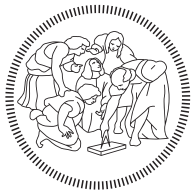


Politecnico di Milano
Department of Mathematics
Doctoral Program In Mathematical Models
and Methods in Engineering

Non-parametric classification and regression techniques for the characterisation of the disease subtypes and the assessment of the temporal evolution of image-based biomarkers

Doctoral Dissertation of:
Riccardo Pascuzzo

Supervisor:
Prof. Simone Vantini

Co-advisor:
Dr. Aymeric Stamm

The Chair of the Doctoral Program:
Prof. Irene Maria Sabadini

Year 2017 – XXX Cycle

Contents

Introduction

The concept of biomarker has evolved during time, from the inclusion of only biological molecules, to the recently updated definition given by the Food and Drug Administration in collaboration with the National Institutes of Health of the US [1]: "*a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions. Molecular, histologic, radiographic, or physiologic characteristics are types of biomarkers*". Thus, it is clear that a biomarker can originate from different modalities (e.g., biological samples, medical images), be qualitatively or quantitatively measured, and used for several purposes (e.g. diagnosis, prognosis, treatment assessment). In addition, according to the aim of the study, the specific research design guides how to collect biomarker measurements (e.g., cross-sectional study, longitudinal study). From the last decades, all these types of biomarker have become more and more commonly adopted in the clinical practice and research, and this has also stimulated in the statistical science an increasing interest in the development of new and advanced tools to properly handle the structure of each dataset [2–4].

In this thesis, we present some non-parametric statistical models and methods that have been developed and adapted to deal with different types of biomarker. In particular, Chapter 1 describes the assessment of the res-

piratory function evolution of Duchenne Muscular Dystrophy (DMD) patients from childhood to adulthood, where spirometric and opto-eletronic plethysmographic biomarker measurements are collected longitudinally at different time points for each subject. We deal with this irregular and subject-specific timing by adopting a regression model based on natural cubic splines with mixed effects, that allows to identify specific time points of respiratory impairment during disease progression, and to investigate possible effects of scoliosis, nocturnal non-invasive mechanical ventilation (NIMV) and steroid therapy. This is a joint work with Prof. Andrea Aliverti and Dr. Maria Antonella Lo Mauro (Department of Bioengineering, Politecnico di Milano) and Dr. Maria Grazia D'Angelo, Dr. Marianna Romei and Dr. Sandra Gandossini (IRCCS Eugenio Medea, Bosisio Parini, Lecco), forthcoming in the *European Respiratory Journal* as an original article.

In Chapter 2, we characterise the sybtypes of the sporadic Creutzfeldt-Jakob disease (sCJD) with imaging biomarkers collected in a cross-sectional design. In this case, the considered biomarkers are the signal hyperintensities of diffusion magnetic resonance imaging (dMRI), that are measured with a semi-quantitative scoring system devised to visually assess the images in different brain regions. After validating this scoring system and its reliability to diagnose the disease, we classify the sCJD patients into their most compatible subtype according to their biomarker measurements, with a classification tree-based method. This is a joint work with Dr. Alberto Bizzi (IRCCS Carlo Besta, Milan) and Dr. Pierluigi Gambetti (Department of Pathology, Case Western Reserve University, School of Medicine, Cleveland OH).

In Chapter 3, using the same data as in Chapter 2, we describe the disease progression in each sCJD subtype by finding the sequence of brain regions

that become detectably hyperintense in DWI. To overcome the missing temporal dependency that is intrinsic in any cross-sectional design, we adapt the recently introduced event-based model [5, 6], a data-driven statistical model that assess the disease evolution in terms of its characterising biomarkers, without relying on a longitudinal dataset. This is a joint work with Dr. Alberto Bizzi (IRCCS Carlo Besta, Milan), Prof. Daniel Alexander, Dr. Neil Oxtoby, Dr. Alexandra Young and Dr. Sara Garbarino (Department of Computer Science and Centre of Medical Image Computing, University College London).

Finally, in Chapter 4 we outline a work still in progress aimed at developing a function-on-function regression model that can deal with temporal dependent biomarkers (e.g., from functional magnetic resonance imaging data). We model a functional response in terms of several functional covariates, and we propose a permutation test to identify sub-regions that exhibit similar statistical differences. Moreover, in case of multiple tests performed at different locations in the same domain (e.g., the voxels of the brain MR image), we extend to a three-dimensional setting the closure multiplicity adjustment method [7, 8] to control the family-wise error rate of the proposed procedure.

Chapter 1

Evolution of Respiratory Function in Duchenne Muscular Dystrophy from Childhood to Adulthood

Duchenne muscular dystrophy (DMD) is an X-linked myopathy resulting in progressive wasting of locomotor and respiratory muscles, with consequent chronic ventilatory failure that is the main cause of death. In these patients it is extremely important, therefore, to measure lung function and respiratory muscle action in order to monitor the progression of the disease, to identify early signs of ventilatory insufficiency, to plan optimal interventions for improving the quality of life and to quantify the effects of novel gene-modifying strategies and pharmacological therapies [9–17].

At present, the outcome measures consider mainly motor function, namely of lower and upper limbs. Specific respiratory outcome measures are needed to objectively evaluate the effects of interventions in DMD, not only regarding spirometric and lung volume indexes but also respiratory muscle function. Forced vital capacity (FVC), when expressed in liters, follows a

pathognomonic pattern characterized by an ascending phase, a plateau and a descending phase during the course of the disease [18–23]. When FVC is instead expressed as percentage of predicted value, it linearly declines with age indicating a progressive increase of lung restriction [18–21], [23–25]. The deterioration of lung function can be stabilized with steroid therapy [13] [26–29], which helps delaying the loss of ambulation and the consequent development of scoliosis, an additional contributor to the restrictive lung pattern [13] [29–35].

Spirometry is recommended by current guidelines for routine lung function evaluation in DMD [9, 10], since FVC has prognostic value for survival [19] and is a useful guidance for treatment [19,23,36,37]. Nevertheless, spirometry has inherent limitations. A high level of patient's cooperation is needed, thus it cannot be applied in early childhood and it becomes difficult in adulthood due to the fatigue induced by repeated maximal maneuvers and/or the presence of macroglossia [38–40]. It provides only a global evaluation of lung restriction, irrespectively of its possible causes, such as alterations in lung, chest wall, respiratory muscles or a combination of them. Moreover, it does not provide any specific information on the impairment of ribcage muscles, diaphragm and abdominal muscles.

The detailed analysis of spontaneous breathing at rest including also thoracoabdominal contributions to tidal volume (breathing pattern), represents a useful approach for noninvasive and non-volitional assessment of respiratory function feasible in all patients [18, 41]. Abdominal contribution to tidal volume progressively decays with age [41] being a strong predictor of nocturnal hypoxemia [42] and inefficient cough [43].

We hypothesized that breathing pattern can provide information complementary to spirometry, regarding the natural course of the disease and the effects of given treatments. The specific aims were to study the natural

evolution of respiratory function in terms of spirometry, lung volumes and breathing pattern from childhood to adulthood, to identify possible key points and to investigate possible effects of scoliosis, nocturnal non-invasive mechanical ventilation (NIMV) and steroid therapy.

1.1 Material and Methods

1.1.1 Patients

This is a 7-year retrospective study of respiratory function in 115 patients, with a defined diagnosis of DMD [11], age ranging from 6 to 24 years and data collected at least at 3 different visits, out of the 167 patients followed at the IRCCS "E.Medea" Insitute (Bosisio Parini, Lecco, Italy). Patients were evaluated once/year until wheelchair bounding and thereafter twice/year, for 574 visits. At each visit, anthropometric and clinical (scoliosis, ambulation, steroid therapy and use of assistive respiratory devices) data were documented. All patients or parents signed a consent form, approved by the Local Ethical Committee accordingly to the declaration of Helsinki. Fifty-six age-matched healthy male subjects were enrolled as control group. The control group was composed by healthy brothers of DMD patients, relatives of the researches and students of the laboratory who volunteered to take part to the study.

1.1.2 Pulmonary function test

At each visit, FVC, Forced Expiratory Volume in 1 sec (FEV1), Peak Expiratory Flow (PEF) and subdivisions of lung volume (Functional Residual Capacity, FRC; Residual volume, RV, and Total Lung Capacity, TLC) by the nitrogen washout technique were measured (Vmax series 22, SensorMedics, Yorba Linda, CA). Spirometric data were presented both as absolute and expressed as percentage of the predicted values [44]. Noc-

turnal oxygen saturation (SpO2) was measured in all patients not under nocturnal non-invasive mechanical ventilation (NIMV) by pulse-oximetry (Nonin, 8500, Quitman, TX) and only recordings >8 hours were considered acceptable.

1.1.3 Assessment of spontaneous breathing pattern at rest

Breathing pattern was measured in supine position using opto-electronic plethysmography (OEP System; BTS, Milan, Italy) and a geometrical model based on 52 markers [45, 46]. After a short period of adaptation to the recording conditions, total and compartmental volumes were continuously measured during five minutes of quiet breathing. An average period of 90 sec of stable breathing was then selected during which the following parameters were calculated breath-by-breath: tidal volume (VT), respiratory rate (RR), minute ventilation, rapid and shallow breathing index (RSBi=RR/VT) and ribcage and abdominal tidal volumes (ΔVRC and ΔVAB, respectively, expressed both in liters and as percentage contribution to VT). Tidal volume was analyzed also normalized according to weight. The ribcage was in turn split into compartments, namely pulmonary ribcage and abdominal ribcage [45, 46]. The volume variations of the two ribcage components were presented both as absolute and percentage values.

1.1.4 Effects of scoliosis, steroids and NIMV

To study the effect of scoliosis, patients were grouped accordingly to the severity of scoliosis: null (NU), mild (MI, Cobb angle $< 20°$), moderate (MO, $20 - 40°$), severe (SE, $> 40°$) and after spinal fusion (SF). To study the effect of steroids, patients were grouped as "naive" (N, never treated or treated for < 1 year), "current" (C, under treatment or stopped since < 1

year before the visit) or "past" (P, not under treatment, but previously treated for > 1 year). To study the effect of NIMV, patients were grouped as "NIMV" (currently under treatment) and "no-NIMV" (patients never treated with NIMV within the period of the study). Patients who underwent NIMV treatment within the period of the study were identified and the data before the start of the treatment constituted a third group called "pre-NIMV".

1.1.5 Statistical data analysis

We used a basis of S natural cubic splines built on the time domain 6-24 years, where $S = 5$ was chosen by minimizing the PRESS index (i.e., Predictive Residual Error Sum of Squares) which is a leave-one-out cross-validation estimate of the average squared prediction error of the fitted model when it is used to predict the value of unobserved data.

Data were analyzed by the following regression model:

$$y_{iG}(t) = \sum_{k=1}^{S} \beta_{kG}\phi_k(t) + \sum_{k=1}^{S} b_{kiG}\phi_k(t) + \epsilon_{iG}(t) \tag{1.1}$$

where: i) $y_{iG}(t)$ is the datum that one would have recorded if the i-th subject in the group G were measured at age $t \in (6, 24)$. The group G can be one of the following: the whole population of DMD patients, the whole population of healthy controls, or one of the categories of the steroid, the scoliosis or the NIMV subgroups; ii) $\sum_{k=1}^{S} \beta_{kG}\phi_k(t)$ indicates the population mean curve; iii) $\sum_{k=1}^{S} b_{kiG}\phi_k(t)$ indicates the subject-specific correction such that $\sum_{k=1}^{S} \beta_{kG}\phi_k(t) + \sum_{k=1}^{S} b_{kiG}\phi_k(t)$ indicates the subject-specific evolution curve, and finally iv) $\epsilon_{iG}(t)$ indicates the session-specific measurement error.

In detail, $\phi_k(t)$ is the value at age t of the k-th natural cubic spline ($k = 1, \ldots, S$), β_{kG} is the regression coefficient related to ϕ_k and to the

group G, b_{kiG} is the random regression coefficient related to ϕ_k and to the i-th subject in the group G, and $\epsilon_{iG}(t)$ is the measurement error at age t related to the i-th subject in the group G. We assumed that $b_{kiG} \sim \mathcal{N}(0, \sigma_k^2)$ for each subject i in the group G, allowing a different variance σ_k^2 for each natural cubic spline, and that the measurement errors $\epsilon_{iG}(t) \sim \mathcal{N}(0, \sigma^2)$ for each subject i in the group G for every time t. Moreover, we assumed all random terms in the models to be independent.

For every acquired datum of the subjects, we computed by maximum-likelihood estimation the population mean curve $\mu_{TOT_G}(t) = \sum_{k=1}^{S} \beta_{kG}\phi_k(t)$ for each time $t \in (6, 24)$ and group G, and its 95% pointwise asymptotic confidence intervals. In addition, we performed asymptotic likelihood-based tests (based on the same set of five spline coefficients) to find whether and at which age t there was a significant difference between the mean curve of a specific group G and the mean curve of another group (or with respect to a known reference curve). Within this procedure, we identified three settings: i) when considering $y_{iG}(t)$ as the percentage of predicted value of a parameter (i.e. FVC, FEV1, ΔVAB), we used 100% as its reference value; ii) when considering $y_{iG}(t)$ as one parameter of the breathing pattern in its original unit of measure, we compared the mean curve of the DMD patients to the mean curve of the control population; iii) instead, when considering $y_{iG}(t)$ as one parameter of the spirometric test in its original unit of measure, first we computed the pairwise difference between the ideal and the observed values of each subject; then we constructed the population mean curve of these differences with the model (1.1) and we tested whether and at which age t the resulting mean curve is significantly non-zero.

Finally, for every acquired datum and for every group G, we assessed the significance of the random effects related to the subject with a likelihood ratio test, which compared the full model with the one that excluded all the

random effects. Significance was determined by $p < 0.05$. The proposed model was implemented in R (version 3.2.3) with the package `lme4`.

1.2 Results

1.2.1 Patients

In Table 1.1, anthropometric and clinical data of the 115 enrolled DMD patients are reported for each age. At the age of 13 yrs all patients but one were wheelchair bound. More than one third of the patients older than 14 yrs were using cough assistive-devices. In the study group, 28 patients were regularly using nocturnal NIMV (since a mean age of 19.1 ± 3.5). The youngest patient using NIMV was 15.6 years old. Scoliosis worsened with age. Steroid treatment was prevalent among younger patients. In the course of the study, 9 patients died, 4 due to respiratory and/or swallowing insufficiency (mean age 16.7 ± 2.3) and 5 to cardiac failure (mean age 17.6 ± 4.5). Median (interquartile range) of the control group age was 16.1 (7.7-22.7) yrs, of the height was 1.7 (1.3-1.8) m, of the weight 60 (29-75) kg and of the body mass index 19.6 (16-23) $\mathrm{kg \cdot m^{-2}}$.

1.2.2 Spirometry, lung volumes and nocturnal oxygen saturation

Evolutions with age of FVC, FEV1 and PEF are shown in Figure 1.1. These parameters, when expressed as percentage of the predicted values, linearly declined by 4.6%/yr (age range 11-22), 5.4%/yr (age range 11-22) and 3.8%/yr (age range 6-20), respectively. All absolute and predicted values were significantly reduced after the age of 7. The reduction of FVC was due to both inspiratory capacity and expiratory reserve volume, being significantly lower than predicted after the age of 6 and 8, respectively (Figure 1.2). The reduction of TLC and FRC began at the age of 7.7 and

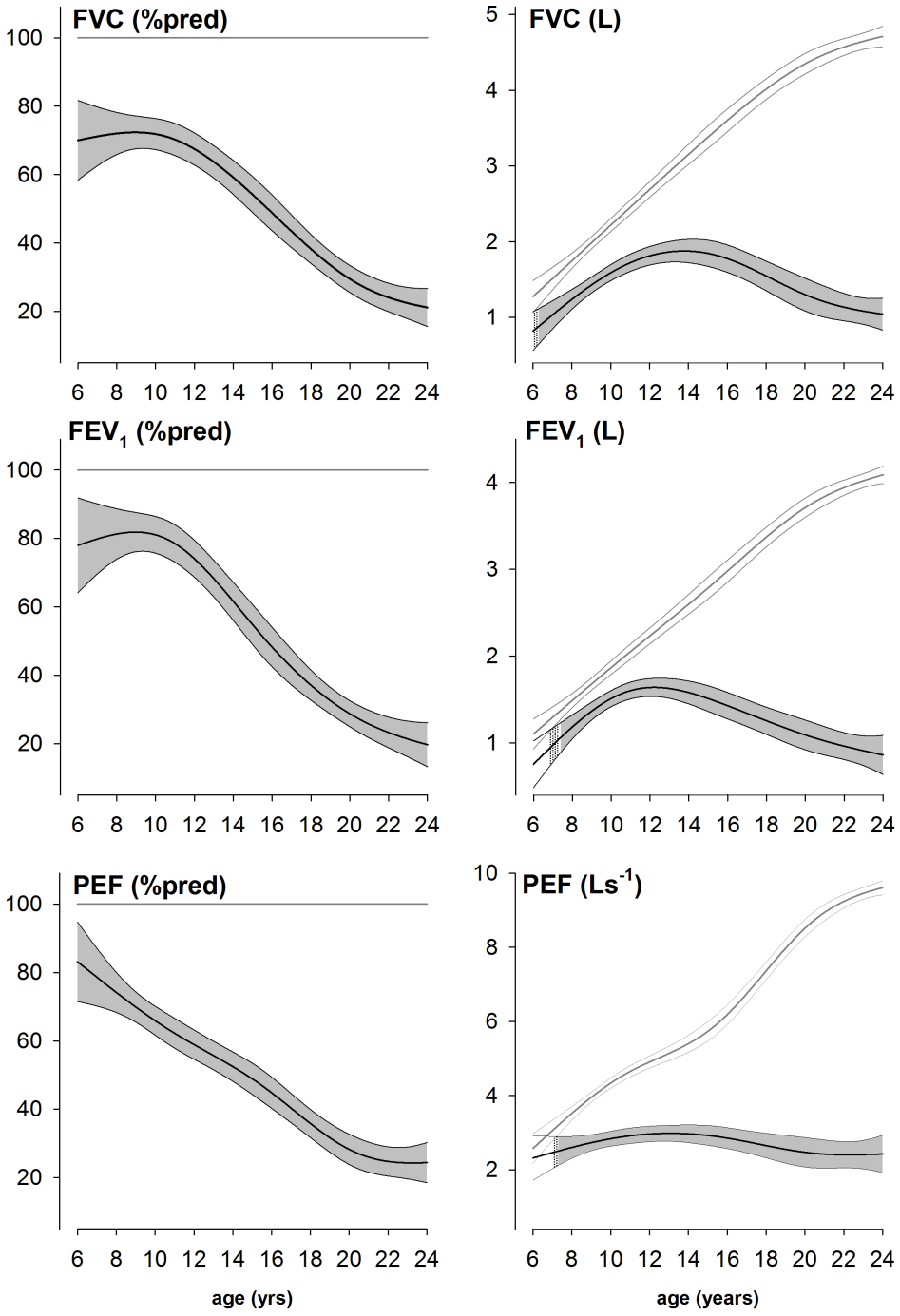


Figure 1.1: Evolution with age of the maximum-likelihood estimation population mean curve (thick line) and its 95% pointwise asymptotic confidence intervals (thinner lines) of forced vital capacity (FVC), forced expiratory volume in one second (FEV1) and peak expiratory flow (PEF) expressed as percentage predicted (A, C, E, respectively) and absolute values (B, D, F, respectively) in DMD patients. Black lines: measured values; grey lines: predicted values. Dotted areas: values significantly different from predicted with $0.01 < p < 0.05$. Grey areas: values significantly different from predicted with $p < 0.01$.

