Genotype-phenotype correlations in valosin-containing protein disease : a retrospective muticentre study

Genotype-phenotype correlations in Valosin Containing Protein Disease: results of an International Multicentre Study

Marianela Schiava[1], Chiseko Ikenaga[2], Rocío Nur Villar-Quiles[3], Marta Caballero[4], Virginia Kimonis[5], Tanya Stojkovic [3], Conrad Weihl[6] and Jordi Díaz-Manera[1,7].

On Behalf of the **VCP International Group**.

1. John Walton Muscular Dystrophy Research Centre. Newcastle University. UK.

2. Johns Hopkins University School of Medicine. USA.

3. Hôpital Pitié-Salpêtrière. Sorbonne Université. Institut de Myologie. Centre de référence des Maladies Neuromusculaires. Paris. France.

4. Neuromuscular Disorders Unit. Neurology Department. Hospital de la Santa Creu i Sant Pau. Barcelona. Spain.

5. Division of Genetic and Genomic Medicine, Department of Pediatrics, University of California, Irvine, CA USA

6. Department of Neurology, Washington University in St. Louis. St. Louis, Missouri, USA

7. Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau de Barcelona.

Corresponding Author:

Prof. Jordi Díaz-Manera MD, PhD. John Walton Muscular Dystrophy Research Centre

Newcastle University Translational and Clinical Research Institute Central Parkway, NE1 3BZ

Newcastle Upon Tyne, UK. Phone: +44(0)1912418757

E-mail:jordi.diaz-manera@newcastle.ac.uk

The VCP International Study Group by countries

Australia

1. Phillipa Lamont

Austria

1. Hakan Cetin and Sigrid Klotz

Belgium

1. Kristl G. Claeys
2. Jan De Bleecker
3. Alicia Alonso, Jonathan Baets, Peter De Jonghe and Willem De Ridder

Brasil

1. Edmar Zanotelli and André Macedo

Canada

1. Jodi Warman

Chile

1. Jorge Alfredo Bevilacqua Rivas, Mario Campero and Nicholas Earle

China

1. Sushan Luo, Wenhua Zhu and Jie Lin

Findland

1. Bjarne Udd, Johanna Palmio

France

1. Aleksandra Nadaj-Pakleza, Jean-Baptiste Chanson, Beatrice Lannes
2. Edoardo Malfatti
3. Pascal Laforet
1. Norma Romero, Teresinha Evagelista, Leonard-Louis Sarah, Bruno Eymard, Guillaume Bassez, Anthony Behin, Marion Masingue, Pascal Richard and Corinne Metay
2. Sarah Souvannanorath, Jean François Deux

Germany

3. Dr Sabine Krause, Prof Dr Benedikt Schöser, Dr Natalia Garcia Angarita, Simone Thiele and Eva Greckl.
4. Matthias Vorgerda and Rudi Kley

Greece

1. George Papadimas and Constantinos Papadopoulos

Hungary

1. Kinga Hadzsiev and Endre Pál

India

1. Sruthi S Nair

Italy

1. Antonio Toscano and Carmelo Rodolico
2. Elena Pegoraro, Luca Bello, Gianni Sorarù and Pietro Riguzzi
3. Giorgio Tasca and Mauro Monforte, Mario Sabatelli and Giulia Bisogni

Japan

1. Ichizo Nishino/ Michio Inoue/ Yukako Nishimori/ Terunori Sano and Yasushi Oya

Netherlands

1. Umesh Badrising
2. Marianne Devisser

Poland

1. Anna Kostera -Pruszczy, Dr Biruta Kierdaszuk and Anna Kaminska

Serbia

1. Stojan Peric, Vidosava Rakocevic-Stojanovic and Zorica Stevic.

South Corea

1. Jin-Hong Shin and Young-Eun Park

Spain

1. Adolfo Lopez De Mundain Arregui/ Roberto Fernández Torron
2. Alba Ramos Fransi/ Alicia Martinez/Giusseppe Lucente
3. Carmen Paradas/ Beatriz Vélez
4. Cristina Dominguez/ Aurelio Hernández Laín
5. Francesc Miralles
6. Jorge Alonso-Pérez/ Rodrigo Alvarez/ Montse Olivé Plana
7. Jorge García García
8. Nuria Muelas / Juan J. Vilches
9. Raúl Domínguez Rubio/ Dr. Velina Nedkova / Monica Povedano / Carla Marco Cazcarra

Sweden

1. Anders Oldfors / Dr Rydelius

UK

1. Ana Topf, Volker Straub, Michela Guglieri, Chiara Marini-Bettolo, Robert Muni Lafora/ Anna Mahewy/ Judith Hutchinson (Newcastle University)
2. Channa Hewamadduma (Sheffield)
3. Timothy L Williams (Newcastle)
4. Kate Taylor /Dr Jill Skeoch (Glasgow)
5. Stefen Bradley (Oxford)
6. Stuart H Ralston, Maria Elena Farrugia (Edinburgh)

USA

1. Lindsay N. Alfano (Ohio)
2. Virginia Kimonis (California)
3. Tom Llyod / William Kelly
4. Grinder Tiffany/ Harrington Elizabeth/ Nance Christopher MD/ Lama Abdel Wahed (Iowa)
5. Hani A. Kushlaf (Cincinatti)
6. Georgios Manousakis
7. Matthew Harms
8. Quinn Colin (Pennsylvania)
9. Chris Weihl

Abstract

Introduction: Valosin-containing protein (VCP) disease is a rare genetic disorder produced by mutations in the *VCP* gene causing a heterogenous presentation including the classic triad of muscle weakness, Paget's disease of bone (PBD) and, Fronto-temporal dementia (FTD). The natural history and genotype - phenotype correlation data available is limited. This study aims to (a) describe the clinical and genetic features of an international cohort of patients with mutations in the VCP gene and (b) investigate potential genotype and phenotype associations.

