienza di Torino—Orthopedic and Trauma Center, Torino, Italy. Giorgio Chevallard: NeuroIntensive Care, Niguarda Hospital, Milan, Italy. Arturo Chiaregato: NeuroIntensive Care, Niguarda Hospital, Milan, Italy. Giuseppe Citerio: School of Medicine and Surgery, Università Milano—Bicocca, Milano, Italy and NeuroIntensive Care, ASST di Monza, Monza, Italy. Hans Clusmann: Department of Neurosurgery, Medical Faculty RWTH Aachen University, Aachen, Germany. Mark Steven Coburn: Department of Anaesthesiology, University Hospital of Aachen, Aachen, Germany. Jonathan Coles: Department of Anesthesia & Neurointensive Care, Cambridge University Hospital NHS Foundation Trust, Cambridge, UK. Jamie D. Cooper: School of Public Health & PM, Monash University and The Alfred Hospital, Melbourne, Victoria, Australia. Marta Correia: Radiology/MRI department, MRC Cognition and Brain Sciences Unit, Cambridge, UK. Endre Czeiter: Department of Neurosurgery, Medical School, University of Pécs, Hungary and Neurotrauma Research Group, János Szentágotthai Research Centre, University of Pécs, Hungary. Marek Czosnyka: Brain Physics Lab, Division of Neurosurgery, Dept of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Claire Dahyot-Fizelier: Intensive Care Unit, CHU Poitiers, Poitiers, France. Paul Dark: University of Manchester NIHR Biomedical Research Centre, Critical Care Directorate, Salford Royal Hospital NHS Foundation Trust, Salford, UK. Véronique De Keyser: Department of Neurosurgery, Antwerp University Hospital and University of Antwerp, Edegem, Belgium. Vincent Degos: Anesthésie-Réanimation, Assistance Publique – Hôpitaux de Paris, Paris, France. Francesco Della Corte: Department of Anesthesia & Intensive Care, Maggiore Della Carità Hospital, Novara, Italy. Hugo den Boogert: Department of Neurosurgery, Radboud University Medical Center, Nijmegen, The Netherlands. Bart Dreier: Department of Neurosurgery, University Hospitals Leuven, Leuven, Belgium. Đula Đilvesi: Department of Neurosurgery, Clinical centre of Vojvodina, Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia. Abhishek Dixit: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Jens Dreier: Center for Stroke Research Berlin, Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany. Guy-Loup Dulière: Intensive Care Unit, CHR Citadelle, Liège, Belgium. Ari Ercole: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Erzsébet Ezer: Department of Anaesthesiology and Intensive Therapy, University of Pécs, Pécs, Hungary. Martin Fabricius: Departments of Neurology, Clinical Neurophysiology and Neuroanaesthesiology, Region Hovedstaden Rigshospitalet, Copenhagen, Denmark. Kelly Foks: Department of Neurology, Erasmus MC, Rotterdam, the Netherlands. Shirin Frisvold: Department of Anesthesiology and Intensive care, University Hospital Northern Norway, Tromsø, Norway. Alex Furmanov: Department of Neurosurgery, Hadassah-hebrew University Medical center, Jerusalem, Israel. Damien Galanaud: Anesthésie-Réanimation, Assistance Publique – Hôpitaux de Paris, Paris, France. Dashiell Gantner: ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. Alexandre Ghuysen: Emergency Department, CHU, Liège, Belgium. Lelde Giga: Neurosurgery clinic, Pauls Stradins Clinical University Hospital, Riga, Latvia. Jagoš Golubovi: Department of Neurosurgery, Clinical centre of Vojvodina, Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia. Pedro A. Gomez: Department of Neurosurgery, Hospital Universitario 12 de Octubre, Madrid, Spain. Benjamin Gravestelijn: Department of Public Health, Erasmus Medical Center-University Medical Center, Rotterdam, The Netherlands. Francesca Grossi: Department of Anesthesia & Intensive Care, Maggiore Della Carità Hospital, Novara, Italy. Deepak Gupta: Department of Neurosurgery, Neurosciences Centre & JPN Apex trauma centre, All India Institute of Medical Sciences, New Delhi-110029, India. Iain Haitsma: Department of Neurosurgery, Erasmus MC, Rotterdam, The Netherlands. Raimund Helbok: Department of Neurology, Neurological Intensive Care Unit, Medical University of Innsbruck, Innsbruck, Austria. Eirik Helseth: Department of Neurosurgery, Oslo University Hospital, Oslo, Norway. Jilke Huijben: Department of Public Health, Erasmus Medical Center-University Medical Center, Rotterdam, The Netherlands. Peter J. Hutchinson: Division of Neurosurgery, Department of Clinical Neurosciences, Addenbrooke's Hospital & University of Cambridge, Cambridge, UK. Stefan Jankowski: Neurointensive Care, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. Faye Johnson: Salford Royal Hospital NHS Foundation Trust Acute Research Delivery Team, Salford, UK. Mladen Karan: Department of Neurosurgery, Clinical centre of Vojvodina, Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia. Angelos G. Kolias: Division of Neurosurgery,

Department of Clinical Neurosciences, Addenbrooke's Hospital & University of Cambridge, Cambridge, UK. Daniel Kondziella: Departments of Neurology, Clinical Neurophysiology and Neuroanaesthesiology, Region Hovedstaden Rigshospitalet, Copenhagen, Denmark. Evgenios Kornaropoulos: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Lars-Owe Koskinen: Department of Clinical Neuroscience, Neurosurgery, Umeå University, Umeå, Sweden. Noémi Kovács: Hungarian Brain Research Program—Grant No. KTIA_13_NAP-A-II/8, University of Pécs, Pécs, Hungary. Ana Kowark: Department of Anaesthesiology, University Hospital of Aachen, Aachen, Germany. Alfonso Lagares: Department of Neurosurgery, Hospital Universitario 12 de Octubre, Madrid, Spain. Steven Laureys: Cyclotron Research Center, University of Liège, Liège, Belgium. Aurelie Lejeune: Department of Anesthesiology-Intensive Care, Lille University Hospital, Lille, France. Fiona Lecky: Centre for Urgent and Emergency Care Research (CURE), Health Services Research Section, School of Health and Related Research (SchARR), University of Sheffield, Sheffield, UK and Emergency Department, Salford Royal Hospital, Salford UK. Didier Ledoux: Cyclotron Research Center, University of Liège, Liège, Belgium. Roger Lightfoot: Department of Anesthesiology & Intensive Care, University Hospitals Southampton NHS Trust, Southampton, UK. Hester Lingsma: Department of Public Health, Erasmus Medical Center-University Medical Center, Rotterdam, The Netherlands. Andrew I.R. Maas: Department of Neurosurgery, Antwerp University Hospital and University of Antwerp, Edegem, Belgium. Alex Manara: Intensive Care Unit, Southmead Hospital, Bristol, Bristol, UK. Hugues Maréchal: Intensive Care Unit, CHR Citadelle, Liège, Belgium. Costanza Martino: Department of Anesthesia & Intensive Care, M. Bufalini Hospital, Cesena, Italy. Julia Mattern: Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany. Catherine McMahon: Department of Neurosurgery, The Walton centre NHS Foundation Trust, Liverpool, UK. David Menon: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Tomas Menovsky: Department of Neurosurgery, Antwerp University Hospital and University of Antwerp, Edegem, Belgium. Benoit Misset: Cyclotron Research Center, University of Liège, Liège, Belgium. Visakh Muralidharan: Karolinska Institutet, INCF International Neuroinformatics Coordinating Facility, Stockholm, Sweden. Lynnette Murray: ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. Ancuta Negru: Department of Neurosurgery, Emergency County Hospital Timisoara, Timisoara, Romania. David Nelson: Department of Physiology and Pharmacology, Section of Perioperative Medicine and Intensive Care, Karolinska Institutet, Stockholm, Sweden. Virginia Newcombe: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. József Nyírádi: János Szentágotthai Research Centre, University of Pécs, Pécs, Hungary. Fabrizio Ortolano: Neuro ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy. Jean-François Payen: Department of Anesthesiology & Intensive Care, University Hospital of Grenoble, Grenoble, France. Vincent Perlbarg: Anesthésie-Réanimation, Assistance Publique – Hôpitaux de Paris, Paris, France. Paolo Persona: Department of Anesthesia & Intensive Care, Azienda Ospedaliera Università di Padova, Padova, Italy. Wilco Peul: Dept. of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands and Dept. of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands. Anna Piippo-Karjalainen: Department of Neurosurgery, Helsinki University Central Hospital. Horia Ples: Department of Neurosurgery, Emergency County Hospital Timisoara, Timisoara, Romania. Inigo Pomposo: Department of Neurosurgery, Hospital of Cruces, Bilbao, Spain. Jussi P. Posti: Division of Clinical Neurosciences, Department of Neurosurgery and Turku Brain Injury Centre, Turku University Hospital and University of Turku, Turku, Finland. Louis Puybasset: Department of Anesthesiology and Critical Care, Pitié-Salpêtrière Teaching Hospital, Assistance Publique, Hôpitaux de Paris and University Pierre et Marie Curie, Paris, France. Andreea Rădoi: Neurotraumatology and Neurosurgery Research Unit (UNINN), Vall d'Hebron Research Institute, Barcelona, Spain. Arminas Ragauskas: Department of Neurosurgery, Kaunas University of technology and Vilnius University, Vilnius, Lithuania. Rahul Raj: Department of Neurosurgery, Helsinki University Central Hospital. Jonathan Rhodes: Department of Anaesthesia, Critical Care & Pain Medicine NHS Lothian & University of Edinburgh, Edinburgh, UK. Sophie Richter: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Saulius Rocka: Department of Neurosurgery, Kaunas University of technology and Vilnius University, Vilnius, Lithuania. Cecilie Roe: Department of Physical Medicine and Rehabilitation, Oslo University Hospital/University of Oslo, Oslo, Norway. Olav Roise: Division of Orthopedics, Oslo University Hospital, Oslo, Norway and

Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. Jeffrey Rosenfeld: National Trauma Research Institute, The Alfred Hospital, Monash University, Melbourne, Victoria, Australia. Christina Rosenlund: Department of Neurosurgery, Odense University Hospital, Odense, Denmark. Guy Rosenthal: Department of Neurosurgery, Hadassah-hebrew University Medical center, Jerusalem, Israel. Rolf Rossaint: Department of Anaesthesiology, University Hospital of Aachen, Aachen, Germany. Sandra Rossi: Department of Anesthesia & Intensive Care, Azienda Ospedaliera Università di Padova, Padova, Italy. Juan Sahuquillo: Neurotraumatology and Neurosurgery Research Unit (UNINN), Vall d'Hebron Research Institute, Barcelona, Spain. Oliver Sakowitz: Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany and Klinik für Neurochirurgie, Klinikum Ludwigsburg, Ludwigsburg, Germany. Renan Sanchez-Porras: Klinik für Neurochirurgie, Klinikum Ludwigsburg, Ludwigsburg, Germany. Oddrun Sandrød: Klinik für Neurochirurgie, Klinikum Ludwigsburg, Ludwigsburg, Germany. Kari Schirmer-Mikalsen: Department of Anaesthesiology and Intensive Care Medicine, St.Olavs Hospital, Trondheim University Hospital, Trondheim, Norway and Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, NTNU, Trondheim, Norway. Rico Frederik Schou: Department of Neuroanesthesia and Neurointensive Care, Odense University Hospital, Odense, Denmark. Charlie Sewalt: Department of Public Health, Erasmus Medical Center-University Medical Center, Rotterdam, The Netherlands. Peter Smielewski: Brain Physics Lab, Division of Neurosurgery, Dept of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Abayomi Sorinola: Department of Neurosurgery, University of Pécs, Pécs, Hungary. Emmanuel Stamatakis: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Ewout W. Steyerberg: Department of Anaesthesiology & Intensive Care, University Hospitals Southampton NHS Trust, Southampton, UK and Dept. of Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands. Nino Stocchetti: Department of Pathophysiology and Transplantation, Milan University, and Neuroscience ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milano, Italy. Nina Sundström: Department of Radiation Sciences, Biomedical Engineering, Umeå University, Umeå, Sweden. Riikka Takala: Perioperative Services, Intensive Care Medicine and Pain Management, Turku University Hospital and University of Turku, Turku, Finland. Viktória Tamás: Department of Neurosurgery, University of Pécs, Pécs, Hungary. Tomas Tamosiutis: Department of Neurosurgery, Kaunas University of Health Sciences, Kaunas, Lithuania. Olli Tenovu: Division of Clinical Neurosciences, Department of Neurosurgery and Turku Brain Injury Centre, Turku University Hospital and University of Turku, Turku, Finland. Matt Thomas: Intensive Care Unit, Southmead Hospital, Bristol, Bristol, UK. Dick Tibboel: Intensive Care and Department of Pediatric Surgery, Erasmus Medical Center, Sophia Children's Hospital, Rotterdam, The Netherlands. Christos Toliás: Department of Neurosurgery, Kings college London, London, UK. Tony Trapani: ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. Cristina Maria Tudora: Department of Neurosurgery, Emergency County Hospital Timisoara, Timisoara, Romania. Andreas Unterberg: Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany. Peter Vajkoczy: Neurologie, Neurochirurgie und Psychiatrie, Charité – Universitätsmedizin Berlin, Berlin, Germany. Egils Valeinis: Neurosurgery clinic, Pauls Stradins Clinical University Hospital, Riga, Latvia. Shirley Vallance: ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. Zoltán Vámos: Department of Anaesthesiology and Intensive Therapy, University of Pécs, Pécs, Hungary. Gregory Van der Steen: Department of Neurosurgery, Antwerp University Hospital and University of Antwerp, Edegem, Belgium. Jeroen T.J.M. van Dijk: Dept. of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands and Dept. of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands. Thomas A. van Essen: Dept. of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands and Dept. of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands. Roel van Wijk: Dept. of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands and Dept. of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands. Alessia Vargiolu: NeuroIntensive Care, ASST di Monza, Monza, Italy. Emmanuel Vega: Department of Anaesthesiology-Intensive Care, Lille

University Hospital, Lille, France. Anne Vik: Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, NTNU, Trondheim, Norway and Department of Neurosurgery, St.Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. Rimantas Vilcinis: Department of Neurosurgery, Kaunas University of Health Sciences, Kaunas, Lithuania. Victor Volovici: Department of Neurosurgery, Erasmus MC, Rotterdam, The Netherlands. Peter Vuleković: Department of Neurosurgery, Clinical centre of Vojvodina, Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia. Eveline Wiegers: Department of Public Health, Erasmus Medical Center-University Medical Center, Rotterdam, The Netherlands. Guy Williams: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Stefan Winzeck: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. Stefan Wolf: Department of Neurosurgery, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany. Alexander Younsi: Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany. Frederick A. Zeiler: Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK and Section of Neurosurgery, Department of Surgery, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada. Agate Ziverte: Neurosurgery clinic, Pauls Stradins Clinical University Hospital, Riga, Latvia. Tommaso Zoerle: Neuro ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy. OzENTER TBI participants and Investigators: Jamie Cooper: School of Public Health & PM, Monash University and The Alfred Hospital, Melbourne, Victoria, Australia. Dashiell Gantner: ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. Russel Gruen: NTU Institute for Health Technologies and Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore. Lynette Murray: ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. Jeffrey V Rosenfeld: National Trauma Research Institute, The Alfred Hospital, Monash University, Melbourne, Victoria, Australia. Dinesh Varma: Department of Radiology, Alfred Health, Melbourne, Victoria, Australia and Department of Surgery, Monash Medical Centre, Melbourne, Victoria, Australia. Tony Trapani: ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. Shirley Vallance: ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. Christopher Maclsaac: Intensive Care Unit, Royal Melbourne Hospital, Melbourne, VIC, Australia. Andrea Jordan: Department of Intensive Care, Royal Melbourne Hospital, Melbourne, Australia.

Author contributions

GC ideated and supervised the project, participated in the data analysis, drafted the manuscript and the supplementary tables, discussed the findings with all the authors and collected the COIs. ER ideated the project, participated in the data analysis and drafted the manuscript and supplementary tables. MP, PR and SG analyzed the data and drafted the manuscript and the supplementary tables. DKM, SM, DJC, AM and EJAW were actively involved in the manuscript drafting and revision. All co-authors gave substantial feedback on the manuscript and approved the final version.

Funding

Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement. European Commission 7th Framework program and the Australian Health and Medical Research Council.

Declarations

Conflicts of interest

GC reports grants and personal fees as Speakers' Bureau Member and Advisory Board Member from Integra and Neuroptics. DKM reports grants from the European Union and UK National Institute for Health Research, during the study; grants, personal fees, and non-financial support from GlaxoSmithKline; personal fees from Neurotrauma Sciences, Lantmaanen AB, Pressura, and Pfizer, outside of the submitted work. DJC reports grants and Fellowship from National Health and Medical Research Council (Australia), Medical Research

Future Fund (Australia) during the study, and personal fees from Pressura, outside of the submitted work.