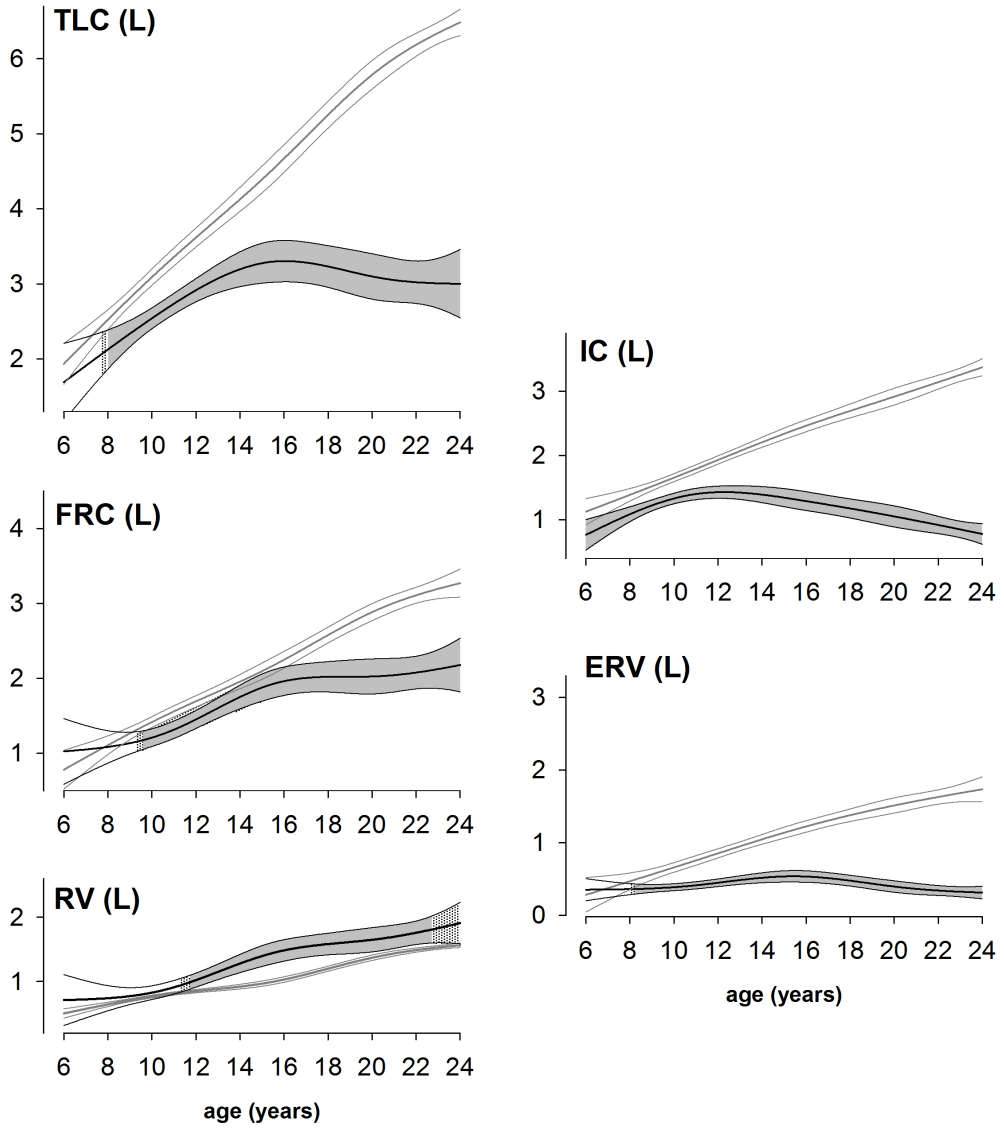


Figure 1.2: Evolution with age of the maximum-likelihood estimation population mean curve (thick line) and its 95% pointwise asymptotic confidence intervals (thinner lines) of total lung capacity (TLC, A), functional residual capacity (FRC, B), residual volume (RV, C), inspiratory capacity (IC, D) and expiratory reserve volume (ERV, E) expressed as absolute values in DMD patients. Black lines: measured values; grey lines: predicted values. Dotted areas: values significantly different from predicted with $0.01 < p < 0.05$. Grey areas: values significantly different from predicted with $p < 0.01$.

Age (yrs)	Visits (N)	Pts (N)	Weight (Kg)	Height (cm)	Wheelchair bound pts (%)	Pts using CAD (%)	Pts under NIV (%)	Current steroid users (%)	**Pts with scoliosis (%)**				
									NU	**M**	**MO**	**S**	**SF**
6	23	23	21.1 (20.6-24.9)	116.5 (112.6-121.5)	9	0	0	78	100	0	0	0	0
7	31	28	25.0 (22.5-30.5)	121.0 (118.5-125.0)	11	4	0	79	96	4	0	0	0
8	28	25	29.0 (25.2-35.0)	127.0 (122.0-131.0)	24	12	0	80	88	12	0	0	0
9	31	28	31.0 (27.0-36.0)	132.0 (129.8-135.5)	43	4	4	68	79	14	4	4	0
10	33	28	37.0 (31.5-45.0)	140.0 (134.0-142.0)	68	11	4	50	54	32	11	4	0
11	37	35	45.0 (38.0-51.5)	144.0 (138.0-149.0)	80	17	0	49	40	34	20	6	0
12	39	34	49.0 (38.0-57.0)	151.0 (138.0-155.5)	85	24	0	26	29	32	18	18	3
13	41	35	54.5 (44.0-66.5)	155.0 (148.0-158.5)	97	29	0	20	26	29	17	23	6
14	42	35	55.0 (50.0-65.0)	160.0 (155.0-165.0)	97	31	3	9	14	17	20	26	23
15	35	27	57.0 (48.3-65.0)	162.0 (157.0-165.0)	96	41	11	4	4	26	33	11	26
16	34	28	56.5 (42.0-68.8)	164.0 (159.3-167.3)	96	57	21	4	7	18	29	25	21
17	39	26	60.0 (47.5-70.0)	165.0 (160.5-168.0)	96	54	27	4	8	19	31	12	31
18	37	26	53.0 (44.0-63.0)	165.0 (160.0-167.0)	100	46	35	4	8	12	23	27	31
19	27	20	58.0 (45.3-73.5)	166.0 (162.8-171.0)	100	60	35	5	10	10	25	35	20
20	37	24	58.0 (46.0-67.9)	165.5 (164.0-168.3)	100	54	33	4	4	17	13	50	17
21	26	18	57.0 (50.0-72.3)	166.5 (165.0-170.3)	100	33	17	6	6	11	17	56	11
22	27	18	54.0 (46.1-64.0)	165.0 (164.0-168.8)	100	83	50	6	0	17	22	44	17
23	19	15	58.0 (42.0-78.0)	167.0 (164.5-176.5)	100	73	33	0	0	27	27	47	0
24	11	8	64.5 (46.5-78.8)	174.5 (165.0-177.3)	100	75	63	0	0	13	38	38	13

Table 1.1: Anthropometric and clinical data of DMD patients. CAD: cough assisted device; NIV: non-invasive ventilation; NU: null; M: mild; MO: moderate; S: severe; SF: spinal fusion.

9.3, respectively, while RV became greater than predicted after the age of 11.3 (Figure 1.2).

In Figure 1.3 the mean curves of nocturnal saturation data are reported in terms of night time spent with SpO2 in the ranges 95-100%, 90-94% and $< 90\%$, number and average time of desaturation events. Peripheral oxygenation during sleep worsened with age. Desaturation events were already present in childhood.

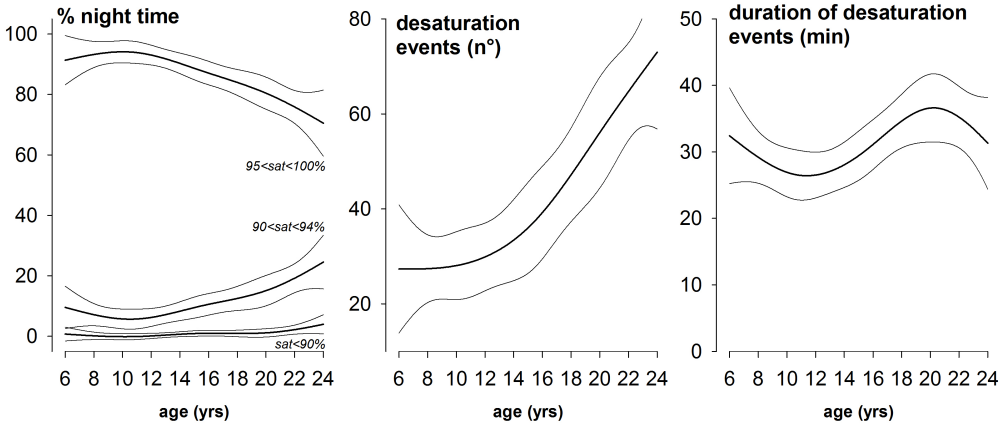


Figure 1.3: Evolution with age of the maximum-likelihood estimation population mean curve (thick line) and its 95% pointwise asymptotic confidence intervals (thinner lines) of the night time spent with nocturnal saturation in the ranges 95-100%, 90-94% and $< 90\%$ (A), of the number (B) and the average time (C) of desaturation events in DMD patients.

1.2.3 Spontaneous breathing pattern at rest

The evolution curves of minute ventilation, RSBi and their two components (i.e., RR and VT) are shown in Figure 1.4. When compared to healthy peers, DMD patients started to hypoventilate at the age of 18.1 because of a reduced VT (after the age of 17.2), being RR similar between the two groups almost all throughout the considered age span. RR became significantly higher than normal after 22.1 yrs and rapid and shallow breathing occurred after the age of 20.7 years. When normalized to actual body weight, tidal volume was similar between DMD and healthy peers in the whole considered age range. The mean curves of ΔVRC and ΔVAB,

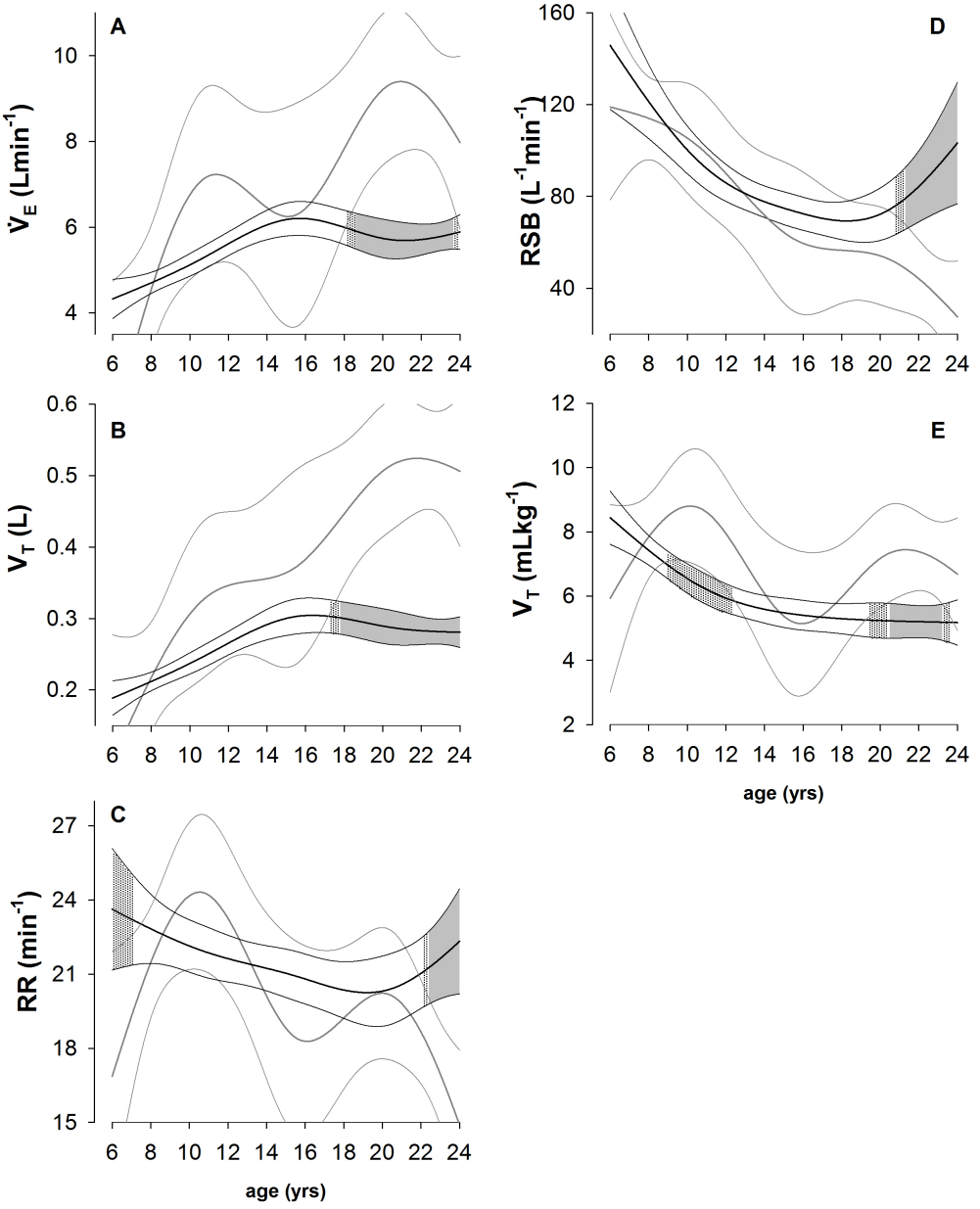


Figure 1.4: Evolution with age of the maximum-likelihood estimation population mean curve (thick line) and its 95% pointwise asymptotic confidence intervals (thinner lines) of minute ventilation ($\dot{V}_E$, A), tidal volume (V_T, B), respiratory rate (RR, C), rapid and shallow breathing index (RSB, D), and tidal volume normalized to body weight (E), during spontaneous breathing in supine position in DMD patients (black lines) and healthy controls (grey lines). Dotted areas: values significantly different from predicted with $0.01 < p < 0.05$. Grey areas: values significantly different from predicted with $p < 0.01$.

both expressed in liters and as %VT, are shown in Figure 1.5. When ex-

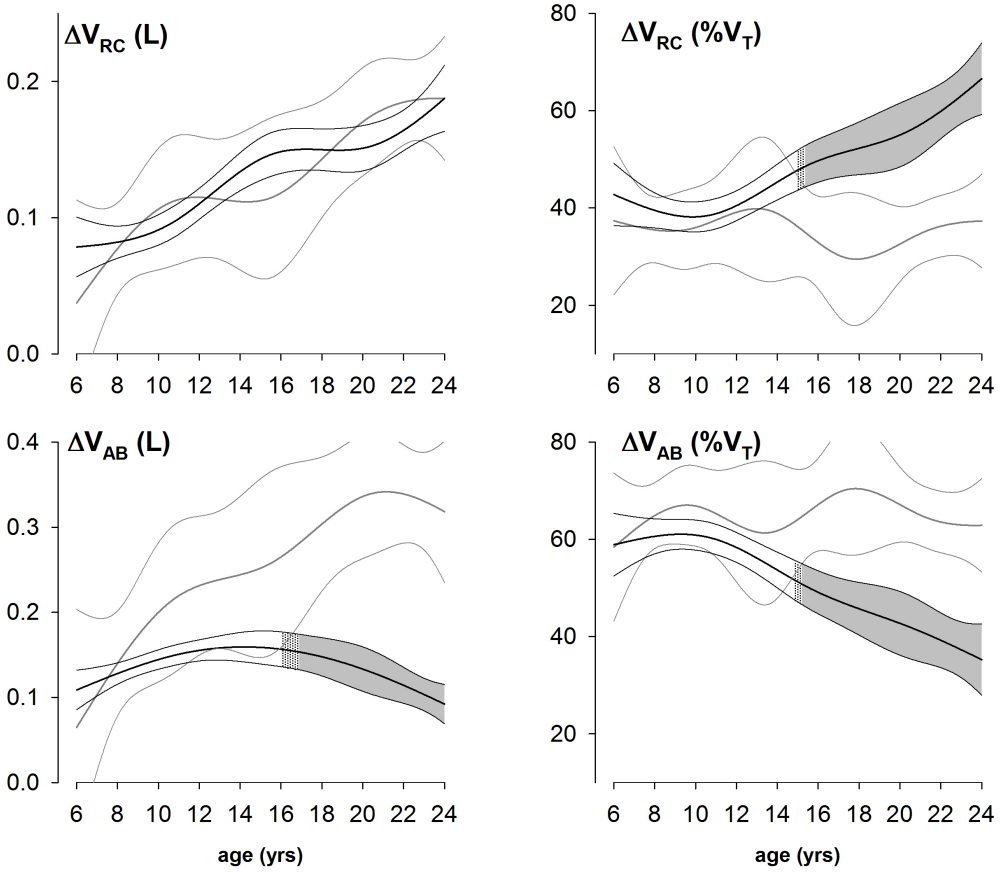


Figure 1.5: Evolution with age of the maximum-likelihood estimation population mean curve (thick line) and its 95% pointwise asymptotic confidence intervals (thinner lines) of ribcage volume variations (ΔVRC) and abdominal volume variation (ΔVAB) expressed in liters (A and C, respectively) and as percentage contribution to tidal volume (B and D, respectively) during spontaneous breathing in supine position in DMD patients (black lines) and in healthy controls (grey lines). Dotted areas: values significantly different from predicted with $0.01 < p < 0.05$. Grey areas: values significantly different from predicted with $p < 0.01$.

pressed in liters, ΔVRC was similar between DMD and healthy controls at all ages, whereas ΔVAB became lower in DMD after the age of 16 yrs. Thoraco-abdominal volume variations in DMD, when instead expressed as %VT, showed a pattern starting to be significantly different than normal after the age of 14.8 years, when ribcage and abdominal contribution became significantly greater and lower than normal. The predominance of rib cage compared to abdomen in the relative contribution to tidal volume progressively increased with age. Similarly, to the ribcage considered as a

whole, the expansion of both pulmonary and abdominal ribcage in DMD was almost similar to healthy subjects. The expansion of the abdominal ribcage became lower than controls starting from the age of 22. The percentage contribution of pulmonary and abdominal ribcage started to be significantly higher than the control group at the age of 13.2 yrs and 19 yrs, respectively.

1.2.4 Effect of scoliosis, steroids and NIMV

The mean curves of FVC (%pred) and ΔVAB (%VT), grouped by the severity of scoliosis, are shown in Figure 1.6. Both parameters were neg-

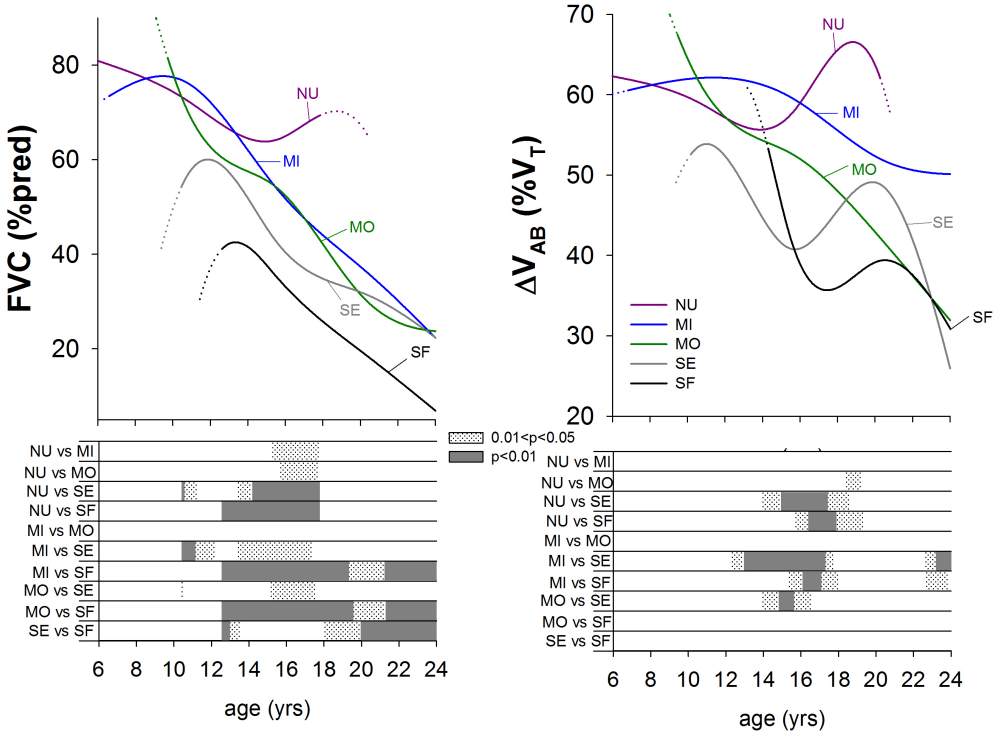


Figure 1.6: Evolution with age of the maximum-likelihood estimation population mean curve of the percentage predicted forced vital capacity (A) and percentage abdominal contribution to tidal volume during spontaneous breathing in supine position (B) grouped by scoliosis classification (NU: null; MI: mild; MO: moderate; SE: severe; SF: spinal fusion). Data are reported in the age range in which the asymptotic confidence intervals is ¡40%. Bottom panels: pairwise comparisons of the population means between all categories of scoliosis (Dotted areas: values significantly different from predicted with $0.01 < p < 0.05$. Grey areas: values significantly different from predicted with $p < 0.01$.)

atively associated with the severity of scoliosis, with patients with severe

scoliosis and spinal fusion showing the lowest values.