Methods: Descriptive retrospective study collecting clinical and genetic data from patients with confirmed mutations in the VCP gene in 52 centres from 24 countries.

Results: We included 234 patients (70% males, mean age 55.54 $\pm$ 9.6 years [y]). Mean age at symptom onset 45.6 $\pm$ 9.3 y, mean diagnostic delay 7.74 $\pm$ 6 y, and mean time of disease progression 11.3 $\pm$ 6.9 y. Disease onset was symmetric lower limb weakness in 50% of the patients progressing towards generalized muscle weakness affecting proximal and distal lower and upper limb muscles. Other clinical features included: respiratory symptoms in 40.3%, PBD in 26.7%, dysautonomia in 21.4%, upper and lower motor neuron signs in 13.3% and 21.85%, and FTD in 13.9% of the patient. Fifty-eight genetic variants were identified being the most frequent the c.464G>A, p.Arg155His in 28% of the patients and the c.463C>T, p.Arg155Cys in 11.1%. Twenty new mutations were identified. The c.463C>T, p.Arg155Cys variant had the earliest age of onset (37.8 $\pm$ 7.6 y) among the 4 most frequent variants and a higher frequency of axial weakness, distal upper limb weakness, scapula winging and mix cognitive. 19.1% of the patients were full time wheelchair users and 4.0% (9/225) were bedridden at a median of 8.5 y and 15 y from onset. Thirty–seven patients died at a mean age of 63.9 $\pm$ 8.1 and at a mean of 15.8 $\pm$ 6.6 y from disease onset, 7 due to respiratory insufficiency and 5 due to rapidly progressive dementia. The presence of a FVC< 50% was associated with being full time

wheelchair user/ bedridden and the presence of a FVC<70% and being full-time wheelchair/bedridden were associated with death.

Conclusion: The heterogeneous clinical features of VCP could resemble other neuromuscular conditions. The c.463C>T p.Arg155Cys variant seems to have an earlier age of onset and more severe phenotype. Presence of FVC<50% is a independent risk factor for loss of ambulation and FVC<70% and FTD are independent risk factor for death.

Introduction

Valosin-containing protein (VCP) is an essential AAA+ (ATPases Associated with diverse cellular Activities) protein involved in numerous cellular functions, including protein degradation mediated by the ubiquitin–proteasome system (UPS), autophagosome maturation, endocytic trafficking and regulation of cellular homeostasis and survival through positive or negative modulation of apoptosis and autophagy[1,2,3]. VCP is ubiquitously expressed and its role in mammalian tissues such as neurons and skeletal muscle is emerging.

Hereditary inclusion-body myopathy with Paget's disease of the bone and frontotemporal dementia (IBMPFD) is an autosomal dominant disorder, with variable penetrance of diverse phenotypic features, produced by mutations in the *VCP* gene[4,5]. The number of mutations reported linked to the disease is progressively increasing.[3,6] An estimated prevalence of 0.66/100,000 has been reported in the UK population[7], although, it precise worldwide prevalence is not known. IBMPFD is a progressive disabling disease characterized a triad of clinical phenotypes: myopathy, Paget's disease of the bone[6], and frontotemporal dementia[8,9,10,11]. Previous publications have cited myopathy as the most common clinical feature affecting 80 to 90% of all patients with an estimate age of onset in the 4th to 5th decade involving the pelvic and shoulder girdle muscles[12]. Muscle weakness progresses and involve respiratory muscles leading to death because of respiratory complications or due to end stage dementia[6,13].

A variety of neurological presentations have been described in patients with mutations in the *VCP* gene including fascio-scapulo-humeral muscular weakness[14], distal myopathy[15], amyotrophic lateral sclerosis (ALS)[16], parkinsonism[17], hereditary spastic paraplegia[18], Charcot-Marie-Tooth disease[19] and, Huntington´s disease[20,13] extending the phenotypic spectrum and making the diagnosis complex, especially if patients do not develop the typical triad, they have atypical symptoms at onset or confusing family history. The broad range of symptoms that patients can develop require multiple medical subspecialities to be involved in the care of

these patients. Moreover, the original nomenclature of IBMPFD is insufficient and has led to the use of multisystem proteinopathy (MSP) as a way of encompassing the disparate phenotypes associated with VCP mutations.

Although disparate tissues are affected in MSP patients (e.g. muscle, brain and bone), they are unified by pathologic features that include the accumulation of ubiquitinated proteins, autophagic debris and TDP-43 inclusions. These pathologies are consistent with findings that VCP disease mutations affect a subset of functions resulting in disease pathogenesis. Indeed, VCP disease mutations disrupt its role within the endolysosomal system impairing autophagy and endocytic trafficking. This loss of VCP cellular function contrasts the increase in ATPase activity seen with most disease mutations. A more rapid turnover of ATP with VCP disease mutations likely represents a structural change in VCP resulting in decreased adaptor binding and ultimately a loss of function.

Information regarding VCP disease natural history and phenotypic variability is limited because of the rarity of the entity[12] which requires a large number of patients to obtain accurate, robust and high quality data. Although some clinical series and case reports have been published[7,8–10,19,21] a detailed genotype - phenotype correlation has not been established yet.

In this scenario, a multicentric international descriptive retrospective study was developed. This study aims to (a) describe the clinical and genetic features of an international population with mutations in the *VCP* gene and (b) investigate potential genotype and phenotype associations.

Methods

This was a descriptive retrospective study collecting clinical and genetic data from patients´ notes, obtained on regular clinical care visits, from patients with a genetically confirmed diagnosis of VCP disease. Fifty-two centres from 24 countries world-wide participated in the study and provided information about their patients.