Open Access

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 21 June 2022 Accepted: 17 August 2022

Published online: 20 October 2022

References

- McHugh GS, Engel DC, Butcher I, Steyerberg EW, Lu J, Mushkudiani N, Hernández AV, Marmarou A, Maas AI, Murray GD (2007) The IMPACT study results from the prognostic value of secondary insults in traumatic brain injury. *J Neurotrauma* 24(2):287–293. <https://doi.org/10.1089/neu.2006.0031>
- MacIntyre NR (2014) Tissue hypoxia: implications for the respiratory clinician. *Respir Care* 59(10):1590–1596. <https://doi.org/10.4187/respcare.03357>
- Itagaki T, Nakano Y, Okuda N, Izawa M, Onodera M, Imanaka H, Nishimura M (2015) Hyperoxemia in mechanically ventilated, critically ill subjects: incidence and related factors. *Respir Care* 60(3):335–340. <https://doi.org/10.4187/respcare.03451>
- Behnke AR, Johnson FS, Poppen JR, Motley EP (1934) The effect of oxygen on man at pressures from 1 to 4 atmospheres. *Am J Physiol* 110:565–572
- Singer M, Young PJ, Laffey JG, Asfar P, Taccone FS, Skrifvars MB, Meyhoff CS, Radermacher P (2021) Dangers of hyperoxia. *Crit Care* 25(1):440. <https://doi.org/10.1186/s13054-021-03815-y>
- Quintard H, Patet C, Suys T, Marques-Vidal P, Oddo M (2015) Normobaric hyperoxia is associated with increased cerebral excitotoxicity after severe traumatic brain injury. *Neurocrit Care* 22(2):243–250. <https://doi.org/10.1007/s12028-014-0062-0>
- Brenner M, Stein D, Peter HU, Kufera J, Woodford M, Scalea T (2012) Association between early hyperoxia and worse outcomes after traumatic brain injury. *Arch Surg* 147(11):1042–1046. <https://doi.org/10.1001/archsurg.2012.1560>
- Rincon F, Kang J, Vibbert M, Urtecho J, Athar MK, Jallo J (2014) Significance of arterial hyperoxia and relationship with case fatality in traumatic brain injury: a multicentre cohort study. *J Neurol Neurosurg Psychiatry* 85(7):799–805. <https://doi.org/10.1136/jnnp-2013-305505>
- Davis DP, Meade W, Sise MJ, Kennedy F, Simon F, Tominaga G, Steele J, Coimbra R (2009) Both hypoxemia and extreme hyperoxemia may be detrimental in patients with severe traumatic brain injury. *J Neurotrauma* 26(12):2217–2223. <https://doi.org/10.1089/neu.2009.0940>
- de Jonge E, Peelen L, Keijzers PJ, Joore H, de Lange D, van der Voort PHJ, Bosman RJ, de Waal RA, Wesselink R, de Keizer NF (2008) Association between administered oxygen, arterial partial oxygen pressure and mortality in mechanically ventilated intensive care unit patients. *Crit Care* 12(6):R156. <https://doi.org/10.1186/cc7150>
- Hirunpattarasilp C, Shiina H, Na-Ek N, Attwell D (2022) The effect of hyperoxemia on neurological outcomes of adult patients: a systematic review and meta-analysis. *Neurocrit Care*. <https://doi.org/10.1007/s12028-021-01423-w>
- Maas AIR, Menon DK, Steyerberg EW, Citerio G, Lecky F, Manley GT, Hill S, Legrand V, Sorgner A, CENTER-TBI Participants and Investigators (2015) Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI): a prospective longitudinal observational study. *Neurosurgery* 76(1):67–80. <https://doi.org/10.1227/NEU.00000000000000575>
- Steyerberg EW, Wieggers E, Sewalt C, Buki A, Citerio G, De Keyser V, Ercole A, Kunzmann K, Lanyon L, Lecky F, Lingsma H, Manley G, Nelson D, Peul W, Stocchetti N, von Steinbüchel N, Vande Vyvere T, Verheyden J, Wilson L, Maas AIR, Menon DK, CENTER-TBI Participants and Investigators (2019) Case-mix, care pathways, and outcomes in patients with traumatic brain injury in CENTER-TBI: a European prospective, multicentre, longitudinal, cohort study. *Lancet Neurol* 18(10):923–934. [https://doi.org/10.1016/S1474-4422\(19\)30232-7](https://doi.org/10.1016/S1474-4422(19)30232-7)
- Young PJ, Mackle D, Hodgson C, Bellomo R, Bailey M, Beasley R, Deane AM, Eastwood G, Finfer S, Freebairn R, King V, Linke N, Litton E, McArthur C, McGuinness S, Panwar R, ICU-ROX Investigators and the Australian and New Zealand Intensive Care Society Clinical Trials Group (2022) Conservative or liberal oxygen therapy for mechanically ventilated adults with acute brain pathologies: A post-hoc subgroup analysis. *J Crit Care* 71:154079. <https://doi.org/10.1016/j.jccr.2022.154079>
- https://www.monash.edu/__data/assets/pdf_file/0003/1049772/OzENT_ER_V2_ANZIC-RC_Web.pdf
- Huijben JA, Dixit A, Stocchetti N, Maas AIR, Lingsma HF, van der Jagt M, Nelson D, Citerio G, Wilson L, Menon DK, Ercole A, CENTER-TBI Investigators and Participants (2021) Use and impact of high intensity treatments in patients with traumatic brain injury across Europe: a CENTER-TBI analysis. *Crit Care* 25(1):78. <https://doi.org/10.1186/s13054-020-03370-y>
- Czeiter E, Amrein K, Gravesteyn BY, Lecky F, Menon DK, Mondello S, Newcombe VJ, Richter S, Steyerberg EW, Vyvere TV, Verheyden J, Xu H, Yang Z, Maas AIR, Wang KKW, Büki A, CENTER-TBI Participants and Investigators (2020) Blood biomarkers on admission in acute traumatic brain injury: relations to severity, CT findings and care path in the CENTER-TBI study. *EBioMedicine* 56:102785. <https://doi.org/10.1016/j.ebiom.2020.102785>
- Vos PE, Jacobs B, Andriessen TMJC, Lamers KJB, Borm GF, Beems T, Edwards M, Rosmalen CF, Vissers JLM (2010) GFAP and S100B are biomarkers of traumatic brain injury: an observational cohort study. *Neurology* 75:1786–1793. <https://doi.org/10.1212/WNL.0b013e3181fd6d2d>
- Mondello S, Papa L, Buki A, Bullock MR, Czeiter E, Tortella FC, Wang KK, Hayes RL (2011) Neuronal and glial markers are differently associated with computed tomography findings and outcome in patients with severe traumatic brain injury: a case control study. *Crit Care* 15(3):R156. <https://doi.org/10.1186/cc10286>
- Robba C, Poole D, McNett M, Asehnoun K, Bösel J, Bruder N, Chierigato A, Cinotti R, Duranteau J, Einav S, Ercole A, Ferguson N, Guerin C, Siempos II, Kurtz P, Juffermans NP, Mancebo J, Mascia L, McCredie V, Nin N, Oddo M, Pelosi P, Rabinstein AA, Neto AS, Seder DB, Skrifvars MB, Suarez JJ, Taccone FS, van der Jagt M, Citerio G, Stevens RD (2020) Mechanical ventilation in patients with acute brain injury: recommendations of the European Society of Intensive Care Medicine consensus. *Intensive Care Med* 46(12):2397–2410. <https://doi.org/10.1007/s00134-020-06283-0>
- Aggarwal NR, Brower RG (2014) Targeting normoxemia in acute respiratory distress syndrome may cause worse short-term outcomes because of oxygen toxicity. *Ann Am Thorac Soc* 11(9):1449–1453. <https://doi.org/10.1513/AnnalsATS.201407-297PS>
- De Graaff AE, Dongelmans DA, Binnekade JM, de Jonge E (2011) Clinicians' response to hyperoxia in ventilated patients in a Dutch ICU depends on the level of FiO₂. *Intensive Care Med* 37(1):46–51
- Helmerhorst HJF, Arts DL, Schultz MJ, van der Voort PHJ, Abu-Hanna A, de Jonge E, van Westerloo DJ (2017) Metrics of arterial hyperoxia and associated outcomes in critical care. *Crit Care Med* 45(2):187–195. <https://doi.org/10.1097/CCM.0000000000002084>
- Hafner S, Beloncle F, Koch A, Radermacher P, Asfar P (2015) Hyperoxia in intensive care, emergency, and peri-operative medicine: Dr. Jekyll or Mr. Hyde? A 2015 update. *Ann Intensive Care* 5:42. <https://doi.org/10.1186/s13613-015-0084-6>
- Madotto F, Rezoagli E, Pham T, Schmidt M, McNicholas B, Protti A, Panwar R, Bellani G, Fan E, van Haren F, Brochard L, Laffey JG, LUNG SAFE Investigators and the ESICM Trials Group (2020) Hyperoxemia and excess oxygen use in early acute respiratory distress syndrome: insights