The mean curves of FVC (%pred) and ΔVAB (%VT), in the different steroid therapy groups, are shown in Figure 1.7. In currently steroid-

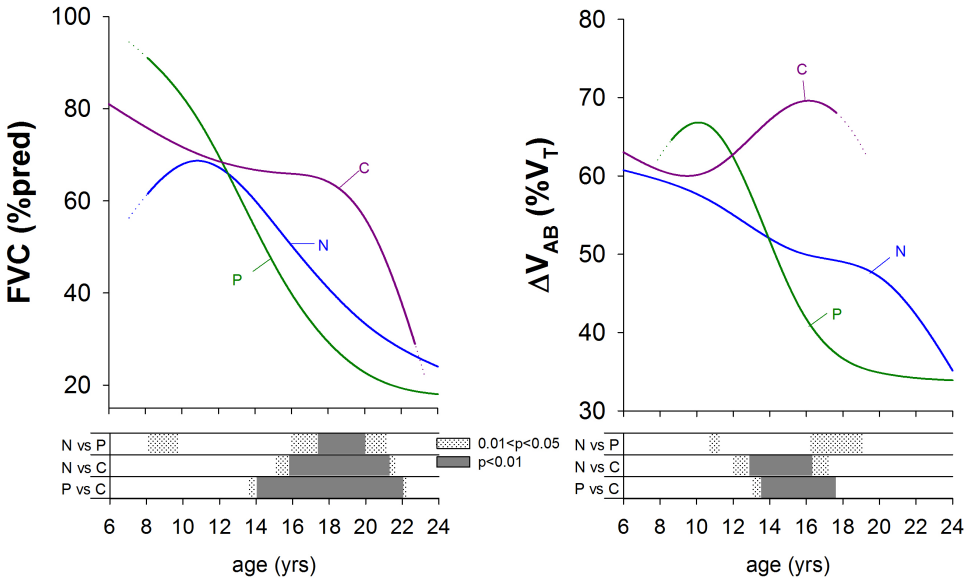


Figure 1.7: Evolution with age of the maximum-likelihood estimation population mean curve of the percentage predicted forced vital capacity (A) and percentage abdominal contribution to tidal volume during spontaneous breathing in supine position (B) grouped by corticosteroids treatment groups (C: currently under treatment; P: "past", i.e. treated in the past for at least one year and not currently receiving steroids; N: naïve, i.e. never treated with steroids). Data are reported in the age range in which the asymptotic confidence intervals is ¡40%. Bottom panels: pairwise comparisons of the population means between all categories of steroid subgroups (Dotted areas: values significantly different from predicted with $0.01 < p < 0.05$. Grey areas: values significantly different from predicted with $p < 0.01$.)

treated patients, FVC (%pred) was significantly higher than naive and past patients in the age range 15.1-21.3 and ΔVAB (%VT) in the age range 13-17.3.

The mean curves of FVC (%pred) and ΔVAB (%VT), grouped by the NIMV use, are shown in Figure 1.8. Compared to no-NIMV group, patients belonging to pre-NIMV group were characterized by lower values of both FVC (%pred) and ΔVAB (%VT) from the age of 13 years. After the start of NIMV, FVC (%pred) and ΔVAB (%VT) in treated (NIMV group) and untreated (no-NIMV) patients remained significantly different until 21.1 and 18.4 years, respectively.

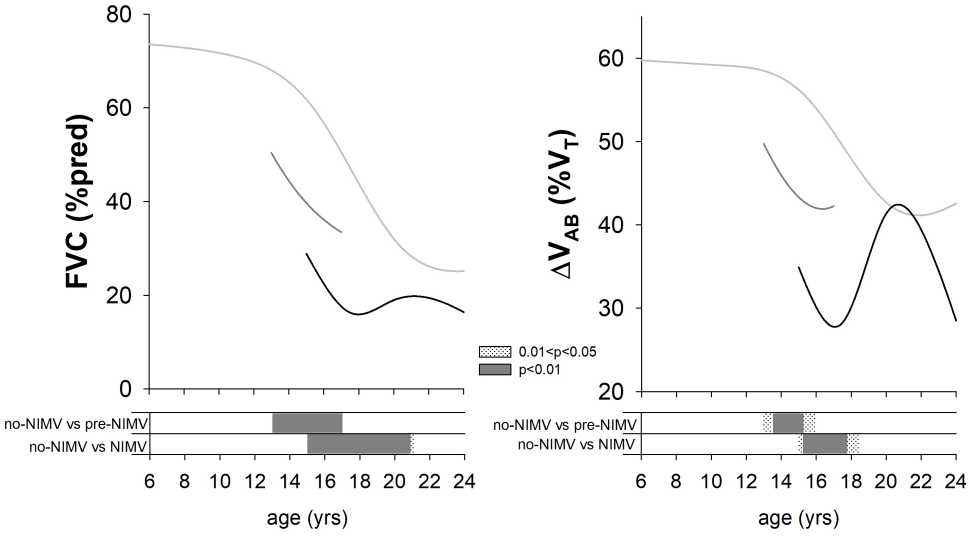


Figure 1.8: Evolution with age of the maximum-likelihood estimation population mean curve of the percentage predicted forced vital capacity (A) and percentage abdominal contribution to tidal volume during spontaneous breathing in supine position (B) grouped by nocturnal non-invasive mechanical ventilation (NIMV) groups (NIMV: currently under treatment; pre-NIMV: patients who underwent NIMV treatment within the period of the study. The curve refers to the visits before the start of NIMV; no-NIMV: patients never treated with NIMV within the period of the study). Pre-NIMV data are reported in the range that contains at least six visits in each age. Bottom panels: pairwise comparisons of the population means between pre-NIMV and no-NIMV and between NIMV and no-NIMV (Dotted areas: values significantly different from predicted with $0.01 < p < 0.05$. Grey areas: values significantly different from predicted with $p < 0.01$.).

1.3 Discussion

In the present study an original and comprehensive description of the evolution of respiratory function in patients with DMD over the age span 6-24 years is provided. Non-invasive measurements were taken on a group of 115 subjects during 574 observations. The main result is that a detailed analysis of breathing pattern is able to provide information regarding specific key time points during the natural course of the disease. While predicted values of FVC, FEV1 and PEF decline since childhood, during spontaneous breathing the following parameters become significantly different than normal in sequential order: contribution of abdominal compartment to tidal volume (lower after 12.3 yrs), tidal volume (lower after 17.3 yrs), minute ventilation (lower after 18.1 yrs) and respiratory rate (higher after 22 yrs). The progressive decline of minute ventilation observed during awake spontaneous breathing at rest, that becomes significant after the age of 18.1, is due to progressively reduced tidal volume caused, in turn, by a decreased abdominal expansion with ribcage volume variations similar to normal all over the considered age span. As a result, the percentage contributions of ribcage and abdomen to tidal volume become increasingly higher and lower, respectively, after the age of 14.8. It must be noted that, when normalized to weight, the differences in tidal volume are no longer present. This is presumably due to the pathological thinness that can develop in adulthood, rather than to a preserved ventilatory pump.

Another novel feature provided by the present study is that both spirometry and lung volumes have been longitudinally measured and analyzed in a large cohort of patients. This allowed assessing the determinants of the decay of FVC, namely the relevant reduction of TLC, IC and ERV since childhood, suggesting that the impairment of both inspiratory and expiratory muscles as a whole represents a clear hallmark of DMD.

It is important to emphasize that both spirometry and lung volume assessment, which are traditionally used to evaluate respiratory impairment in DMD patients, represent volitional tests that require a high level of patient's cooperation that is not always possible. In addition, these tests do not evaluate the patient under his normal conditions, represented by spontaneous breathing at rest, but only during maximal maneuvers. In addition, they provide global indexes of respiratory system function, without being able to differentiate between decreased respiratory system compliance and/or increased respiratory muscle weakness. Finally, they cannot be specific also in differentiating the relative impairment of each respiratory muscle functional group, namely ribcage muscles, diaphragm and abdominal muscles.

During inspiration, rib cage expansion is the result of the action of all inspiratory rib cage muscles, namely the scalene, the external intercostals, the parasternals and the sternocleidomastoid whilst abdominal expansion is due to the action of the diaphragm. Therefore, although in our DMD patients the rib cage expansion remains within the normal range, it can be supposed that inspiratory rib cage muscles are also impaired, being unable to compensate for the insufficient action of the diaphragm that becomes unable to provide an adequate tidal volume. In this scenario, the impairment of the inspiratory muscles is unbalanced with the diaphragm impacted more substantially than the inspiratory rib cage muscles.

A further evidence of the earlier impairment of the diaphragm with respect to the ribcage muscles derives from the analysis of the abdominal ribcage expansion that becomes significantly lower only at late ages. This is probably due to the fact that this compartment is submitted to both inspiratory ribcage muscles and diaphragm. Therefore, diaphragmatic impairment is masked in the age range 15-22 by the relatively preserved ribcage muscles

becoming evident also in this compartment only after the age of 22.

The diaphragmatic action in DMD can be affected in opposite directions by the presence of scoliosis and by corticosteroid therapy. Regarding scoliosis, we have now demonstrated that the worsening of scoliosis exacerbates not only the restrictive lung pattern, but also the action of the diaphragm as well. The reduced contribution of abdominal compartment to tidal volume with increasing severity is a result somehow surprising, as the main effects were expected to be on the rib cage. However, we can speculate that the presence of scoliosis determines a compression of the abdomen and increases diaphragmatic load, shifting chest wall expansion toward the ribcage. Regarding steroids treatment, our results are in agreement with those reported by several recent studies [13,14,23,47], which demonstrated that currently corticosteroid user patients have better spirometric values than the other two groups. In the present study we report that, in addition, there is a significant effect on abdominal contribution to tidal volume in the age range between 13 and 17.3 years.

A limitation of the study is related to the lack of complete availability of measurements at the extremities of the considered age range of DMD patients, i.e. children younger than 6 and adults older than 24. In these patients, although breathing pattern was easily evaluated by opto-electronic plethysmography, the availability and/or reliability of spirometry and lung volumes was very poor due to the lack of collaboration in the youngest and the presence of macroglossia and/or facial muscular weakness in the oldest. In addition, most of the patients under current steroid treatment were the younger ones, with very few old patients. The reason of this bias is that in the clinical practice steroid treatment is gradually stopped after the loss of ambulation because of the side effects [11, 28, 48].

Nevertheless, the study has several strengths. This is the largest data

set to date including serial spirometry, lung volumes, nocturnal oxygen saturation and breathing pattern assessment, with data collected at different time points for each subject in the interval 6-24 years. Although this did not allow a direct time-matched data comparison, we dealt with this irregular and subject-specific timing by adopting a regression model based on natural cubic splines with mixed effects. Similar models have been employed in the longitudinal data analysis literature and used in a wide range of applications [49, 50]. Our proposed regression model allows to obtain estimated values at any time point of our domain taking into account the possible effects of scoliosis and steroids (i.e. fixed effects), the specific temporal evolution of each subject (i.e. random effects), and the session-specific measurement errors (i.e. error term).

The study has clinical implications. Firstly, we have shown that after the age of 14.8 yrs the diaphragm shows early evidence of weakness, starting to be unable to contribute to an adequate tidal volume and minute ventilation. Noteworthy, in the same period the percentage of patients under NIMV becomes significant (Table 1). Moreover, we have shown that those patients who started NIMV during the period of the study were characterized by a faster decline of FVC (%pred) and ΔVAB (%VT). After the start of NIMV, FVC (%pred) and ΔVAB (%VT) continued to decline for about 3 years, then these parameters increased until the age of 21 years, and then declined again, similarly to untreated patients.

On the basis of these results, we believe that the specific involvement of the diaphragm should be considered in addition to the FVC decline for the definition of treatment guidelines, including the timing of starting NIMV [9–12]. Our results, therefore, suggest that technological developments should be addressed in order to provide simple but accurate measurement of abdominal kinematics during spontaneous breathing that

could be available in all clinical centers.

In addition, we have shown original evidence that scoliosis represents a burden to the dystrophic diaphragm of these patients and that steroid treatment has efficacy on spirometry and diaphragm's contribution to tidal volume, supporting the relevance of steroid treatment as part of DMD patients' care.

In conclusion, the evolution curves of spirometry, lung volumes and breathing pattern parameters here presented might be considered to better define outcome measures in the forthcoming clinical trials both in non-ambulant and ambulant DMD patients.

Chapter 2

Diagnosis of Subtypes with Diffusion Magnetic Resonance Imaging in Sporadic Creutzfeldt-Jakob Disease

Early and accurate diagnosis is increasingly recognised as critical to achieve an effective treatment of neurodegenerative diseases. This notion applies especially to prion diseases that often have a rapid course. MRI has recently gained a high degree of reliability in the diagnosis of prion diseases with the introduction of diffusion-weighted imaging (DWI) in the diagnostic criteria, as well as the increasing experience acquired by image readers despite the rarity of prion diseases [51–53]. Diagnostic sensitivity and specificity values of over 90% have recently been reported [51, 54–56]. Another recent test, the Real-Time Quaking-Induced-Conversion (RT-QuIC), has attained high degree of sensitivity and specificity in the diagnosis of prion diseases based on the detection of minute amounts of disease-associated prion protein (PrPD) in body fluids and olfactory epithelium non-invasively obtained [57,58]. However, magnetic resonance imaging (MRI) remains the

test of choice at the initial clinical evaluation when the diagnosis is open and requires timely identification of treatable conditions [53].

The clinical diagnosis of prion diseases is further compounded by the diversity of their phenotypes [56,59,60]. Sporadic Creutzfeldt-Jakob diseases (sCJD) alone comprises five distinct clinical and histopathological phenotypes or subtypes, which are associated with distinctive pairings of the genotypes at the methionine (M) and valine (V) polymorphic codon 129 (MM, MV, and VV) with the PrPD types 1 and 2. The five histopathological phenotypes are characterised by the anatomical distribution and/or type of the spongiform change, distribution and shape of PrPD deposits and severity of astrogliosis and loss of neurons [61].

The distinction of sCJD subtypes is clinically relevant not only because individual subtypes may significantly differ in disease duration, but also because they may respond differently to drug treatment, once effective treatments will be developed. This suggests that treatment may need to be tailored to individual sCJD subtypes.

However, the only way to have a confirmation of the sCJD subtype is by tissue examination either at biopsy or, more definitely, at autopsy. Until recently, no clinical test had been shown to reliably diagnose sCJD subtype, with the only exception of the RT-QuIC test, that has been reported to distinguish MM and VV sCJD subtypes when the prion protein 129 genotype is known [62]. Regarding the use of imaging biomarkers, few attempts have been made to distinguish sCJD subtypes on MRI examination [63,64]. Although significant differences were observed between some sCJD subtypes, no comprehensive attempt was made to adapt these findings to the clinical diagnosis. Morover, the resuls of those works on the subtype identification were based on a classification scheme of 1999 [61], while recently a revision of the classification of sCJD histopathological phenotypes has been

provided [60]. Establishing in vivo biomarkers that could predict sCJD subtype would thus be relevant for patient management and clinical trials. MRI is an ideal candidate for providing such biomarker in vivo because it evaluates the whole brain rather than just a few small specimens, as neuropathology does. MRI has great potential to be established as the method of choice to assess regional extension of brain lesions in the early stage of prion diseases. Among different MRI modalities, diffusion MRI (dMRI) has been shown the most sensitive to peculiar brain microstructural alterations such as spongiosis and PrPD deposition. dMRI measures microstructure by using the random walk of water molecules as a probe. Entrapment of water within vacuoles and/or slowing of water diffusivity in the interstitium enriched with prion protein aggregates is likely responsible for the signal hyperintensity detected with diffusion-weighted imaging (DWI) in affected regions of the brain [65, 66].

In this study, variations of the MRI signal were rated and compared over 12 brain anatomical regions in a cohort of patients in which the histological, genetic and PrPD examinations allowed for the definitive diagnoses of prion diseases as for form, type and subtype of prion disease. We report on the diagnostic accuracy and lesion distribution of the MRI examinations carried out on 306 patients with sCJD. One hundred twenty three patients in which prion disease was definitively rule out were used as controls. The MRI rating profiles generated with this procedure revealed significant differences which enable the diagnosis of individual sCJD subtypes.

2.1 Materials and methods

2.1.1 Study design

We considered a retrospective and a prospective study for this analysis. The cohort of subjectsd́ata belonging to the retrospective study was

collected by the National Prion Disease Pathology Surveillance Center (NPDPSC) of Cleveland (Ohio), that received the biological material (i.e. encephalon, CSF, blood) of patients to diagnose the prion disease for these subjects. Subsequently a MRI was requested and it was sent to a neuroradiologist (AB) for a consultation. Blindly to the preliminary diagnosis, AB reviewed and assessed the MRI. The prospective study is composed of all the suspected cases of prion disease, for which the doctor of the patient asked the NPDPSC for consultation to confirm the diagnosis. Thus, a MRI was sent to the Center and, often, also biological material while the patient was still alive. Then the NPDPSC tried to retrieve the encephalon after the patient's death to make the brain cutting and the definite diagnosis on tissue with neuropathological examination and Western Blot technique. In this study arm there is the possibility of not having a confirmed diagnosis (positive or negative) on a relatively high number of subjects (about 30% in our cohort).

Patients eligible for inclusion were subjects with suspected prion disease, having at least one brain MRI collected by the NPDPSC. For both the retrospective and the prospective study, we excluded patients if they did not have at least a DWI (or ADC) of the brain and if they did not have a confirmed diagnosis of sCJD or a confirmed diagnosis of non prion disease, made through a neuropathological examination. Here, the subtype identification is based on a recent classification of sCJD histopathological phenotypes [60].

2.1.2 Test methods

We proposed a semi-quantitative index test to grade the hyperintensities of 12 brain regions of the DWI magnetic resonance images collected in

our study on a four-point (i.e., 0-3) ordinal scale. This proposed index test, used to classify the levels of hyperintensities of the DWI signal, is as follows: a score "zero" is assigned to a brain region where clearly there are no hyperintensities; a score "one" is assigned to a brain region where the presence of a hyperintensity is questionable; a score "two" is assigned to a brain region where the signal is obviously hyperintense in a restricted area of the region considered; a score "three" is assigned to a brain region where the hyperintensity is clearly present and it is spread over a large part of the brain region considered. A neuroradiologist (AB) blind to the subject diagnosis was enrolled to score all the MRI of our study, according to this index test. The ordered set of the scores given to the MRI of a subject defines the lesion profile that describes the visual assessment of the hyperintensities pattern of that subject. As reference standards in prion disease diagnosis, we considered the measurements of 14-3-3 and tau proteins concentration in the CSF, since these two tests are quite often assessed in the common clinical practice to support the diagnose of sCJD. We also considered the recently introduced second generation Real-Time Quaking-Induced Conversion (RT-QuIC) technique, that has been reported having high sensitivity of detecting prion seeds in CSF specimens from CJD patient [67]. Our proposed index test is considered positive when at least one of the 12 regions is graded with a score greater than or equal to "two", otherwise negative. The tau protein test is considered positive when the measurement is greater than 1,150 pg/mL, otherwise negative. The 14-3-3 protein test is qualitative and identifies three diagnostic categories of positive, negative, and ambiguous outcomes. We adopted the same decision criteria described in [68] to classify the results of both tau and 14-3-3 protein tests. Moreover, we considered RT-QuIC responses as positive or negative following the criteria given in [69].