Inclusion criteria for the audit were: (i) patients with a pathogenic (P)/ likely pathogenic (LP) mutation/s as defined by the American College of Medical and Genomic Genetics[22] and as reported in the LOVD database[23] in the *VCP* gene (transcript reference NM_007126.3) and (ii) enough data available in the clinical notes to answer clinical questions about age of disease onset, diagnosis and clinical progression, signs/symptoms at onset or during disease´s progression, ambulatory status and ancillary test results if performed. Patients having a Variant of Uncertain Significance (VUS) were allowed if an *in silico* analysis labelled the variant as P/LP. All genetic variants identified were centrally reviewed and curated by experts in genetics at JWMDC.

Caldicott approval was obtained from the Newcastle upon Tyne Hospitals Register Audit and Institutional Review Boards approvals were obtained from the LMU Klinikum, Ludwing-Maximilians-University, Munich, Germany, Washington University School of Medicine and the Johns Hopkins Hospital in USA.

Data sources

All participating centres completed a survey for each patient followed at their centre. The following data were collected: age of disease onset, age at last assessment, age at genetic diagnosis, DNA mutation and effect on protein, mutation effect (missense, nonsense, frame shift, aberrant splicing, other, non-specified), presence of muscle weakness, pattern of muscle weakness, cardiac impairment, respiratory impairment, cognitive impairment (mild cognitive impairment, Alzheimer disease, FDT), presence of parkinsonism or Parkinson disease, upper

motor neuron and lower motor neuron signs, presence of PDB, neuropathy; ambulatory status; family history and ancillary tests results (spirometry Vital Force Capacity (FVC) absolute value in litres and percentage of predicted, electromyography and nerve conduction velocity, echocardiography left ventricular ejection fraction –LVEF- percentage of predicted, muscle biopsy, muscle MRI, and serum creatin kinase (CK) and alkaline phosphatase values). We also collected data on muscle biopsy and muscle MRI results that will be reported in subsequent papers.

The age of sign/symptom development was recorded as reported by the patient, a relative or the clinician responsible for patient´s care, whoever noticed the sign/symptom first. The age of disease onset (in years) was defined as the age at which the first sign/symptom was noticed. To analyse disease progression the median time from disease onset to ambulation with assistance was recorded defined as ambulation with a stick/cane, wheelchair for outdoor activities, wheelchair for outdoor and indoor activities, or being bedridden. The age of loss of ambulation was defined as the time when patients required a full-time wheelchair for both outdoor and indoor activities or were bedridden.

Cardiac involvement was defined by a left ventricular ejection fraction lower than 55%, the existence of morphological abnormalities in the ventricular walls evaluated by echocardiography or the existence of cardiac conduction defects evaluated by an electrocardiogram.

The use of non-invasive or invasive ventilation assistance at last assessment was recorded as well as the age at which they were prescribed. A severe respiratory impairment was defined as a FVC less than 50% of predicted or the use of invasive ventilation.

In addition, the following disease times in years were collected: age at last clinical assessment (in years as recorded in clinical notes), time of disease progression (difference between the age at last clinical assessment or death and the age of disease onset), age at genetic diagnosis (age

at a genetic test confirming a variant in the *VCP* gene), time to genetic diagnostic (difference between the age at a genetic diagnosis and the age of disease onset), age of death (as recorded in clinical notes) and time to death (difference between the age of death and the age of disease onset).

Statistics

Quantitative variables were analysed using the Kolmogorov-Smirnov or Shapiro-Wilko tests as appropriate to verify their normal distribution. Data was expressed as number and percentage for categorical variables, as mean ± SD and range for quantitative variables following a normal distribution and as median and first and third interquartile for quantitative variables not following a normal distribution. The Chi-squared or Fisher exact tests were used for the association between signs/symptoms and the four most frequent mutation types. Bonferroni correction for multiple comparisons was applied. Analysis of covariance, with adjustment for age at last assessment, was used to compare means of signs/symptoms onset among the 4 most frequent mutation types.

We performed a two-step analysis to select which variables were associated with being a full-time wheelchair user or death. First, a Chi-squared or Fisher exact test as appropriate was used for the association between signs/symptoms and each outcome. Pearson correlation was used to examine the correlations between age of full-time wheelchair user or age of death with the ages of the variables associated with those outcomes. Second, those variables that showed a significantly different distribution among groups (considered as $p<0.05$) were included in a binary logistic regression analysis (using the enter method). Following that, a Cox proportional hazard regression model (enter method) was performed to identify the variables associated with a risk of being a full-time wheelchair user or death. Finally, a Kaplan Meier survival analysis was carried to determine the time to being a full-time wheelchair user or death by the

variables identified through the previous analysis. The level of significance allowed was $p<0.05$. Statistics analysis was performed using SPSS software version 20 from IBM.

Data availability

The data that support the findings of this audit are available from the corresponding author, upon reasonable request.

Results

A total of 255 patients from 194 families from 24 countries were collected. Twenty-one patients were excluded from the analysis. Fifteen patients did not have any symptoms at last assessment, 3 had insufficient clinical information, 2 patients were duplicated and 1 had a homozygous presentation[24]. Non-symptomatic individuals were relatives of patients with symptoms that were diagnosed for genetic counselling. Following exclusion, 234 symptomatic patients were included in the analysis. Seventy percent of the population (163/234) were males and the mean age at enrolment was 56.7 (SD 9.6) years old. Demographics and the times to disease milestones are described in Table 1. The number of patients, families and variants by country are described in eTable 1

Genetic variants

Fifty-eight different variants in the *VCP* gene were reported. All patients were heterozygous for one genetic variant only. All variants were single nucleotide changes except for one small deletion-insertion (c.431_432delGAinsAC, p.Arg144His). Most of the variants were in exon 5 (63% of the patients, 147/234 and 30% of the variants, 17/57) and exon 3 (14% of the patients, 32/234 and 19% of the variants, 11/57). The four most frequent variants identified were: c.464G>A, p.Arg155His (28.6%, 67/234); c.463C>T, p.Arg155Cys (11.1%, 26/234); c.476G>A, p.Arg159His (7.7%, 18/234); and c.277C>T, p.Arg93Cys (7.3%, 17/234). Nineteen variants not previously clinically characterized were identified and will be described carefully in an upcoming manuscript (Schiava et al, in preparation). The list of variants, their frequency and location are described in eTable 2.

