Longitudinal observational study investigating outcome measures for clinical trials in Inclusion Body Myositis

Gina Sangha,[1*] Bohao Yao,[2*] Daniel Lunn,[2] Iwona Skorupinska,[3,4] Louise Germain,[3,4] Damian Kozyra,[3] Matt Parton,[3,4] James Miller,[5] Michael G Hanna,[3,4] David Hilton-Jones,[1] Jane Freebody,[1] Pedro M Machado[3,4,6]

[1]*Department of Neurology, Oxford University Hospitals NHS Foundation Trust, Oxford, UK*

[2]*Department of Statistics, University of Oxford, Oxford, UK*

[3]*Department of Neuromuscular Diseases, Institute of Neurology, University College London, London, UK*

[4]*Queen Square Centre for Neuromuscular Diseases, University College Hospitals NHS Foundation Trust, London, UK*

[5]*Department of Neurology, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle, UK*

[6]*Centre for Rheumatology, Division of Medicine, University College London, London, UK*

***Shared first authors**

Corresponding author:

Pedro M Machado MD PhD

Department of Neuromuscular Diseases

Queen Square Institute of Neurology

University College London

1st Floor, Russell Square House

10-12 Russell Square, WC1B 5EH London

Email: p.machado@ucl.ac.uk

Word count: 3428

Word count abstract: 262

Number of tables: 3

Number of figures: 5 (1 colour, 4 mono)

Number of supplementary tables: 7

Number of supplementary figures: 5

Keywords: Clinical trials, Cohort studies, Natural history studies, Outcome research, Muscle disease, Inclusion Body Myositis.

Abstract

Objective:

To describe decline in muscle strength and physical function in patients with sporadic Inclusion Body Myositis (IBM).

Methods:

Manual muscle testing (MMT), quantitative muscle testing (QMT) and disability scoring using the IBM Functional Rating Scale (IBMFRS) were undertaken for 181 patients for up to 7.3 years. The relationship between MMT, QMT and IBMFRS composite scores and time from onset were examined using linear mixed effects models adjusted for gender and age of disease onset. Adaptive LASSO regression analysis was used to identify muscle groups that best predicted the time elapsed from onset. Cox proportional hazards regression was used to evaluate time to use of a mobility aid.

Results:

Multilevel modelling of change in percentage MMT, QMT and IBMFRS score over time yielded an average decline of 3.7% (95%CI: 3.1 to 4.3), 3.8% (95%CI: 2.7 to 4.9) and 6.3% (95%CI: 5.5 to 7.2) per year, respectively. The decline, however, was not linear, with steeper decline in the initial years. Older age of onset was associated with a more rapid IBMFRS decline ($p=0.007$), but did not influence the rate of MMT/QMT decline. Combination of selected muscle groups allowed for generation of single measures of patient progress (MMT and QMT factors). Median (interquartile range) time to using a mobility aid was 5.4 (3.6 to 9.2) years, significantly affected by greater age of onset (HR=1.06, 1.04 to 1.09, $p<0.001$).

Conclusion:

This prospective observational study represents the largest IBM cohort to date. Measures of patient progress evaluated in this study accurately predict disease progression in a reliable and useful way to be used in trial design.

Introduction

Sporadic inclusion body myositis (IBM) is an inflammatory and degenerative myopathy of uncertain aetiology first appropriately described in 1978.[1] IBM presents in middle or late age with a predilection for males. Its distinctive clinical signature includes selective weakness of quadriceps and forearm flexor muscles of insidious onset and progressive development, becoming more widespread. Few studies have prospectively assessed the natural history of the condition.[2-8] Median or average time to using assistive devices is 5 to 16 years for using a stick/cane and 12 to 20 years for using a wheelchair. Various rates of annual decline in muscle strength and physical function have been reported, ranging from 1% to 28% per year, depending on the outcome measure assessed.[2-12]

Despite increased trial activity and recent advances in our understanding of IBM, an evidence base for effective treatment is lacking with no current licensed treatment. Outcome measures in IBM have included the assessment of physical function (e.g. IBM Functional Rating Scale (IBMFRS) and IBM Physical Functioning Assessment (sIFA)), muscle strength (e.g. manual muscle testing (MMT), quantitative muscle testing (QMT) and hand-held dynamometry), mobility/endurance (e.g. 6 minute walk distance (6MWD) and timed up and go (TUG) test) and imaging characteristics including MRI thigh muscle volume and fat fraction.[13, 14] However, consistency among clinical trials and observational studies is lacking, creating methodological difficulties and hindering drug development and meta-analytic techniques.

This study is a merge of data from three prospective natural history studies across three major neuromuscular units and will therefore prove to be the largest prospective IBM observational study to date. Our principal aims were to describe the natural progression of IBM and to develop reliable statistical models of progression over time that can be used in clinical trial design and as outcome measures in the research setting to test the efficacy of interventions.