2.1.3 Statistical analysis

The MRI of 797 patients were blindly reviewed and scored by one neuroradiologist (AB) according to the index test proposed in the previous section. Four hundred twenty nine patients with a definitive diagnosis and diffusion weighted MR imaging were retained for the study: 306 patients with the diagnosis of sCJD and 123 with a non prion disease diagnosis were pathologically confirmed. Demographics of the sCJD patients are reported in Table 2.1.

Subtype	MRI	FN	Age at death (years) Median (Range)	Disease duration (months) Mean (SD)	Path. proven
MM1	76	4	65 (39-91)	3.5 (3.1)	108
MM2C	28	1	63 (46-89)	14.7 (11.4)	29
MM1+2C	23	0	64 (45-86)	8.3 (12.1)	37
MV1	20	3	68 (49-88)	8 (7.8)	20
MV2C	15	0	68 (46-89)	17.6 (13.4)	15
MV2K	23	1	63 (47-85)	12.4 (7.4)	23
MV2C+2K	14	1	64 (56-82)	12.1 (7.4)	14
MV1+2C	11	1	69 (57-77)	16.6 (11.1)	11
MV1+2K	8	0	62 (54-75)	13.7 (5.1)	9
MV1+2C+2K	3	0	62 (62-63)	11.7 (4.8)	3
VV1	16	1	60 (35-88)	9 (6.5)	16
VV2	55	6	65 (45-80)	5.6 (2.3)	55
VV1+2	4	1	63 (60-72)	7 (4)	4
sCJD-nos	10	0	72 (64-80)	8.8 (10.9)	11

Table 2.1: Demographics for the sCJD patients involved in our study, divided by subtype. "MRI": number of cases with available DWI or ADC; "FN": number of false negative MRI; "Path. proven": number of cases with confirmed diagnosis (some of them could not have an available MRI, see the MRI column); "sCJD-nos": sCJD not otherwise specified, meaning that a definite diagnosis of sCJD is pathologically confirmed, but it is not yet available the subtype classification.

We assessed the validity and reliability of our proposed index test on these data calculating its diagnostic reliability (i.e. computing accuracy, sensitivity, specificity) and its inter-rater reliability. As a further step, we selected the 217 patients with positive MRI and pure sCJD subtype (e.g. MM1, VV2 and not MM1+2, VV1+2), and pairwise compared their mean lesion profiles to find the brain regions that mainly characterise each subtype. Finally, based on the individual MRI profile scores of these 217 patients with pure sCJD subtype, we built a classifier to assist the radiologist/neurologist in predicting the pure sCJD subtype while assessing the

MRI of a new patient. Further details about the methodology are given in the following lines.

First, the comparative analysis of MRI diagnostic reliability (i.e., comparison of accuracy, sensitivity and specificity) with those of CSF tests used in routine diagnosis of prion diseases was carried out in a subset of 215 patients (170 sCJD and 45 non prion patients) that had undergone MRI, 14-3-3 and tau tests. Since 14-3-3 test can give an ambiguous outcome, we assessed the accuracy, the sensitivity and the specificity of this test considering two scenarios: i) the ambiguous cases were considered as false negatives or false positives respectively in the CJD group and in the non prion group, obtaining a lower bound for accuracy, sensitivity and specificity; ii) the ambiguous cases were considered as true negative or true positive respectively in the non prion group and in the CJD group, obtaining an upper bound for accuracy, sensitivity and specificity. This choice of calculating an interval estimation of the diagnostic statistics allowed us not to discard many observations from the comparison (61 over 215 [28%]), and at the same time including the uncertainty of this test (i.e. the ambiguous outcome). We compared also the diagnostic accuracy of the DWI scoring and RT-QuIC test on a subgroup of 39 patients (33 sCJD and 6 non prion) having both exams available. The relatively small number of patients with second generation RT-QuIC is due to its introduction in the clinical practice only in the last few years. We computed for each diagnostic statistic (i.e., accuracy, sensitivity, specificity) an exact binomial 95% confidence interval using the method of Clopper and Pearson [70, 71], for which the nominal coverage is guaranteed since its construction is based directly on the binomial distribution.

Second, in order to assess the inter-rater reliability (IRR) of this MRI scoring system, other two neuroradiologists (MG, RL) blind to the subject

diagnosis were enrolled to score the images collected in our study, according to the 4-point scoring system described above. In total, a common pool of 128 MR images (80 sCJD and 48 non prion patients) were scored by all the three raters. We assessed the IRR using a two-way mixed, consistency, average-measures intra-class correlation (ICC) coefficient [72] to assess the degree that raters provided consistency in their ratings across subjects, being the ICC one of the most commonly-used statistics appropriate to assess IRR for ordinal scales [73]. Following the cutoffs given by [74], we considered the IRR being poor for ICC values less than 0.40, fair for values between 0.40 and 0.59, good for values between 0.60 and 0.74, and excellent for values between 0.75 and 1.0.

Third, we selected 233 MRI examined patients who harbored a PrPD type pure sCJD subtype (e.g. MM1, VV2 and not MM1+2, VV1+2) and we analysed and pairwise compared the profiles of these groups - excluding the 16 patients with a false negative MRI result - to identify the discriminant brain regions with significantly different scores. We performed the analysis at two levels: with the permutational Hotelling T-squared test for two samples, we first tested for the global difference between the mean scores of two given subtypes (e.g. MM1 vs. VV2); this test is the multivariate extension of the t-test for two independent samples, adopted when the number of response variables are at least two (e.g. in our case, the 12 brain region scores). Since the assumptions of normally distributed data are not met in our case (i.e. we have ordinal data on the scale 0-3), we used a permutational version of the standard T-squared test to obtain a non-parametric inference tool. Subsequently, consistently with the previous hypothesis test, we performed a non-parametric univariate test (i.e. Wilcoxon rank-sum test for ordinal data) for each of the 12 brain regions scored, to find the areas that exhibited the most significant differences be-

tween the two groups. We adjusted the p-values of these univariate tests for multiple comparison with the Holm-Sidak correction, that provides a conservative familywise error control [75, 76] when the tests are not independent, as in our case where we compared the same group multiple times (e.g. MM1 vs VV2, MM1 vs. MV1).

Finally, to predict the subtype of a new sCJD patient with MRI lesion profile, we adopted a method for tree-based classification known as CART (i.e., Classification And Regression Tree [77]) that provides a classification algorithm in the form of a decision tree without imposing a specific assumption on the underlying distribution of the data. We used the Gini impurity index to choose at each step the candidate splits used to grow the tree, and we pruned the fully grown tree such that the predicted classification error computed by a leave-one-out cross-validation procedure was minimised.

2.2 Results

2.2.1 Diagnostic reliability

Three neuroradiologists agreed on the final diagnosis of prion disease on 112 of 128 (87.5%) patients. Remarkably the agreed examinations included six false negative but no false positive. Among the 16 patients with discordant diagnosis there were 12 sCJD patients and 4 non prion subjects. The inter-rater reliability (IRR) of the scoring system on DWI was excellent for most of the regions, with ICC values greater than 0.75 for the striatum, thalamus, neocortex and part of the limbic system (and greater than 0.9 for precuneus, parietal caudate and putamen); ICC was good only for the cerebellum and for the hippocampus (see Table 2.2).

The first MRI exam of 429 patients (306 sCJD and 123 non-prion subjects) revealed 0.93 diagnostic accuracy, 0.91 sensitivity and 0.98 speci-

Brain region	ICC	IRR rating
Frontal	0.88 (0.84-0.91)	Excellent
Temporal	0.87 (0.80-0.92)	Excellent
Precuneus	0.91 (0.88-0.94)	Excellent
Parietal	0.93 (0.90-0.95)	Excellent
Occipital	0.80 (0.72-0.86)	Excellent
Cingulate	0.84 (0.78-0.88)	Excellent
Insula	0.84 (0.78-0.88)	Excellent
Hippocampus	0.60 (0.47-0.71)	Good
Caudate	0.94 (0.92-0.96)	Excellent
Putamen	0.93 (0.90-0.95)	Excellent
Thalamus	0.85 (0.79-0.89)	Excellent
Cerebellum	0.68 (0.56-0.77)	Good

Table 2.2: ICC for the 12 brain regions scored by the three raters (95% confident interval in brackets), with a qualitative rating of the IRR that follows Cicchetti (1994).

ficity. In seven patients the MRI became positive in a follow-up examination, therefore when considering all MRI examinations acquired for each patient, the accuracy and sensitivity increased to 0.95 and 0.94, respectively, while specificity remained unchanged. At first MRI examinations, the prospective arm of the study (218 patients) showed 0.91 accuracy, 0.89 sensitivity and 0.98 specificity, while the retrospective study (211 patients) showed 0.96 accuracy, 0.95 sensitivity and 0.97 specificity.

Analyses of 215 patients, including 170 and 45 subjects with confirmed diagnoses of sCJD and non prion disease respectively, pointed to MRI as the most reliable test with 0.94 accuracy, 0.94 sensitivity and 0.98 specificity, followed by tau in CSF with 0.80, 0.84 and 0.67, respectively while 14-3-3 results were even lower (Table 2.3).

	MRI	14-3-3	Tau
Accuracy	0.94 (0.90-0.97)	0.76 (0.68-0.82) (0.55-0.82)	0.80 (0.74-0.85)
Sensitivity	0.94 (0.89-0.97)	0.87 (0.80-0.93) (0.65-0.91)	0.84 (0.77-0.89)
Specificity	0.98 (0.88-1)	0.24 (0.10-0.43) (0.16-0.51)	0.67 (0.51-0.80)

Table 2.3: Comparison of the diagnostic statistics for the DWI scoring system, 14-3-3 and tau tests, evaluated on a common pool of 215 subjects (of which 170 sCJD patients). For each statistic, a 95% confidence interval is computed using the method of Clopper and Pearson (Newcombe, 1998c). For the 14-3-3 protein test, it is provided also another interval to take into account the presence of "ambiguous" results, as stated in the previous section.

Diagnostic reliability was also compared with that of second generation RT-QuIC in 33 sCJD cases and 6 non prion disease patients. In this limited

population, MRI accuracy, sensitivity and specificity were 0.92, 0.91 and 1.0 and 0.80, 0.76 and 1.0 for RT-QuIC (Table 2.4).

	MRI	RT-QuIC
Accuracy	0.92 (0.79-0.98)	0.80 (0.63-0.91)
Sensitivity	0.91 (0.76-0.98)	0.76 (0.58-0.89)
Specificity	1.00 (0.54-1.00)	1.00 (0.54-1.00)

Table 2.4: Comparison of the diagnostic statistics for the MRI scoring system and the second generation RT-QuIC test, evaluated on a common pool of 39 subjects (of which 33 sCJD patients). For each statistic, a 95% confidence interval is computed using the method of Clopper and Pearson.

2.2.2 DWI topographic lesion distribution in 5 sCJD subtypes

The mean scores of 12 brain regions for each of the sCJD subtypes are plotted in Figure 2.1.

In MM1 patients, DWI signal hyperintensity mean scores were very high either in the most of the neocortex with exception of the occipital visual cortex and in the cingulate, they were moderate in the insula and striatum. The hippocampus, thalamus and cerebellum scores were very low confirming that these regions are rarely affected in this subtype. The parietal cortex was the region with the highest mean score.

In MV1, DWI signal hyperintensity mean scores and lesion distribution were not statistically different from MM1, however few differences were found. The cingulate and the caudate were the two regions with the highest score followed by the parietal cortex and putamen. Thalamus and cerebellum obtained a relatively low scores.

In MM2, DWI scores were very high in the most of the neocortex with exception of the occipital visual cortex, they were moderate in cingulate and insula, low in hippocampus and very low in striatum, thalamus and cerebellum. Again the parietal cortex was the region with the highest mean score.

In MV2C, DWI scores were very high in most of the neocortex with a

moderate score in the occipital cortex, they were moderate in cingulate and insula, low in hippocampus and striatum, and very low in thalamus and cerebellum. The parietal cortex was the region with the highest mean score.

In VV1, DWI scores were high in the limbic regions (insula, cingulate and hippocampus) and in most of the neocortex with exception of the occipital visual cortex, they were moderate in striatum, very low in thalamus and cerebellum. The insula was the region with the highest mean score.

In MV2K, DWI scores were very different from the previously described subtypes. Scores were very high in striatum and thalamus, moderate in cingulate, insula and cerebellum, while they were very low in the neocortex and hippocampus. The caudate was by far the region with the highest mean score.

In VV2, DWI scores were very similar to MV2K but different from the other subtypes. Scores were very high in caudate, moderate in putamen, thalamus, cerebellum and cingulate, low in the insula and hippocampus and very low in the neocortex. The caudate was by far the region with the highest mean score.

Then, as a second step of the analysis, we pairwise compared the mean score lesion profile of each subtype, globally with the permutation T-squared Hotelling test and region-by-region with a Wilcoxon rank-sum test. MM1 was significantly different from MM2, as well as MV1 from MV2C, in the caudate and putamen ($p<0.001$ in both cases). MV1 was significantly different from MV2K in the thalamus and parietal cortex ($p=0.0054$ and $p=0.0349$ respectively). MV2C was significantly different from MV2K in the temporal, precuneus, parietal, caudate, putamen and thalamus (all p-values <0.001). VV1 was significantly different from VV2 in the insula, precuneus and thalamus (all p-values <0.001), from MM1 in the insula

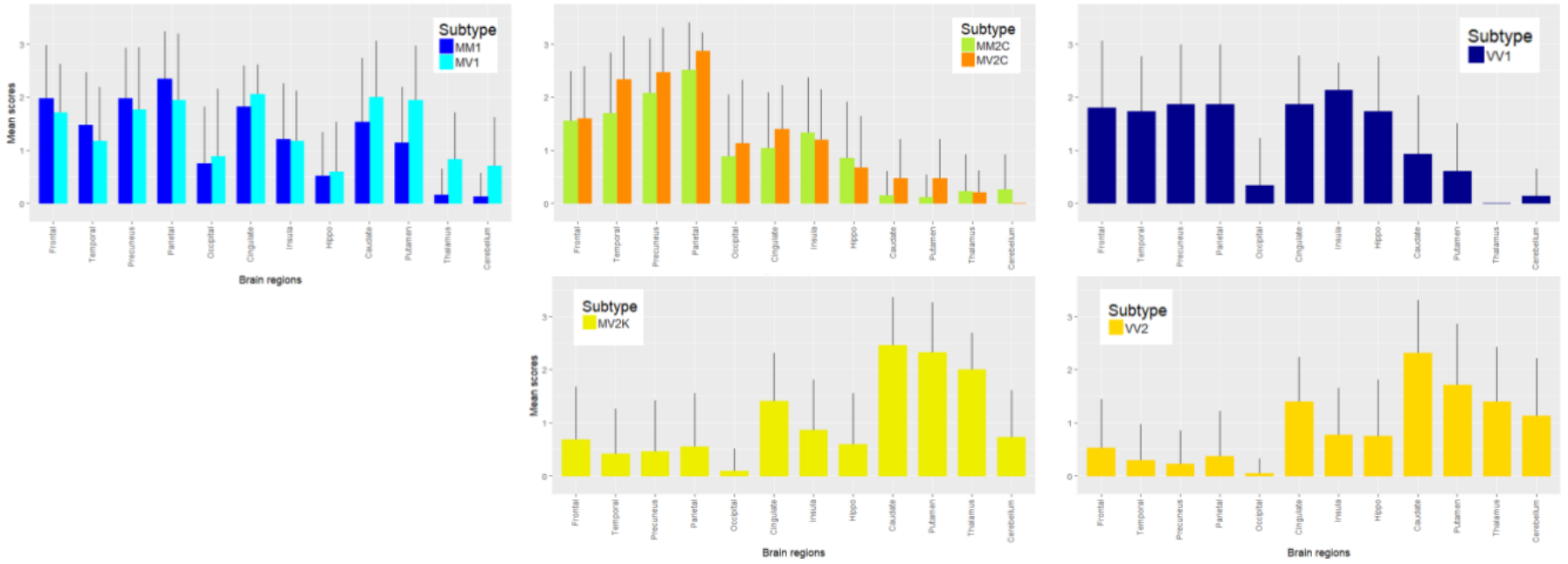


Figure 2.1: The mean scores of the 12 brain regions are represented as bars, and a segment represent the corresponding standard deviation. The brain regions are ordered as follows: from left to right, frontal, temporal, precuneus, parietal, occipital, cingulate, insula, hippocampus, caudate, putamen, thalamus and cerebellum. From top left to top right are represented the sCJD subtypes MM1 and MV1, MM2C and MV2C, and VV1. From bottom left to bottom right are represented MV2K and VV2.

(p=0.0149), from MV1 in the putamen (p=0.0315), and from MV2K in the temporal cortex (p=0.0245), precuneus (p=0.0445), insula (p=0.0003), caudate (0.0124), putamen (p=0.0017) and thalamus (p<0.0001).

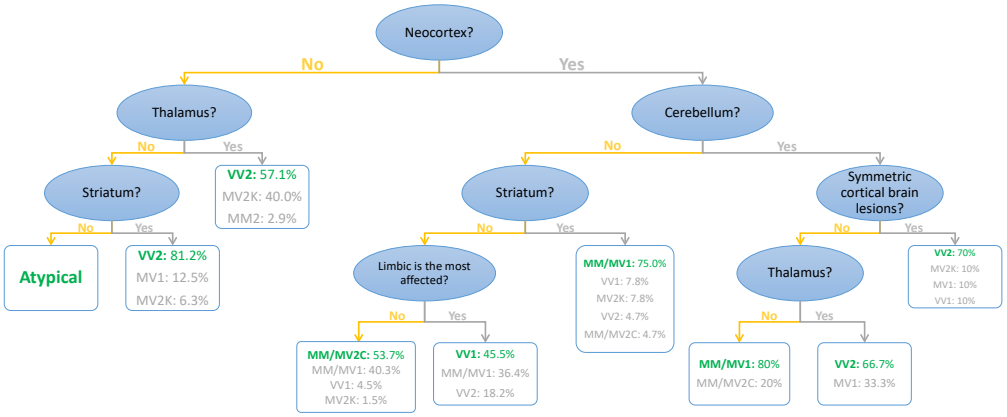
The lesion profiles of the two most common subtype of sCJD, MM1 and VV2, had significantly different mean scores in 8 brain regions: the whole neocortex, caudate, thalamus and cerebellum (all p-values < 0.001, but for caudate p=0.009), with the exception of the limbic regions and putamen.

The following comparisons did not show any significant different mean score among brain regions: MM1 vs MV1, MM2C vs MV2C, MM2C vs VV1, MV2C vs VV1, MV2K vs VV2. For simplicity, when the information about the genotype at codon 129 is missing, the two MM1 and MV1, MM2C and MV2C pairs are combined and thereafter identified as MM/MV1 and MM/MV2C, respectively, as it is common practice in clinical and histopathological diagnoses of sCJD.