Clinical features

Data regarding the symptom at onset was available for 226 patients. Muscle weakness was reported as first symptom in 91% (205/226) of patients. Disease onset was symmetric lower

limb impairment in 50% of the cases (113/226) followed by either a proximal symmetric upper limb weakness (9.3%, 21/226) or a combination of upper and lower limb girdle weakness (9.3%, 21/226, Fig 1). Eight percent (18/234) of the patients showed asymmetric weakness either in the upper or the lower limbs at the beginning of symptoms. Disease onset was slowly progressive in 89% (196/220) of the patients and subacute in the remaining 11% (24/220). Mean CK value at onset (n=37) was 2540 UI/L (IQ1 199.0 – IQ3410.5; min 48 – max 1822).

The frequency of signs and symptoms identified at last assessment is shown in Figure 2. The median time of development of each clinical feature from disease onset is shown in Figure 3. Age of patients at each clinical feature is shown in eTable 3.

At last assessment, all patients had muscle weakness except for one patient who manifested with isolated PDB. The final pattern of muscle involvement consisted of a generalized muscle weakness with proximal and distal lower and upper limb weakness. Scapular winging or axial weakness were reported in half of the population.

Paget Disease of the Bone (PDB) was the most frequent symptom (28.2%, 64/224) after muscle weakness. Among patients with PDB, 77% (49/64) had bone lesions compatible with the diagnosis detected on a bone radiography. The location of the bone lesions was available in 43 of the PDB patients: dorsal and lumbar spine (29/43), hip (15/43), pelvic bone (15/43), skull (11/43) and femur (6/43). An elevated serum alkaline phosphatase levels (ALP) was identified in 55% (35/64) of patients. The median ALP value was 153.50 UI/L (IQ1 80.75 – IQ3 364.25; minimum 39; maximum 1901). Bone pain was reported in the 48% (31/64) of the cases.

The presence of cognitive impairment was identified in 25.5% (59/231) of patients with FTD being the most frequent pattern followed by mixed cognitive impairment. Only one patient was clinically diagnosed with Alzheimer disease, a 62-year old Caucasian man carrying the variant c.271A>C, p.Asn91His. Thirty-two patients were also diagnosed of depression, with 16/32 unrelated to an associated to cognitive impairment.

Regarding respiratory symptoms, dyspnea on exertion was reported in 25.3% (56/221) of the patients, nocturnal hypoventilation in 15.6% (34/218) and recurrent respiratory infection in 3.2% (7/221). Percentage predicted FVC was available for 116 patients and was below 80% in 52.6% of the cases (61/116). The FVC percentage distribution among these 116 cases was: 31.1% (19/61) between 70-80%, 21.3% (13/61) between 60-70%, 21.3% (13/61) between 60-50% and 26.2% (16/61) below 50%. Twelve percent (27/224) patients required part time Non-Invasive Ventilation (NIV), 4.0% (9/224) full time NIV and 0.9% (2/224) used invasive ventilation.

Cardiac impairment was seen in 7.3% of the patients (17/234). Changes compatible with hypertrophic cardiomyopathy and with dilated cardiomyopathy in the echocardiogram were reported in 7 and 5 cases respectively. Four patients had changes in the ECG including atrial fibrillation, right or left bundle branch block and a permanent tachycardia. None of the patients required a pacemaker or a defibrillator. The left ventricular ejection fraction (LVEF) value, as measured by transthoracic ultrasound, was available for 28 patients. No patients had a LVEF less than 50%, in 4 it was between 50-55% and in the remaining the LVEF was above 55%.

Features consisting with dysautonomia were reported in the 21.4% (42/196) of the patients. Only one symptom of dysautonomia was seen in the 66.7% (28/42, 12 urinary incontinence, 7 constipation, 6 diarrhoea, and 3 erectile dysfunction), two symptoms in the 26.2% (11/42) and 3 symptoms in the 7.1% (3/42). Among patients with dysautonomia, 9 had a concomitant diagnosis of polyneuropathy. Three of the patients with urinary incontinence (n=22) had a concomitant diagnosis of FDT.

Motor neuron involvement was identified in 25.0% of patients (54/216). 48.1% (26/54) showed lower motor neuron signs exclusively, 37.0% (20/54) lower and upper motor neuron

signs and 14.8% (8/54) upper motor neuron signs exclusively. Six percent of the patients (13/212) fulfilled ALS criteria as reported by the clinicians.

Of the 18.0% (39/217) of the cohort that showed bulbar signs, isolated dysphagia was reported in 59.0% (23/39), 35.9% (14/39) had dysphagia and dysarthria and only two patients had exclusively dysarthria. Of these 39 patients, 6 had motor neuron involvement and 5 fulfilled ALS criteria.

The classic triad of IBM, FTD and PDB was reported only in 7 patients (2.9%). Seventeen patients had IBM and isolated FTD (7.2%) and 38 patients showed IBM associated with isolated PDB (16.2%). The number of patients who had muscle weakness, PDB and ant type of cognitive impairment was 19 (8.1%). The frequency of clinical phenotype reported in other family members of the index cases is described in eFigure1.

Neurophysiology

The results of 182 nerve conduction studies and needle EMG were available. Of the 15.36% (35/229) of patients with clinical polyneuropathy, the NCS results were available in 33 cases: 14 patients showed a sensory-motor neuropathy, 10 a sensory neuropathy and 9 a motor neuropathy. The NCS pattern was pure axonal in 20 cases, axonal but with intermediate conduction velocities in 12 and conduction velocities compatible with demyelination in 3 cases. All the patients with neuropathy had muscle weakness and the 42.8% (15/35) had lower motor neuron signs associated.