Material and methods

Study population

Patients with a diagnosis of IBM made by a neuromuscular expert and fulfilling predefined criteria,[15,16,17] were consecutively recruited into this longitudinal observational cohort study. Patients were enrolled across three centres; the University College London (UCL) Queen Square Centre for Neuromuscular Diseases, Oxford University Hospitals NHS Foundation Trust and Newcastle upon Tyne Hospitals NHS Foundation Trust. Subjects' consent and ethical approval was obtained (REC references 10/H0721/28, 12/LO/1557 and 98/8/86).

Assessments

Data was collected in a standardised manner at enrolment with regards to baseline demographics, age of onset, disease duration and clinical characteristics using a pre-defined protocol (London) or a reduced/modified version of the protocol (Oxford and Newcastle)[3].

Muscle strength was assessed by MMT[13, 18] and by maximum voluntary isometric contraction testing (MVICT) using the QMT system designed by Computer Source, Atlanta, Georgia, USA.[13, 19]

MMT was performed on 23 muscle groups in London (right and left side): facial muscles (Orbitalis Oculi and Orbicularis Oris, with the lowest score being recorded), neck flexors and extensors, shoulder abductors, elbow flexors and extensors, wrist flexors and extensors, long finger flexors, short finger flexors, finger extensors, first dorsal interosseus, abductor digiti minimi, abductor pollicis brevis, hip flexors and extensors, knee flexors and extensors, ankle dorsiflexors, plantar flexors, inverters and everters and extensor hallucis longus. MMT was performed on 15 muscle groups in Oxford (right and left side): facial muscles, neck flexors and extensors, shoulder abductors, elbow flexors and extensors, wrist flexors and extensors, finger extensors, hip flexors and extensors, knee flexors and extensors, ankle dorsiflexors and plantar flexors. Assessment of finger flexion in Oxford was measured using overall grip strength and therefore excluded from the analyses as it did not allow the separation between short and long finger flexors strength. MMT was not performed in Newcastle. Strength of each muscle group was graded utilizing a modified MRC score that was converted to a 13-point scale for the purposes of constructing a composite score: grade 0=0, 1=1, 2-=1.67, 2= 2, 2+=2.33, 3-=2.67, 3=3, 3+=3.33, 4-=3.67, 4=4, 4+=4.33, 5-=4.67, 5=5.[20-23] Scores across

right and left were averaged for individual muscle groups (as this approach increases the "signal-to-noise" ratio and reduces measurement error), and then scores were summed across muscle groups to form a composite. This was converted to a percentage of maximal available score for the purposes of cross-site comparison.

MVICT was performed in London on 6 muscle groups (right and left side): elbow flexors and extensors, knee flexors and extensors, ankle dorsiflexors and grip. Each muscle was tested twice and the maximum force (in kg) generated by the patient from the two trials was recorded for each muscle group. Individual MVICT scores were standardised by converting scores to a percentage of the predicted score for each muscle group using regression equations generated from isometric strength data from normal individuals.[24] Composite MVICT score was calculated by averaging the standardised MVICT score across all muscle groups tested.

Disability was measured at all sites using the IBMFRS that is intended only for patients with IBM. It includes 10 items (swallowing, handwriting, cutting food and handling utensils, fine motor tasks, dressing, hygiene, turning in bed and adjusting covers, changing position from sitting to standing, walking, and climbing stairs) graded on a Likert scale from 0 (being unable to perform) to 4 (normal). Composite IBMFRS score was calculated as the sum of the 10 items giving a value between 0 (complete dependency) and 40 (independent functioning) and converted to percentage of maximal score.[13, 25-27] Time at which patients first required a mobility aid from the onset of disease was recorded.

Statistical analysis

Data are expressed as number (%) for categorical variables and mean (minimum, maximum; standard deviation)/median (interquartile range) for continuous variables. Pearson's product moment correlation was used to evaluate linear relationships between baseline MMT, MVICT and IBMFRS scores. Trajectories of composite MMT, MVICT and IBMFRS scores were modelled using linear mixed effects models, with a random effect for both centre and participant identifiers, and including gender and age of onset as covariates. Estimates are displayed with 95% confidence intervals. We used adaptive LASSO regression analysis[28] to identify muscle groups that best predict the time elapsed from onset. Cox proportional hazards regression was used to evaluate time to use of a mobility aid. Missing data was handled using multiple imputation. All statistical analyses were performed using R version

3.3.1 (R foundation for statistical computing, Vienna, Austria). Further details about the statistical analysis can be found in the supplementary material.[29-31]

Results

Study population

A total of 181 patients were recruited, 110 in London, 46 in Oxford and 25 in Newcastle. Table 1 summarises the baseline characteristics of the cohort. Follow up time ranged from zero years (one visit only) to 7.3 years. One hundred and thirty patients (72%) attended at least one follow up visit. Of these, the median and mean follow up times were 1.3 and 2.0 years, respectively. Most patients were male (67%) with a mean age of onset of 58.4 years and mean disease duration at enrolment of 8.7 years at recruitment. Mean diagnostic delay was 5.3 years.

