Identification of principal pathological criteria for the diagnosis of inclusion body myositis

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-004552
Article Type:	Research
Date Submitted by the Author:	25-Nov-2013
Complete List of Authors:	Brady, Stefen; MRC Centre for Neuromuscular Diseases, Division of Neuropathology Squier, Waney Sewry, Caroline Hanna, Mike Hilton-Jones, David Holton, Janice; UCL Institute of Neurology, Department of Molecular Neuroscience and MRC Centre for Neuromuscular Diseases
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Diagnostics
Keywords:	NEUROLOGY, Adult neurology < NEUROLOGY, Neuromuscular disease < NEUROLOGY, Neuropathology < PATHOLOGY

Identification of principal pathological criteria for the diagnosis of inclusion body myositis

Corresponding author:

Dr Janice L Holton

Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

janice.holton@ucl.ac.uk

Tel: 00 44 (0)20 3448 4239

Fax: 00 44 (0)20 3448 4486

Authors:

Stefen Brady[1], Waney Squier[2], Caroline Sewry[3,4], Michael Hanna[1], David Hilton-Jones[5], Janice L Holton[6]

[1]MRC Centre for Neuromuscular Diseases, UCL Institute of Neurology and National Hospital for Neurology, Neurosurgery, Queen Square, London, UK.

[2]Department of Neuropathology, University of Oxford, John Radcliffe Hospital, Oxford, UK.

[3]Dubowitz Neuromuscular Centre, Institute of Child Health and Great Ormond Street Hospital for Children, London, UK.

[4]Wolfson Centre of Inherited Neuromuscular Diseases, RJAH Orthopaedic Hospital, Oswestry, UK.

[5]Nuffield Department of Clinical Neurosciences (Clinical Neurology), University of Oxford, John Radcliffe Hospital, Oxford, UK.

[6]Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

Keywords: Neurology, adult neurology, neuromuscular disease, neuropathology

Running title: Pathological criteria for inclusion body myositis

Word count: 2990

ABSTRACT

Objectives

The current pathological diagnostic criteria for sporadic inclusion body myositis (IBM) lack sensitivity. Using immunohistochemical techniques abnormal protein aggregates have been identified in IBM, including some associated with neurodegenerative disorders. Our objective was to investigate the diagnostic utility of a number of markers of protein aggregates together with mitochondrial and inflammatory changes in IBM.

Design

Retrospective cohort study. The sensitivity of pathological features was evaluated in cases of Griggs' definite IBM. The diagnostic potential of the most reliable features was then assessed in clinically-typical IBM with rimmed vacuoles and clinically-typical IBM without rimmed vacuoles and IBM mimics - vacuolar myopathies and steroid-responsive inflammatory myopathies.

Setting

Specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London.

Results

Individual pathological features, in isolation, lacked sensitivity and specificity. However, the morphology and distribution of p62 aggregates in IBM were characteristic and in a myopathy with rimmed vacuoles, the combination of characteristic p62 aggregates and increased sarcolemmal and internal MHC Class I expression or endomysial cytotoxic T-cells were diagnostic for IBM (sensitivity 93% and specificity 100%). In an inflammatory myopathy lacking rimmed vacuoles, the presence of mitochondrial changes was 100% sensitive and 73% specific for IBM; characteristic p62 aggregates were specific, but lacked sensitivity.

Conclusions

We propose an easily applied diagnostic algorithm for the pathological diagnosis of IBM. Additionally our findings support the hypothesis that many of the pathological features considered typical of IBM develop occur later in the disease, explaining their poor sensitivity at disease

presentation and emphasising the need for a revised pathological criteria to supplement the clinical criteria in the diagnosis of IBM.

STRENGTHS AND LIMITATIONS

The present study is a multicentre retrospective evaluation of the diagnostic utility of pathological findings for differentiating IBM from myopathies important in the differential diagnosis – myopathies containing rimmed vacuoles and steroid-responsive inflammatory myopathies.

The main strength of our study was the systematic detailed analysis of well defined cases. This enabled us to determine the sensitivity and specificity of individual pathological features and produce an easily applied pathological diagnostic algorithm for IBM for use in clinical practice.

Study limitations include the small number of cases and the retrospective design. Further prospective studies are now required in larger cohorts of patients.

INTRODUCTION

Sporadic inclusion body myositis (IBM) is the commonest acquired myopathy in those aged over 50 years.[1] Although classified as an idiopathic inflammatory myopathy, muscle biopsy reveals both degenerative and inflammatory features. The widely used Griggs diagnostic criteria require the presence of several pathological findings,[2] namely rimmed vacuoles, an inflammatory infiltrate with invasion of non-necrotic fibres by mononuclear inflammatory cells (partial invasion), and either amyloid deposits or 15-18 nm tubulofilaments identified by electron microscopy (EM). Although these features in combination are highly specific for IBM, individually they occur in other myopathies, including some important in the differential diagnosis for IBM.[3-7] Moreover, cases of clinically-typical IBM have been reported where the combination of these pathological features is absent causing diagnostic difficulty.[8-11]

Over the last two decades, pathological accumulation of many different proteins has been reported in muscle fibres in IBM.[12] Proteins typically associated with neurodegenerative diseases such as β-amyloid (Aβ), hyperphosphorylated tau and ubiquitin and newer neurodegenerative markers such as p62 and transactivation response DNA binding protein-43 (TDP-43) have been identified, as well as proteins associated with myofibrillar myopathies (MFM), including desmin and αB-crystallin. However, not all observations have been consistently reproduced.[13,14] Mitochondrial changes have also been proposed for inclusion in IBM diagnostic criteria,[15]. Clear guidelines for the incorporation of immunohistochemical findings and mitochondrial changes into diagnostic criteria for IBM have not been established.[16]

Previously, we have shown that the characteristic pattern of weakness associated with IBM is indicative of the diagnosis, even if Griggs pathological features are absent.[11] However, it is not invariably found at presentation. Here we sought to identify which pathological features, other than the Griggs pathological criteria, add further support to the diagnosis of IBM. We systematically investigated which pathological features are present in Griggs pathologically-definite IBM and then established the diagnostic utility of these features in cases of IBM lacking the Griggs criteria, using myopathies considered in the differential diagnosis of IBM as controls.

MATERIALS AND METHODS

The study received ethical approval from the Departments of Research and Development at Oxford University Hospitals NHS Trust, Oxford and University College London Hospitals NHS Foundation Trust, London.

Cases

All patients were followed by specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London. Biopsies were taken for diagnostic purposes and prior to any treatment.

Methods for demonstrating pathological features in IBM, additional to those defined by the Griggs criteria, were determined in six Griggs pathologically-definite cases of IBM. Cases with no clinical or pathological evidence of neuromuscular disease were used as controls. The diagnostic utility of the pathological features identified was assessed in two groups of clinically-typical IBM; one with rimmed vacuoles on muscle biopsy (IBM+RV; n=15), the other without rimmed vacuoles on muscle biopsy (IBM-RV; n=9). Disease controls were cases of steroid-responsive inflammatory myopathies [polymyositis and dermatomyositis; (PM&DM); n=11], protein accumulation myopathies with rimmed vacuoles (PAM; n=7). Clinical characteristics and inclusion criteria are summarised in Supplementary tables 1 and 2. Tissue from brains donated to the Queen Square Brain Bank for Neurological Disorders was used as positive controls for protein aggregate staining.