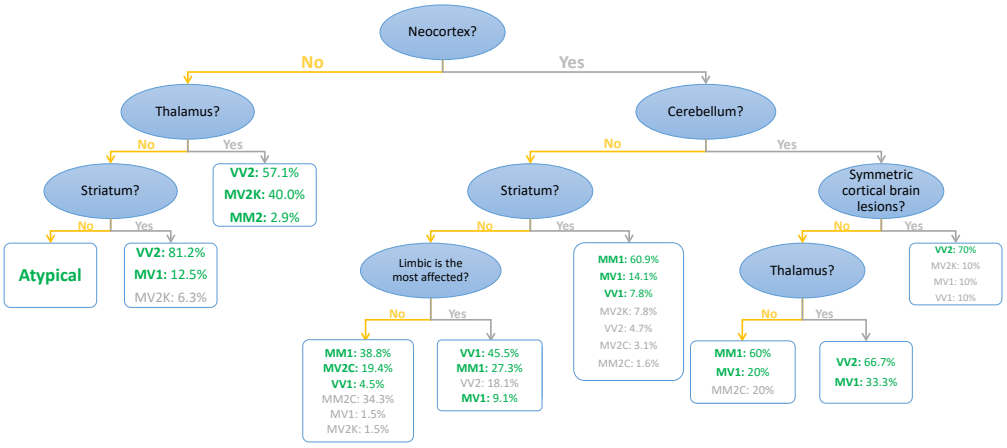
From the data reported above two main subsets of subtypes with similar lesion profiles emerge: the first subset is characterised by high scores and consistent involvement of the neocortex and lower scores and relative sparing of striatum, thalamus and cerebellum, and includes MM/MV1, MM/MV2C and VV1; the second subset is characterised by high scores and consistent involvement of striatum, thalamus and cerebellum and lower scores and relative sparing of the neocortex, and includes MV2K and VV2.

2.2.3 Subtype classification

A third step of the analysis consisted in generating a diagnostic algorithm to assist the radiologist/neurologist in predicting the subtype in individual cases with the MRI lesion profile. This algorithm was inspired by the results of a classification tree built on the scores of the 217 true positive sCJD patients with a pure subtype of our dataset. In details, we proposed



(a) Codon not available.



(b) Codon available.

Figure 2.2: Decision trees representing the output of the performed classification analysis. Round blue ovals represent the splitting criteria used to move down the tree: for example, the first is "Neocortex?" and it is a short question for "Is the neocortex affected?". If the answer is positive, goes to the right branch, otherwise to the left. The white-filled squares at the end of the bottom of the tree represent the terminal nodes of the algorithm (i.e., distinct phenotypes), in which all the subtypes that follow that branch of the tree are listed, ordered by their probability of being in that terminal node. In green, it is shown the predicted subtype that our proposed algorithm assign to a subject that falls into that terminal node. In (a) the codon is not available, so we always predict one subtype at each node (e.g., a subject with neocortex not affected and thalamus affected is classified as VV2), while in (b) the codon is available and therefore in each terminal node we can predict up to three subtypes, according to the codon 129 polymorphism of the subject (e.g., a subject of the group MV with neocortex not affected and thalamus affected is classified as MV2K).

a diagnostic algorithm that provides, for a new subject, the probability of belonging to each of the five sCJD pure subtypes. This algorithm produced 9 distinct phenotypes, as illustrated in Figure 2.2a, that characterise the five sCJD pure subtypes (i.e. MM/MV1, MM/MV2C, MV2K, VV1, VV2). The accuracy in predicting the correct subtype is about 62%, considering the subtypes with the highest probability in each phenotype as the predicted outcome of the algorithm for that phenotype (e.g. a patient with neocortex not affected but with an involvement of the thalamus is classified as VV2). We remark that this accuracy is calculated weighting the misclassification errors by their prevalence in the sCJD population. Key steps in the algorithm are the presence or not of DWI hyperintensity in the neocortex, striatum, thalamus and cerebellum, as well as the occurrence of symmetric vs. asymmetric lesions in individual cases and whether the region showing the most intense hyperintensity was the limbic one.

To further improve the accuracy of the proposed algorithm and to distinguish subtypes with similar phenotypes (e.g. MV2K vs VV2), we recommend analysing the genotype at codon 129 (see Figure 2.2b). With this new information, the weighted prediction accuracy of our algorithm increases to 90%. However, in some phenotypes the availability of the codon is not enough to completely discriminate the subtypes (e.g. MM1 and MM2C with the neocortex affected and the striatum spared), thus we suggest to exploit the different duration time of each subtype: for example, MM1 is the fastest among the sCJD (its mean duration is about 3 months) while MM2C has a significantly longer mean duration (about 15 months). If examined within 3 months from the symptoms onset, a patient with the neocortex affected and the striatum spared will be more likely to be a MM1, due to its prevalence in the relative frequencies of the disease. On the other end, a patient with the same lesion profile but examined more

than 3 months from the symptoms onset will be more likely classified as a MM2C.

2.3 Discussion

Our study of 429 patients, the largest cohort to date comprising MRI data, has achieved 0.91 sensitivity and 0.98 specificity with excellent IRR in most brain regions matching the highest MRI diagnostic values previously reported [51, 52, 56]. Furthermore, our MRI diagnostic values surpassed those of the cerebrospinal fluid test and the RT-QuIC performed on the same patients. While this finding is likely due to the small size of this subset, our MRI diagnostic values are comparable to the RT-QuIC values recently reported for the most advanced version of this test on a large number of patients and far exceed diagnostic values of previous tests like tau and 14-3-3 either performed on the same patients as in our study or reported in the literature [62, 68]. The high MRI diagnostic values are especially relevant because MRI abnormalities are often present at the early stages of the disease and in at least two instances MRI was reported positive 2 and 3 months before the disease became symptomatic [56, 78–80].

A major goal of this study was determining whether the five phenotypes of sCJD, which account for approximately 90% of all human prion diseases and are easily distinguishable because of the distinct types and distributions of the brain lesions, could be also identified on MRI examination based on the combinations of the variations in severity and brain distribution of the abnormal DWI signal [60]. To accomplish this task, in this study we provided a diagnostic algorithm in the form of a classification tree, that related each sCJD subtype to one or more of nine distinct phenotypes.

The first split in the tree was identified by the involvement of the neocortical regions. The phenotypes without an affected neocortex were three: an atypical one, where the main regions characterising the sCJD (i.e., neocortex and striatum) were not affected, and other two composed of mainly both MV2K and VV2 sCJD subtypes. These two subtypes have similar MRI pattern distributions and our test did not find any statistical difference between them, in agreement with other findings in literature [81, 82]. However, the likelihood of MV2K subtype drastically decreases when the thalamus is not affected.

On the other branch, the prevalent subtypes were MM/MV1, MM/MV2C and VV1, with the exception of two phenotypes characterised by the involvement of the cerebellum (i.e., the last two terminal nodes on the right in Figure 2.2) in which the VV2 subtype is the most likely. Among the phenotypes with abnormal signal detected in the neocortex and not in the cerebellum, the most likely was MM/MV1, then MM/MV2C and lastly the rarer VV1 sCJD subtype. However, it is worth nothing that in the case of a strong involvement of the limbic region, the VV1 is the most likely. MM/MV1 and MM/V2C are easily separable in the case of both neocortex and striatum affected, where MM/MV1 is definitely the most likely.

If the genotype at the codon 129 is available, each resulting phenotype of our classification tree could be then related up to three different subtypes, i.e. the most probable subtype for MM, MV and VV respectively. For instance, in the left branch of the tree, VV2 appears with high probability together with MV2K or, if the thalamus is not affected, with MV1: in these cases, knowing the genotype (VV or MV) extremely increases the classification accuracy, allowing to subtype also almost every MV subject. Instead, the same strategy does not further improve the accuracy when

only the neocortex is affected: the problem is that in this case the subtypes MM1 and MM2C (sharing the same codon 129 polymorphism) show overlapping features at the MRI assessment, thus the two strains are almost equally likely and it is impossible to distinguish between them. A possible feature that could help in separating the two MM subtypes, could be found looking at the duration of the disease, since MM1 is characterised by a faster impairment of the patient conditions with respect to MM2C, thus if a patient is still alive after several months from the symptoms onset, MM2C is more likely. However, MM2C is still unidentifiable from MM1 if the MRI exam is taken at an early stage of the disease.

To conclude, we have shown that the overall diagnostic accuracy of MRI to diagnose prion diseases early in the symptomatic phase of the disorder is excellent and it is higher than CSF tests such as RT-QuIC, tau and 14-3-3. Moreover, we provided a simple and effective diagnostic algorithm by a non-parametric classifier (i.e., the classification tree) that relies on MRI assessment to predict the sCJD subtype of a patient, with an even higher prediction accuracy if the codon 129 polymorphism is considered. The results of the classification tree and the pairwise comparisons between subtypes have led to a better characterisation of the sCJD subtypes in terms of their features detectable through MRI. This work provides potential new diagnostic tools for the early diagnosis of the sCJD patients and subtype identification, that should be validated reproducing the same methodology on a new dataset. Indeed, a limitation of the study is that the same dataset was used for training the model and assess its accuracy, although a cross-validation technique was adopted to prune the classification tree and reduce the overfitting. Moreover, a data collection of new patients has started recently, in order to build a new test set on which we could validate our proposed classification model. Further analysis on the characterisation

of the sCJD subtypes are done in Chapter 3, where we aim at describing the disease progression for each subtype. Finally, another limitation of this study consist in the exclusion from the analysis of the complex case of the co-occurence of the two type of prion protein in the same patient: the analysis of the brain lesions detectable with MRI and the classification of these mixed genotypes would be of interest for future works.

Chapter 3

An Application of the Event-Based Model in Sporadic Creutzfeldt-Jakob Disease

In the previous chapter, we have shown the good performances of MRI for the diagnosis and characterisation of the sporadic Creutzfeldt-Jakob disease (sCJD) subtypes. In particular, we have seen that each subtype has distinctive features that should be always considered when analysing the whole spectrum of sCJD. In this chapter, we want to move further and explore the disease progression in sCJD with the use of imaging biomarkers such as signal abnormalities on DWI. From the results described in the previous chapter, it is clear that we have to consider each sCJD suptype separately, since it is likely that in different strains the propagation could be different.

Recent studies have dealt with the assessment of the disease evolution in sCJD using DWI signal abnormalities as biomarkers of the disease: [83] analysed the MRI of a group of 37 patients, of whom 21 having a confirmed

sCJD diagnosis postmortem by autopsy, while the others were diagnosed as probable sCJD. They reported increasing signal abnormality in cortical and subcortical regions at follow-up with respect to the baseline, most substantially in the striatum. However, they did not take into account the subtype classification in the analysis, nor provided a real staging of the disease based on the signal abnormalities. Another study [84] has analysed retrospectively 36 sCJD patients with MM genotype, of whom only 5 with a definite diagnosis of sCJD based on brain autopsy, while the others were considered probable sCJD. They classified the subjects a priori into four disease stages based on their clinical state: vague symptomatic, possible CJD, probable CJD and chronic vegetative state. They found that at the first stage, almost all patients had cortical hyperintensities, with a predominance in the neocortical regions; at the last stage, all the patients have DWI hyperintensities in the cortical and subcortical regions, meaning that signal abnormality were spreading from the cortex to the subcortical regions, as the stages progress. However, a limitation of the study was ignoring the type of PrPD in each subject and, even if an evolution in the brain pattern of signal abnormality was detected, it still relied on a priori staging system, that limits the temporal resolution of progression models. Here, we consider a novel data-driven model to assess the disease progression in each sCJD subype: the event-based model [5, 6] has been recently introduced to study the evolution of Alzheimer's and Huntington's disease. It does not rely on a priori staging system, but reconstructs the disease progression directly from the data.

3.1 Materials and methods

3.1.1 Patients

The data collection of the subjects involved in our analysis is described in Chapter 2, where MRI studies (T2WI, FLAIR, diffusion-weighted MRI), pathology subtype (codon 129 polymorphism, PrPD type) and demographical information are collected from more than 400 patients with autopsy (or biopsy) confirmed diagnosis of prion disease or alternative conditions. For the purposes of this study, we select all the patients with a confirmed diagnosis of sporadic CJD having a pure subtype (i.e., MM1, MV1, MM2C, MV2C, MV2K, VV1, VV2), and all the subjects with a non-prion disease confirmed diagnosis as a control group.

3.1.2 Image assessment

A distinctive characteristic of sCJD patients is the regional appearance of the DWI signal hyperintensities in the MRI studies performed after the symptoms onset. The visual assessment of such diffusion images is one of the most adopted pre-test for early diagnosis of CJD in many clinical centres, with a high diagnostic reliability as reported by several authors: see [85] for a recent review on DWI accuracy in sCJD. In our study, we consider 12 brain regions that are known to be specifically related to the presence and the effects of the misfolded prion protein [64]: five neocortical regions (frontal, parietal, precuneus, temporal, occipital), three regions belonging to the limbic system (cingulate gyrus, insula and hippocampus), basal ganglia (caudate and putamen), thalamus and cerebellum. The signal hyperintensity in DWI of these regions are considered as separate candidate biomarkers to describe the disease progression in each sCJD pure subtype. The visual assessment of the diffusion MR images of all the patients is based on the scoring system proposed in Chapter 2, as a semi-quantitative

method to evaluate the hyperintensities of DWI: an integer ordinal score is assigned to each of the 12 brain regions selected, with values ranging from "zero" (i.e., absence of hyperintensity) to "three" (i.e., high level of hyperintensity and spread over the region considered). A diffusion MR image of a subject is considered positive for the diagnosis of prion disease if there is at least one brain region with a score equal or greater than "two". Otherwise, if all the assigned scores are less than or equal to "one", the image is considered negative for the prion disease diagnosis.

3.1.3 General theory of the event-based model

In this section, we describe the theory of the EBM, adopting the same notation as in [5]. The EBM is a data-driven model that provides a description of the disease progression in terms of a sequence of events. In this framework, an event is the switching from a normal state to an abnormal state for a biomarker of a patient. We consider a set of N events $E_1, \dots, E_N$ related to N corresponding biomarkers of the disease. The EBM finds the most likely ordering of the events, denoted with $\bar{S} = (\bar{s}(1), \dots, \bar{s}(N))$ (i.e., a permutation of the integers $1, \dots, N$), given the biomarkers measurements $\{X_{ij} : i = 1, \dots, M \text{ and } j = 1, \dots, N\}$ of M subjects. The ordering $\bar{S}$ is obtained by maximizing the data likelihood

$$P(X|S) = \prod_{i=1}^{M} \sum_{k=0}^{N} P(k) \left(\prod_{j=1}^{k} P(x_{i,s(j)}|E_{s(j)}) \prod_{j=k+1}^{N} P(x_{i,s(j)}|\neg E_{s(j)}) \right), \tag{3.1}$$

where $P(k)$ is the prior probability of being at stage k in the generic ordering S and $P(x_{i,s(j)}|E_{s(j)})$ and $P(x_{i,s(j)}|\neg E_{s(j)})$ are the likelihood of a measurement $x_{i,s(j)}$ given that for the i-th subject the event $E_{s(j)}$ has or has not occurred, respectively. We assume that $P(k)$ is uniform, that is equivalent to consider all the stages a priori equally probable, and that

the measurements (conditionally to the occurrence of any event) follow a discrete distribution, since in our case all the $x_{i,s(j)}$ are integers from 0 to 3.

3.1.4 Estimation of the event-based model

Similarly to [5], we assume that none of the E_j has occurred for the subjects in the control group, whereas we do not know a priori whether and which event has occurred for the sCJD patients of a given subtype. Thus, to compute the data likelihood 3.1, we fit a mixture model whose components correspond to $P(x_{i,s(j)}|\neg E_{s(j)})$ and $P(x_{i,s(j)}|E_{s(j)})$: we fit the former component only to the data from the controls, and then we fix the estimated parameters while fitting the mixture to the complete data set, using the Expectation Maximization algorithm. The choice of the discrete distributions is discussed in the next section.

The final step to estimate the most likely ordering $\bar{S}$ of the events is the application of a Markov Chain Monte Carlo (MCMC) algorithm to sample from the posterior distribution $P(S|X)$ on the possible orderings S, under the assumption that, a priori, all orderings are equiprobable. We run this procedure for $1,000,000$ iterations, obtaining a list of candidate orderings, from which we find $\bar{S}$ by selecting the ordering with maximum likelihood among the MCMC samples. Moreover, we quantify the variability of the most likely ordering through the positional variance diagram [5], where the entry of each position (j,k) represents the likelihood that E_j appears in the k-th position of the sequence, computed from the set of MCMC samples. As a further step, we provide a more conservative estimation of the uncertainty by calculating new maximum likelihood event sequences for 100 bootstrapping data samples: for each new sample, first we refit the mixture model and then we obtain a new characteristic sequence with the

same procedure as before. Thus, we calculate a second positional variance diagram, whose entries represent the proportion of bootstrap samples in which E_j appears in the k-th position of the sequence $\bar{S}$.

The EBM relies on the fundamental assumptions that the ordering of the events is common for all the subjects in the cohort. Thus, to reduce the heterogeneity among the patients, we apply the EBM to each of the seven sCJD subtypes separately, since they are characterised by clinical and neuropathological differences. However, since many authors, e.g. [60], agree on considering MM1 and MV1 as a unique subtype called MM/MV1, as well as MM/MV2C as the union of MM2C and MV2C subtypes, we apply the EBM also on these two merged subtypes. Then, we compare the EBM sequences of events among all subtypes, highlighting qualitative differences (or similarities) in the disease progression for each strain.

3.1.5 Mixture models for the data likelihood

The EBM relies also on the assumption of independence of measurements from different subjects, thus, we select only one MR image per subject. For these reasons, we retain from the control group only one MR image of each non-prion subject having all negative MRI. We use these measurements to directly estimate the event distribution $P(x|\neg E_j)$: we choose a Bernoulli distribution to model this component, since the observed values of every biomarker are 0 or 1 in this group. Regarding the event distribution $P(x|E_j)$, we retain only the first positive MRI of each sCJD patient to model this component. The other follow-ups, if present, are used to assess the longitudinal consistency of the patient staging in each sCJD subtype. Since in this group the values of x can range from 0 to 3, we choose a uniform discrete distribution with support on the set $\{0, 1, 2, 3\}$, that assigns the same probability to every possible value of the score.

3.1.6 Patient staging

We assign a disease stage for the i-th subject of each sCJD subtype, identifying the stage that maximizes the probability of the observed biomarker measurements for that patient, given the characteristic order $\bar{S}$ found by the EBM for that subtype, i.e. the stage

$$\begin{aligned} k^* &= \underset{k}{\operatorname{argmax}}\, P(X_i|\bar{S}, k) \\ &= \underset{k}{\operatorname{argmax}}\, P(k) \prod_{j=1}^{k} P(x_{i,s(j)}|E_{s(j)}) \prod_{j=k+1}^{N} P(x_{i,s(j)}|\neg E_{s(j)}) \end{aligned} \tag{3.2}$$

where the notation in 3.2 is the same as in 3.1, as well as the assumptions made on the distributions. The possible stage ranges from 0 (i.e., no biomarkers are abnormal) to N (i.e., all the biomarkers are abnormal). If a patient is assigned to a stage k, then our model predicts that all events appearing before the event k in the sequence $\bar{S}$ have occurred, while the others have not yet occurred. The stage assigned to the patients should be interpreted as the most compatible stage with their biomarker measurements, since the fitting to the model is not exact.

3.1.7 Longitudinal validation

Due to the short time duration of the sCJD and its fast spread in the brain of the patients, we have few follow-ups in our study, some of them unchanged with respect to the baseline. Moreover, the images were acquired at variable time points, since the data come from several medical centres with no common design of experiment. We evaluate the consistency of patient staging with the available longitudinal measurements of the pure subtypes sCJD patients, that have some changes across time points: we identify 43 suitable sCJD patients (of 71 total) with a first follow-up MRI scan, seven of them having also a second follow-up and only one patient

with a third one. We compare the stage assigned to the first positive MRI scan (i.e., baseline EBM stage) with the EBM stages assigned at the corresponding follow-ups.

3.2 Results

3.2.1 Subjects

In this study, we include in the control group 120 subjects having a non-prion confirmed diagnosis and negative MRI, with scores 0 or 1 for each brain region. Regarding the sCJD patients, we select 217 cases with confirmed pure subtype sCJD diagnosis, their first positive MRI (i.e., baseline) and, if available, the follow-ups showing changes with respect to the baseline. Further information about the sCJD patients are illustrated in Table 3.1.