Regarding the needle EMG pattern, 46.7% (85/182) of the patients showed an exclusively myopathic pattern, 20.9% (38/182) an exclusively neurogenic pattern and 20.3% (37/182) a mixed myopathic/neurogenic. The presence of spontaneous activity was reported in 44.5% (81/182).

Ambulatory status

At last assessment, 23.1% (52/225) of the patients required a full-time wheelchair for ambulation or were bedridden, at a median time of 8.5 (range 1-25) and 15 (range 7-23) years from disease onset respectively. The mean age at each ambulatory status and the median ambulatory status time from disease onset are shown in table 2.

Causes of death

Thirty-seven patients included in this study died during the follow-up at a mean age of 63.9 years (range 45-81) and 15.8 (range 2-31) years after the onset of symptoms. Mean age of death and mean time from disease onset to death are described in Table 1. Cause of death was available for 14 patients: 7 due to respiratory insufficiency, 5 due to rapidly progressive dementia, 1 myocardial infarction and 1 infectious COVID-related.

Genotype-phenotype associations

The frequency of signs/symptoms in the four most frequent variants are shown in eTable4. The frequency of females with the c.463C>T (p.Arg155Cys) variant was particularly high (56.7%) compared with the mean of 30% observed in the whole cohort. We did not identify differences in the distribution of muscle weakness, except for axial weakness, distal upper limb weakness and scapular winging that were more frequent in the variant c.463C>T (p.Arg155Cys), Fig 4. There were no differences in the frequency of cognitive symptoms across these mutations except for mixed cognitive impairment that was more common in the variant c.463C>T (p.Arg155Cys). We did not observe differences in the frequency of PDB, respiratory or cardiac involvement. The frequency of Parkinsonism was higher in variant c.476G>A (p.Arg159His) than c.464G>A (p.Arg155His; 16.7%, 3/18 vs 0.0%, 0/63 respectively, Chi Square test, adjusted p value 0.015).

After adjustment for age at last assessment, the variant c.463C>T (p.Arg155Cys) had an earlier age at first symptom onset than variants c.476G>A (p.Arg159His) and c.277C>T (p.Arg93Cys)

(37.8±7.6 years vs 49.6±9.0 and 52.4±5.7 years respectively, Analysis of Covariance, F 4.68 p=0.004 with Bonferroni correction for multiple comparisons). Similarly, variant c.464G>A (p.Arg155His) had an earlier age at first symptom than variant c.277C>T (p.Arg93Cys; 42.6±6.7 vs 37.8±7.6 years respectively, Analysis of Covariance F 4.685, p=0.04 with Bonferroni correction for multiple comparisons).

No significant differences were found in the mean ages of onset of respiratory impairment, cardiac impairment, FTD, dysautonomia, ambulatory status at last assessment and age of death among the 4 most frequent variants after adjusting for age at last assessment (ANCOVA test p>0.05 in all cases).

Variables associated with loss of ambulation and death.

We carried out a univariate analysis to test whether the presence of a severe reduced FVC, cardiac impairment, FTD, motor neuron signs, Parkinsonism, polyneuropathy, axial weakness, proximal and/or distal lower limb weakness were associated with requiring a full time wheelchair and/or being bedridden. The following variables were significantly associated with being a full time wheelchair user: FTD (37.5% 13/32 vs 20.4% 39/191, Chi-square test, p=0.03), FVC<50% (68.8% 11/16 vs 20.2% 20/99, Fisher exact test, p<0.01), distal lower limb weakness (28.2% 46/163 vs 11.8% 6/51, Chi-square test, p=0.017), proximal lower limb weakness (25.6% 51/199 vs 4.2% 1/24, Fisher exact test, p=0.019), axial weakness (37.5% 12/32 vs 20.4% 39/191, Chi-square test, p=0.03), and dysautonomia (38.1% 16/42 vs 15.8%23/146, Chi-square test, p=0.002). We identified a significant correlation between age at which the previous signs or symptoms were developed as well as the age of first symptom with the age being wheelchair user (Pearson correlation coefficients: FTD r=0.93; FVC<50% r=0.93; dysautonomia r=0.99; proximal lower limb weakness r=0.82, distal lower limb weakness r=0.82, and age at fist symptom r=0.82; p<0.001 in all cases).These variables were included in a binary logistic regression analysis after which only the variable FVC< 50% remained as a the one associated

with being wheelchair user or being bedridden (eTable 5). To determine which of these variables were associated with a higher risk of being full-time wheelchair user a Cox regression analysis was done after which only the variable FVC<50% remained as a the one representing a risk for this outcome (eTable 5). The time to being a full-time wheelchair user by FVC<50% was analysed through a Kaplan-Meier estimator and showed significant differences as represented in Fig 5.

Regarding the variables associated with death, we tested if the presence of reduced FVC value, cardiac impairment, dysphagia, dysarthria, drop head, axial weakness, FTD, motor neuron signs and/or being a full-time wheelchair user and/or being bedridden could be associated with a higher frequency of death. The following variables were significantly associated with death: FTD (37.0% 10/27 vs 15.7% 26/166 Fisher exact test, p=0.015), FVC<50% in comparison to patients with an FVC>80% (40% 6/15 vs 12.4% 1/42, Fisher exact test, p=0.001), dysphagia (31% 11/35 vs 16.3% 24/147, Chi-square test, p=0.042), drop-head (41.7% 10/24 vs 15.3% 24/147, Fisher exact test, p=0.002) and being a full-time wheelchair user/bedridden (36.2% 17/47 vs 13.4% 19/142, Chi-square test, p=0.001). The age at which the previous signs or symptoms were developed as well as the age of symptom onset correlated with the age of death (Pearson correlation coefficients: FDT r=0.958; FVC<50% r=0.944; dysphagia r=0.978; full time wheelchair user/bedridden r=0.844, and age at first symptom r=0.71; p<0.001 in all cases). These variables were included in a binary logistic regression analysis after which only the following variables remained associated with death: having a reduced FVC<50%, FVC<60–69% and being full-time wheelchair/bedridden (eTable 6). A Cox regression analysis showed that the variables FVC< 50%, FVC 60 – 69%, FTD and age at first symptom represented a risk for death (eTable 5). The time to death by FVC value and FTD was analysed through a Kaplan-Meier estimator, Fig 5.