Tissue sections were stained with haematoxylin and eosin (H&E), combined cytochrome oxidase (COX) succinate dehydrogenase (SDH) histochemistry and for amyloid using alkalinised Congo red, crystal violet and thioflavin S. Immunohistochemistry (IHC) was performed using eight µm frozen sections. Sections were fixed, if required, incubated in 0.5% hydrogen peroxide then 5% normal goat serum and then systematically stained for: 1) proteins classically associated with neurodegenerative disease: tau and hyperphosphorylated tau, ubiquitin, Aβ and α-synuclein; 2) proteins more recently reported in neurodegenerative disease: p62, TDP-43, fused in sarcoma protein (FUS) and valosin containing protein (VCP); 3) nuclear membrane proteins: lamin A/C and emerin; 4) proteins associated with MFM: desmin, myotilin and αB-crystallin; and 5) inflammatory cells and major histocompatibility complex class I (MHC Class I): T-cells, helper T-cells, cytotoxic T-cells, B-cells and macrophages. Primary antibody binding was visualised using Dako REAL™ EnVision™ Detection System. Details of commercial antibodies and conditions used are provided in Supplementary Table 3. IHC was performed using a positive and negative control.

Definitions and quantification

The total number of fibres and the number undergoing partial invasion, containing rimmed vacuoles, protein aggregates and COX-negative SDH-positive (COX-/SDH+) fibres were quantified using ImagePro version 6.2 (Media Cybernetics), to ensure that the whole biopsy was systematically analysed. Only transversely-orientated fibres not undergoing necrosis or regeneration were quantified. Tissue sections stained with Congo red were visualised under fluorescent and polarised light. Areas of fluorescence were examined using rhodamine red (excitation 512-546 nm and emission 600-640 nm) and fluorescein isothiocyanate (excitation 440-480 nm and emission 527-530 nm) filters to exclude auto-fluorescence. Supplementary Table 4 provides definitions of the pathological features assessed. The inflammatory infiltrate and MHC Class I staining were analysed using a modified version of the semi-quantitative juvenile dermatomyositis score-tool (Supplementary Figure 1).[17] Assessments were performed blind to clinical details and diagnosis by a single individual (SB). Ten per cent of slides were re-counted to assess intra-observer reliability and 336 slides were assessed independently by two observers (SB and JLH) to determine inter-observer reliability.

Statistical analysis

Statistical analyses were performed using GraphPad PRISM version 5. Continuous and categorical variables were compared using Mann Whitney *U*-test and chi-squared or Fisher's exact test respectively. Spearman's rank order correlation was used to determine the strength and direction of associations between pathological findings. Linear regression was used to determine relationships between clinical features and pathological findings. Test characteristics were calculated using receiver operating characteristic (ROC) curves and 2x2 contingency tables. A test was considered diagnostic when sensitivity >75% and specificity >95% or sensitivity >95% and specificity >75%. Intra-observer and inter-observer agreement was calculated using Bland-Altman plots and Cohen's kappa statistic (κ). Repeat counts were within 95% confidence intervals using Bland-Altman plots and κ was ≥0.7 indicating good intra-observer and good or excellent inter-observer reliability. Statistical significance was set at $p<0.05$.

RESULTS

Pathological findings in Griggs' pathologically-definite IBM

p62, TDP-43, ubiquitin, myotilin and αB-crystallin immunoreactive aggregates were present in all six IBM cases but not in normal controls (Figures 1A-E). p62 and αB-crystallin immunoreactive aggregates were present in a greater percentage of fibres than the pathological features required in the Griggs criteria ($p<0.05$) (Figure 2). Despite their abundance, αB-crystallin immunoreactive aggregates were difficult to quantify due to a significant variability in their morphology. No immunoreactive deposits were observed in IBM cases or normal controls with antibodies to tau and phosphorylated tau, Aβ, α-synuclein, desmin, emerin, lamin A/C, FUS or VCP. Alkalinised Congo red staining was more sensitive than crystal violet and thioflavin S staining for observing amyloid aggregates (Figure 1F). Tissue sections containing congophilic deposits identified under fluorescence light showed no apple-green birefringence under polarised light. Mitochondrial changes and increased sarcolemmal and sarcoplasmic MHC Class I staining were observed in all six IBM cases, but not in normal controls. The inflammatory infiltrate was predominantly composed of endomysial cytotoxic T-cells and macrophages, with relatively few B-cells.

Quantitative analysis of pathological features in IBM and disease controls

Having shown that p62, TDP-43, ubiquitin and myotilin aggregates, congophilic deposits, MHC Class I and inflammatory cells were prevalent in Griggs' pathologically-definite IBM, the presence of these abnormalities, together with mitochondrial changes were assessed in IBM+RV, IBM-RV and disease controls.

The percentage of fibres containing p62, TDP-43, myotilin and ubiquitin aggregates and congophilic deposits were greater in IBM+RV than in IBM-RV; there was no difference in the number of COX-/SDH+ fibres (Figure 3A-F). Protein aggregates were observed in morphologically-normal fibres and in fibres exhibiting Griggs' pathological features. p62 and TDP-43 positive aggregates were present in a greater percentage of fibres in IBM+RV compared to PAM; however, there were no differences in

the percentage of fibres containing myotilin and ubiquitin aggregates or congophilic deposits. The percentage of fibres containing p62, TDP-43 and ubiquitin aggregates or congophilic deposits were similar in IBM-RV and PM&DM; however, myotilin aggregates were present in a greater percentage of fibres in PM&DM and COX-/SDH+ fibres were more abundant in IBM-RV. Analysis of the inflammatory infiltrate (T-cells, B-cells and macrophages) in the endomysium, perimysium and perivascular areas revealed that there were greater numbers of inflammatory cells in the endomysium and perimysium in IBM+RV than in PAM (p<0.03). The distribution and intensity of the inflammatory infiltrate in IBM-RV and PM&DM was similar.

Diagnostic utility of pathological features in IBM and disease controls

To mimic the diagnostic difficulty encountered in clinical practice, the ability of each test to differentiate between myopathies containing rimmed vacuoles (IBM+RV and PAM) and between inflammatory myopathies (IBM–RV and PM&DM) was assessed.

Diagnostic utility determined using receiver-operating characteristic curves

Individually, the presence of p62 immunoreactive inclusions and COX-/SDH+ fibres had the highest sensitivity and specificity for differentiating IBM+RV from PAM, (Supplementary Figure 2) (Table 1). Differentiating between IBM–RV and PM&DM, myotilin positive inclusions or COX-/SDH+ fibres had the highest sensitivity and specificity for IBM-RV (Supplementary Figure 3) (Table 1). Only the presence of myotilin positive inclusions satisfied criteria to be considered suitable as a diagnostic test (<0.01% of fibres containing myotilin aggregates had a sensitivity of 100% and specificity of 82% for IBM-RV).

Table 1 Test characteristics

Test feature	IBM+RV *v.* PAM				IBM-RV *v.* PM&DM			
	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity
Rimmed vacuoles	0.60	>0.28	0.53	0.71	-	-	-	-
p62 aggregates	0.87	>0.48	0.87	0.86	0.60	>0.21	0.22	0.91
TDP-43 aggregates	0.80	>0.34	0.80	0.86	0.53	<0.01	0.89	0.18
Ubiquitin aggregates	0.68	>0.18	0.53	0.85	0.64	<0.01	1.00	0.27
Myotilin aggregates	0.55	<0.25	1.00	0.29	0.91	<0.01	1.00	0.82
Congophilic deposits	0.56	>0.24	0.73	0.71	0.56	<0.03	0.11	0.82
COX-/SDH+ fibres	0.87	>0.04	0.86	0.86	0.93	>0.1	0.78	0.91

Table shows the area under the curve and optimum cut-off for each test with the accompanying sensitivity and specificity. AUC = Area under the curve.