- from the LUNG SAFE study. *Crit Care* 24(1):125. <https://doi.org/10.1186/s13054-020-2826-6>
26. Aggarwal NR, Brower RG, Hager DN, Thompson BT, Netzer G, Shanholtz C, Lagakos A, Checkley W, National Institutes of Health Acute Respiratory Distress Syndrome Network Investigators (2018) Oxygen exposure resulting in arterial oxygen tensions above the protocol goal was associated with worse clinical outcomes in acute respiratory distress syndrome. *Crit Care Med* 46(4):517–524. <https://doi.org/10.1097/CCM.00000000000002886>
 27. Alali AS, Temkin N, Vavilala MS, Lele AV, Barber J, Dikmen S, Chesnut RM (2019) Matching early arterial oxygenation to long-term outcome in severe traumatic brain injury: target values. *J Neurosurg* 132(2):537–544
 28. Weeden M, Bailey M, Gabbe B, Pilcher D, Bellomo R, Udy A (2021) Functional outcomes in patients admitted to the intensive care unit with traumatic brain injury and exposed to hyperoxia: a retrospective multicentre cohort study. *Neurocrit Care* 34(2):441–448. <https://doi.org/10.1007/s12028-020-01033-y>
 29. Rockswold SB, Rockswold GL, Zaun DA, Liu J (2013) A prospective, randomized Phase II clinical trial to evaluate the effect of combined hyperbaric and normobaric hyperoxia on cerebral metabolism, intracranial pressure, oxygen toxicity, and clinical outcome in severe traumatic brain injury. *J Neurosurg* 118(6):1317–1328. <https://doi.org/10.3171/2013.2>
 30. Bækgaard JS, Abback PS, Boubaya M, Moyer JD, Garrigue D, Raux M, Champigneulle B, Dubreuil G, Pottecher J, Laitselart P, Laloum F, Bloch-Queyrat C, Adnet F, Paugam-Burtz C, Traumabase Study Group (2020) Early hyperoxemia is associated with lower adjusted mortality after severe trauma: results from a French registry. *Crit Care* 24(1):604. <https://doi.org/10.1186/s13054-020-03274-x>
 31. Smith JL (1899) The pathological effects due to increase of oxygen tension in the air breathed. *J Physiol* 24(1):19–35. <https://doi.org/10.1113/jphysiol.1899.sp000746>
 32. Hedley-Whyte J, Winter PM (1967) Oxygen therapy. *Clin Pharmacol Ther* 8(5):696–737. <https://doi.org/10.1002/cpt.196785696>
 33. Urner M, Calfee CS, Fan E (2021) Titrating oxygen therapy in critically ill patients. *JAMA* 326(10):911–913. <https://doi.org/10.1001/jama.2021.9843>
 34. Angus DC (2020) Oxygen therapy for the critically ill. *N Engl J Med* 382(11):1054–1056. <https://doi.org/10.1056/NEJMe2000800>
 35. Frank L, Bucher JR, Roberts RJ (1978) Oxygen toxicity in neonatal and adult animals of various species. *J Appl Physiol* 45:699–704
 36. Davis WB, Rennard SJ, Bitterman PB, Crystal RG (1983) Pulmonary oxygen toxicity. Early reversible changes in human alveolar structures induced by hyperoxia. *N Engl J Med* 309(15):878–883. <https://doi.org/10.1056/NEJM198310133091502>
 37. Mantell LL, Lee PJ (2000) Signal transduction pathways in hyperoxia-induced lung cell death. *Mol Genet Metab* 71(1–2):359–370. <https://doi.org/10.1006/mgme.2000.3046>
 38. Wu J, Hafner C, Schramel JP, Kaun C, Krychtiuk KA, Wojta J, Boehme S, Ullrich R, Tretter EV, Markstaller K, Klein KU (2016) Cyclic and constant hyperoxia cause inflammation, apoptosis and cell death in human umbilical vein endothelial cells. *Acta Anaesthesiol Scand* 60(4):492–501. <https://doi.org/10.1111/aas.12646>
 39. Wang X, Wang Y, Kim HP, Nakahira K, Ryter SW, Choi AMK (2007) Carbon monoxide protects against hyperoxia-induced endothelial cell apoptosis by inhibiting reactive oxygen species formation. *J Biol Chem* 282(3):1718–1726. <https://doi.org/10.1074/jbc.M607610200>
 40. Bak Z, Sjöberg F, Rousseau A, Steinvall I, Janerot-Sjöberg B (2007) Human cardiovascular dose-response to supplemental oxygen. *Acta Physiol (Oxf)* 191(1):15–24. <https://doi.org/10.1111/j.1748-1716.2007.01710.x>
 41. Casey DP, Joyner MJ, Claus PL, Curry TB (2013) Vasoconstrictor responsiveness during hyperbaric hyperoxia in contracting human muscle. *J Appl Physiol* (1985) 114(2):217–224. <https://doi.org/10.1152/jappphysiol.01197.2012>
 42. Mak S, Egri Z, Tanna G, Colman R, Newton GE (2002) Vitamin C prevents hyperoxia-mediated vasoconstriction and impairment of endothelium-dependent vasodilation. *Am J Physiol Heart Circ Physiol* 282(6):H2414–H2421. <https://doi.org/10.1152/ajpheart.00947.2001>
 43. McNulty PH, Robertson BJ, Tulli MA, Hess J, Harach LA, Scott S, Sinoway LI (2007) Effect of hyperoxia and vitamin C on coronary blood flow in patients with ischemic heart disease. *J Appl Physiol* (1985) 102(5):2040–2045. <https://doi.org/10.1152/jappphysiol.00595.2006>
 44. Yusa T, Beckman JS, Crapo JD, Freeman BA (1987) Hyperoxia increases H₂O₂ production by brain in vivo. *J Appl Physiol* (1985) 63(1):353–358. <https://doi.org/10.1152/jappphysiol.1987.63.1.353>
 45. Brueckl C, Kaestle S, Kerem A, Habazettl H, Krombach F, Kuppe H, Kuebler WM (2006) Hyperoxia-induced reactive oxygen species formation in pulmonary capillary endothelial cells in situ. *Am J Respir Cell Mol Biol* 34(4):453–463. <https://doi.org/10.1165/rcmb.2005-0223OC>
 46. Hafner C, Wu J, Tiboldi A, Hess M, Mitulovic G, Kaun C, Krychtiuk KA, Wojta J, Ullrich R, Tretter EV, Markstaller K, Klein KU (2017) Hyperoxia induces inflammation and cytotoxicity in human adult cardiac myocytes. *Shock* 47(4):436–444. <https://doi.org/10.1097/SHK.0000000000000740>
 47. Nortje J, Coles JP, Timofeev I, Fryer TD, Aigbirhio FI, Smielewski P, Outtrim JG, Chatfield DA, Pickard JD, Hutchinson PJ, Gupta AK, Menon DK (2008) Effect of hyperoxia on regional oxygenation and metabolism after severe traumatic brain injury: preliminary findings. *Crit Care Med* 36(1):273–281. <https://doi.org/10.1097/01.CCM.0000292014.60835.15>
 48. Magnoni S, Ghisoni L, Locatelli M, Caimi M, Colombo A, Valeriani V, Stocchetti N (2003) Lack of improvement in cerebral metabolism after hyperoxia in severe head injury: a microdialysis study. *J Neurosurg* 98(5):952–958. <https://doi.org/10.3171/jns.2003.98.5.952>
 49. Girardis M, Busani S, Damiani E, Donati A, Rinaldi L, Marudi A, Morelli A, Antonelli M, Singer M (2016) Effect of conservative vs conventional oxygen therapy on mortality among patients in an intensive care unit: the oxygen-ICU randomized clinical trial. *JAMA* 316(15):1583–1589. <https://doi.org/10.1001/jama.2016.11993>
 50. Asfar P, Schortgen F, Boissramé-Helms J, Charpentier J, Guérot E, Megarbane B, Grimaldi D, Grelon F, Anguel N, Lasocki S, Henry-Lagarrigue M, Gonzalez F, Legay F, Guitton C, Schenck M, Doise JM, Devaquet J, Van Der Linden T, Chatellier D, Rigaud JP, Dellamonica J, Tamion F, Meziani F, Mercat A, Dreyfuss D, Seegers V, Radermacher P, HYPER2S Investigators; REVA Research Network (2017) Hyperoxia and hypertonic saline in patients with septic shock (HYPER2S): a two-by-two factorial, multicentre, randomised, clinical trial. *Lancet Respir Med* 5(3):180–190. [https://doi.org/10.1016/S2213-2600\(17\)30046-2](https://doi.org/10.1016/S2213-2600(17)30046-2)
 51. Chu DK, Kim LH, Young PJ, Zamiri N, Almenawer SA, Jaeschke R, Szczeklik W, Schünemann HJ, Neary JD, Alhazzani W (2018) Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA): a systematic review and meta-analysis. *Lancet* 391(10131):1693–1705. [https://doi.org/10.1016/S0140-6736\(18\)30479-3](https://doi.org/10.1016/S0140-6736(18)30479-3)
 52. Barrot L, Asfar P, Mauny F, Winiszewski H, Montini F, Badie J, Quenot JP, Pili-Floury S, Bouhemad B, Louis G, Souweine B, Collange O, Pottecher J, Levy B, Puyraveau M, Vettoretti L, Constantin JM, Capellier G, LOCO2 Investigators and REVA Research Network (2020) Liberal or conservative oxygen therapy for acute respiratory distress syndrome. *N Engl J Med* 382(11):999–1008. <https://doi.org/10.1056/NEJMoa1916431>
 53. Mackle D, Bellomo R, Bailey M, Beasley R, Deane A, Eastwood G, Finfer S, Freebairn R, King V, Linke N, Litton E, McArthur C, McGuinness S, Panwar R, Young P, ICU-ROX Investigators the Australian and New Zealand Intensive Care Society Clinical Trials Group (2020) Conservative oxygen therapy during mechanical ventilation in the ICU. *N Engl J Med* 382(11):989–998. <https://doi.org/10.1056/NEJMoa1903297>
 54. Schjørring OL, Klitgaard TL, Perner A, Wetterslev J, Lange T, Siegemund M, Bäcklund M, Keus F, Laake JH, Morgan M, Thormar KM, Rosborg SA, Bisgaard J, Erntgaard AES, Lynnerup AH, Pedersen RL, Crescioli E, Gielstrup TC, Behzadi MT, Poulsen LM, Estrup S, Laigaard JP, Andersen C, Mortensen CB, Brand BA, White J, Jarnvig IL, Møller MH, Quist L, Bestle MH, Schønemann-Lund M, Kamper MK, Hindborg M, Hollinger A, Gebhard CE, Zellweger N, Meyhoff CS, Hjort M, Bech LK, Grøfte T, Bundgaard H, Østergaard LHM, Thyø MA, Hildebrandt T, Uslu B, Sølling CG, Møller-Nielsen N, Brøchner AC, Borup M, Okkonen M, Dieperink W, Pedersen UG, Andreassen AS, Buus L, Aslam TN, Winding RR, Scheffold JC, Thorup SB, Iversen SA, Engstrøm J, Kjær MN, Rasmussen BS, HOT-ICU Investigators (2021) Lower or higher oxygenation targets for acute hypoxemic respiratory failure. *N Engl J Med* 384(14):1301–1311. <https://doi.org/10.1056/NEJMoa2032510>
 55. Gelissen H, de Grooth HJ, Smulders Y, Wils EJ, de Ruijter W, Vink R, Smit B, Röttgering J, Atmowihardjo L, Girbes A, Eibers P, Tuinman PR, Oudemans-van Straaten H, de Man A (2021) Effect of low-normal vs high-normal oxygenation targets on organ dysfunction in critically ill patients: a randomized clinical trial. *JAMA* 326(10):940–948. <https://doi.org/10.1001/jama.2021.13011>