Discussion

We report the clinical and genetic features of the largest series of patients with mutations in the VCP gene reported so far. This study was only possible thanks to the collaboration of clinicians from 24 countries on four different continents. The study provides valuable information regarding the phenotypic spectrum of the disease and data on disease progression across the world. The large number of patients included allowed the identification of 19 variants not previously described and to establish genotype-phenotype correlations. The data described in this paper is also of value to design natural history studies and/or clinical trials.

VCP gene was found to be the cause of IBMPFD in 2004 when six missense pathogenic variants were found in 61 affected individuals of 13 different families[25]. This first report already described that patients could develop three main clinical presentations: a myopathy, PDB and FTD. Since then, several published case report or patients' cohorts have expanded the phenotypic spectrum of the disease that now includes a plethora of syndromes affecting central nervous system[17,26,18,27,] motor neurons[16,1], sensory and/or motor peripheral nerves[19] and the skeletal muscle[15,7,28,14]. Our study confirms that clinical presentations at disease's onset are very heterogeneous although almost all patients will either present with muscle weakness or develop muscle weakness during the disease. The pattern of muscle weakness at onset is variable and can include both proximal or distal muscles of the lower and/or upper limbs turning VCP into a challenging diagnosis since patients can be classified as limb girdle muscle weakness, distal myopathy or even scapula-peroneal syndrome[29]. In addition, in this study an 8% (18/234) of the patients showed an asymmetric muscle weakness presentation extending the diagnostic challenge to motor neuron diseases and motor neuropathies. Regardless the pattern of weakness at onset, during disease progression weakness is generalized and all muscles are affected including axial muscles and scapula winging in half of the cases. Progression of muscle weakness is quicker than expected if compared with other adult onset inherited neuromuscular conditions with disease's progression data published[30]. In

our cohort, 23.1% of the patients were not ambulant at last assessment after a median duration of 8.5 years (7-23 y). This suggests that not only muscle weakness can affect more severely a larger number of muscles crucial for gait earlier in disease's progression but also the multisystemic nature of the disease with CNS and respiratory impairment could contribute to an earlier loss of ambulation. In fact, the presence of an FVC<50%, FTD and dysautonomia were associated with loss of ambulation in the univariate analysis, suggesting that other factors besides limb muscle weakness should be considered when interpreting disease progression.

Interestingly, we observed a high variability of CNS/PNS involvement in our cohort that included FTD, Alzheimer's disease or mixed cognitive decline as well as parkinsonism, motor neuron disease, ALS, peripheral neuropathy and dysautonomia. Dysautonomia was present in 21.4% of the patients and could present with different symptoms such as diarrhoea and faecal or urinary incontinence. The frequency is similar to previously published series[7]. Recently, more awareness of the multisystemic nature of VCP was raised so clinicians are now paying more attention to these symptoms. Unfortunately, we did not collect data about the results on ancillary tests or manoeuvres assessing the autonomous nervous system that could have provided a more accurate frequency of this clinical feature[31]. This data support further studies not only in the frequency of dysautonomia in VCP but also in the pathophysiology behind this symptom in order to determine whether it is due to an isolated or a combination or central, peripheral or autonomic nervous system involvement[11].

In a similar way, we also noticed a higher-than-expected frequency of peripheral neuropathy that affected 15.3% of patients in our cohort. In this case we also observed a variety of patterns, including pure motor or sensory motor axonal or demyelinating neuropathies. This study was not designed to collect detailed data on peripheral neuropathy symptoms, but we think that further studies could be informative, and clinicians should be aware of the high

prevalence of the involvement of the peripheral nervous system that can also improve VCP patients' care.

We have identified 58 variants in the VCP gene in patients with clinical symptoms compatible with MSP. All mutations, including a deletion and insertion resulted in a missense mutation within the VCP protein. Exon 5 contained 63% of the variants in our cohort although other exons were also affected such as exon 3. The four most frequent variants identified were responsible of 54.7% patients in our cohort and were seen in almost all countries that participated in this study. We identified 19 new non-previously reported variants that were considered as pathogenic or likely pathogenic after the *in-silico* tests. Fifteen of the 58 variants in 164 patients replaced an arginine with a different aminoacid at different domains of the VCP protein: the N-domain (12/15), the linking 1 domain (2/15) or the ATPase D2 domain (1/15). Arginine seems to have a key role in maintaining the structure and function of VCP protein. For example, an interaction between the N-domain and the D1-domain is required for the formation of the hemaxeric ring where the hydrolytic activity takes place and variants changing arginine by other aminoacids seems to disrupt this interaction[32]. Arginine is also a highly conserved residue in the D1 domain where it plays a critical role maintaining the hexameric structure of VCP which is required for binding polyubiquitinated proteins[33].

We have also identified genotype-phenotype correlations that are informative for care and clinical research planning. We did not identify any specific variants that were exclusively associated to an specific group of symptoms. However, among the two most frequent variants, patients harbouring the variant c.463C>T (p.Arg155Cys) had an earlier onset of the disease and showed higher frequency of axial weakness, distal upper limb weakness, scapular winging and mixed cognitive impairment. We also identified risk factors that were associated to the disease progression. As expected, lower limb weakness was associated with being a full time wheelchair user/bedridden but interestingly the presence of axial weakness, dysautonomia and FVC<50% were factors that influenced on ambulation status as well. Moreover, FTD,

respiratory involvement with FVC<50%, bulbar involvement in terms of dysphagia and/or dysarthria were associated to earlier age of death. However, after a multivariate analysis respiratory involvement with FVC<50% remained as the only risk factor associated with loss of ambulation while FVC<70% and FTD where both associated with risk of death. Based on this data it seems that FTD and especially respiratory involvement are symptoms that speed up the progression of the disease reducing the time until wheelchair and putting patients at risk of death and deserve special attention by clinicians in order to increase the quality of care provided.