Diagnostic utility determined by comparing proportion of affected cases in each diagnostic group

In the aforementioned experiments, the number of fibres within each muscle biopsy was quantified. However, this is impractical for routine clinical use. Thus, the proportions of affected cases in each group were compared (Table 2). This revealed that neither staining for protein aggregates nor congophilic deposits could differentiate between IBM+RV and PAM. The pathological findings in IBM-RV and PM&DM were also similar, except that the absence of myotilin immunoreactive aggregates was sensitive and specific for IBM-RV. COX-/SDH+ fibres were also suggestive of IBM-RV; one or more COX-/SDH+ fibres had a sensitivity of 100% and specificity 73% for IBM-RV.

Increased MHC Class I expression lacked specificity. However, strong (diffuse sarcolemmal and sarcoplasmic) MHC Class I up-regulation was diagnostic for IBM+RV, differentiating it from PAM, as were the presence of either endomysial T-cell or helper T-cell scores >1 or an endomysial cytotoxic T-cell score >0. Partial invasion was specific for IBM+RV, but lacked sensitivity. Greater numbers of perimysial T-cells, cytotoxic T-cells and endomysial B-cells were observed in PM&DM than in IBM–RV ($p \leq 0.02$), however, this was not diagnostically useful. There was no difference in the proportion of cases with fibres undergoing partial invasion between IBM–RV and PM&DM.

Table 2 Comparison of the proportion of positive cases in each group

Pathological features	IBM+RV	PAM	IBM+RV *v.* PAM		IBM–RV	PM&DM	IBM-RV *v.* PM&DM		IBM+RV *v.* IBM-RV
	n (%)	n (%)	Sensitivity	Specificity	n (%)	n (%)	Sensitivity	Specificity	*p* value
Number of cases	15 (100)	7 (100)			9 (100)	11 (100)			
Aggregated proteins, n (%)									
p62	15 (100)	6 (86)	1.00	0.14	4 (44)	3 (27)‡	0.40	0.72	0.003*
TDP-43	13 (87)	5 (71)	0.87	0.29	1 (11)	2 (18)‡	0.11	0.82	0.001*
Ubiquitin	11 (73)	4 (57)	0.73	0.43	0 (0)	3 (27)‡	0.00	0.73	0.001*
Myotilin	10 (67)	5 (71)	0.67	0.29	0 (0)	9 (82)	0.00	0.18	0.002*
Congophilic deposits	13 (87)	7 (100)	0.87	0.00	1 (11)	0 (0)	0.11	1.00	0.001*
COX-/SDH+ fibres†, n (%)									
Any	12 (86)	2 (29)	0.80	0.71	9 (100)	3 (27)	1.00	0.73	0.5
Inflammatory features, n (%)									
MHC Class I up-regulation	15 (100)	3 (43)	1.00	0.57	9 (100)	11 (100)	1.00	0.00	1.00
Strong MHC Class I up-regulation	14 (93)	0 (0)	0.93	1.00	9 (100)	10 (91)	1.00	0.09	0.53
Partial invasion	10 (67)	0 (0)	0.67	1.00	3 (33)	2 (18)	0.33	0.82	0.11
Endomysial CD3+ T-cell score >1	13 (87)	0 (0)	0.87	1.00	4 (44)	7 (64)	0.44	0.36	0.02*
Endomysial CD4+ helper T-cell score >1	12 (80)	0 (0)	0.80	1.00	2 (22)	5 (45)	0.22	0.46	0.01*
Endomysial CD8+ cytotoxic T-cell score >0	14 (93)	0 (0)	0.93	1.00	4 (44)	5 (45)	0.44	0.54	0.02*
Endomysial CD68+ macrophage score >1	12 (80)	0 (0)	0.80	1.00	4 (44)	8 (73)	0.44	0.17	0.07

†In IBM with rimmed vacuoles *n*=14. ‡Pathological features present in DM, but not PM cases. *Statistically significant results.

Because IBM-RV is more pathologically akin to PM than DM, analyses were repeated comparing IBM-RV and PM cases (*n*=6). No p62, TDP-43 or ubiquitin immunoreactive aggregates were observed in PM cases and the diagnostic utility of tests for differentiating between IBM-RV and PM yielded similar results to prior analyses between IBM-RV and PM&DM.

Diagnostic utility of categorising the pattern of p62 staining

The pattern of p62 staining could be categorised into four distinct groups (Figure 1G-J). Aggregates observed in IBM were present in vacuolated and non-vacuolated fibres and were strongly stained, discreet and clearly delineated, round or angular and typically located subsarcolemmal, perinuclear and peri-vacuolar (pattern I). This pattern was observed in every IBM case with p62 aggregates, one (9%) case of DM and three (43%) cases of PAM (hereditary IBM, dystrophinopathy and genetically undefined MFM). Patterns II and III were associated with PAM, particularly myotilinopathy (n=2; 67%), and DM (n=2; 40%) respectively. Pattern IV occurred in a genetically undefined case of MFM. No differences were observed in the morphology of TDP-43, myotilin or ubiquitin aggregates between biopsies.

Clinicopathological correlation

In IBM+RV, IBM-RV and pathologically-definite IBM, there were no correlations in individual biopsies between pathological features. No relationships were identified between the pathological findings and age at symptom onset, age at biopsy, disease duration or serum creatine kinase. The same results were obtained when the IBM groups were analysed separately and as one.

Proposed diagnostic algorithm

Based on our pathological findings, we propose a diagnostic algorithm for differentiating IBM from its disease mimics (Figure 4).

The algorithm was tested in a further 23 cases that fulfilled the criteria for IBM+RV (*n*=12) and IBM-RV (*n*=11). The algorithm correctly diagnosed 20 (87%) cases: 12 (100%) cases of IBM+RV and eight (73%) cases of IBM-RV. In IBM-RV, COX-/SDH+ fibres were present in 8 (73%) cases, pattern I p62 aggregates in 8 (73%) cases and both in 6 (55%) cases.

DISCUSSION

While Griggs' pathological criteria have been accepted as diagnostic of IBM, many patients who, observed over time undoubtedly have IBM, lack one or more of the Griggs pathological features at presentation, and even on repeat biopsy.[8,11] Despite IBM being associated with a characteristic pattern of finger flexor and knee extensor weakness, not all patients have this pattern at disease onset, and muscle biopsy remains an important tool in differentiating IBM from its mimics. We sought to determine which additional pathological features support a diagnosis of IBM, demonstrating that characteristic p62 immunoreactive aggregates, strong MHC Class I upregulation, endomysial cytotoxic T-cell score >0 and COX-/SDH+ fibres are features with sufficient sensitivity and specificity to differentiate IBM from pathologically similar myopathies and we propose an easily applied pathological algorithm for the diagnosis of IBM (Figure 4).

In agreement with previous studies, we observed p62,[18] TDP-43,[19] ubiquitin [13] and αB-crystallin [20] immunoreactive aggregates and a predominantly endomysial inflammatory infiltrate [3] in Griggs pathologically-definite IBM. Diagnostic pathological studies of IBM have concentrated on differentiating IBM from other inflammatory myopathies and two recent quantitative studies have found that TDP-43, p62 and LC3 may be of diagnostic use.[21,22] However, in these studies only a fraction of each biopsy was analysed i.e. 200 fibres. We have found this limited quantification does not correlate with the percentage of affected fibres in a biopsy nor does it reflect the way in which a muscle biopsy is assessed. Additionally, studies have lacked vacuolar myopathy control cases as it is believed that the inflammatory changes present in IBM enable it to be easily differentiated from other vacuolar myopathies.[22] However, inflammatory changes are not infrequently observed in muscular

dystrophies and the degree of inflammatory change necessary to confidently diagnose IBM is currently unknown.