-
56. Abdelhak A, Foschi M, Abu-Rumeileh S, Yue JK, D'Anna L, Huss A, Oeckl P, Ludolph AC, Kuhle J, Petzold A, Manley GT, Green AJ, Otto M, Tumani H (2022) Blood GFAP as an emerging biomarker in brain and spinal cord disorders. *Nat Rev Neurol*. <https://doi.org/10.1038/s41582-021-00616-3>
 57. Frankel M, Fan L, Yeatts SD, Jeromin A, Vos PE, Wagner AK, Wolf BJ, Pauls Q, Lunney M, Merck LH, Hall CL, Palesch YY, Silbergleit R, Wright DW (2019) Association of very early serum levels of S100B, glial fibrillary acidic protein, ubiquitin C-terminal hydrolase-L1, and spectrin breakdown product with outcome in ProTECT III. *J Neurotrauma* 36(20):2863–2871. <https://doi.org/10.1089/neu.2018.5809>
 58. Kochanek PM, Bramlett HM, Shear DA, Dixon CE, Mondello S, Dietrich WD, Hayes RL, Wang KK, Poloyac SM, Empey PE, Povlishock JT, Mountney A, Browning M, Deng-Bryant Y, Yan HQ, Jackson TC, Catania M, Glushakova O, Richieri SP, Tortella FC (2016) Synthesis of findings, current investigations, and future directions: operation brain trauma therapy. *J Neurotrauma* 33(6):606–614. <https://doi.org/10.1089/neu.2015.4133>
 59. Browning M, Shear DA, Bramlett HM, Dixon CE, Mondello S, Schmid KE, Poloyac SM, Dietrich WD, Hayes RL, Wang KK, Povlishock JT, Tortella FC, Kochanek PM (2016) Levetiracetam treatment in traumatic brain injury: operation brain trauma therapy. *J Neurotrauma* 33(6):581–594. <https://doi.org/10.1089/neu.2015.4131>
 60. Okonkwo DO, Shutter LA, Moore C, Temkin NR, Puccio AM, Madden CJ, Andaluz N, Chesnut RM, Bullock MR, Grant GA, McGregor J, Weaver M, Jallo J, LeRoux PD, Moberg D, Barber J, Lazaridis C, Diaz-Arrastia RR (2017) Brain oxygen optimization in severe traumatic brain injury phase-II: a phase II randomized trial. *Crit Care Med* 45(11):1907–1914. <https://doi.org/10.1097/CCM.0000000000002619>
 61. Chesnut R, Aguilera S, Buki A, Bulger E, Citerio G, Cooper DJ, Arrastia RD, Diring M, Figaji A, Gao G, Geocadin R, Ghajar J, Harris O, Hoffer A, Hutchinson P, Joseph M, Kitagawa R, Manley G, Mayer S, Menon DK, Meyfroidt G, Michael DB, Oddo M, Okonkwo D, Patel M, Robertson C, Rosenfeld JV, Rubiano AM, Sahuquillo J, Servadei F, Shutter L, Stein D, Stocchetti N, Taccone FS, Timmons S, Tsai E, Ullman JS, Vespa P, Videtta W, Wright DW, Zammit C, Hawryluk GWJ (2020) A management algorithm for adult patients with both brain oxygen and intracranial pressure monitoring: the Seattle International Severe Traumatic Brain Injury Consensus Conference (SIBICC). *Intensive Care Med* 46(5):919–929. <https://doi.org/10.1007/s00134-019-05900-x>

Appendix B

Simulations R code

```
## Complete case data generation
sim.data.joint<-function(old.results=list(),
  file,
  nSIM,#number of simulations
  fixed_trajectory,
  betaEvent_assoc, #True association parameter (log hazard
    ratio) in the event submodel
  n, #Number of individuals
  max_fuptime, #The maximum follow-up time in whatever the
    desired time units are.
  betaLong_intercept, #True intercept in the longitudinal
    submodel
  betaLong_continuous, #True coefficient for the continuous
    covariate in the longitudinal submode
  betaLong_binary, #True coefficient for the binary
    covariate in the longitudinal submode
  betaLong_linear , #True coefficient for the fixed effect
    linear term in the longitudinal submodel
  betaLong_quadratic, #True coefficient for the fixed
    effect quadratic term in the longitudinal submodel
    when fixed_trajectory = "quadratic"
  betaEvent_binary, #True coefficient (log HR) for the
    binary covariate in the event submodel
  betaEvent_continuous, #True coefficient (log HR) for the
    continuous covariate in the event submodel
  betaLong_aux,
  mean_Z2, #mean of Normal distribution for continuous
    baseline covariate
  sd_Z2, # mean of Normal distribution for continuous
    baseline covariate
  max_yobs, #max number of obs. per patient
  random_trajectory,#The desired type of trajectory in the
    random effects part of the longitudinal model
  balanced, # A logical, specifying whether the timings of
    the longitudinal measurements should be balanced
    across individuals
  b_sd,#vector or matrix for variance of RE
  b_rho,#scalar for correlation among RE
  return_eta=T,
  censor){
```

```

data<-list()
for(s in c(1:nSIM)){
  #creating the dataset
  mdat1 <- simjm::simjm(M = 1,
                        fixed_trajectory = fixed_trajectory,
                        betaEvent_assoc = betaEvent_assoc,
                        n = n,
                        max_fuptime = max_fuptime,
                        betaLong_intercept = betaLong_intercept,
                        betaLong_continuous = betaLong_continuous,
                        betaLong_binary = betaLong_binary,
                        betaLong_linear = betaLong_linear,
                        betaLong_quadratic = betaLong_quadratic,
                        betaEvent_binary = betaEvent_binary,
                        betaEvent_continuous = betaEvent_continuous,
                        betaLong_aux =betaLong_aux ,
                        mean_Z2 = mean_Z2,
                        sd_Z2 = sd_Z2,
                        random_trajectory = random_trajectory,
                        balanced = balanced,
                        b_sd = b_sd,
                        b_rho = b_rho,
                        return_eta=T,
                        max_yobs = max_yobs)

  #Random Censoring
  if(censor){
    mdat1[[1]] <- mdat1[[1]] %>%
      mutate(random.cens=runif(id,0,max_fuptime*10),
             tempoevento=pmin(eventtime,random.cens),
             stato=ifelse((tempoevento < eventtime) & status==1,
                           stato<-0,
                           stato<-status),
             status=stato,
             eventtime=tempoevento)

    mdat1[[2]] <- merge(mdat1[[2]],mdat1[[1]][,c("id","eventtime","status
          ")],by="id")
    names(mdat1[[2]])[c(9,10)]<-c("eventtime","status")

    mdat1[[2]] <- within(mdat1[[2]][mdat1[[2]]$tij <= mdat1[[2]]$eventtime
          ,], rm(eventtime.x,status.x))
  }

  ##count number of observation for each pz
  mdat1[[2]] <- mdat1[[2]] %>% group_by(id) %>% mutate(n = row_number())

  data[[s]]<-list(mdat1[[1]],mdat1[[2]])
}
try(save(data,file= paste0("nSIM",nSIM,
                           "_n",n,
                           "_FU",max_fuptime,
                           "_nobs",max_yobs,
                           "_",fixed_trajectory,