Among the limitations of our study, we highlight its retrospective design that is associated with several missing data, loss of accuracy in terms of time of symptoms development and an underestimation of mortality rates that could happen if a patient deceased after the last assessment and the clinician was not aware. Similarly, many symptoms such as bone, cardiac and cognitive impairment might be underestimated since no ancillary test were required to be performed to be included in this study. Despite these limitations, we consider this data of high quality to inform about the clinical phenotype and progression of the disease. We have only included patients with genetic confirmation while relatives that were not studied with genetic tests, even if they were symptomatic, were not included. Although this increases the certainty, it reduces the number of patients and limits the variety of the clinical phenotypic heterogeneity identified within families. It is possible that some of the symptoms are underestimated as we have mainly contacted colleagues working in neuromuscular clinics but not neurologists working primarily on dementia, autonomous nervous system, general neurologists, or other specialists such as rheumatologist managing patients with PDB. Moreover, most of the patients were being followed by specialist in muscle diseases, and there are only a few that were diagnosed in ALS or peripheral nerve clinics.

In summary, the data presented here expands the spectrum of VCP disease and provides valuable information about disease's progression. The large number of patients included allowed us to identify FTD, dysautonomia and respiratory involvement as risk factors for early loss of ambulation and FTD, respiratory involvement, dysphagia, dysarthria and drophead as risk factors for death. This study has only been possible by the generous collaboration of many clinicians involved in the diagnosis and follow-up of VCP patients working in countries in Europe, Asia, South America, North America and Oceania. The fact that patients come from different areas has allowed us to identify up to 19 new mutation and characterize disease symptoms in different clinical settings enriching the data set and standardizing the description of the disease across countries.

References

1. Johnson JO, Mandrioli J, Benatar M, et al. Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. *Neuron*. 2010;68(5):857-864. doi:10.1016/j.neuron.2010.11.036

2. Sun X, Qiu H. Valosin-containing protein, a calcium-associated atpase protein, in endoplasmic reticulum and mitochondrial function and its implications for diseases. *Int J Mol Sci*. 2020;21(11). doi:10.3390/ijms21113842

3. Yeo BK, Yu SW. Valosin-containing protein (VCP): structure, functions, and implications in neurodegenerative diseases. *Animal Cells Syst (Seoul)*. 2016;20(6):303-309. doi:10.1080/19768354.2016.1259181

4. Kimonis VE, Kovach MJ, Waggoner B, et al. Clinical and molecular studies in a unique family with autosomal dominant limb-girdle muscular dystrophy and Paget disease of bone. *Genet Med*. 2000;2(4):232-241. doi:10.1097/00125817-200007000-00006

5. Kovach MJ, Waggoner B, Leal SM, et al. Clinical delineation and localization to chromosome 9p13.3-p12 of a unique dominant disorder in four families: Hereditary inclusion body myopathy, paget disease of bone, and frontotemporal dementia. *Mol Genet Metab*. 2001;74(4):458-475. doi:10.1006/mgme.2001.3256

6. Nalbandian A, Donkervoort S, Dec E, et al. The multiple faces of valosin-containing protein-associated diseases: Inclusion body myopathy with Paget's disease of bone, frontotemporal dementia, and amyotrophic lateral sclerosis. *J Mol Neurosci*. 2011;45(3):522-531. doi:10.1007/s12031-011-9627-y

7. S Figueroa-Bonaparte, J Hudson RB, T Polvikoski, T Williams, A Töpf EH, D Hilton-Jones, R Petty, T A Willis, C Longman, C F Dougan, M J Parton, M G Hanna, R Quinlivan MEF, M Guglieri, K Bushby, V Straub, H Lochmüller TE. Mutational spectrum and phenotypic variability of VCP-related neurological disease in the UK. 2016;87(6):680-681. doi:10.1136/jnnp-2015-310362

8. Mehta SG, Khare M, Ramani R, et al. Genotype-phenotype studies of VCP-associated inclusion body myopathy with Paget disease of bone and/or frontotemporal dementia. *Clin Genet*. 2013;83(5):422-431. doi:10.1111/cge.12000

9. Stojkovic T, Hammouda EH, Richard P, et al. Clinical outcome in 19 French and Spanish patients with valosin-containing protein myopathy associated with Paget's disease of bone and frontotemporal dementia. *Neuromuscul Disord*. 2009;19(5):316-323. doi:10.1016/j.nmd.2009.02.012

10. Plewa J, Surampalli A, Wencel M, et al. A cross-sectional analysis of clinical evaluation in 35 individuals with mutations of the valosin-containing protein gene. *Neuromuscul Disord*. 2018;28(9):778-786. doi:10.1016/j.nmd.2018.06.007

11. T D Miller, A P Jackson, R Barresi, C M Smart, M Eugenicos, D Summers, S Clegg VS, Stone J. Inclusion body myopathy with Paget disease and frontotemporal dementia (IBMPFD): clinical features including sphincter disturbance in a large pedigree. *J Neurol Neurosurg Psychiatry*. 2009;80(5):583-584. doi:10.1136/jnnp.2008.148676

12. Ikenaga C, Findlay AR, Seiffert M, et al. Phenotypic diversity in an international Cure VCP Disease registry. *Orphanet J Rare Dis*. 2020:1-9. doi:10.1186/s13023-020-01551-0

13. Weihl CC, Pestronk A, Kimonis VE. Valosin-containing protein disease: Inclusion body

myopathy with Paget's disease of the bone and fronto-temporal dementia. *Neuromuscul Disord*. 2009;19(5):308-315. doi:10.1016/j.nmd.2009.01.009