To mimic the typical diagnostic conundrums encountered in clinical practice, we evaluated the ability of the pathological findings to differentiate IBM+RV from other vacuolar myopathies and IBM-RV from steroid-responsive inflammatory myopathies. We found that quantitative analysis of protein aggregates, congophilic deposits and COX-/SDH+ fibres was of limited diagnostic use. Analysing the biopsies dichotomously and using a semi-quantitatively score-tool revealed that increased MHC Class I labelling was sensitive for IBM making it a good initial screening test, its absence excluding the diagnosis. In agreement with an earlier study, we found p62 aggregates identified the largest number of affected fibres in IBM.[23] Additionally, as a novel finding, the morphology and distribution of p62 aggregates was characteristic in IBM. This characteristic pattern of p62 immunoreactive aggregates was 100% sensitive for IBM+RV; their absence from a biopsy containing rimmed vacuoles effectively ruling-out a diagnosis of IBM. We confirmed that the most diagnostically useful pathological findings in IBM+RV were evidence of an immune mediated process; strong MHC Class I staining or an endomysial cytotoxic T-cell score >0 were both diagnostic. Having identified either of these features in a biopsy containing rimmed vacuoles no extra diagnostic certainty was gained from observing partial invasion, COX-/SDH+ fibres or congophilic deposits.

The most discriminative pathological tests for differentiating between IBM–RV and PM&DM were COX/SDH staining and myotilin IHC. Consistent with a recent study,[9] we found the absence of mitochondrial changes to be strong evidence against a diagnosis of IBM. There was no difference in the median age between IBM-RV and PM&DM cases to account for the difference observed in COX-/SDH+ fibres. The presence of myotilin and ubiquitin immunoreactive aggregates appeared to rule out a diagnosis of IBM-RV. However, we believe the presence of these features in IBM+RV indicates that they are unlikely to be diagnostically reliable features for differentiating between IBM-RV and steroid-responsive inflammatory myopathies. Pattern I p62 immunoreactive aggregates were only

present in 44% of the initial IBM-RV cases tested, but they were not observed in PM cases and were very rare in DM. Their presence in a further eight out of 11 (73%) cases of IBM-RV that were assessed suggests that p62 IHC warrants further investigation and validation in a larger, independent series.

Almost all pathological features - protein aggregates, congophilic deposits and inflammation - were more abundant in IBM+RV than IBM-RV. Despite using slightly different inclusion criteria, similar differences have been reported between pathologically-typical and pathologically-atypical IBM.[21] However, we found no differences in the number of COX-/SDH+ fibres, the degree of MHC Class I upregulation, the morphology and distribution of p62 immunoreactive aggregates or the pattern of the inflammation between IBM+RV and IBM-RV, supporting our clinical observations that these are the same disease. We believe that the pathological differences between IBM+RV and IBM-RV are, in part, due to differences in disease duration. Two studies have shown that rimmed vacuoles are more common in patients who are older at the time of muscle biopsy,[24,11] suggesting that they are associated with chronologically more advanced disease. Therefore, the pathological findings which are more abundant in IBM+RV and thought to be typical of IBM may instead be indicative of chronologically more advanced disease explaining their limited sensitivity at disease presentation. However, possibly due to the number of cases analysed, we were unable to confirm a relationship between pathological features and clinical findings.

A robust clinicopathological definition of IBM is of paramount importance for diagnosis and for selection and entry of patients into clinical trials. We have shown that certain pathological findings are more abundant than those included in the current pathologically-focussed diagnostic criteria. Moreover, p62 immunoreactive deposits, increased MHC Class I expression, endomysial cytotoxic T-cells and COX-/SDH+ fibres have sufficient sensitivity and specificity to enable the histological differentiation of IBM from disease mimics, supporting their inclusion as diagnostic criteria for IBM alongside clinical criteria. Using our diagnostic algorithm, we found there would be little additional

diagnostic security in identifying partial invasion, performing EM or staining for amyloid deposits. Finally, mitochondrial changes and MHC Class I up-regulation were the most consistent findings in our IBM cases suggesting that they are central to the pathogenesis and that further investigation and therapeutic intervention should be directed towards these features.

REFERENCES

1. Needham M, James I, Corbett A, Day T, Christiansen F, Phillips B, et al. Sporadic inclusion body myositis: phenotypic variability and influence of HLA-DR3 in a cohort of 57 Australian cases. J Neurol Neurosurg Psychiatr. 2008;79:1056–60.
2. Griggs RC, Askanas V, DiMauro S, Engel A, Karpati G, Mendell JR, et al. Inclusion body myositis and myopathies. Ann Neurol. 1995;38:705–13.
3. Arahata K, Engel AG. Monoclonal antibody analysis of mononuclear cells in myopathies. I: Quantitation of subsets according to diagnosis and sites of accumulation and demonstration and counts of muscle fibers invaded by T cells. Ann Neurol. 1984;16:193–208.
4. Mhiri C, Gherardi R. Inclusion body myositis in French patients. A clinicopathological evaluation. Neuropathol Appl Neurobiol. 1990;16:333–44.
5. Villanova M, Kawai M, Lübke U, Oh SJ, Perry G, Six J, et al. Rimmed vacuoles of inclusion body myositis and oculopharyngeal muscular dystrophy contain amyloid precursor protein and lysosomal markers. Brain Res. 1993;603:343–7.
6. Van der Meulen MF, Hoogendijk JE, Moons KG, Veldman H, Badrising UA, Wokke JH. Rimmed vacuoles and the added value of SMI-31 staining in diagnosing sporadic inclusion body myositis. Neuromuscul Disord. 2001;11:447–51.
7. Ferrer I, Olivé M. Molecular pathology of myofibrillar myopathies. Expert Rev Mol Med. 2008;10:e25.
8. Amato AA, Gronseth GS, Jackson CE, Wolfe GI, Katz JS, Bryan WW, et al. Inclusion body myositis: clinical and pathological boundaries. Ann Neurol. 1996;40:581–6.

9. Chahin N, Engel AG. Correlation of muscle biopsy, clinical course, and outcome in PM and sporadic IBM. Neurology. 2008;70:418–24.

10. Benveniste O, Guiguet M, Freebody J, Dubourg O, Squier W, Maisonobe T, et al. Long-term observational study of sporadic inclusion body myositis. Brain. 2011;134:3176–84.

11. Brady S, Squier W, Hilton-Jones D. Clinical assessment determines the diagnosis of inclusion body myositis independently of pathological features. J Neurol Neurosurg Psychiatr. 2013; 16.

12. Greenberg SA. Theories of the Pathogenesis of Inclusion Body Myositis. Current Rheumatology Reports. 2010;12:221–8.

13. Sherriff FE, Joachim CL, Squier MV, Esiri MM. Ubiquitinated inclusions in inclusion-body myositis patients are immunoreactive for cathepsin D but not β-amyloid. Neuroscience Letters. 1995;194:37–40.

14. Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. BMJ. 2009;339:b2680.