Baseline measures of clinical, quantitative and functional muscle strength

Median baseline MMT composite score was similar between London and Oxford patients and between males and females (Table 1 and online supplementary table 1); 99.2/115 (86.3%) in London patients and 65/75 (86.6%) in Oxford patients. Lowest median MMT scores were observed for long finger flexors (3.3) and knee extensors (3.5), followed by other small muscles of the hand, forearm muscles and hip flexors (online supplementary table 2) consistent with muscles previously described as being the most affected in the disease. Median baseline composite MVICT score in London patients was 30.6% of predicted score indicating significant weakness at baseline. Baseline score was lower in males (29.4%) than females (34.6%) (Table 1). Figure 1 demonstrates the distribution of scores at baseline in MMT and percentage of predicted MVICT scores in 155 and 82 patients, respectively. Median scores for all muscles tested both manually and quantitatively are decreased at baseline compared to maximal/normal strength scores. Median percentage of predicted MVICT scores were below 40% for all muscle groups tested at baseline, with the lowest score for knee extensors (17.3%).

Median baseline composite IBMFRS score was 29 out of a maximal score of 40 indicating a degree of functional disability at baseline. This was similar across sites (supplementary Table 1) and between males and females (Table 1). The greatest functional disability related to

standing from a sitting position (2/4), walking (2/4) and climbing stairs (1/4), implying weakness of proximal lower limb muscles. Median scores for all the other tasks; dressing, hygiene, fine motor control, swallowing, handwriting, handling utensils and turning in bed were 3/4 suggesting that most patients experienced some level of difficulty with these tasks.

Figure 2 demonstrates the relationship at baseline between composite IBMFRS and MMT (strong positive correlation; $R=0.7$, $p<0.001$), IBMFRS and MVICT (moderate positive correlation; $R=0.54$, $p<0.001$), IBMFRS and disease duration (moderate negative correlation; $R=-0.47$, $p<0.001$), and MVICT and MMT (moderate positive correlation; $R=0.6$, $p<0.001$).

Changes in physical function and muscle strength from baseline

Composite MMT, MVICT and IBMFRS scores declined over time from onset of disease (Figure 3). Multilevel modelling of change in percentage MMT score, composite MVICT score and percentage IBMFRS score over time yielded an average decline of 3.7% (95% CI: 3.1 to 4.3), 3.8% (95% CI: 2.7 to 4.9) and 6.3% (95% CI: 5.5 to 7.2) per year over 10 years. The decline, however, was not linear (Figure 4). Each measure progressed faster in the initial years. Table 2 demonstrates predicted change over 10 years from onset of disease using these models. The rate of decline of composite MMT, IBMFRS or MVICT scores over time from onset of disease did not vary significantly between males and females (online supplementary table 3). Older age of onset was associated with a more rapid decline in IBMFRS ($p=0.007$), but did not influence the rate of decline of MMT or MVICT composite scores.

We sought to identify muscle groups which contributed most to disease progression in two domains; MMT and QMT. We used adaptive LASSO regression for variable selection to identify these muscle groups and then formulated MMT and QMT factors consisting of weighted sums of contributing muscle groups, with weights derived from canonical correlation of muscle groups with time, adjusted for gender and age of onset (online supplementary table 4 and online supplementary figures 1 and 2). This was done for London data only because QMT was only performed in London, and because different muscles were tested between centres. The models fitted the measured data well (online supplementary Figures 3 and 4). The MMT factor comprised average values for five muscle groups (MMT-5): neck extensors, elbow flexors, abductor pollicis brevis, long finger flexors and knee extensors. The QMT factor comprised of average values for ankle dorsiflexors and grip (QMT-2). These muscle groups contributed most to predictive ability of the model for disease

progression over time. The maximum value for the MMT and QMT factors are 5 and 100 respectively. The average decline of the MMT factor is 0.293, 0.587, 0.758 and 1.01 for 1, 3, 5 and 10 years respectively since the onset of the disease. The average decline of the QMT factor is 10.3, 20.7, 26.7 and 35.7 for 1, 3, 5 and 10 years respectively since the onset of the disease. For the exact results, please refer to online supplementary table 5.

Online supplementary table 6 illustrates sample size calculations for a double-blind placebo-controlled trial (1:1 ratio) based on the IBMFRS progression rate and variance, and for different power, trial duration and treatment effect size assumptions. Larger variances were found for MMT and QMT factors, which would have resulted in larger sample sizes if these measures were used as primary trial endpoints (data not shown).

Time to use of mobility aid

Data regarding the use of a mobility aid was recorded for 139 patients. Of these, 106 patients (76%) required the use of a mobility aid over the course of the study. Table 3 provides details for the type of mobility aid participants required; 88.7% of patients required a stick, followed by ankle foot orthoses (8.5%), frame (1.9%) and wheelchair (0.9%).