Subtype	Positive MRI	Age at death (years) Median (Range)	Disease duration (months) Mean (SD)	Follow-ups
MM1	72	65 (39-91)	3.5 (3.1)	16
MM2C	27	63 (46-89)	14.7 (11.4)	6
MV1	17	68 (49-88)	8 (7.8)	3
MV2C	15	68 (46-89)	17.6 (13.4)	2
MV2K	22	63 (47-85)	12.4 (7.4)	6
VV1	15	60 (35-88)	9 (6.5)	1
VV2	49	65 (45-80)	5.6 (2.3)	9

Table 3.1: Description of the sCJD patients per subtype. Positive MRI: the number of patients with a positive diffusion MRI (i.e., baseline). Follow-ups: the number of patients having a follow-up that presents some changes with respect to the corresponding baseline.

3.2.2 Event sequences

The results on the maximum likelihood sequence and its uncertainty are shown in the positional variance diagrams in Figures (a-c),(a-c) and (a-c), separately for each sCJD pure subtype, and in Figure (d-f),(d-f) and (d-f) their corresponding positional variance diagrams when the bootstrap samples are considered. All the diagrams show on the y-axis the events ordered according to the maximum likelihood sequence. For each subtype, all the

regions not affected in the patient group (i.e., the regions in which no patient's image has a score greater than 1) are excluded from the ordering and thus do not appear in the figures. Finally, it is worth noting that a general consistence between the two positional variance diagrams can be observed for each sCJD subtype.

MM1, MV1 and MM/MV1

The MM1 strain is characterised by a first involvement of the cingulate gyrus and neocortical structures, followed by the striatum (i.e. caudate and putamen). The other limbic grey-matter structures (i.e. insula and hippocampus) appear more lately in the sequence, followed by thalamus and cerebellum (see Figures a and d). Two regions of the limbic grey-matter (i.e. insula and cingulate) are the first to become abnormal in the MV1 subtype, followed by the frontal neocortex and the striatum (Figures b and e). Then, the thalamus appears in the sequence, followed by the other neocortical regions and eventually the cerebellum and hippocampus. Considering together the MM1 and MV1 subtypes (Figures c and f), the ordering is exactly the same as in the MM1 group alone, due to the fact that the sample size in the two populations is unbalanced (72 MM1 patients and 17 MV1 patients).

MM2C, MV2C and MM/MV2C

The MM2C strain is characterised by a first involvement of the regions of the limbic system and subsequently the neocortical structures. Thalamus, cerebellum and striatum appear only at the end of the sequence (see Figures 3.2a and 3.2d). The MV2C sequence is similar to the one of MM2C subtype, besides a shift of the location of the temporal region that appears later in MV2C, and the absence of thalamus and cerebellum (Figures 3.2b

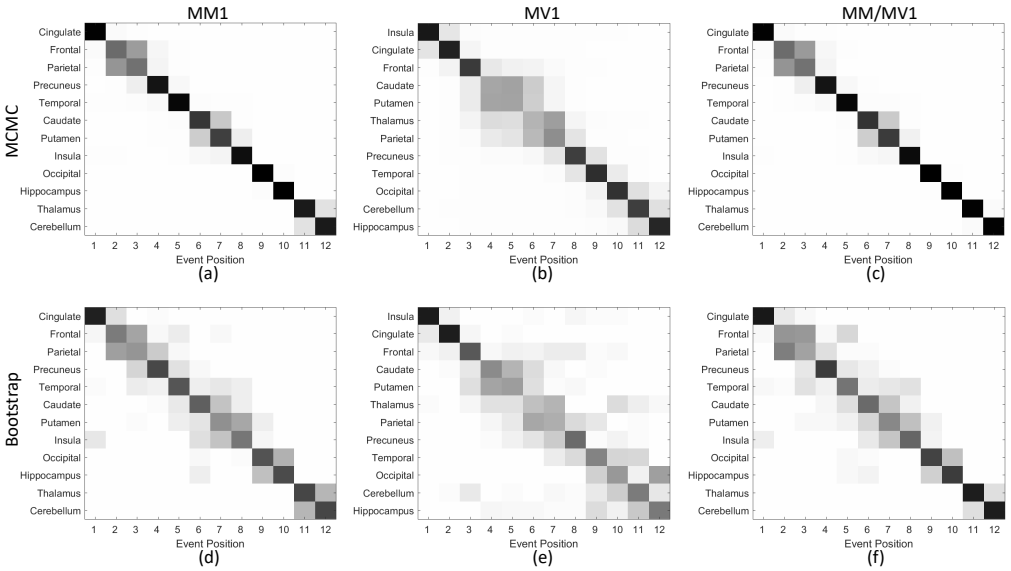


Figure 3.1: Positional variance diagrams showing the ordering and the uncertainty of event sequences in the sCJD subtypes MM1 (a, d), MV1 (b, e) and MM/MV1 (c, f): (ac) are obtained by using the MCMC samples within the EBM, (df) by the bootstrap samples of the data. Each entry of a diagram represents the proportion of MCMC samples, in (ac), or bootstrap samples, in (df), in which events occur at a specific position in the sequence on the x-axis. The darker the square, the closer the proportion is to 1.

and 3.2e). Considering together the MM2C and MV2C subtypes (Figures 3.2c and 3.2f), the events are ordered in a similar way, with the only exception that the insula occurs later in the merged group (albeit with a high uncertainty) and the involvement of thalamus and cerebellum appears only at the end.

MV2K

The MV2K strain is characterised by a first involvement of regions of the limbic system and subsequently the striatum, thalamus and cerebellum. The neocortical regions appear only at the end of the sequence (see Figure 3.3c and 3.3f).

VV1

The VV1 strain is characterised by a first involvement of the regions of the limbic system, followed by the neocortical regions. The striatum and

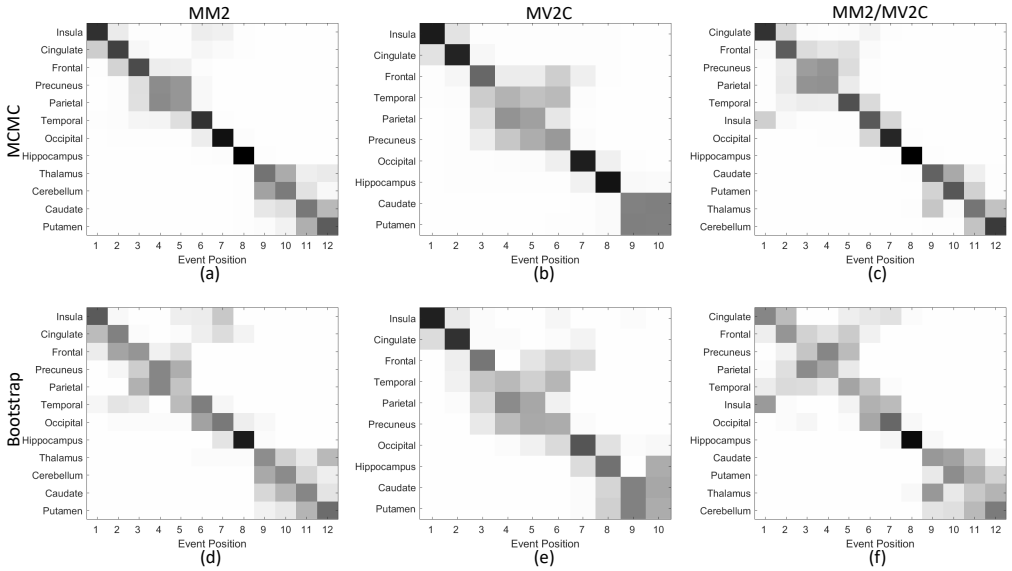


Figure 3.2: Positional variance diagrams showing the ordering and the uncertainty of event sequences in the sCJD subtypes MM2C (a, d), MV2C (b, e) and MM/MV2C (c, f): (ac) are obtained by using the MCMC samples within the EBM, (df) by the bootstrap samples of the data. Each entry of a diagram represents the proportion of MCMC samples, in (ac), or bootstrap samples, in (df), in which events occur at a specific position in the sequence on the x-axis. The darker the square, the closer the proportion is to 1.

cerebellum appear only at the end of the sequence, while the thalamus is absent, since it is negative for every VV1 patient. (see Figure 3.3a and 3.3d).

VV2

The VV2 strain is characterised by a first involvement of the cingulate, followed by striatum, thalamus and cerebellum. The other limbic regions (i.e., hippocampus and insula) and all the neocortical regions appear only later, at the end of the sequence (see Figure 3.3b and 3.3e).

3.2.3 Longitudinal staging

In Figure 3.4, we compare the EBM stage of each subject at baseline with the EBM stage at follow-ups, separately for all the subtypes. Each plot is composed of the corresponding positional variance diagram obtained by the EBM, on which a scatterplot is superimposed: the size of the dots

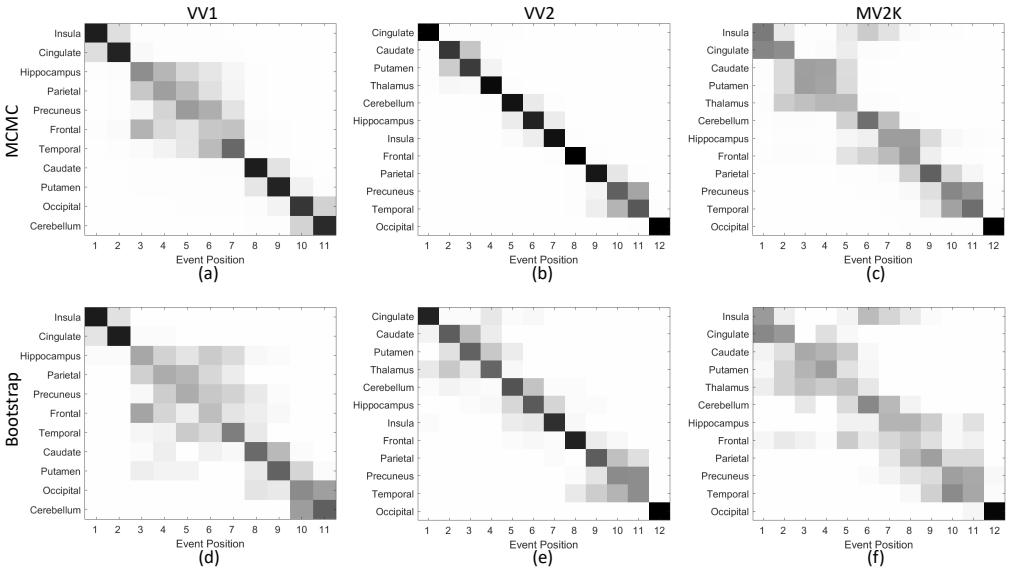


Figure 3.3: Positional variance diagrams showing the ordering and the uncertainty of event sequence in the sCJD subtypes VV1 (a, d), VV2 (b, e) and MV2K (c, f): (ac) are obtained by using the MCMC samples within the EBM, (df) by the bootstrap samples of the data. Each entry of a diagram represents the proportion of MCMC samples, in (ac), or bootstrap samples, in (df), in which events occur at a specific position in the sequence on the x-axis. The darker the square, the closer the proportion is to 1.

represents the number of subjects having those particular EBM stages at baseline (x-coordinate) and at follow-up (y-coordinate). For every group, the longitudinal consistency of the patient staging appears to be excellent: in almost every subject the EBM stage increases or remains unchanged at each follow-up. There is only one case having a decreasing stage at follow-up: it is a VV2 patient with a poor quality MRI scan, for which the reliability of the visual assessment should be questioned. All the other points in every plot are on or above the diagonal, generally within the grey shaded area that depicts the uncertainty estimated by the EBM.

3.3 Discussion

In order to determine the characteristic sequence of biomarkers becoming abnormal in the sCJD, we have revised the EBM [5,6] for use with discrete data coming from a semi-quantitative assessment of DWI brain images. Due to the model assumptions, requiring a homogenous population, we

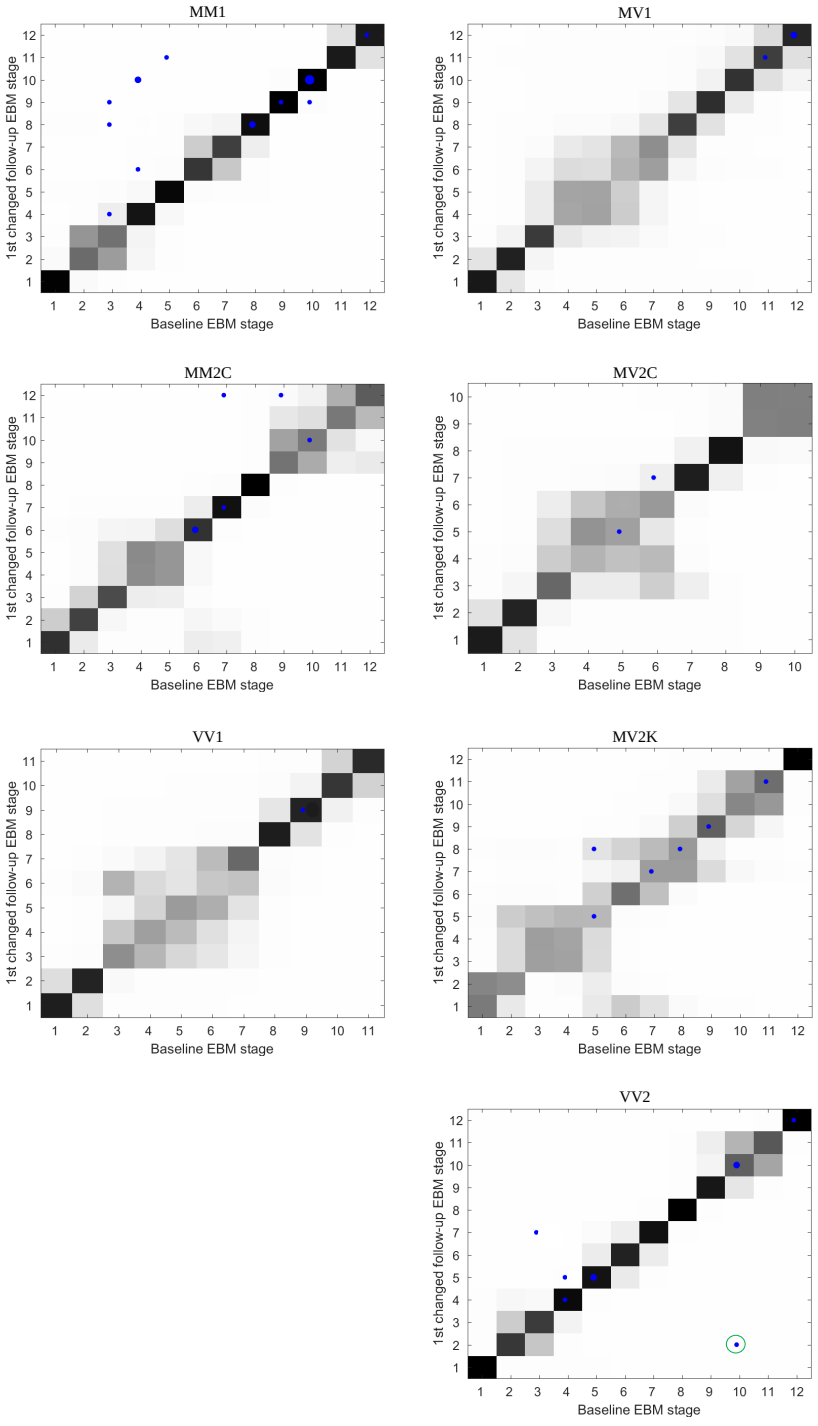


Figure 3.4: Longitudinal consistency of patient staging in each sCJD subtype group over a variable follow-up period. The size of the dot is proportional to the number of subjects having those particular EBM stages at baseline (x-coordinate) and at follow-up (y-coordinate). The largest dot, at (10,10) in the MM1 plot, represents 4 subjects and the smallest dots represent one subject. The EBM stage is longitudinally consistent for the subjects whose points are on or above the line $y = x$ and/or within the grey shaded area, that depicts the uncertainty in the sequence estimated by the EBM. The VV2 patient whose point (circled in green) is below the line $y = x$ has a poor quality image at the follow-up, for which the reliability of the visual assessment should be questioned.

have run the analysis on the data of each sCJD subtype separately. We have found that the cingulate cortex is one of the first regions appearing in the characteristic sequence of each subtype. Then, the patterns emerging from the orderings can be divided in three sets: neocortex affected before striatum (MM1, MM2, MV2C and VV1); striatum affected before neocortex (MV2K and VV2); neocortical regions and striatum appearing together (only MV1, whose ordering seems to have higher uncertainty with respect to the others). In the first group, the striatum appears right after the neocortex for MM1 and VV1, but only at the end for MM2 and MV2C. Moreover, VV1 is the only subtype having all the limbic regions appearing at the beginning. On the other hand, the subtypes in the second group have very similar orderings, besides the insula that is in the middle of the sequence for VV2, while it is the first region for MV2K, even if with high uncertainty.

Moreover, it is worth noting that the uncertainty has been calculated directly from the EBM or via bootstrap sampling: the former underestimates the uncertainty since the variability of the biomarker distribution is not taken into account, while the latter tends to overestimate the uncertainty since each bootstrap sample does not include all the cases. The results from both methods should be used together to better understand the variability of the orderings and to compare the homogeneity of the subgroups: for instance, the MV1 ordering (especially looking at the bootstrap results) appears to be the most variable among all the subtypes, indicating a higher heterogeneity within the cases of this subgroup with respect to the others. In general, the heterogeneity of a group could be explained in term of both natural variability among subjects, and possible mixture of different types of misfolded prion protein [86] or even different neurodegenerative diseases (e.g., Alzheimer disease) concurring in the same patient [87]. Besides the

homogeneity of the group considered, other factors could contribute to the uncertainty in a sequence of events: for instance, when two or more events occur close in time, it is unlikely that the data points are acquired exactly in the middle to separate them, especially when we deal with a rare disease like sCJD, for which the sample size within each subtype is small. Moreover, we do not take into account when the events are highly correlated (e.g., caudate and putamen form a single functional structure, the dorsal striatum), since in the present formulation of the EBM the covariance among the measurements is ignored. As future work, it would be important to evaluate the relationship among measurements in order to fit this correlation within the model, relaxing the hypothesis of independence of each measurement on event occurrence.

Important areas for next study on the progression of sCJD can be tracked looking at the general disease progression modelling state-of-the-art, that has been recently reviewed [88], where the authors highlight possible future directions of research. In particular, a promising new extension of the EBM has been recently published and applied to sporadic Alzheimer's disease (Young et al., 2015), relaxing the assumption of a unique event ordering in the population. The authors start considering a generalised Mallows model, which allows subjects to have a different latent ordering with respect to a central event sequence. Then, they develop a Dirichlet process mixture of generalised Mallows models, which is a generative clustering model that gathers together subjects with similar event sequences. Such a model could be applied also to the sCJD population, without distinguishing the subtypes a priori and trying to find clusters of patients according to the ordering of their disease progression. Another interesting approach would be to enhance existing network models [89] to examine changes of the brain's anatomical connectivity network due to sCJD and

then use relevant network metrics to create an EBM sequence of the regional connectivity across the brain: see [90] for a similar application in Alzheimer disease. This could offer the unique opportunity to evaluate the degree of support that image-based network propagation model can provide in other diseases, in particular dementia, in which evidence continues to emerge of "prion-like" propagation [91–93]. Our results of the EBM provide for the first time data-driven models of sCJD progression within each pure subtype: the characteristic orderings are based on a cross-sectional dataset and their longitudinal consistency has been validated comparing the stages at follow-ups with the baseline. However, the interpretation of these results should be made cautiously: like in any data-driven model, the events are ordered by the EBM taking into account when the measurement of the biomarkers significantly change between cases and controls [5]. Thus, this may not correspond to the underlying order of the pathology, but more realistically to the order in which a measurement become detectably abnormal [94, 95].