15. Needham M, and Mastaglia FL. (2007). Inclusion body myositis: current pathogenetic concepts and diagnostic and therapeutic approaches. Lancet Neurol. 6, 620–631.

16. Benveniste O, Hilton-Jones D. International Workshop on Inclusion Body Myositis held at the Institute of Myology, Paris, on 29 May 2009. Neuromuscular Disorders. 2010;20:414–21.

17. Wedderburn LR, Varsani H, Li CKC, Newton KR, Amato AA, Banwell B, et al. International consensus on a proposed score system for muscle biopsy evaluation in patients with juvenile dermatomyositis: a tool for potential use in clinical trials. Arthritis Rheum. 2007;57:1192–201.

18. Nogalska A, Terracciano C, D'Agostino C, King Engel W, Askanas V. p62/SQSTM1 is overexpressed and prominently accumulated in inclusions of sporadic inclusion-body myositis muscle fibers, and can help differentiating it from polymyositis and dermatomyositis. Acta Neuropathol. 2009;118:407–413.

19. Weihl CC, Temiz P, Miller SE, Watts G, Smith C, Forman M, Hanson PI, Kimonis V, Pestronk A. TDP-43 accumulation in inclusion body myopathy muscle suggests a common pathogenic

mechanism with frontotemporal dementia. J. Neurol. Neurosurg. Psychiatr. 2008;79:1186–1189.

20. Banwell BL, Engel AG. Alpha B-crystallin immunolocalization yields new insights into inclusion body myositis. Neurology. 2000;54:1033–1041.

21. Dubourg O, Wanschitz J, Maisonobe T, Béhin A, Allenbach Y, Herson S, and Benveniste O. Diagnostic value of markers of muscle degeneration in sporadic inclusion body myositis. Acta Myol. 2011. 30, 103–108.

22. Hiniker A, Daniels BH, Lee HS, Margeta M. Comparative utility of LC3, p62 and TDP-43 immunohistochemistry in differentiation of inclusion body myositis from polymyositis and related inflammatory myopathies. Acta Neuropathologica Communications. 2013. 1:29.

23. D'Agostino C, Nogalska A, Engel WK, Askanas V. In sporadic inclusion-body myositis muscle fibres TDP-43-positive inclusions are less frequent and robust than p62-inclusions, and are not associated with paired helical filaments. Neuropathol Appl Neurobiol. 2010; Available from: http://www.ncbi.nlm.nih.gov/pubmed/20626631.

24. Momma K, Noguchi S, Malicdan MCV, Hayashi YK, Minami N, Kamakura K, et al. Rimmed vacuoles in Becker muscular dystrophy have similar features with inclusion myopathies. PLoS ONE. 2012;7:e52002.

ACKNOWLEDGEMENTS

SB is funded by the Myositis Support Group. JLH is supported by the Reta Lila Weston Institute for Neurological Studies and the Myositis Support Group. This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme.

CONTRIBUTORSHIP STATEMENT

Dr Stefen Brady - Acquisition of data, analysis and interpretation of data and drafting of manuscript.

Dr Waney Squier - Critical revision of manuscript for important intellectual content.

Prof. Caroline Sewry - Study concept and design and critical revision of manuscript for important intellectual content.

Prof. Mike Hanna - Critical revision of manuscript for important intellectual content.

Dr David Hilton-Jones - Critical revision of manuscript for important intellectual content.

Dr Janice Holton - Study concept and design, critical revision of manuscript for important intellectual content and study supervision.

DATA SHARING

All additional data can be found in supplementary tables and figures.

Supplementary Figure 1 IBM inflammatory score-tool

Score tool modified from the published juvenile dermatomyositis inflammatory (JDM) score tool [17] to specifically assess the type, degree and distribution of inflammation in IBM. The inflammatory domain was augmented to include T-cells, T-cell subtypes, B-cells and macrophages. MHC Class I staining was expanded to include three patterns of labelling. The vascular, muscle fibre and connective tissue domains which are present in the JDM score tool were not included.

Figure 1 Protein aggregates and congophilic deposits in IBM

Stained cryostat sections, showing fibres, often in clusters, containing protein aggregates stained for p62 (A), TDP-43 (B), ubiquitin (C), αB-crystallin (D) and myotilin (E). Protein aggregates were present throughout fibres, and were observed in apparently normal fibres, vacuolated fibres and fibres surrounded by inflammatory infiltrates. In fibres containing TDP-43 aggregates, myonuclear TDP-43 staining was frequently reduced (B). Congophilic deposits were observed in vacuolated fibres using epifluorescence (F). Tissue sections were examined using both the rhodamine red and fluorescein isothiocyanate filters to exclude areas of auto-fluorescence (arrow). Combined fluorescent image is shown. Four patterns of immunoreactivity were observed in IBM and disease controls stained for p62 using IHC (G)(H)(I)(J). Pattern I (G) - strongly stained, discreet and clearly delineated, round or

angular aggregates, variable in number and size within a muscle fibre but rarely filling it and predominantly located subsarcolemmal, but also perinuclear and adjacent to vacuoles. Pattern II (H) - large aggregates of variable staining intensity. Pattern III (I) - fine granular aggregates dispersed throughout the fibre. Pattern IV (J) - fine granules and wisps of p62 immunoreactivity set within weakly basophilic inclusions.

Scale bar represents 50 µm in A and D; 25 µm in B, C and E-J.

Figure 2 Percentage of muscle fibres containing protein aggregates and Griggs' pathological features

Box and whisker plot illustrating the percentage of muscle fibres containing pathological abnormalities contained in the Griggs criteria and protein aggregates in Griggs' pathologically-definite IBM. Fibres containing aggregates immunoreactive for p62 and αB-crystallin were more frequent than those containing the current diagnostic pathological features (red bars) ($p<0.05$). Protein aggregates recognised by all antibodies were found in a significantly larger number of fibres than partial invasion ($p<0.02$).

Figure 3 Percentage of fibres containing protein aggregates and COX-/SDH+ fibres in each group

Box and whisker plots illustrating the percentage of fibres in each diagnostic category containing p62 (A), TDP-43 (B), myotilin (C) and ubiquitin (D) immunoreactive aggregates, congophilic deposits (E) and COX-/SDH+ fibres (F). All protein aggregates were present in a greater percentage of fibres in IBM+RV than in IBM–RV. There was no difference in the percentage of COX-/SDH+ muscle fibres between these groups. IBM+RV biopsies had a greater percentage of fibres containing p62 (A) and TDP-43 (B) immunoreactive aggregates and COX-/SDH+ fibres (F) than PAM. Pathological findings were similar in IBM-RV and PM&DM, with no differences in the percentage of fibres containing p62 (A), TDP-43 (B) and ubiquitin (D) immunoreactive aggregates or congophilic deposits (E). However, there was a greater percentage of COX-/SDH+ fibres (F) in IBM–RV than PM&DM and a greater

percentage of fibres containing myotilin immunoreactive aggregates (C) in PM&DM than IBM-RV. *Statistically significant results.

Supplementary Figure 2 Sensitivity and specificity of rimmed vacuoles, protein aggregates and mitochondrial changes in IBM+RV compared to PAM

Receiver operating characteristic curves for each test including the area under the curve and optimum cut-off with its associated sensitivity and specificity for rimmed vacuoles (A), myotilin (B), ubiquitin (C), TDP-43 (D), p62 (E) immunoreactive deposits, congophilic deposits (F) and COX-/SDH+ fibres (G). COX/SDH HC staining was the most discriminative test for differentiating IBM+RV and PAM (G). However, there was little difference between COX/SDH HC staining, TDP-43 and p62 IHC staining and none were sufficiently discriminative to be considered diagnostic. AUC = Area under the curve.