Median (interquartile range) time to using a mobility aid was 5.4 (3.6 to 9.2) years. Using Cox (proportional hazards) regression we investigated the effect of the covariates gender and age of onset on survival without the use of a mobility aid (Figure 5). Gender had no statistically significant effect on survival without a mobility aid (p=0.93, hazard ratio=0.98 (0.64-1.51)). The impact of age of onset, however, was highly significant (HR=1.06 per 1 year increase in age of onset (95% CI: 1.04-1.09); p=<0.001) (online supplementary table 7). Time to use of a mobility aid decreased with later ages of disease onset (online supplementary Figure 5).

Discussion

This prospective observational study is the largest to date with the aim of delineating the natural history of IBM and identifying useful outcome measures for clinical trials. Participants were observed to have decreased strength and physical function at baseline on enrolment into the study and the three outcome measures that were studied (MMT, MVICT and IBMFRS) demonstrated a decline over time from onset of disease. We observed moderate

to strong positive correlations between composite MMT, MVICT and IBMFRS scores at baseline.

MMT is commonly used in myositis trials as a primary or secondary clinical endpoint and in clinical practice to assess disease progress, reported as a summed or composite score of tested muscle groups.[13, 21, 27] The International Myositis Assessment and Clinical Studies Group (IMACS) have defined and validated a modified MMT score comprising 8 proximal, distal and axial muscle groups (MMT-8) in patients with inflammatory myopathies.[22] This group however, does not constitute those specifically affected in IBM and so the utility of this MMT-8 tool in IBM might be limited given the selective pattern of weakness observed in this disease.

We aimed to identify the muscle groups which were the most consistent markers of disease progression in the cohort by MMT and QMT. Results from adaptive LASSO regression of MMT data revealed that the combination neck extensors, elbow flexors, long finger flexors abductor pollicis brevis and knee extensors were the minimum combination that would provide predictive ability for progression of MMT over time (MMT-5). This combination consists of muscle groups known to be most affected in IBM – including distal upper limb muscles and proximal lower limb muscles. Weakness of neck extensors is common in neuromuscular disorders including IBM, manifesting as dropped head syndrome.[32] This data might suggest that this clinical feature of the disease is more prominent than previously considered and could contribute to earlier detection and diagnosis of IBM.

Results from adaptive LASSO regression of QMT data revealed that the QMT factor only required ankle dorsiflexors and grip (QMT-2) to provide predictive ability for QMT progression over time and adding other muscle groups increased the cross-validation error, suggesting overfitting. The contribution of ankle dorsiflexors to the QMT factor reflects common involvement of tibialis anterior in this disease. Other reports have revealed greater weakness in this group than proximal lower limb muscles in certain patients suggesting quantitative measurement of the combination of ankle dorsiflexion and grip muscles might be sufficient and sensitive in clinical practice.[33] Exclusion of grip from the analysis resulted in all other muscles being combined into the QMT factor, suggesting that grip was sufficient to explain the variability in elbow flexors, elbow extensors, knee flexors and knee extensors over time from onset of disease thus highlighting the importance of grip strength in disease progression. By combining tested muscle groups into a single factor we are able to extend the clinical practicability and applicability of quantitative myometry.

The models we have generated may be used to determine sample sizes required in therapeutic trials. The number of years taken for a given change in IBMFRS, MMT factor or QMT factor as a primary outcome measure can be predicted enabling clinical trial design. The IBMFRS model had less variance compared to the other models, therefore resulting in smaller sample size calculations, and suggesting that the IBMFRS has greater responsiveness and should therefore be preferentially used as primary endpoint in clinical trials. The faster initial progression borne out in the models may allow for detection of early improvement/stabilisation/slower decline if they are used as clinical trial outcome measures, reducing trial duration needs. They might also provide clues to pathogenesis. All measures of muscle strength in this study progressed non-linearly over time, with a steeper initial decline which could suggest an initial insult hastens progression, followed by a degenerative process resulting in the maintenance of steady deterioration in physical and functional strength.[34] An important area for further development in IBM is the identification of monitoring and treatment biomarkers that are sensitive to change and that may allow a reduction in sample size and/or trial duration, namely in early phase and proof-of-concept clinical trials. Quantitative MRI (e.g. fat fraction, remaining cross-sectional muscle area, and muscle volume) has shown potential in this regard.[14]

As prior studies have shown, the IBMFRS is a useful tool for evaluating the progression of disease[13, 25] and is the primary outcome in the on-going trial of Arimoclomol (NCT02753530; NCT04049097), a drug that amplifies the production of heat shock proteins.[35] Our study further validates the use of IBMFRS as a useful primary endpoint. The greatest disability related to use of proximal lower limb muscles (walking, climbing stairs, standing from sitting), but participants also struggled with use of the hands and some difficulty with other tasks such as swallowing. Older age of disease onset was associated with a steeper decline in IBMFRS score. This likely reflects domains of frailty other than weakness as such a relationship was not observed for MMT or composite MVICT score. Such domains include cognitive impairment and chronic conditions.