```

```

"_alpha",ifelse(betaEvent_assoc>0,as.character(
  gsub('[.]', '_', betaEvent_assoc)),paste0("
  negative",as.character(gsub('[.]', '_', abs(
  betaEvent_assoc))))),
"_beta0",ifelse(betaLong_intercept>0,as.character
  (gsub('[.]', '_', betaLong_intercept)),paste0("
  negative",as.character(gsub('[.]', '_', abs
  (betaLong_intercept))))),
"_betaC",ifelse(betaLong_continuous>0,as.
  character(gsub('[.]', '_',
  betaLong_continuous)),paste0("negative",as.
  character(gsub('[.]', '_', abs(
  betaLong_continuous))))),
"_betaQ",ifelse(fixed_trajectory=="quadratic",as.
  character(gsub('[.]', '_', betaLong_quadratic
  )), "_"),
"_betaB",ifelse(betaLong_binary>0,as.character(
  gsub('[.]', '_', betaLong_binary)),paste0("
  negative",as.character(gsub('[.]', '_', abs(
  betaLong_binary))))),
"_betaT",ifelse(betaLong_linear>0,as.character(
  gsub('[.]', '_', betaLong_linear)),paste0("
  negative",as.character(gsub('[.]', '_', abs(
  betaLong_linear))))),
"_bal",as.character(balanced),
"_stdRE",as.character((gsub('[.]', '_', b_sd[1])
  )),
"_sdMeasErr",betaLong_aux,
".RData")))
}

```

Intermittent data generation

```

sim.data.joint.miss<-function(data,
  file,
  nSIM,#number of simulations
  fixed_trajectory,
  betaEvent_assoc, #True association parameter (
  log hazard ratio) in the event submodel
  n, #Number of individuals
  max_fuptime, #The maximum follow-up time in
  whatever the desired time units are.
  betaLong_intercept, #True intercept in the
  longitudinal submodel
  betaLong_continuous, #True coefficient for the
  continuous covariate in the longitudinal
  submode
  betaLong_binary,
  betaLong_linear,
  betaLong_quadratic,
  betaEvent_binary,
  betaEvent_continuous,
  betaLong_aux,
  mean_Z2,
  sd_Z2,
  max_yobs,
  missing,

```

```

MR,
odds,
cutoffquantile){

datamiss<-list()
data.2 <- list()
data.2.1 <- list()

for(s in 1:nSIM){

#creating the dataset

if(missing!="none"){
data.2[[s]] <- data[[s]][[2]][which(data[[s]][[2]]$n<=2),]
data[[s]][[2]]<-missMethods::delete_MNAR_1_to_x(data[[s]][[2]][which
(data[[s]][[2]]$n>2),], p=MR, cols_mis="Xij_1",cutoff_fun =
function(i){quantile(i,prob=cutoffquantile)},x = odds)
data.2.1[[s]]<- rbind(data.2[[s]],data[[s]][[2]])
data[[s]][[2]]<- data.2.1[[s]]
data[[s]][[2]] <-data[[s]][[2]][with(data[[s]][[2]] , order(id, tij)
), ]
#data[[s]][[2]]<-missMethods::delete_MNAR_1_to_x(data[[s]][[2]], p=
MR, cols_mis="Xij_1",cutoff_fun = function(i){quantile(i,prob=
cutoffquantile)},x = odds)
for(i in 1:nrow(data[[s]][[2]])){
if(is.na(data[[s]][[2]][i,"Xij_1"])){
data[[s]][[2]][i,"Yij_1"]<-NA
}
}
}

datamiss[[s]]<-list(data[[s]][[1]],data[[s]][[2]])
}
try(save(datamiss,file= paste0("nSIM",nSIM,
"_n",n,
"_FU",max_fuptime,
"_nobs",max_yobs,
"_",ifelse(fixed_trajectory=="quadratic","q","1")
,
"_alpha",ifelse(betaEvent_assoc>0,as.character(
gsub('[.]', '_ ', betaEvent_assoc)),paste0("
negative",as.character(gsub('[.]', '_ ', abs(
betaEvent_assoc))))),
"_beta0",ifelse(betaLong_intercept>0,as.character
(gsub('[.]', '_ ', betaLong_intercept)),paste0(
"negative",as.character(gsub('[.]', '_ ', abs(
betaLong_intercept))))),
"_betaC",ifelse(betaLong_continuous>0,as.
character(gsub('[.]', '_ ',
betaLong_continuous)),paste0("negative",as.
character(gsub('[.]', '_ ', abs(
betaLong_continuous))))),
"_betaB",ifelse(betaLong_binary>0,as.character(
gsub('[.]', '_ ', betaLong_binary)),paste0("
negative",as.character(gsub('[.]', '_ ', abs(

```

```

        betaLong_binary))))),
    "_betaT",ifelse(betaLong_linear>0,as.character(
      gsub('[.]', '_', betaLong_linear)),paste0("
      negative",as.character(gsub('[.]', '_', abs(
      betaLong_linear))))),
    "_stdI",as.character((gsub('[.]', '_', b_sd[1]))),
    ,
    "_stdS",as.character((gsub('[.]', '_', b_sd[2]))),
    ,
    "_sdMeasErr",betaLong_aux,
    "_miss",missing,
    "_MR",as.character((gsub('[.]', '_', MR))),
    "_cutoff",as.character((gsub('[.]', '_',
      cutoffquantile))),
    "_odds",as.character((gsub('[.]', '_', odds))),
    ".RData"))
}

## "Lost assesment" data generation
sim.data.joint.carrymiss<-function(data,
    file,
    nSIM,
    fixed_trajectory,
    betaEvent_assoc,
    n,
    betaLong_intercept,
    betaLong_binary,
    betaLong_linear,
    betaLong_quadratic,
    betaEvent_continuous,
    betaLong_aux,
    mean_Z2,
    sd_Z2,
    max_yobs,
    b_sd,#vector or matrix for variance of RE
    b_rho,
    missing,
    MR,
    odds,
    cutoffquantile,
    carried_miss){

data.miss.carried<-list()
for(s in (1:nSIM)){
  #creating the dataset
  if(carried_miss){
    data[[s]][[2]]$cumsumNA <- ave(is.na( data[[s]][[2]]$Yij_1), data
      [[s]][[2]]$id, FUN = cumsum)
    data[[s]][[2]] <- data[[s]][[2]][-which( data[[s]][[2]]$cumsumNA
      >0),]
  }

  print(paste(s))

  data.miss.carried[[s]]<-list(data[[s]][[1]],data[[s]][[2]])

```

```

}
try(save(data.miss.carried,file= paste0("nSIM",nSIM,
    "_n",n,
    "_FU",max_fuptime,
    "_nobs",max_yobs,
    "_",ifelse(fixed_trajectory=="quadratic","q","l")
    ,
    "_alpha",ifelse(betaEvent_assoc>0,as.character(
      gsub('[.]', '_ ', betaEvent_assoc)),paste0("
      negative",as.character(gsub('[.]', '_ ', abs(
      betaEvent_assoc))))),
    "_beta0",ifelse(betaLong_intercept>0,as.character
      (gsub('[.]', '_ ', betaLong_intercept)),paste0("
      negative",as.character(gsub('[.]', '_ ', abs(
      betaLong_intercept))))),
    "_betaC",ifelse(betaLong_continuous>0,as.
      character(gsub('[.]', '_ ',
      betaLong_continuous)),paste0("negative",as.
      character(gsub('[.]', '_ ', abs(
      betaLong_continuous))))),
    "_betaB",ifelse(betaLong_binary>0,as.character(
      gsub('[.]', '_ ', betaLong_binary)),paste0("
      negative",as.character(gsub('[.]', '_ ', abs(
      betaLong_binary))))),
    "_betaT",ifelse(betaLong_linear>0,as.character(
      gsub('[.]', '_ ', betaLong_linear)),paste0("
      negative",as.character(gsub('[.]', '_ ', abs(
      betaLong_linear))))),
    "_stdI",as.character((gsub('[.]', '_ ', b_sd[1])))
    ,
    "_stdS",as.character((gsub('[.]', '_ ', b_sd[2])))
    ,b('[.]', '_ ', sd_Z2)),
    "_sdMeasErr",betaLong_aux,
    "_miss",missing,
    "_MR",as.character((gsub('[.]', '_ ', MR))),
    "_cutoff",as.character((gsub('[.]', '_ ',
      cutoffquantile))),
    "_odds",as.character((gsub('[.]', '_ ', odds))),
    "_carried",carried_miss,
    ".RData"))))
}

## Model fitting on generated data (we report the case of JMs with B-
splines approximated baseline hazard, but other specifications can be
achieved by changing the "method" in JM::jointmodel(), accordingly

sim.joint.spline<- function(old.results=list(),file=file,data,formulaLME,
  nSIM,carried_miss,fixed_trajectory){

  resultsJMspline <- old.results
  nevent <- c()
  nsubj <- c()
  resultsCOX <- old.results
  for(s in c(1:nSIM)){