Chapter 4

Function-on-Function Regression Models with Spatio-Temporal Domains

The functional data analysis (FDA) paradigm has changed the statistical science of the XXI century: when data are collected to describe a continuous process, it is natural to treat them as functional data (e.g., curves), to preserve and exploit the continuity property of the phenomenon. This new idea has led to the development of new methods that generalised the standard tools of multivariate statistics. Here we focus on the extension of the multiple regression model, considering the case in which both the response variable and the covariates are functions. This kind of model is known as function-on-function regression model in the FDA literature and there have been several works that have analysed its properties, see for instance the book [96]. Such a model is potentially useful to compare how functional responses of different groups react to a set of external functional inputs, for example in the case of functional magnetic resonance imaging (fMRI). In this case, the blood oxygenation level dependent (BOLD) signal that is registered through the fMRI in the brain of a patient is intrinsically

varying in space and time (i.e., the functional response). The BOLD signal is usually related to one or multiple stimuli (i.e., functional covariates) in a task-based fMRI experiment, in order to find brain regions that exhibit statistical differences among groups of patients. In the next sections we describe the function-on-function regression model and the corresponding inferential statistical analysis, on a domain setting compatible with the characteristics of fMRI data (i.e., 3-dimensional space and finite time interval).

4.1 Function-on-Function Regression Model

Let $s \in \mathcal{S} \subset \mathbb{R}^3$ be a point in the 3-dimensional domain $\mathcal{S}$, and $t \in [a, b]$ an instant in the time interval $[a, b]$. Assume that the functional response $y(s, t)$ is a function of both space and time (e.g., it represents the BOLD signal registered in a brain subject at point s at time t) and that the functional covariates $\mathbf{x}(t) = (x_1(t), \ldots, x_p(t))$ are only temporal dependent (e.g., it represents the external stimuli used in task-based fMRI experiment, not dependent on the location s). After collecting a sample $\{\mathbf{x}_i(t), y_i(s, t)\}_{i=1,\ldots,n}$ of our variables in every $s \in \mathcal{S}$ and $t \in [a, b]$ for each subject i, we can write the function-on-function regression model as follows:

$$y_i(s,t) = \beta_0(s) + \sum_{j=1}^{p} \beta_j(s) x_{ij}(t) + \epsilon_i(s,t), \quad \forall s \in \mathcal{S}, \forall t \in [a,b] \text{ and } \forall i = 1, \ldots, n \tag{4.1}$$

where the response and the covariates are evaluated on the same interval and the value of the response at time t depends only on the values of the covariates at the same time t. The regressor coefficient β_j are only function of the space locations, assuming that the linear relationship between the response and the covariates is fixed in time (e.g., in the fMRI framework, this means to consider the brain response to a specific stimulus being the

same if measured at two different time points). In the functional data analysis literature, equation (4.1) is a function-on-scalar regression model in s, while it is a constant coefficient function-on-function regression model in t because the regressors $x_i(t)$ do not depend on s. Note that, in the latter case, each location s can be analysed separately for fitting purposes: in this framework, it is easy to obtain the functional representation of the estimated regression coefficients $\hat{\boldsymbol{\beta}}(s) = (\hat{\beta}_0(s), \hat{\beta}_1(s), \ldots, \hat{\beta}_p(s))$ with a ordinary least squares (OLS) method. We now discuss how to perform hypothesis tests on the functional regression parameters.

4.2 Inference for the concurrent model

In [97] it is proposed a non-parametric procedure for testing a functional-on-scalar linear model. This method can be easily extended to our model since it is fitted using the same OLS method. Considering a general combination matrix $C \in \mathbb{R}^{(q\times(p+1))}$ that specify $1 \leq q \leq p+1$ linear combinations of the $p+1$ functional regression coefficients, the general hypothesis can be formulated as follows:

$$\begin{cases} H_{0,C} : C\boldsymbol{\beta}(s) = \mathrm{c}_0(s) & \forall s \in \mathcal{S} \\ H_{0,C} : C\boldsymbol{\beta}(s) \neq \mathrm{c}_0(s) & \text{for some } s \in \mathcal{S} \end{cases} \tag{4.2}$$

where $c_0(s) = (c_{01}(s), \ldots, c_{0q}(s))'$ is a vector of fixed functions in $L^2(\mathcal{S}) \cap C^0(\mathcal{S})$. If we reject the null hypothesis H_0 in (4.2), it would be of interest finding the regions on $\mathcal{S}$ that exhibit significant differences. The natural idea is to test the hypothesis in (4.2) at each s separately, but theoretically this should be done for an uncountable infinite number of points. Nevertheless, if we manage to control the family-wise error rate that originates from this family of tests, we can in principle solve the problem. In [97], the authors adopted the procedure introduced in [98] that provided a control

of the interval-wise error rate on a mono-dimensional space. A theoretical extension of this procedure to a three-dimensional setting is outlined in the following steps.

First, perform the test (4.2) for a given closed interval $\mathcal{R} \subseteq \mathcal{S}$, for all the $s \in \mathcal{R}$, computing the statistic

$$T_C^{\mathcal{R}} = \int_{\mathcal{R}} T_C(s)ds, \tag{4.3}$$

where $T_C(s) = (C\hat{\boldsymbol{\beta}}(s) - \mathrm{c}_0(s))'(C\hat{\boldsymbol{\beta}}(s) - \mathrm{c}_0(s))$.

Second, compute the p-value $p_C^{\mathcal{R}}$ of the previous test for a given $\mathcal{R}$ with the functional permutation test implemented in [97], that is based on the Freedman and Lane permutation scheme. Then, define the adjusted p-value function at point s as the supremum p-value of all the test on the intervals $\mathcal{R}$ that contains s:

$$\widetilde{p}_C(s) = \sup_{\mathcal{R} \ni s} p_C^{\mathcal{R}}, \quad s \in \mathcal{S}.$$

Third, select the subregions SR_α in $\mathcal{S}$ where any of the null hypotheses are rejected at level α, defined as $SR_\alpha = \{s \in \mathcal{S} : \widetilde{p}_C(s) \leq \alpha\}$.

4.3 Multiple testing procedure in a 3-dimensional domain in FDA and perspectives

From the previous section, we obtain for each $s \in \mathcal{S}$ a specific adjusted p-value function $\widetilde{p}_C(s)$ of the test (4.2) that allows to find the significant regions SR_α where to reject the null hypothesis at level α. However, computational issues emerge when the number of points is high, due to the fact that the method relies on a permutational approach. To limit this problem, an appealing solution come from the works [7, 8]: the authors proposed an inferential procedure in the functional ANOVA framework

based on splitting the domain into mutually exclusive and exhaustive sub-intervals selected a priori. The procedure still relies on permutation tests, but the number of applications of permutation tests can be reduced in practice if the number of partitions of the domain is low. This method has been applied in a functional ANOVA framework, considering a mono-dimensional domain. Thus, we first have to extend this procedure to a multi-dimensional domain to select sub-regions where significant differences are observed between groups, and then adjust for multiple tests. The extension to a three-dimensional domain appears quite straightforward on a theoretical point of view, and it resembles the extension we proposed in the previous section, with the exception that we do not test for each possible $\mathcal{R}$ (that theoretically could have been defined as having any possible geometrical shape, e.g. cubes, spheres), but only on the predefined subregions of the domain. Moreover, in this case the adjustment for multiplicity is not based on [98], but on the closure principle as described in [7, 8], that provides a control of the family-wise error rate within each region in the weak sense (i.e., the control is guaranteed only when all the null hypothesis are true). The conclusions of the test depend on the initially chosen partition, thus, the application of this procedure is recommended when the domain $\mathcal{S}$ has natural sub-regions, for example in the case of the human brain regions. Further analysis are needed to explore the computational feasibility of both the approaches sketched above, and to provide a method that can guarantee a strong control of the family-wise error rate (i.e., allowing all the null hypotheses to be true or false). These tasks will be addressed in future works, with the aim of applying this methodology on a case study based on real fMRI data.

Bibliography

[1] FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource, https://www.ncbi.nlm.nih.gov/books/NBK326791/ , 2017.

[2] Douglas G. Altman. Statistics in medical journals: Some recent trends. *Statistics in Medicine*, 19(23):3275–3289, 2000.

[3] Lauren D Arnold, Melissa Braganza, Rondek Salih, and Graham A Colditz. Statistical trends in the Journal of the American Medical Association and implications for training across the continuum of medical education. *PloS one*, 8(10):e77301, 2013.

[4] Yasunori Sato, Masahiko Gosho, Kengo Nagashima, Sho Takahashi, James H Ware, and Nan M Laird. Statistical Methods in the Journal An Update. *New England Journal of Medicine*, 376(11):1086–1087, 2017.

[5] Hubert M. Fonteijn, Marc Modat, Matthew J. Clarkson, Josephine Barnes, Manja Lehmann, Nicola Z. Hobbs, Rachael I. Scahill, Sarah J. Tabrizi, Sebastien Ourselin, Nick C. Fox, and Daniel C. Alexander. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. *NeuroImage*, 60(3):1880–1889, 2012.

[6] Alexandra L. Young, Neil P. Oxtoby, Pankaj Daga, David M. Cash, Nick C. Fox, Sebastien Ourselin, Jonathan M. Schott, and Daniel C. Alexander. A data-driven model of biomarker changes in sporadic Alzheimer's disease. *Brain*, 137(9):2564–2577, 2014.

[7] O. A. Vsevolozhskaya, M. C. Greenwood, G. J. Bellante, S. L. Powell, R. L. Lawrence, and K. S. Repasky. Combining functions and the closure principle for performing follow-up tests in functional analysis of variance. *Computational Statistics and Data Analysis*, 67:175–184, 2013.

[8] Olga Vsevolozhskaya, Mark Greenwood, and Dmitri Holodov. Pairwise comparison of treatment levels in functional analysis of variance with application to erythrocyte hemolysis. *Annals of Applied Statistics*, 8(2):905–925, 2014.

[9] Jonathan D. Finder, David Birnkrant, John Carl, Harold J. Farber, David Gozal, Susan T. Iannaccone, Thomas Kovesi, Richard M. Kravitz, Howard Panitch, Craig Schramm, Mary Schroth, Girish Sharma, Lisa Sievers, Jean M. Silvestri, and Laura Sterni. Respiratory care of the patient with duchenne muscular dystrophy: ATS consensus statement. In *American Journal of Respiratory and Critical Care Medicine*, volume 170, pages 456–465, may 2004.

[10] Katharine Bushby, Richard Finkel, David J Birnkrant, Laura E Case, Paula R Clemens, Linda Cripe, Ajay Kaul, Kathi Kinnett, Craig McDonald, Shree Pandya, James Poysky, Frederic Shapiro, Jean Tomezsko, and Carolyn Constantin. Diagnosis and management of Duchenne muscular dystrophy, part 2: implementation of multidisciplinary care, feb 2010.

[11] Katharine Bushby, Richard Finkel, David J Birnkrant, Laura E Case, Paula R Clemens, Linda Cripe, Ajay Kaul, Kathi Kinnett, Craig McDonald, Shree Pandya, James Poysky, Frederic Shapiro, Jean Tomezsko, and Carolyn Constantin. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management, jan 2010.

[12] David J. Birnkrant, Katharine M.D. Bushby, Raouf S. Amin, John R. Bach, Joshua O. Benditt, Michelle Eagle, Jonathan D. Finder, Maninder S. Kalra, John T. Kissel, Anastassios C. Koumbourlis, and Richard M. Kravitz. The respiratory management of patients with Duchenne muscular dystrophy: A DMD care considerations working group specialty article. *Pediatric Pulmonology*, 45(8):739–748, jul 2010.

[13] Erik K. Henricson, R. Ted Abresch, Avital Cnaan, Fengming Hu, Tina Duong, Adrienne Arrieta, Jay Han, Diana M. Escolar, Julaine M. Florence, Paula R. Clemens, Eric P. Hoffman, and Craig M. Mcdonald. The cooperative international neuromuscular research group Duchenne natural history study: Glucocorticoid treatment preserves clinically meaningful functional milestones and reduces rate of disease progression as measured by manual muscle testing and othe. *Muscle and Nerve*, 48(1):55–67, jul 2013.

[14] Craig M. Mcdonald, Erik K. Henricson, R. Ted Abresch, Jay J. Han, Diana M. Escolar, Julaine M. Florence, Tina Duong, Adrienne Arrieta, Paula R. Clemens, Eric P. Hoffman, and Avital Cnaan. The Cooperative International Neuromuscular Research Group Duchenne Natural History Study-A Longitudinal Investigation In The Era Of Glucocorticoid Therapy: Design Of Protocol And The Methods Used. *Muscle and Nerve*, 48(1):32–54, jul 2013.

[15] Katharine Bushby, Richard Finkel, Brenda Wong, Richard Barohn, Craig Campbell, Giacomo P. Comi, Anne M. Connolly, John W. Day, Kevin M. Flanigan, Nathalie Goemans, Kristi J. Jones, Eugenio Mercuri, Ros Quinlivan, James B. Renfroe, Barry Russman, Monique M. Ryan, Mar Tulinius, Thomas Voit, Steven A. Moore, H. Lee Sweeney, Richard T. Abresch, Kim L. Coleman, Michelle Eagle, Julaine Florence, Eduard Gappmaier, Allan M. Glanzman, Erik Henricson, Jay Barth, Gary L. Elfring, Allen Reha, Robert J. Spiegel, Michael W. O'donnell, Stuart W. Peltz, and Craig M. Mcdonald. Ataluren treatment of patients with nonsense mutation dystrophinopathy. *Muscle and Nerve*, 50(4):477–487, oct 2014.

[16] Jerry R Mendell, Nathalie Goemans, Linda P Lowes, Lindsay N Alfano, Katherine Berry, James Shao, Edward M Kaye, and Eugenio Mercuri. Longitudinal effect of eteplirsen versus historical control on ambulation in Duchenne muscular dystrophy. *Annals of Neurology*, 79(2):257–271, feb 2016.

[17] Gunnar M Buyse, Thomas Voit, Ulrike Schara, Chiara S M Straathof, M Grazia D'Angelo, Günther Bernert, Jean Marie Cuisset, Richard S Finkel, Nathalie Goemans, Craig M McDonald, Christian Rummey, and Thomas Meier. Efficacy of idebenone on respiratory function in patients with Duchenne muscular dystrophy not using glucocorticoids (DELOS): A double-blind randomised placebo-controlled phase 3 trial. *The Lancet*, 385(9979):1748–1757, may 2015.

[18] Sonia Khirani, Adriana Ramirez, Guillaume Aubertin, Michèle Boulé, Chrystelle Chemouny, Véronique Forin, and Brigitte Fauroux. Respiratory muscle decline in duchenne muscular dystrophy. *Pediatric Pulmonology*, 49(5):473–481, may 2014.

[19] Margaret F Phillips, Rosaline C M Quinlivan, Richard H T Edwards, and Peter M A Calverley. Changes in spirometry over time as a prognostic marker in patients with duchenne muscular dystrophy. *American Journal of Respiratory and Critical Care Medicine*, 164(12):2191–2194, dec 2002.

[20] Andreas Hahn, John R Bach, Anne Delaubier, Alice Renardel-Irani, Claire Guillou, and Yves Rideau. Clinical implications of maximal respiratory pressure determinations for individuals with Duchenne muscular dystrophy. *Archives of Physical Medicine and Rehabilitation*, 78(1):1–6, jan 1997.

[21] Jerome Gayraud, Michele Ramonatxo, François Rivier, Véronique Humberclaude, Basil Petrof, and Stefan Matecki. Ventilatory parameters and maximal respiratory pressure changes with age in duchenne muscular dystrophy patients. *Pediatric Pulmonology*, 45(6):552–559, jun 2010.

[22] S. E. Tangsrud, I L Petersen, K. C. LØdrup Carlsen, and K H Carlsen. Lung function in children with Duchenne's muscular dystrophy. *Respiratory Medicine*, 95(11):898–903, nov 2001.

[23] Véronique Humbertclaude, Dalil Hamroun, Kamel Bezzou, Carole Bérard, Odile Boespflug-Tanguy, Christine Bommelaer, Emmanuelle Campana-Salort, Claude Cances, Brigitte Chabrol, Marie Christine Commare, Jean Marie Cuisset, Capucine De Lattre, Claude Desnuelle, Bernard Echenne, Cécile Halbert, Olivier Jonquet, Annick Labarre-Vila, Marie Ange N'Guyen-Morel, Michel Pages, Jean Louis Pepin, Thierry Petitjean, Jean Pouget, Elisabeth Ollagnon-Roman, Christian Richelme, Franois Rivier, Sabrina Sacconi, Vincent Tiffreau, Carole Vuillerot, Marie Christine Picot, Mireille Claustres, Christophe

Béroud, and Sylvie Tuffery-Giraud. Motor and respiratory heterogeneity in Duchenne patients: Implication for clinical trials. *European Journal of Paediatric Neurology*, 16(2):149–160, mar 2012.

[24] O H Mayer, R S Finkel, C Rummey, M J Benton, A M Glanzman, J Flickinger, B. M. Lindstrom, and T Meier. Characterization of pulmonary function in Duchenne Muscular Dystrophy. *Pediatric Pulmonology*, 50(5):487–494, may 2015.

[25] Rolando Roberto, Anto Fritz, Yolanda Hagar, Braden Boice, Andrew Skalsky, Hosun Hwang, Laurel Beckett, Craig McDonald, and Munish Gupta. The Natural History of Cardiac and Pulmonary Function Decline in Patients With Duchenne Muscular Dystrophy. *Spine*, 36(15):E1009–E1017, jul 2011.

[26] Gunnar M Buyse, Nathalie Goemans, Marleen Van Den Hauwe, and Thomas Meier. Effects of glucocorticoids and idebenone on respiratory function in patients with duchenne muscular dystrophy. *Pediatric Pulmonology*, 48(9):912–920, sep 2013.

[27] W Douglas Biggar, Michele Gingras, Darcy L Fehlings, Vivien A Harris, and Catherine A Steele. Deflazacort treatment of Duchenne muscular dystrophy. *Journal of Pediatrics*, 138(1):45–50, jan 2001.

[28] W D Biggar, V A Harris, L Eliasoph, and B Alman. Long-term benefits of deflazacort treatment for boys with Duchenne muscular dystrophy in their second decade. *Neuromuscular Disorders*, 16(4):249–255, apr 2006.

[29] Richard T Moxley, Shree Pandya, Emma Ciafaloni, Deborah J Fox, and Kim Campbell. Change in Natural History of Duchenne Muscular Dystrophy With Long-term Corticosteroid Treatment: Implications

for Management. *Journal of Child Neurology*, 25(9):1116–1129, sep 2010.

[30] J R Mendell, R T Moxley, R C Griggs, M H Brooke, G M Fenichel, J P Miller, W King, L Signore, S Pandya, and J Florence. Randomized, double-blind six-month trial of prednisone in Duchenne's muscular dystrophy. *The New England journal of medicine*, 320(24):1592–7, jun 1989.

[31] G.M. Fenichel, J.M. Florence, A Pestronk, J.R. Mendell, R.T. Moxley, R.C. Griggs, M.H. Brooke, J.P. Miller, J Robison, W King, L. Signore, S. Pandya, J. Schierbecker, and B. Wilson. Long-term benefit from prednisone therapy in Duchenne muscular dystrophy. *NEUROLOGY*, 41(12):1874–1877, dec 1991.

[32] R C Griggs, R T Moxley, J R Mendell, G M Fenichel, M H Brooke, A Pestronk, and J P Miller. Prednisone in Duchenne dystrophy. A randomized, controlled trial defining the time course and dose response. Clinical Investigation of Duchenne Dystrophy Group. *Archives of neurology*, 48(4):383–8, apr 1991.

[33] Anne M Connolly, Jeanine Schierbecker, Renee Renna, and Julaine Florence. High dose weekly oral prednisone improves strength in boys with Duchenne muscular dystrophy. *Neuromuscular Disorders*, 12(10):917–925, dec 2002.