The frequent lower limb muscle weakness in this disease explains the requirement, for most patients, for a mobility aid over the course of their illness. Survival without a mobility aid was unaffected by gender, but significantly affected by age of disease onset, with later onset resulting in a shorter delay to requiring a mobility aid. This provides important information for patients about the natural history of the condition, with a steep decline in survival without a mobility aid between 5-15 years of disease onset. It also has implications for social

placement, care and equipment needs as the disability from the disease progresses and allows appropriate timing of interventions. Most patients in this cohort required a stick (88.7%), followed by ankle foot orthoses (8.5%), illustrating the involvement of ankle dorsiflexors in disease progression.

This study does have several limitations. The most significant limitation was in the practical elements of data collection across three centres. There was no prior standardisation exercise and inter-observer reliability of assessments could not be calculated. There was variability in follow-up time, consistency and completeness of data collection. As a result, only patients with complete baseline and follow-up data were included in the models. The differences in the number of patients undergoing each analysis in different centres further limits the generalisability of conclusions drawn from this study, which would require validation in additional cohorts. Finally, there was an inability to control for differences such as baseline comorbidities and medications, limiting the interpretative ability of the data because of possible confounders. To mitigate this, individual patient effects were incorporated into the multilevel models. Furthermore, no medication has been found to be effective in IBM.

Nevertheless, the importance of demonstrating the natural history of IBM, including the significant heterogeneity in clinical presentation and progression, adds to the external validity of this data in clinical practice. This is especially important with regards to the evolution of clinically-oriented classification criteria, as there remains to be a delay in diagnosis of 5-6 years and data-driven classification criteria may facilitate earlier recognition. Reasons for this diagnostic delay are the characteristics of the disease itself, namely middle-late age of onset, slow progression and overlapping clinical presentation with other neuromuscular disorders resulting in misdiagnosis (39% of patients in this study had an alternative initial diagnosis), and the reduced awareness/experience amongst community and secondary care specialists of this rare condition.

In summary, this study set out to delineate the natural history of disease and to assess the behaviour of outcome measures for clinical trials of therapeutics in IBM. MMT, QMT and the IBMFRS models were able to detect disease progression over time. The models generated in this study are useful in the assessment of IBM and may be used as outcome measures as well as in trial design, including sample size calculation and patient stratification. Future work to further validate these tools will be useful to support the design of large randomised trials powered to assess treatment effects.

Contributors: All authors critically revised the manuscript for important intellectual content. Specific roles included study design (GS, DH-J, PMM), data collection (IS, LG, DK, MP, JM, MGH, HH-J, JF, PMM), statistical analyses (GS, BY, DL) and drafting of the first version of the manuscript (GS, DH-J, PMM).

Funding: Study supported by grants from the National Institute for Health Research (NIHR) Rare Diseases Translational Research Collaboration (BRC148/NS/MH) and Myositis UK (grant reference not available). Pedro M. Machado is supported by the NIHR University College London Hospitals (UCLH) Biomedical Research Centre (BRC). The views expressed are those of the authors and not necessarily those of the (UK) National Health Service (NHS), the NIHR, or the (UK) Department of Health.

Competing interests: GS reports no disclosures. BY reports no disclosures. DL reports no disclosures. IS reports no disclosures. LG reports no disclosures. DK reports no disclosures. MP reports no disclosures. JM reports no disclosures. MGH reports no disclosures. DH-J reports no disclosures. JF reports no disclosures. PMM has received grants and/or honoraria from Abbvie, BMS, Celgene, Eli Lilly, Janssen, MSD, Novartis, Orphazyme, Pfizer, Roche and UCB.

Data availability statement: All analyses are described in the manuscript. For additional information, please contact the corresponding author. Data are available upon reasonable request.

Acknowledgements: In memory of the late Dr Daniel Lunn, without whom this work would not be possible.