```



```

data[[s]][[2]] <- data[[s]][[2]] %>% tidyr::drop_na()
data[[s]][[1]] <- filter(data[[s]][[1]], data[[s]][[1]]$id %in% unique(
  data[[s]][[2]]$id)
#####
##COX##
resultsCOX[[s]]<-list()
resultsCOX[[s]]$estimatesCOX<-matrix(NA,nrow=1,ncol=1,dimnames=list(c("
  COX"),c("alpha")))
resultsCOX[[s]]$lower<-matrix(NA,nrow=1,ncol=1,dimnames=list(c("COX"),c
  ("CIlower")))
resultsCOX[[s]]$upper<-matrix(NA,nrow=1,ncol=1,dimnames=list(c("COX"),c
  ("CIupper")))

temp2 <- survival::tmerge(data[[s]][[1]], data[[s]][[1]], id=id, endpt =
  event(eventtime, status))

temp3 <- survival::tmerge(temp2,data[[s]][[2]],
  id=id,
  Yij_1 = tdc(tij,Yij_1))

Ex.Cox <- try(survival::coxph(survival::Surv(tstart, tstop, endpt) ~
  Yij_1 ,data=temp3,x=T,y=T))# control = coxph.control(timefix = FALSE
  )

if(class(Ex.Cox)!="try-error"){
  resultsCOX[[s]]$estimatesCOX<-Ex.Cox$coefficients

  resultsCOX[[s]]$sd<-sqrt(Ex.Cox$var)
  sd<-sqrt(Ex.Cox$var)
  resultsCOX[[s]]$lower<-Ex.Cox$coefficients-1.96*sd
  resultsCOX[[s]]$upper<-Ex.Cox$coefficients+1.96*sd
}else{
  resultsCOX[[s]]$estimatesCOX<-matrix(NA,nrow=1,ncol=1,dimnames=list(c
  ("COX"),c("alpha")))
  resultsCOX[[s]]$lower<-matrix(NA,nrow=1,ncol=1,dimnames=list(c("COX"),
  c("CIlower")))
  resultsCOX[[s]]$upper<-matrix(NA,nrow=1,ncol=1,dimnames=list(c("COX"),
  c("CIupper")))
}

resultsJMspline[[s]]<-list()
resultsJMspline[[s]]$LMM_int<-matrix(NA,nrow=1,ncol=1,dimnames=list(c("
  LMM"),c("Intercept")))
resultsJMspline[[s]]$LMM_beta1<-matrix(NA,nrow=1,ncol=1,dimnames=list(c
  ("LMM"),c("tij")))
resultsJMspline[[s]]$LMM_betaBinary<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("LMM"),c("Z1")))
resultsJMspline[[s]]$LMM_betaContinuous<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("LMM"),c("Z2")))
resultsJMspline[[s]]$LMM_beta_1lower<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("LMM"),c("tij_lower")))
resultsJMspline[[s]]$LMM_beta_1upper<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("LMM"),c("tij_upper")))
resultsJMspline[[s]]$LMM_betaBinarylower<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("LMM"),c("Z1_lower")))

```

```

resultsJMspline[[s]]$LMM_betaBinaryupper<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("LMM"),c("Z1_upper")))
resultsJMspline[[s]]$LMM_betaContinuouslower<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("LMM"),c("Z2_lower")))
resultsJMspline[[s]]$LMM_betaContinuousupper<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("LMM"),c("Z2_upper")))
resultsJMspline[[s]]$estimatesJM<-matrix(NA,nrow=1,ncol=1,dimnames=list(
  c("joint"),c("alpha")))
resultsJMspline[[s]]$stderr<-matrix(NA,nrow=1,ncol=1,dimnames=list(c("
  joint"),c("stderr")))
resultsJMspline[[s]]$stderralpha<-matrix(NA,nrow=1,ncol=1,dimnames=list(
  c("joint"),c("stderralpha")))
resultsJMspline[[s]]$alphalower<-matrix(NA,nrow=1,ncol=1,dimnames=list(c(
  "joint"),c("alphalower")))
resultsJMspline[[s]]$alphaupper<-matrix(NA,nrow=1,ncol=1,dimnames=list(c(
  "joint"),c("alphaupper")))
resultsJMspline[[s]]$estimatesJM_beta0<-matrix(NA,nrow=1,ncol=1,dimnames
  =list(c("joint"),c("beta0")))
resultsJMspline[[s]]$stderr_beta0<-matrix(NA,nrow=1,ncol=1,dimnames=list
  (c("joint"),c("stderr_beta0")))
resultsJMspline[[s]]$beta_0lower<-matrix(NA,nrow=1,ncol=1,dimnames=list(
  c("joint"),c("beta0_lower")))
resultsJMspline[[s]]$beta_0upper<-matrix(NA,nrow=1,ncol=1,dimnames=list(
  c("joint"),c("beta0_upper")))
resultsJMspline[[s]]$estimatesJM_beta1<-matrix(NA,nrow=1,ncol=1,dimnames
  =list(c("joint"),c("beta1")))
resultsJMspline[[s]]$stderr_beta1<-matrix(NA,nrow=1,ncol=1,dimnames=list
  (c("joint"),c("stderr_beta1")))
resultsJMspline[[s]]$beta_1lower<-matrix(NA,nrow=1,ncol=1,dimnames=list(
  c("joint"),c("beta1_lower")))
resultsJMspline[[s]]$beta_1upper<-matrix(NA,nrow=1,ncol=1,dimnames=list(
  c("joint"),c("beta1_upper")))
resultsJMspline[[s]]$estimatesJM_betaBinary<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("betaBinary")))
resultsJMspline[[s]]$stderr_betaBinary<-matrix(NA,nrow=1,ncol=1,dimnames
  =list(c("joint"),c("stderr_betaBinary")))
resultsJMspline[[s]]$betaBinarylower<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("betaBinary_lower")))
resultsJMspline[[s]]$betaBinaryupper<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("betaBinary_upper")))
resultsJMspline[[s]]$estimatesJM_betaContinuous<-matrix(NA,nrow=1,ncol
  =1,dimnames=list(c("joint"),c("betaContinuous")))
resultsJMspline[[s]]$stderr_betaContinuous<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("stderr_betaContinuous")))
resultsJMspline[[s]]$betaContinuouslower<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("betaContinuous_lower")))
resultsJMspline[[s]]$betaContinuousupper<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("betaContinuous_upper")))
resultsJMspline[[s]]$estimatesJM_stdIntercept <- matrix(NA,nrow=1,ncol
  =1,dimnames=list(c("joint"),c("stdIntercept")))
resultsJMspline[[s]]$estimatesJM_stdSlope <- matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("stdSlope")))

cox.fit<-survival::coxph(survival::Surv(time=eventtime,status) ~ 1, data
  = data[[s]][[1]], x=TRUE)

```

```

nevent <- c(nevent,cox.fit$nevent)
nsubj <- c(nsubj,cox.fit$n)
print(paste("sim ",s))

##fit the longitudinal submodel for the biomarker
if(fixed_trajectory=="quadratic"){
  lme.fit<-try(lme(Yij_1 ~ I(tij^2)+tij + Z2 + Z1, random = ~ 1 + tij|
    id, data = data[[s]][[2]],
    control = lmeControl(opt = 'optim',tolerance = 1e-16)
  ))
}
if(fixed_trajectory=="linear"){
  lme.fit<-try(lme(Yij_1 ~ tij + Z2 + Z1, random = ~ 1 + tij| id, data
    = data[[s]][[2]],
    control = lmeControl(opt = 'optim',tolerance = 1e-16)
  ))
}
resultsJMspline[[s]]$LMM_int<-lme.fit$coefficients$fixed["(Intercept)
"]
resultsJMspline[[s]]$LMM_beta1<-lme.fit$coefficients$fixed["tij"]
resultsJMspline[[s]]$LMM_betaBinary<-lme.fit$coefficients$fixed["Z1"]
resultsJMspline[[s]]$LMM_betaContinuous<-lme.fit$coefficients$fixed["
Z2"]
resultsJMspline[[s]]$LMM_beta_1lower<-lme.fit$coefficients$fixed["tij
"]-1.96*sqrt(diag(vcov(lme.fit)))[["tij"]
resultsJMspline[[s]]$LMM_beta_1upper<-lme.fit$coefficients$fixed["tij
"]+1.96*sqrt(diag(vcov(lme.fit)))[["tij"]
resultsJMspline[[s]]$LMM_betaBinarylower<-lme.fit$coefficients$fixed["
Z1"]-1.96*sqrt(diag(vcov(lme.fit)))[["Z1"]
resultsJMspline[[s]]$LMM_betaBinaryupper<-lme.fit$coefficients$fixed["
Z1"]+1.96*sqrt(diag(vcov(lme.fit)))[["Z1"]
resultsJMspline[[s]]$LMM_betaContinuouslower<-lme.
fit$coefficients$fixed["Z2"]-1.96*sqrt(diag(vcov(lme.fit)))[["Z2"]
resultsJMspline[[s]]$LMM_betaContinuousupper<-lme.
fit$coefficients$fixed["Z2"]+1.96*sqrt(diag(vcov(lme.fit)))[["Z2"]
resultsJMspline[[s]]$LMM_int<-lme.fit$coefficients$fixed["(Intercept)
"]
resultsJMspline[[s]]$LMM_beta1<-lme.fit$coefficients$fixed["tij"]

resultsJMspline[[s]]$LMM_beta_1lower<-lme.fit$coefficients$fixed["tij
"]-1.96*sqrt(diag(vcov(lme.fit)))[["tij"]
resultsJMspline[[s]]$LMM_beta_1upper<-lme.fit$coefficients$fixed["tij
"]+1.96*sqrt(diag(vcov(lme.fit)))[["tij"]

if(class(lme.fit)=="try-error"){
  if(fixed_trajectory=="quadratic"){
    lme.fit<-try(lme(Yij_1 ~ I(tij^2)+tij, random = ~ 1+tij| id, data
      = data[[s]][[2]],control = lmeControl(opt = 'optim',tolerance =
        1e-16,msMaxIter=100)),silent = TRUE)#list(opt="optim")
  }
  if(fixed_trajectory=="linear"){
    lme.fit<-try(lme(Yij_1 ~ tij, random = ~ 1+tij| id, data = data[[s
      ]][[2]],control = lmeControl(opt = 'optim',tolerance = 1e-16,
        msMaxIter=100)),silent = TRUE)#list(opt="optim")
  }
}