Supplementary Figure 3 Sensitivity and specificity of protein aggregates and mitochondrial changes in IBM-RV compared to PM&DM

Receiver operating characteristic curves for each test showing the area under the curve and optimum cut-off with its sensitivity and specificity for myotilin (A), ubiquitin (B), TDP-43 (C), p62 (D) immunoreactive deposits, congophilic deposits (E) and COX-/SDH+ fibres (F). COX/SDH histochemical staining (F) and myotilin (G) IHC were the most discriminative tests for differentiating IBM-RV and PM&DM. AUC = Area under the curve.

Figure 4 Proposed diagnostic algorithm for IBM based on pathological findings

Flow diagram showing a proposed pathway for diagnosing IBM based on the pathological findings. Increased MHC Class I staining was observed in all cases of IBM and pattern I p62 aggregates in all cases of IBM+RV making them good initial screening tests. Their absence rules-out a diagnosis of IBM and IBM+RV respectively. The presence of an endomysial cytotoxic T-cell score >0 or strong MHC Class I staining in a biopsy with rimmed vacuoles and p62 aggregates secures a diagnosis of IBM+RV. Differentiating IBM-RV and PM&DM pathologically is more challenging. The presence of

COX-/SDH+ fibres is not specific to IBM-RV. However, an absence of COX-/SDH+ fibres effectively rules-out a diagnosis of IBM-RV. Pattern I p62 aggregates may enable IBM to be differentiated from PM when present. However, they may lack sensitivity for IBM-RV, therefore their absence does not rule out the diagnosis.

Supplementary Table 1 Clinical characteristics

Characteristic	G-IBM	IBM+RV	IBM-RV	PM&DM	PAM	IBM+RV*	IBM-RV*
Number of cases	6	15	9	11	7	12	11
Male:female	5:1	10:5	4:5	4:3	4:7	10:2	9:2
Median age at symptom onset, years (IQR)	69 (66-70)	54 (49-67)	62 (48-68)	55 (34-65)	46 (24-54)	58 (55-73)	60 (57-72)
Median age at muscle biopsy, years (IQR)	77 (68-78)	64 (59-71)	68 (47-74)	55 (34-65)	54 (29-59)	66 (62-77)	70 (63-74)
Median duration of symptoms, years (IQR)	5 (3-9)	5 (4-7)	3 (2-8)	0 (0-0)	5 (3-9)	5 (4-7)	4 (3-7)
Mean creatine kinase, IU/L, mean (±SD)	377 (±213)	1748 (±1348)	926 (±800)	6744 (±5875)	739 (±320)	662 (±360)	466 (±338)
Mean number of muscle fibres per biopsy	2929 (±1357)	1463 (±954)	1795 (±990)	3534 (±1934)	2749 (±1357)	NA	NA

G-IBM = Griggs' pathologically-definite IBM; IQR = Interquartile range; SD = Standard deviation; NA = Not applicable. *Cases used to test proposed diagnostic flow-chart.

Supplementary Table 2 Clinical inclusion criteria

Diagnostic category	Criteria
G-IBM	Patients fulfilling Griggs' definite criteria (rimmed vacuoles, inflammatory infiltrate with partial invasion of fibres and 15-18 nm tubulofilaments on EM) with prominent finger flexor and knee extensor weakness and CK <12 x ULN.
IBM+RV	Age at symptom onset >45 years, symptoms present for >12 months, finger flexion strength less than shoulder abduction strength and knee extension weakness greater than hip flexion weakness, CK ≤15 x ULN and a muscle biopsy revealing rimmed vacuoles on H&E or GT stained sections without features inconsistent with IBM on a standard diagnostic histological assessment.
IBM-RV	Clinical features and CK as detailed under IBM+RV. Rimmed vacuoles absent on H&E and GT stained sections and without features inconsistent with IBM on a standard diagnostic histological assessment.
PAM	Genetically or clinically and pathologically confirmed cases of PAM with typical rimmed vacuoles present on muscle biopsy. Cases included myotilinopathy (n=2), hIBM with compound heterozygous mutations in GNE (n=1), IBMPFD with mutation in VCP (n=1), genetically unconfirmed cases of myofibrillar myopathy (n=2), and dystrophinopathy with deletion of exons 45-47 (n=1).
PM&DM	Subacute onset of limb girdle weakness, significantly raised CK, inflammatory cell infiltrate present on muscle biopsy and a sustained unequivocal clinical and biochemical response to steroid immunosuppression. DM cases also had to have cutaneous manifestations consistent with the diagnosis.
Normal controls	Patients investigated for cramps or fatigue, normal clinical examination performed by a muscle specialist, normal CK, normal neurophysiological assessment and normal muscle biopsy.

G-IBM = Griggs' pathologically-definite IBM; IBM+RV = Clinically-typical IBM with rimmed vacuoles; IBM-RV = Clinically typical IBM lacking rimmed vacuoles; PAM = Protein accumulation myopathies with rimmed vacuoles; PM&DM = Steroid-responsive inflammatory myopathies; hIBM = Hereditary inclusion body myopathy; IBMPFD = Inclusion body myopathy with Paget's disease and frontotemporal dementia; CK = Creatine kinase; GT = Gomori trichrome; ULN = Upper limit of normal.

Supplementary Table 3 Antibodies and optimum staining conditions

Antibody	Source	Clone	Control tissue	Fixative	Dilution	Primary incubation conditions
p62	BD Transduction	3/P62	AD brain	A	1:400	1 hour, RT
TDP-43	Proteintech	NA	FTLD-TDP brain	PFA	1:800	24 hours, 4°C
Tau*	Dako	NA	AD brain	A	1:1600	1 hour, RT
Phosphorylated tau**	Autogen Bioclear	AT8	AD brain	A	1:1600	1 hour, RT
Ubiquitin	Dako	NA	AD brain	A	1:100	1 hour, RT
Aβ	Dako	6F/3D	AD brain	PFA and FA	1:100	1 hour, RT
α-synuclein	Abcam	4D6	MSA brain	PBS	1:800	1 hour, RT
FUS	Novus Biologicals	NA	FTLD-FUS brain	A	1:2000	1 hour, RT
Desmin	Dako	D33	Normal muscle	A	1:50	24 hours, 4°C
Myotilin	Novocastra	RSO34	Normal muscle	A	1:500	24 hours, 4°C
αB-crystallin	Novocastra	G2JF	CBD brain	A	1:300	1 hour, RT
VCP	Abcam	5	Normal muscle	A	1:100	1 hour, RT
Lamin A/C	Novocastra	636	Normal muscle	A	1:50	1 hour, RT
Emerin	Novocastra	4G5	Normal muscle	A	1:400	1 hour, RT
MHC Class I	Novocastra	W6/32	Normal muscle	A	1:25	24 hours, 4°C
CD3 (T-cells)	Novocastra	UCHT1	Tonsil	A	1:100	1 hour, RT
CD4 (Helper T-cells)	Novocastra	4B12	Tonsil	A	1:400	1 hour, RT
CD8 (Cytotoxic T-cells)	Novocastra	1A5	Tonsil	A	1:50	1 hour, RT
CD20 (B-cells)	Novocastra	L26	Tonsil	A	1:400	1 hour, RT
CD68 (Macrophages)	Novocastra	KP1	Tonsil	A	1:1600	1 hour, RT

NA = Not applicable; AD = Alzheimer's disease; FTLD-TDP = Frontotemporal lobar degeneration with TDP-43 positive inclusions; MSA = Multiple system atrophy; FTLD-FUS = Frontotemporal lobar degeneration with FUS positive inclusions; CBD = Corticobasal degeneration; A = Acetone; PFA = 4% Paraformaldehyde; FA = Formic acid; PBS = Phosphate buffered saline; RT = Room temperature. Antibodies were directed at * amino acids 243-441 irrespective of phosphorylation and ** phosphorylated Ser202.