Twitter: @pedrommcmachado

ORCID IDs: Pedro M Machado http://orcid.org/0000-0002-8411-7972

References

1. Carpenter S, Karpati G, Heller I, Eisen A. Inclusion body myositis: a distinct variety of idiopathic inflammatory myopathy. Neurology 1978;28:8-17.
2. Alfano LN, Yin H, Dvorchik I, et al. Modeling functional decline over time in sporadic inclusion body myositis. Muscle Nerve 2017;55:526-531.
3. Cortese A, Machado P, Morrow J, et al. Longitudinal observational study of sporadic inclusion body myositis: implications for clinical trials. Neuromuscul Disord 2013;23:404-412.
4. Cox FM, Titulaer MJ, Sont JK, Wintzen AR, Verschuuren JJ, Badrising UA. A 12-year follow-up in sporadic inclusion body myositis: an end stage with major disabilities. Brain 2011;134:3167-3175.
5. Felice KJ, North WA. Inclusion body myositis in Connecticut: observations in 35 patients during an 8-year period. Medicine (Baltimore) 2001;80:320-327.
6. Hogrel JY, Allenbach Y, Canal A, et al. Four-year longitudinal study of clinical and functional endpoints in sporadic inclusion body myositis: implications for therapeutic trials. Neuromuscul Disord 2014;24:604-610.
7. Rose MR, McDermott MP, Thornton CA, Palenski C, Martens WB, Griggs RC. A prospective natural history study of inclusion body myositis: implications for clinical trials. Neurology 2001;57:548-550.
8. Oldroyd AGS, Lilleker JB, Williams J, Chinoy H, Miller JAL. Long-term strength and functional status in inclusion body myositis and identification of trajectory subgroups. Muscle Nerve 2020.
9. Badrising UA, Maat-Schieman ML, van Houwelingen JC, et al. Inclusion body myositis. Clinical features and clinical course of the disease in 64 patients. J Neurol 2005;252:1448-1454.
10. Benveniste O, Guiguet M, Freebody J, et al. Long-term observational study of sporadic inclusion body myositis. Brain 2011;134:3176-3184.
11. Lindberg C, Persson LI, Bjorkander J, Oldfors A. Inclusion body myositis: clinical, morphological, physiological and laboratory findings in 18 cases. Acta Neurol Scand 1994;89:123-131.
12. Needham M, James I, Corbett A, et al. Sporadic inclusion body myositis: phenotypic variability and influence of HLA-DR3 in a cohort of 57 Australian cases. J Neurol Neurosurg Psychiatry 2008;79:1056-1060.
13. Rider LG, Aggarwal R, Machado PM, et al. Update on outcome assessment in myositis. Nat Rev Rheumatol 2018;14:303-318.
14. Morrow JM, Sinclair CD, Fischmann A, et al. MRI biomarker assessment of neuromuscular disease progression: a prospective observational cohort study. Lancet Neurol 2016;15:65-77.
15. Griggs RC, Askanas V, DiMauro S, et al. Inclusion body myositis and myopathies. Ann Neurol 1995;38:705-713.
16. Hilton-Jones D, Miller A, Parton M, Holton J, Sewry C, Hanna MG. Inclusion body myositis: MRC Centre for Neuromuscular Diseases, IBM workshop, London, 13 June 2008. Neuromuscul Disord 2010;20:142-147.

17. Rose MR, Group EIW. 188th ENMC International Workshop: Inclusion Body Myositis, 2-4 December 2011, Naarden, The Netherlands. Neuromuscul Disord 2013;23:1044-1055.
18. Rider LG, Werth VP, Huber AM, et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: Physician and Patient/Parent Global Activity, Manual Muscle Testing (MMT), Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (C-HAQ), Childhood Myositis Assessment Scale (CMAS), Myositis Disease Activity Assessment Tool (MDAAT), Disease Activity Score (DAS), Short Form 36 (SF-36), Child Health Questionnaire (CHQ), physician global damage, Myositis Damage Index (MDI), Quantitative Muscle Testing (QMT), Myositis Functional Index-2 (FI-2), Myositis Activities Profile (MAP), Inclusion Body Myositis Functional Rating Scale (IBMFRS), Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), Cutaneous Assessment Tool (CAT), Dermatomyositis Skin Severity Index (DSSI), Skindex, and Dermatology Life Quality Index (DLQI). Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S118-157.
19. Tawil R, McDermott MP, Mendell JR, Kissel J, Griggs RC. Facioscapulohumeral muscular dystrophy (FSHD): design of natural history study and results of baseline testing. FSH-DY Group. Neurology 1994;44:442-446.
20. Harris-Love MO, Shrader JA, Koziol D, et al. Distribution and severity of weakness among patients with polymyositis, dermatomyositis and juvenile dermatomyositis. Rheumatology (Oxford) 2009;48:134-139.
21. Rider LG, Giannini EH, Harris-Love M, et al. Defining Clinical Improvement in Adult and Juvenile Myositis. J Rheumatol 2003;30:603-617.
22. Rider LG, Koziol D, Giannini EH, et al. Validation of manual muscle testing and a subset of eight muscles for adult and juvenile idiopathic inflammatory myopathies. Arthritis Care Res (Hoboken) 2010;62:465-472.
23. Clarke C, Institute of Neurology QS, National Hospital for N, Neurosurgery. Neurology : a queen square textbook. Chichester, England: Wiley Blackwell, 2016.
24. Muscular weakness assessment: use of normal isometric strength data. The National Isometric Muscle Strength (NIMS) Database Consortium. Arch Phys Med Rehabil 1996;77:1251-1255.
25. Jackson CE, Barohn RJ, Gronseth G, Pandya S, Herbelin L, Muscle Study G. Inclusion body myositis functional rating scale: a reliable and valid measure of disease severity. Muscle Nerve 2008;37:473-476.
26. Ramdharry G, Morrow J, Hudgens S, et al. Investigation of the psychometric properties of the Inclusion Body Myositis Functional Rating Scale using Rasch analysis. Muscle Nerve 2019.
27. Rider LG, Ruperto N, Pistorio A, et al. 2016 ACR-EULAR adult dermatomyositis and polymyositis and juvenile dermatomyositis response criteria-methodological aspects. Rheumatology (Oxford) 2017;56:1884-1893.
28. Fan Y, Li R. Variable Selection in Linear Mixed Effects Models. Ann Stat 2012;40:2043-2068.
29. Kirkwood B, Sterne J. Essential Medical Statistics, eTextbook.: Hoboken, NJ: Wiley-Blackwell, 2010.

30. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52:591-611.
31. Venables WN, Ripley BD. Modern Applied Statistics with S, Fourth ed: Springer, 2002.
32. Alhammad RM, Naddaf E. Myopathies presenting with head drop: Clinical spectrum and treatment outcomes. Neuromuscul Disord 2020;30:128-136.
33. Dimachkie MM, Barohn RJ. Inclusion body myositis. Neurol Clin 2014;32:629-646, vii.
34. Benveniste O, Stenzel W, Hilton-Jones D, Sandri M, Boyer O, van Engelen BG. Amyloid deposits and inflammatory infiltrates in sporadic inclusion body myositis: the inflammatory egg comes before the degenerative chicken. Acta Neuropathol 2015;129:611-624.
35. Ahmed M, Machado PM, Miller A, et al. Targeting protein homeostasis in sporadic inclusion body myositis. Sci Transl Med 2016;8:331ra341.

TABLES

Table 1 Baseline clinical and demographic data	
Variable *(number of patients with available data)*	All patients (n=181)
Sex (n = 181) Male Female	 122 (67.4) 59 (32.6)
Age at enrolment (years) (n=176)	66.9 (42.0, 86,3; 9.3) / 67.3 (61.2 – 74.1)
Time of follow-up (years) (n=181)	2.0 (0.0, 7.3; 1.9) / 1.3 (0 - 3.5)
Age of onset (years) (n = 170)	58.4 (16.4, 82.4; 10.5) / 59.3 (50.8 – 65.9)
Disease duration (years) (n = 170)	8.7 (1.2, 30.6; 5.2) / 7.6 (5.1 – 10.8)
Diagnostic delay (years) (n = 165)	5.3 (0.0, 19.0; 3.9) / 4.6 (2.5 – 7.0)
Site of onset(n = 171)* Proximal muscles of lower limbs Distal muscles of lower limbs Proximal muscles of upper limbs Distal muscles of upper limbs Bulbar muscles Other	 113 (66.1) 19 (11.1) 2 (1.2) 38 (22.2) 13 (7.6) 1 (0.6)
Nature of onset (n = 147)* Weakness Pain Falls Other	 122 (83.0) 7 (4.8) 30 (20.4) 7 (4.8)
Assymmetry At onset *(n = 140)* At any time *(n = 137)*	 89 (63.6) 108 (78.8)
Initial alternative diagnosis (n = 140) None Polymyositis or other inflammatory myopathy Other myopathy Other neurological disease Other disease	 86 (61.4) 31 (22.1) 8 (5.7) 11 (7.9) 4 (2.9)
Baseline MMT composite score (n = 155) *London (n = 109, max = 115)* *Oxford (n = 46, max = 75)* *Baseline percent of maximal score (%)* *London (n = 109, max = 115)*	 96.2 (53.1, 112.6; 12.3) / 99.2 (91.5 – 104.3) M: 93.9 (32.0, 114; 15.1) / 99.1 (85.2 – 104.5) F: 91.0 (66.8, 111.8; 10.6) / 93.0 (82.2 – 98.1) *62.8 (31.5, 74.4; 8.9) / 65.0 (60.5 – 69.0)* *M: 60.9 (27.0, 74.4; 9.3)/ 62.0 (58.0 – 67.5)* *F: 61.4 (31.5, 73.0; 8.6) / 63.2 (56.0 – 68.8)* 83.6 (46.1, 97.9; 10.7) / 86.2 (79.6 – 90.7) M: 81.6 (27.8, 99.1; 13.1) / 86.2 (74.1 – 90.9) F: 79.2 (58.0, 97.2; 9.2) / 80.9 (71.5 – 85.3)

Oxford (n = 46,max = 75)	*83.8 (42.0, 99.1; 11.9) / 86.7 (80.7 – 92.0)* *M: 81.2 (36.0, 99.1; 12.5) / 82.7 (77.4 – 90.0)* *F: 81.9 (42.0, 97.3; 11.5) / 84.3 (74.7 – 91.7)*
Baseline composite percent predicted MVICT score (%) (n = 65)	32.6 (13.0, 89.0; 14.0) / 30.6 (22.3 – 40.9) M: 31.7 (13.0, 89.0; 14.0) / 29.4 (22.0 – 39.0) F: 35.2 (13.7, 64.8; 14.1) / 34.6 (24.9 – 43.0)
Baseline composite IBMFRS score (n = 181) (max = 40) *Baseline percent of maximal score (%)*	*27.1 (5.0, 40.0; 6.6) / 29.0 (22.0 – 32.0)* *M: 27.8 (10.0, 40.0; 6.4) / 29.0 (23.3 – 33.0)* *F: 25.7 (5.0, 36.0; 6.8) / 26.0 (22.0 – 31.0)* *67.8 (12.5, 100.0; 16.5) / 72.5 (55.0 – 80.0)* *M: 69.5 (25.0, 100.0; 16.0) / 72.5 (58.1 – 82.5)* *F: 64.3 (12.5, 90.0; 17.0) / 65.0 (55.0 – 77.5)*
Data are expressed as number (n) and (%) for categorical variables and mean (minimum, maximum; standard deviation)/median (interquartile range) for continuous variables. *Participants may have presented with multiple sites or symptoms at onset. F, females; IBMFRS, Inclusion Body Myositis Functional Rating Scale; M, males; MMT, manual muscle testing; MVICT, Maximum Voluntary Isometric Contraction Testing.	