[34] F Shapiro, D Zurakowski, T Bui, and B T Darras. Progression of spinal deformity in wheelchair-dependent patients with Duchenne muscular dystrophy who are not treated with steroids: Coronal plane (scoliosis) and sagittal plane (kyphosis, lordosis) deformity. *Bone and Joint Journal*, 96 B(1):100–105, jan 2014.

[35] J D Hsu. The development of current approaches to the management of spinal deformity for patients with neuromuscular disease. *Seminars in neurology*, 15(1):24–8, mar 1995.

[36] Leanne M. Gauld and Alison Boynton. Relationship between peak cough flow and spirometry in Duchenne muscular dystrophy. *Pediatric Pulmonology*, 39(5):457–460, may 2005.

[37] M Toussaint, M Chatwin, and P Soudon. Review Article: Mechanical ventilation in Duchenne patients with chronic respiratory insufficiency: clinical implications of 20 years published experience. *Chronic Respiratory Disease*, 4(3):167–177, aug 2007.

[38] American Thoracic Society/European Respiratory Society. ATS/ERS Statement on respiratory muscle testing. *American journal of respiratory and critical care medicine*, 166(4):518–624, aug 2002.

[39] M R Miller, R Crapo, J Hankinson, V Brusasco, F Burgos, R Casaburi, A Coates, P Enright, C P M van der Grinten, P Gustafsson, R Jensen, D C Johnson, N MacIntyre, R McKay, D Navajas, O F Pedersen, R Pellegrino, G Viegi, and Wanger J. General considerations for lung function testing. *European Respiratory Journal*, 26(1):153–161, jul 2005.

[40] M. R. Miller, J Hankinson, Vito Brusasco, F Burgos, R Casaburi, A Coates, R Crapo, P Enright, C. P.M. van der Grinten, P Gustafsson, R Jensen, D C Johnson, N. MacIntrye, R McKay, D Navajas, O F Pedersen, R Pellegrino, G Viegi, and J. Wagner. Standardisation of spirometry, aug 2005.

[41] A. Lo Mauro, Maria Grazia D'Angelo, Marianna Romei, F. Motta, D. Colombo, G. P. Comi, A. Pedotti, E. Marchi, A. C. Turconi,

N. Bresolin, and Andrea Aliverti. Abdominal volume contribution to tidal volume as an early indicator of respiratory impairment in Duchenne muscular dystrophy. *European Respiratory Journal*, 35(5):1118–1125, apr 2010.

[42] M. Romei, M. G. D'Angelo, A. Lomauro, S. Gandossini, S. Bonato, E. Brighina, E. Marchi, G. P. Comi, A. C. Turconi, A. Pedotti, N. Bresolin, and A. Aliverti. Low abdominal contribution to breathing as daytime predictor of nocturnal desaturation in adolescents and young adults with Duchenne Muscular Dystrophy. *Respiratory Medicine*, 106(2):276–283, feb 2012.

[43] Antonella LoMauro, Marianna Romei, Maria Grazia D'Angelo, and Andrea Aliverti. Determinants of cough efficiency in Duchenne muscular dystrophy. *Pediatric pulmonology*, 49(4):357–65, apr 2014.

[44] PH Quanjer, GJ Tammeling, JE Cotes, OF Pedersen, R Peslin, and JC Yernault. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl.*, 16:5–40, mar 1993.

[45] S J et Al. Cala. Chest wall and lung volume estimation by optical reflectance motion analysis. *Journal of applied physiology*, 81(6):2680–2689, dec 1996.

[46] Andrea Aliverti, Raffaele Dellacà, Paolo Pelosi, Davide Chiumello, Luciano Gattinoni, and Antonio Pedotti. Compartmental analysis of breathing in the supine and prone positions by optoelectronic plethysmography. *Annals of Biomedical Engineering*, 29(1):60–70, jan 2001.

[47] Darlene L Machado, Elaine C Silva, Maria B D Resende, Celso R F Carvalho, Edmar Zanoteli, and Umbertina C Reed. Lung function

monitoring in patients with duchenne muscular dystrophy on steroid therapy. *BMC Research Notes*, 5(1):435, aug 2012.

[48] R. T. Moxley, S. Ashwal, S. Pandya, A. Connolly, J. Florence, K. Mathews, L. Baumbach, C. McDonald, M. Sussman, and C. Wade. Practice Parameter: Corticosteroid treatment of Duchenne dystrophy: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. *Neurology*, 64(1):13–20, jan 2005.

[49] J A Rice and C O Wu. Nonparametric mixed effects models for unequally sampled noisy curves. *Biometrics*, 57(1):253–259, mar 2001.

[50] Lloyd J. Edwards, Paul W. Stewart, James E. MacDougall, and Ronald W. Helms. A method for fitting regression splines with varying polynomial order in the linear mixed model. *Statistics in Medicine*, 25(3):513–527, feb 2006.

[51] P. Vitali, E. MacCagnano, E. Caverzasi, R. G. Henry, A. Haman, C. Torres-Chae, D. Y. Johnson, B. L. Miller, and M. D. Geschwind. Diffusion-weighted MRI hyperintensity patterns differentiate CJD from other rapid dementias. *Neurology*, 76(20):1711–1719, may 2011.

[52] Marc Manix, Piyush Kalakoti, Miriam Henry, Jai Thakur, Richard Menger, Bharat Guthikonda, and Anil Nanda. Creutzfeldt-Jakob disease: updated diagnostic criteria, treatment algorithm, and the utility of brain biopsy. *Neurosurgical Focus*, 39(5):E2, nov 2015.

[53] Simona Gaudino, Emma Gangemi, Raffaella Colantonio, Annibale Botto, Emanuela Ruberto, Rosalinda Calandrelli, Matia Martucci, Maria Gabriella Vita, Carlo Masullo, Alfonso Cerase, and Cesare Colosimo. Neuroradiology of human prion diseases, diagnosis and differential diagnosis. *La radiologia medica*, 122(5):369–385, may 2017.

[54] Geoffrey S Young, Michael D Geschwind, Nancy J Fischbein, Jennifer L Martindale, Roland G Henry, Songling Liu, Ying Lu, Stephen Wong, Hong Liu, Bruce L Miller, and William P Dillon. Diffusion-weighted and fluid-attenuated inversion recovery imaging in Creutzfeldt-Jakob disease: High sensitivity and specificity for diagnosis. *American Journal of Neuroradiology*, 26(6):1551–1562, 2005.

[55] B. Meissner, K. Kallenberg, P. Sanchez-Juan, A. Krasnianski, U. Heinemann, D. Varges, M. Knauth, and I. Zerr. Isolated Cortical Signal Increase on MR Imaging as a Frequent Lesion Pattern in Sporadic Creutzfeldt-Jakob Disease. *American Journal of Neuroradiology*, 29(8):1519–1524, aug 2008.

[56] Gianfranco Puoti, Alberto Bizzi, Gianluigi Forloni, Jiri G. Safar, Fabrizio Tagliavini, and Pierluigi Gambetti. Sporadic human prion diseases: Molecular insights and diagnosis. *The Lancet Neurology*, 11(7):618–628, 2012.

[57] Ryuichiro Atarashi, Katsuya Satoh, Kazunori Sano, Takayuki Fuse, Naohiro Yamaguchi, Daisuke Ishibashi, Takehiro Matsubara, Takehiro Nakagaki, Hitoki Yamanaka, Susumu Shirabe, Masahito Yamada, Hidehiro Mizusawa, Tetsuyuki Kitamoto, Genevieve Klug, Amelia McGlade, Steven J Collins, and Noriyuki Nishida. Ultrasensitive human prion detection in cerebrospinal fluid by real-time quaking-induced conversion. *Nature Medicine*, 17(2):175–178, feb 2011.

[58] Christina D Orrú, Matilde Bongianni, Giovanni Tonoli, Sergio Ferrari, Andrew G Hughson, Bradley R Groveman, Michele Fiorini, Maurizio Pocchiari, Salvatore Monaco, Byron Caughey, and Gianluigi Zanusso. A Test for CreutzfeldtJakob Disease Using Nasal Brushings. *New England Journal of Medicine*, 371(6):519–529, aug 2014.

[59] Pierluigi Gambetti, Ignazio Cali, Silvio Notari, Qingzhong Kong, Wen Quan Zou, and Witold K. Surewicz. Molecular biology and pathology of prion strains in sporadic human prion diseases. *Acta Neuropathologica*, 121(1):79–90, jan 2011.

[60] Piero Parchi, Laura De Boni, Daniela Saverioni, Mark L. Cohen, Isidro Ferrer, Pierluigi Gambetti, Ellen Gelpi, Giorgio Giaccone, Jean Jacques Hauw, Romana Höftberger, James W. Ironside, Casper Jansen, Gabor G. Kovacs, Annemieke Rozemuller, Danielle Seilhean, Fabrizio Tagliavini, Armin Giese, and Hans A. Kretzschmar. Consensus classification of human prion disease histotypes allows reliable identification of molecular subtypes: An inter-rater study among surveillance centres in Europe and USA. *Acta Neuropathologica*, 124(4):517–529, 2012.

[61] Piero Parchi, Armin Giese, Sabina Capellari, Paul Brown, Walter Schulz-Schaeffer, Otto Windl, Inga Zerr, Herbert Budka, Nicolas Kopp, Pedro Piccardo, Sigrid Poser, Amyn Rojiani, Nathalie Streichemberger, Jean Julien, Claude Vital, Bernardino Ghetti, Pierluigi Gambetti, and Hans Kretzschmar. Classification of Sporadic Creutzfeldt-Jakob Disease Based on Molecular and Phenotypic Analysis of 300 Subjects. *Ann Neurol*, 46:224–233, 1999.

[62] Aaron Foutz, Brian S. Appleby, Clive Hamlin, Xiaoqin Liu, Sheng Yang, Yvonne Cohen, Wei Chen, Janis Blevins, Cameron Fausett, Han Wang, Pierluigi Gambetti, Shulin Zhang, Andrew Hughson, Curtis Tatsuoka, Lawrence B. Schonberger, Mark L. Cohen, Byron Caughey, and Jiri G. Safar. Diagnostic and prognostic value of human prion detection in cerebrospinal fluid. *Annals of Neurology*, 81(1):79–92, jan 2017.

[63] A. Krasnianski, K. Kallenberg, D. A. Collie, B. Meissner, W. J. Schulz-Schaeffer, U. Heinemann, D. Varges, D. M. Summers, H. A. Kretzschmar, T. Talbot, R. G. Will, and I. Zerr. MRI the classical MM1 and the atypical MV2 subtypes of sporadic CJD: An inter-observer agreement study. *European Journal of Neurology*, 15(8):762–771, 2008.

[64] B. Meissner, K. Kallenberg, P. Sanchez-Juan, D. Collie, D. M. Summers, S. Almonti, S. J. Collins, P. Smith, P. Cras, G. H. Jansen, J. P. Brandel, M. B. Coulthart, H. Roberts, B. Van Everbroeck, D. Galanaud, V. Mellina, R. G. Will, and I. Zerr. MRI lesion profiles in sporadic Creutzfeldt-Jakob disease. *Neurology*, 72(23):1994–2001, 2009.

[65] Matteo Figini, Daniel C Alexander, Veronica Redaelli, Fabrizio Fasano, Marina Grisoli, Giuseppe Baselli, Pierluigi Gambetti, Fabrizio Tagliavini, and Alberto Bizzi. Mathematical models for the diffusion magnetic resonance signal abnormality in patients with prion diseases. *NeuroImage: Clinical*, 7:142–154, 2015.

[66] Raffaele Lodi, Piero Parchi, Caterina Tonon, David Manners, Sabina Capellari, Rosaria Strammiello, Rita Rinaldi, Claudia Testa, Emil Malucelli, Barbara Mostacci, Giovanni Rizzo, Giulia Pierangeli, Pietro Cortelli, Pasquale Montagna, and Bruno Barbiroli. Magnetic resonance diagnostic markers in clinically sporadic prion disease: A combined brain magnetic resonance imaging and spectroscopy study. *Brain*, 132(10):2669–2679, 2009.

[67] Alessia Franceschini, Simone Baiardi, Andrew G. Hughson, Neil Mckenzie, Fabio Moda, Marcello Rossi, Sabina Capellari, Alison Green, Giorgio Giaccone, Byron Caughey, and Piero Parchi. High

diagnostic value of second generation CSF RT-QuIC across the wide spectrum of CJD prions. *Scientific Reports*, 7(1):10655, dec 2017.

[68] Clive Hamlin, Gianfranco Puoti, Sally Berri, Elliott Sting, Carrie Harris, Mark Cohen, Charles Spear, Alberto Bizzi, Sara M. Debanne, and Douglas Y. Rowland. A comparison of tau and 14-3-3 protein in the diagnosis of Creutzfeldt-Jakob disease. *Neurology*, 79(6):547–552, 2012.

[69] Christina D. Orrù, Bradley R Groveman, Andrew G Hughson, Gianluigi Zanusso, Michael B Coulthart, and Byron Caughey. Rapid and sensitive RT-QuIC detection of human creutzfeldt-jakob disease using cerebrospinal fluid. *mBio*, 6(1):e02451–14, jan 2015.

[70] C. J. Clopper and E. S. Pearson. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika*, 26(4):404, dec 1934.

[71] Robert G Newcombe. Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Statistics in Medicine*, 17(8):857–872, apr 1998.

[72] Kenneth O. McGraw and S. P. Wong. Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1):30–46, 1996.

[73] Kevin A Hallgren. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1):23–34, 2012.

[74] Domenic V. Cicchetti. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4):284–290, 1994.

[75] Zbynek Sidak. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *Journal of the American Statistical Association*, 62(318):626, jun 1967.

[76] S. Paul Wright. Adjusted P-Values for Simultaneous Inference. *Biometrics*, 48(4):1005, dec 1992.

[77] Leo Breiman, Jerome Friedman, Charles J. Stone, and R.A. Olshen. *Classification and regression trees Regression trees*. The Wadsworth and Brooks-Cole statistics-probability series. Taylor & Francis, 1984.

[78] Yuka Terasawa, Koji Fujita, Yuishin Izumi, and Ryuji Kaji. Early detection of familial Creutzfeldt-Jakob disease on diffusion-weighted imaging before symptom onset. *Journal of the Neurological Sciences*, 319(1-2):130–132, aug 2012.

[79] Simone Baiardi, Anna Magherini, Sabina Capellari, Veronica Redaelli, Anna Ladogana, Marcello Rossi, Fabrizio Tagliavini, Maurizio Pocchiari, Giorgio Giaccone, and Piero Parchi. Towards an early clinical diagnosis of sporadic CJD VV2 (ataxic type). *Journal of Neurology, Neurosurgery & Psychiatry*, pages 1–9, 2017.

[80] Diego Cardoso Fragoso, Augusto Lio da Mota Gonçalves Filho, Felipe Torres Pacheco, Bernardo Rodi Barros, Ingrid Aguiar Littig, Renato Hoffmann Nunes, Antônio Carlos Martins Maia Júnior, and Antonio J. da Rocha. Imaging of Creutzfeldt-Jakob Disease: Imaging Patterns and Their Differential Diagnosis. *RadioGraphics*, 37(1):234–257, 2017.

[81] Piero Parchi, Rosaria Strammiello, Armin Giese, and Hans Kretzschmar. Phenotypic variability of sporadic human prion disease and its molecular basis: Past, present, and future, jan 2011.

[82] Piero Parchi and Daniela Saverioni. Molecular pathology, classification, and diagnosis of sporadic human prion disease variants. *Folia Neuropathologica*, 50(1):20–45, 2012.

[83] Laura Eisenmenger, Marie-Claire Porter, Christopher J Carswell, Andrew Thompson, Simon Mead, Peter Rudge, John Collinge, Sebastian Brandner, Hans R Jäger, and Harpreet Hyare. Evolution of Diffusion-Weighted Magnetic Resonance Imaging Signal Abnormality in Sporadic Creutzfeldt-Jakob Disease, With Histopathological Correlation. *JAMA Neurology*, 73(1):76, jan 2016.

[84] So Young Park, Min Jeong Wang, Jae Won Jang, Young Ho Park, Jae Sung Lim, Young Chul Youn, Jungeun Kim, and Sangyun Kim. The Clinical Stages of Sporadic Creutzfeldt-Jakob Disease with Met/Met Genotype in Korean Patients. *European Neurology*, 75(5-6):213–222, 2016.

[85] Federico Caobelli, Milena Cobelli, Claudio Pizzocaro, Marco Pavia, Silvia Magnaldi, and Ugo Paolo Guerra. The Role of Neuroimaging in Evaluating Patients Affected by Creutzfeldt-Jakob Disease: A Systematic Review of the Literature, jan 2015.

[86] Piero Parchi, Rosaria Strammiello, Silvio Notari, Armin Giese, Jan P M Langeveld, Anna Ladogana, Inga Zerr, Federico Roncaroli, Patrich Cras, Bernardino Ghetti, Maurizio Pocchiari, Hans Kretzschmar, and Sabina Capellari. Incidence and spectrum of sporadic Creutzfeldt-Jakob disease variants with mixed phenotype and co-occurrence of PrPSc types: An updated classification. *Acta Neuropathologica*, 118(5):659–671, 2009.

[87] Iván Fernández-Vega, Javier Ruiz-Ojeda, Ramon A Juste, Maria Geijo, Juan Jose Zarranz, Jose Luis Sánchez Menoyo, Ikerne Vicente-

Etxenausia, Jennifer Mediavilla-García, and Isabel Guerra-Merino. Coexistence of mixed phenotype Creutzfeldt-Jakob disease, Lewy body disease and argyrophilic grain disease plus histological features of possible Alzheimer's disease: A multi-protein disorder in an autopsy case. *Neuropathology*, 35(1):56–63, feb 2015.

[88] Neil P. Oxtoby and Daniel C. Alexander. Imaging plus X. *Current Opinion in Neurology*, 30(4):371–379, 2017.

[89] Ashish Raj, Amy Kuceyeski, and Michael Weiner. A Network Diffusion Model of Disease Progression in Dementia. *Neuron*, 73(6):1204–1215, mar 2012.

[90] Neil P Oxtoby, Sara Garbarino, Nicholas Christopher Firth, Jason D Warren, Jonathan M Schott, and Daniel C Alexander. Data Driven Sequence of Changes to Anatomical Brain Connectivity in Sporadic Alzheimer's Disease. *Frontiers in Neurology*, 8:580, 2017.

[91] Bess Frost and Marc I. Diamond. Prion-like mechanisms in neurodegenerative diseases. *Nature Reviews Neuroscience*, 11(3):155–9, dec 2009.

[92] Stanley B Prusiner. A Unifying Role for Prions in Neurodegenerative Diseases. *Science*, 336(6088):1511–1513, jun 2012.

[93] Yasser Iturria-Medina, Vladimir Hachinski, and Alan C. Evans. The vascular facet of late-onset Alzheimers disease. *Current Opinion in Neurology*, page 1, oct 2017.

[94] Sanjay Mittal, Peter Farmer, Peter Kalina, Peter B Kingsley, and John Halperin. Correlation of diffusion-weighted magnetic resonance imaging with neuropathology in Creutzfeldt-Jakob disease. *Archives of neurology*, 59(1):128–134, jan 2002.

[95] Stéphane Haïk, Didier Dormont, Baptiste A. Faucheux, Claude Marsault, and Jean Jacques Hauw. Prion protein deposits match magnetic resonance imaging signal abnormalities in Creutzfeldt-Jakob disease. *Annals of Neurology*, 51(6):797–799, jun 2002.

[96] J O Ramsay and B W Silverman. *Functional data analysis*. Springer, New York, 2005.

[97] Konrad Abramowicz, Charlotte Häger, Alessia Pini, Lina Schelin, Sara Sjöstedt De Luna, and Simone Vantini. Nonparametric inference for functional-on-scalar linear models applied to knee kinematic hop data after injury of the anterior cruciate ligament. *MOX-Report No. 30/2016, MOX, Politecnico di Milano*.

[98] Alessia Pini and Simone Vantini. Interval-wise testing for functional data. *Journal of Nonparametric Statistics*, 29(2):407–424, apr 2017.