14. Sacconi S, Camaño P, de Greef JC, et al. Patients with a phenotype consistent with facioscapulohumeral muscular dystrophy display genetic and epigenetic heterogeneity. *J Med Genet*. 2012;49(1):41-46. doi:10.1136/jmedgenet-2011-100101

15. Palmio J, Sandell S, Suominen T, et al. Distinct distal myopathy phenotype caused by VCP gene mutation in a finnish family. *Neuromuscul Disord*. 2011;21(8):551-555. doi:10.1016/j.nmd.2011.05.008

16. González-Pérez P, Cirulli ET, Drory VE, et al. Novel mutation in VCP gene causes atypical amyotrophic lateral sclerosis. *Neurology*. 2012;79(22):2201-2208. doi:10.1212/WNL.0b013e318275963b

17. S. Spinaa, A. D. Van Laarb, J. R. Murrella, R. L. Hamiltonc, J. K. Koflerc, F. Eppersona MR, Farlowd, O. L. Lopezb, J. Quinlane, S. T. DeKoskyb and BG. Phenotypic variability in three families with valosin-containing protein mutation. *Eur J Neuro*. 2013;February(20(2)):251-258. doi:10.1111/j.1468-1331.2012.03831.x.

18. De Bot ST, Schelhaas HJ, Kamsteeg EJ, Van De Warrenburg BPC. Hereditary spastic paraplegia caused by a mutation in the VCP gene. *Brain*. 2012;135(12). doi:10.1093/brain/aws201

19. Gonzalez MA, Feely SM, Speziani F, et al. A novel mutation in VCP causes Charcot-Marie-Tooth Type 2 disease. *Brain*. 2014;137(11):2897-2902. doi:10.1093/brain/awu224

20. Mariani LL, Tesson C, Charles P, et al. Expanding the spectrum of genes involved in Huntington disease using a combined clinical and genetic approach. *JAMA Neurol*. 2016;73(9):1105-1114. doi:10.1001/jamaneurol.2016.2215

21. Shinjo SK, Oba-Shinjo SM, Lerario AM, Marie SKN. A Brazilian family with inclusion body myopathy associated with Paget's disease of bone and frontotemporal dementia linked to the VCP pGly97Glu mutation. *Clin Rheumatol*. 2018;37(4):1129-1136. doi:10.1007/s10067-017-3913-1

22. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424. doi:10.1038/gim.2015.30

23. Song C, Wang Q, Li CCH. ATPase activity of p97-valosin-containing protein (VCP): D2 mediates the major enzyme activity, and D1 contributes to the heat-induced activity. *J Biol Chem*. 2003;278(6):3648-3655. doi:10.1074/jbc.M208422200

24. De Ridder W, Azmi A, Clemen CS, et al. Multisystem proteinopathy due to a homozygous p.Arg159His VCP mutation: A tale of the unexpected. *Neurology*. 2020;94(8):e785-e796. doi:10.1212/WNL.0000000000008763

25. Watts GDJ, Wymer J, Kovach MJ, et al. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet*. 2004;36(4):377-381. doi:10.1038/ng1332

26. Schröder R, Watts GDJ, Mehta SG, et al. Mutant valosin-containing protein causes a

novel type of frontotemporal dementia. *Ann Neurol*. 2005;57(3):457-461. doi:10.1002/ana.20407

27. Nakamura T, Kawarabayashi T, Koh K, Takiyama Y, Ikeda Y, Shoji M. Spastic paraplegia with paget's disease of bone due to a VCP gene mutation. *Intern Med*. 2021;60(1):141-144. doi:10.2169/internalmedicine.4617-20

28. Liewluck T, Milone M, Mauermann ML, Castro-Couch M, Cerhan JH, Murthy NS. A novel VCP mutation underlies scapuloperoneal muscular dystrophy and dropped head syndrome featuring lobulated fibers. *Muscle and Nerve*. 2014;50(2):295-299. doi:10.1002/mus.24290

29. Jerath NU. Resolving a Multi-Generational Neuromuscular Mystery in a Family Presenting with a Variable Scapuloperoneal Syndrome in a c.464G>A, p.Arg155His VCP Mutation . *Case Rep Genet*. 2019;2019:1-7. doi:10.1155/2019/2403024

30. Mercuri E, Muntoni F. Muscular dystrophies. *Lancet*. 2013;381(9869):845-860. doi:10.1016/S0140-6736(12)61897-2

31. Freeman R, Chapleau MW. *Testing the Autonomic Nervous System*. Vol 115. 1st ed. Elsevier B.V.; 2013. doi:10.1016/B978-0-444-52902-2.00007-2

32. Niwa H, Ewens CA, Tsang C, Yeung HO, Zhang X, Freemont PS. The role of the N-domain in the atpase activity of the mammalian AAA ATPase p97/VCP. *J Biol Chem*. 2012;287(11):8561-8570. doi:10.1074/jbc.M111.302778

33. Wang Q, Song C, Irizarry L, Dai R, Zhang X, Li CH. Multifunctional Roles of the Conserved Arg Residues in the Second Region of Homology of p97 / Valosin-containing. 2005;280(49):40515-40523. doi:10.1074/jbc.M509636200

Abbreviations

VCP: Valosin Containing Protein Disease

ALS: Amyotrophic Lateral Sclerosis

FTD: Fronto Temporal Dementia

IBM: Inclusion Body Miopathy

IBMPFD: Inclusion Body Miopathy with Paget Disease of the Bone and Fronto-temporal Dementia.

UMN: Upper Motor Neuron Signs

LMN: Lower Motor Neuron Signs

LGMD: Limb Girdle Muscular Dystrophy

NIV: Non Invasive Ventilation

PDB: Paget's Disease Of The Bone

FVC: Vital Force Capacity