Table 2. Predicted change of MMT percentage score, MVICT composite score and IBMFRS percentage score per year from onset of disease

Year	MMT percentage score (95% CI)	MVICT composite score (95% CI)	IBMFRS percentage score (95% CI)
1	-10.7 (-12.4 to -9.0)	-11.0 (-14.1 to -7.8)	-18.4 (-20.9 to -15.8)
2	-17.0 (-19.6 to -14.3)	-17.4 (-22.3 to -12.4)	-29.1 (-33.1 to -25.1)
3	-21.4 (-24.8 to -18.0)	-21.9 (-28.2 to -15.6)	-36.7 (-41.8 to -31.6)
4	-24.8 (-28.8 to -20.9)	-25.4 (-32.7 to -18.2)	-42.6 (-48.5 to -36.7)
5	-27.7 (-32.0 to -23.3)	-28.3 (-36.4 to -20.2)	-47.4 (-54.0 to -40.9)
6	-30.0 (-34.8 to -25.3)	-30.8 (-39.5 to -22.0)	-51.5 (-58.6 to -44.4)
7	-32.1 (-37.2 to -27.0)	-32.9 (-42.3 to -23.5)	-55.1 (-62.6 to -47.5)
8	-33.9 (-39.3 to -28.6)	-34.7 (-44.7 to -24.8)	-58.2 (-66.2 to -50.2)
9	-35.5 (-41.2 to -29.9)	-36.4 (-46.8 to -26.0)	-61.0 (-69.4 to -52.6)
10	-37.0 (-42.9 to -31.2)	-37.9 (-48.7 -to-27.1)	-63.5 (-72.2 to -54.7)

IBMFRS, Inclusion Body Myositis Functional Rating Scale; MMT, manual muscle testing; MVICT, Maximum Voluntary Isometric Contraction Testing.

Table 3. Types of mobility aid (n=106)	
Mobility aid	*Number (%)*
Ankle foot orthosis	9 (8.5)
Stick	94 (88.7)
Frame	2 (1.9)
Wheelchair	1 (0.9)

FIGURES AND LEGENDS

Figure 1. Boxplots illustrating baseline (A) manual muscle testing (MMT) scores and (B) standardized maximum voluntary isometric contraction testing (MVICT) scores for individual muscle groups in 155 and 82 IBM patients, respectively.

Fig_1.pdf

Figure 2. Relationship between percentage manual muscle testing (MMT), maximum voluntary isometric contraction testing (MVICT) and percentage Inclusion Body Myositis Functional Rating Scale (IBMFRS) scores at baseline. Moderate to strong positive correlations were observed between percentage IBMFRS and MMT score (A), percentage IBMFRS and composite MVICT score (B) and percentage MMT score and composite MVICT score at baseline (D). A moderate negative correlation was observed between percentage IBMFRS score and disease duration (C). R = Pearson's correlation coefficient.

Fig_2.pdf

Figure 3. Progression of composite scores. Spaghetti plots demonstrate decline in composite manual muscle testing (MMT) score as percentage of maximal score (A) maximum voluntary isometric contraction testing **(**MVICT) score (B) and Inclusion Body Myositis Functional Rating Scale (IBMFRS) as a percentage of maximal score (C) over time from onset of disease (each line represents an individual patient). Linear regression lines displayed.

Fig_3.pdf

Figure 4. Predicted change in Percentage Manual Muscle Testing (MMT) score (A), Composite maximum voluntary isometric contraction testing (MVICT) score (B) and Percentage Inclusion Body Myositis Functional Rating Scale (IBMFRS) score over time from onset of disease using multilevel models. Observed change is non-linear, with faster rate of decline in the initial years. Upper and lower 95% confidence intervals are demonstrated.

Fig_4.pdf

Figure 5. Time to use of a mobility aid. Survival curves for time to needing mobility aid, using Cox (proportional hazards) regression model (n=139) Covariates include gender and age of onset. Upper and lower 95% confidence intervals depicted.

Fig_5.pdf

Supplementary material - Figures FI

Supplementary material - Statistics F

Supplementary material - Tables FIN