```

```

resultsJMspline[[s]]$LMM_int<-lme.fit$coefficients$fixed["(Intercept
)"]
resultsJMspline[[s]]$LMM_beta1<-lme.fit$coefficients$fixed["tij"]

resultsJMspline[[s]]$LMM_beta_1lower<-lme.fit$coefficients$fixed["
tij"]-1.96*sqrt(diag(vcov(lme.fit)))["tij"]
resultsJMspline[[s]]$LMM_beta_1upper<-lme.fit$coefficients$fixed["
tij"]+1.96*sqrt(diag(vcov(lme.fit)))["tij"]
resultsJMspline[[s]]$LMM_betaBinarylower<-lme.fit$coefficients$fixed
["Z1"]-1.96*sqrt(diag(vcov(lme.fit)))["Z1"]
resultsJMspline[[s]]$LMM_betaBinaryupper<-lme.fit$coefficients$fixed
["Z1"]+1.96*sqrt(diag(vcov(lme.fit)))["Z1"]
resultsJMspline[[s]]$LMM_betaContinuouslower<-lme.
fit$coefficients$fixed["Z2"]-1.96*sqrt(diag(vcov(lme.fit)))["Z2
"]
resultsJMspline[[s]]$LMM_betaContinuousupper<-lme.
fit$coefficients$fixed["Z2"]+1.96*sqrt(diag(vcov(lme.fit)))["Z2
"]
}
}

jointFit.simjm.pc.spline <- try(JM::jointModel(lme.fit, cox.fit, timeVar
= "tij", method = "spline-PH-aGH"))

if(class(jointFit.simjm.pc.spline)!="try-error"){
  resultsJMspline[[s]]$estimatesJM <- jointFit.simjm.pc.
  spline$coefficients$alpha
  resultsJMspline[[s]]$estimatesJM_beta0 <- jointFit.simjm.pc.
  spline$coefficients$betas["(Intercept)"]
  resultsJMspline[[s]]$estimatesJM_beta1 <- jointFit.simjm.pc.
  spline$coefficients$betas["tij"]
  resultsJMspline[[s]]$estimatesJM_betaBinary <- jointFit.simjm.pc.
  spline$coefficients$betas["Z1"]
  resultsJMspline[[s]]$estimatesJM_betaContinuous <- jointFit.simjm.pc.
  spline$coefficients$betas["Z2"]
  resultsJMspline[[s]]$estimatesJM_stdIntercept <- sqrt(jointFit.simjm.
pc.spline$coefficients$D[1,1])
  resultsJMspline[[s]]$estimatesJM_stdSlope <- sqrt(jointFit.simjm.pc.
spline$coefficients$D[2,2])

  resultsJMspline[[s]]$stderr<-try(sqrt(diag(solve(jointFit.simjm.pc.
spline$Hessian))),silent = TRUE)

  resultsJMspline[[s]]$stderralpha<-resultsJMspline[[s]]$stderr["T.alpha
"]
  resultsJMspline[[s]]$alphalower<- tryCatch(jointFit.simjm.pc.
spline$coef$alpha-1.96*resultsJMspline[[s]]$stderralpha,error=
function(err) NA)
  resultsJMspline[[s]]$alphaupper<- tryCatch(jointFit.simjm.pc.
spline$coef$alpha+1.96*resultsJMspline[[s]]$stderralpha,error=
function(err) NA)

  resultsJMspline[[s]]$stderr_beta0 <- resultsJMspline[[s]]$stderr["Y.(
Intercept)"]

```

```

resultsJMspline[[s]]$beta_0lower <- tryCatch(jointFit.simjm.pc.
  spline$coef$betas["(Intercept)"]-1.96*resultsJMspline[[s]]
  $stderr_beta0, error=function(err) NA)
resultsJMspline[[s]]$beta_0upper <- tryCatch(jointFit.simjm.pc.
  spline$coef$betas["(Intercept)"+1.96*resultsJMspline[[s]]
  $stderr_beta0, error=function(err) NA)

resultsJMspline[[s]]$stderr_beta1 <- resultsJMspline[[s]]$stderr["Y.
  tij"]
resultsJMspline[[s]]$beta_1lower <- tryCatch(jointFit.simjm.pc.
  spline$coef$betas["tij"]-1.96*resultsJMspline[[s]]$stderr_beta1,
  error=function(err) NA)
resultsJMspline[[s]]$beta_1upper <- tryCatch(jointFit.simjm.pc.
  spline$coef$betas["tij"+1.96*resultsJMspline[[s]]$stderr_beta1,
  error=function(err) NA)

resultsJMspline[[s]]$stderr_betaBinary <- resultsJMspline[[s]]$stderr
  ["Y.Z1"]
resultsJMspline[[s]]$betaBinarylower <- tryCatch(jointFit.simjm.pc.
  spline$coef$betas["Z1"]-1.96*resultsJMspline[[s]]
  $stderr_betaBinary, error=function(err) NA)
resultsJMspline[[s]]$betaBinaryupper <- tryCatch(jointFit.simjm.pc.
  spline$coef$betas["Z1"]*resultsJMspline[[s]]$stderr_betaBinary,
  error=function(err) NA)

resultsJMspline[[s]]$stderr_betaContinuous <- resultsJMspline[[s]]
  $stderr["Y.Z2"]
resultsJMspline[[s]]$betaContinuouslower <- tryCatch(jointFit.simjm.pc
  .spline$coef$betas["Z2"]-1.96*resultsJMspline[[s]]
  $stderr_betaContinuous, error=function(err) NA)
resultsJMspline[[s]]$betaContinuousupper <- tryCatch(jointFit.simjm.pc
  .spline$coef$betas["Z2"+1.96*resultsJMspline[[s]]
  $stderr_betaContinuous, error=function(err) NA)
}else{
resultsJMspline[[s]]<-list()
resultsJMspline[[s]]$estimatesJM<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("alpha")))
resultsJMspline[[s]]$stderralpha<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("stderralpha")))
resultsJMspline[[s]]$alphalower<-matrix(NA,nrow=1,ncol=1,dimnames=list
  (c("joint"),c("alphalower")))
resultsJMspline[[s]]$alphaupper<-matrix(NA,nrow=1,ncol=1,dimnames=list
  (c("joint"),c("alphalower")))

resultsJMspline[[s]]$estimatesJM_beta0<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("beta0")))
resultsJMspline[[s]]$stderr_beta0<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("stderr_beta0")))
resultsJMspline[[s]]$beta_0lower<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("beta0_lower")))
resultsJMspline[[s]]$beta_0upper<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("beta0_upper")))

resultsJMspline[[s]]$estimatesJM_beta1<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("beta1")))

```

```

resultsJMspline[[s]]$stderr_beta1<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("stderr_beta1")))
resultsJMspline[[s]]$beta_0lower<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("beta1_lower")))
resultsJMspline[[s]]$beta_1upper<-matrix(NA,nrow=1,ncol=1,dimnames=
  list(c("joint"),c("beta1_upper")))

resultsJMspline[[s]]$estimatesJM_betaBinary<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("betaBinary")))
resultsJMspline[[s]]$stderr_betaBinary<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("stderr_betaBinary")))
resultsJMspline[[s]]$beta_Binarylower<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("betaBinary_lower")))
resultsJMspline[[s]]$beta_Binaryupper<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("betaBinary_upper")))

resultsJMspline[[s]]$estimatesJM_betaContinuous<-matrix(NA,nrow=1,ncol
=1,dimnames=list(c("joint"),c("betaContinuous")))
resultsJMspline[[s]]$stderr_betaContinuous<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("stderr_betaContinuous")))
resultsJMspline[[s]]$beta_Continuouslower<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("betaContinuous_lower")))
resultsJMspline[[s]]$beta_Continuousupper<-matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("betaContinuous_upper")))
resultsJMspline[[s]]$estimatesJM_stdIntercept <- matrix(NA,nrow=1,ncol
=1,dimnames=list(c("joint"),c("stdIntercept")))
resultsJMspline[[s]]$estimatesJM_stdSlope <- matrix(NA,nrow=1,ncol=1,
  dimnames=list(c("joint"),c("stdSlope")))
}
}
try(save(resultsJMspline,
  nevent,
  nsubj,
  resultsCOX,
  file = file))
}

```