Supplementary Table 4 Definitions of pathological features

Pathological feature	Definition
Rimmed vacuoles	Irregular vacuole with a granular basophilic rim or containing granular basophilic material when stained with H&E or stained red in the GT. Both H&E and GT stained sections were reviewed before concluding the absence of rimmed vacuoles.
Inflammatory infiltrate and partial invasion	Inflammatory cells must show a nucleus fully circumscribed by a ring of positive staining. T-cells and B-cells must have a lymphoid morphology. Partial invasion was defined as unequivocal invasion of an otherwise structurally normal fibre by one or more inflammatory cells on H&E stained sections or sections stained using IHC.
Protein aggregates	Area of definite staining within a transversely orientated muscle fibre. Diffuse staining affecting the whole of a fibre was not counted nor were protein aggregates in necrotic fibres or regenerating fibres.
Congophilic deposits	Assessed using polarising and fluorescence microscopes. Positive staining using a polarising microscope was defined as congophilic deposits within a muscle fibre that exhibited apple-green birefringence under polarised light. Positive staining with a fluorescence microscope was defined as fluorescent material within a muscle fibre only visible under the rhodamine red filter. Areas of auto-fluorescence were excluded by visualising areas of fluorescence with both rhodamine red and FITC filters.

GT = Gomori trichrome; FITC = Fluorescein isothiocyanate.

IBM Inflammatory Score Tool

Case Number:_____ Date:_____

	Score	Description
T-cells (CD3)		For each inflammatory cell type in the endomysial, perimysial and perivascular locations score positive infiltrating cells as follows: if none or <4 cells in a x20 field- score 0; if >4 cells in a x20 field and/or 1 cluster (where a cluster is ≥10 cells) - score 1; if >2 clusters in the entire biopsy, and/or diffusely infiltrating cells (i.e.> 20 cells in a x20 field) - score 2.
CD3+ endomysial infiltration	0, 1, 2	
CD3+ perimysial infiltration	0, 1, 2	
CD3+ perivascular infiltration	0, 1, 2	
Helper T-cells (CD4)		
CD4+ endomysial infiltration	0, 1, 2	
CD4+ perimysial infiltration	0, 1, 2	
CD4+ perivascular infiltration	0, 1, 2	
Cytotoxic T-cells (CD8)		
CD8+ endomysial infiltration	0, 1, 2	
CD8+ perimysial infiltration	0, 1, 2	
CD8+ perivascular infiltration	0, 1, 2	
B-cells (CD20)		
CD20+ endomysial infiltration	0, 1, 2	
CD20+ perimysial infiltration	0, 1, 2	
CD20+ perivascular infiltration	0, 1, 2	
Macrophages (CD68)		
CD68+ endomysial infiltration	0, 1, 2	
CD68+ perimysial infiltration	0, 1, 2	
CD68+ perivascular infiltration	0, 1, 2	
MHC Class I	0, 1, 2	For the whole biopsy score as follows: normal (capillary staining only) - score 0; if increased: i) mildly (weak diffuse sarcolemmal staining or scattered positive muscle fibres) - score 1; ii) strongly increased (diffuse definite sarcoplasmic and sarcolemmal increase in staining) score 2.

Supplementary Figure 1 IBM inflammatory score-tool
Score tool modified from the published juvenile dermatomyositis inflammatory (JDM) score tool [17] to specifically assess the type, degree and distribution of inflammation in IBM. The inflammatory domain was augmented to include T-cells, T-cell subtypes, B-cells and macrophages. MHC Class I staining was expanded to include three patterns of labelling. The vascular, muscle fibre and connective tissue domains which are present in the JDM score tool were not included.

188x255mm (300 x 300 DPI)

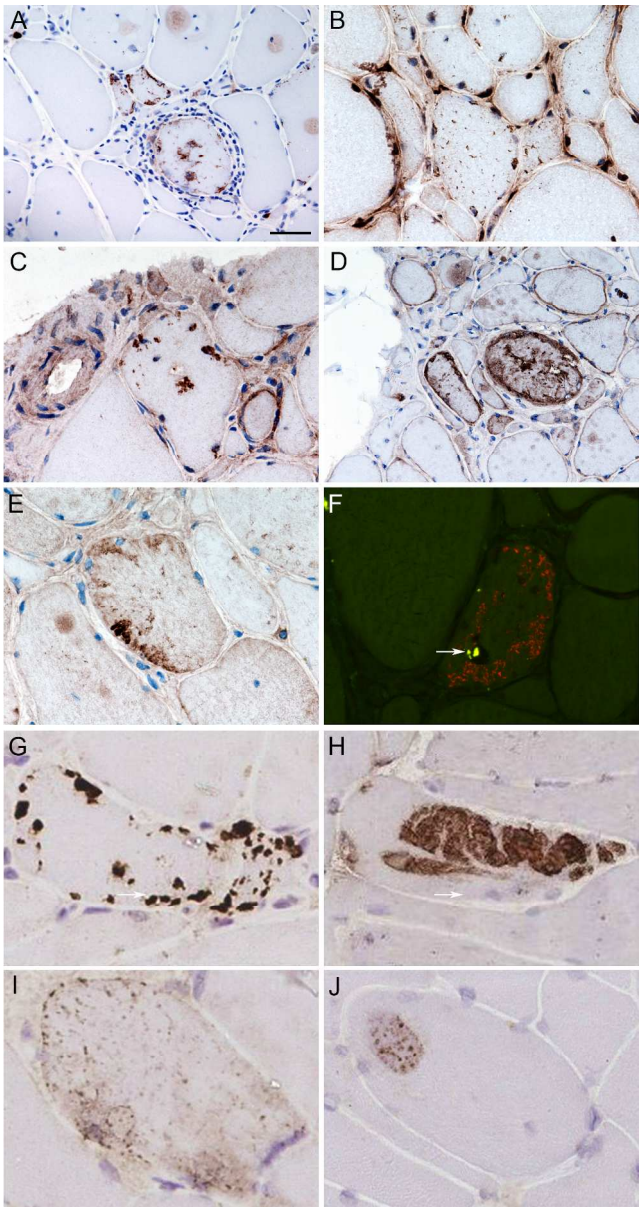


Figure 1 Protein aggregates and congophilic deposits in IBM
Stained cryostat sections, showing fibres, often in clusters, containing protein aggregates stained for p62 (A), TDP-43 (B), ubiquitin (C), αB-crystallin (D) and myotilin (E). Protein aggregates were present throughout fibres, and were observed in apparently normal fibres, vacuolated fibres and fibres surrounded by inflammatory infiltrates. In fibres containing TDP-43 aggregates, myonuclear TDP-43 staining was frequently reduced (B). Congophilic deposits were observed in vacuolated fibres using epifluorescence (F). Tissue sections were examined using both the rhodamine red and fluorescein isothiocyanate filters to exclude areas of auto-fluorescence (arrow). Combined fluorescent image is shown. Four patterns of immunoreactivity were observed in IBM and disease controls stained for p62 using IHC (G)(H)(I)(J). Pattern I (G) - strongly stained, discreet and clearly delineated, round or angular aggregates, variable in number and size within a muscle fibre but rarely filling it and predominantly located subsarcolemmal, but also perinuclear and adjacent to vacuoles. Pattern II (H) - large aggregates of variable staining intensity. Pattern III (I) - fine granular aggregates dispersed throughout the fibre. Pattern IV (J) - fine granules and wisps of

p62 immunoreactivity set within weakly basophilic inclusions.
Scale bar represents 50 μm in A and D; 25 μm in B, C and E-J.

161x305mm (300 x 300 DPI)

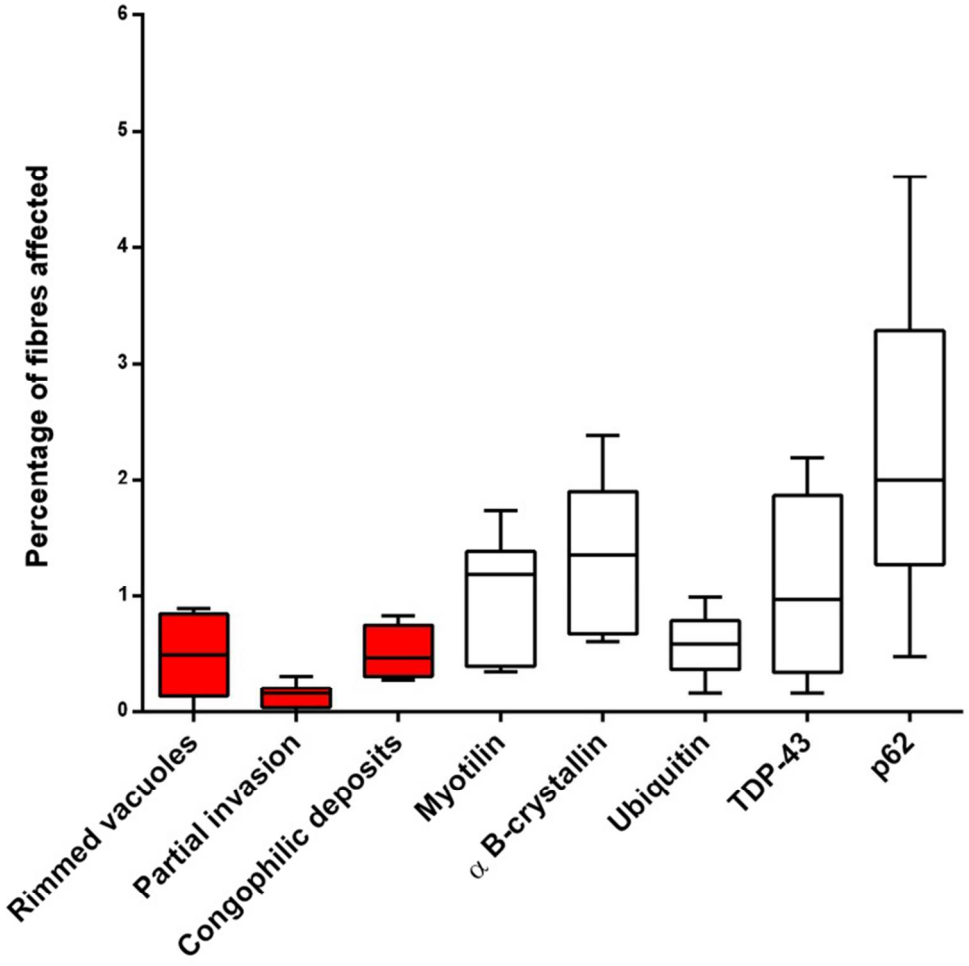


Figure 2 Percentage of muscle fibres containing protein aggregates and Griggs' pathological features
Box and whisker plot illustrating the percentage of muscle fibres containing pathological abnormalities contained in the Griggs criteria and protein aggregates in Griggs' pathologically-definite IBM. Fibres containing aggregates immunoreactive for p62 and αB-crystallin were more frequent than those containing the current diagnostic pathological features (red bars) ($p<0.05$). Protein aggregates recognised by all antibodies were found in a significantly larger number of fibres than partial invasion ($p<0.02$).

97x99mm (300 x 300 DPI)

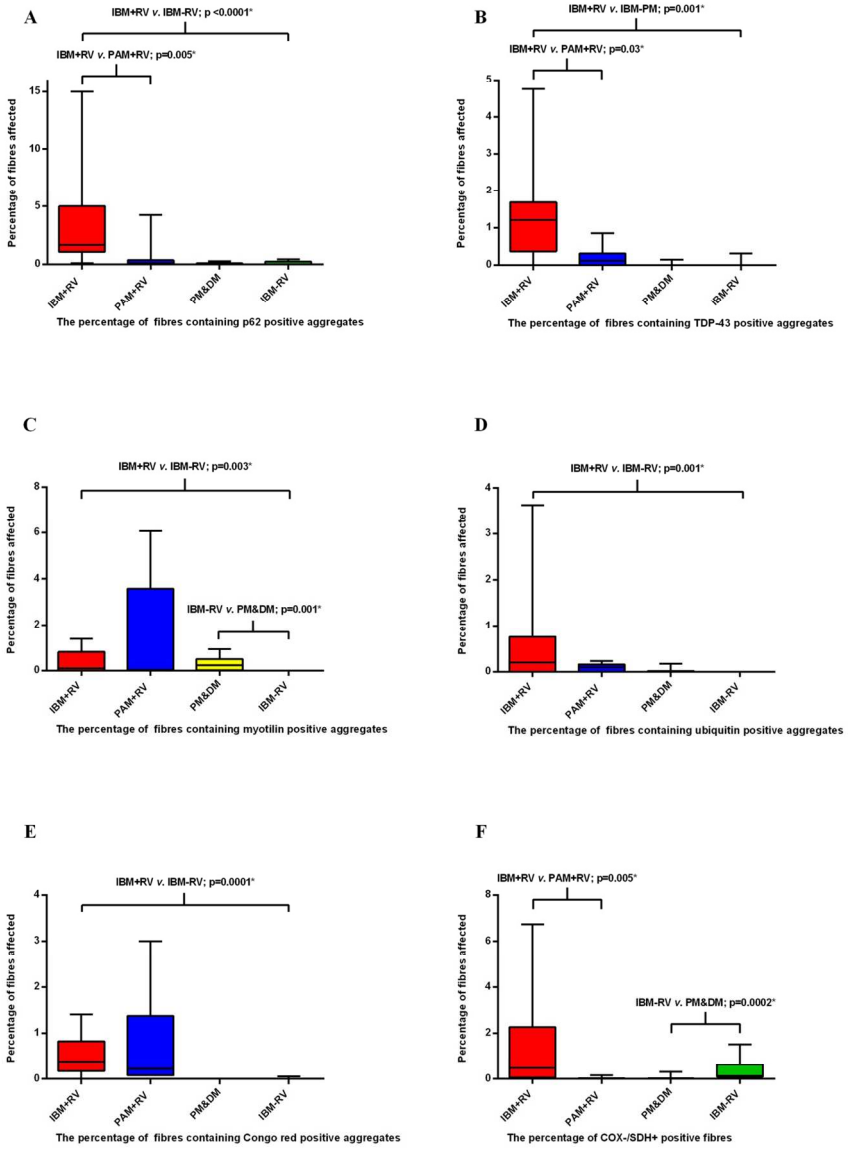
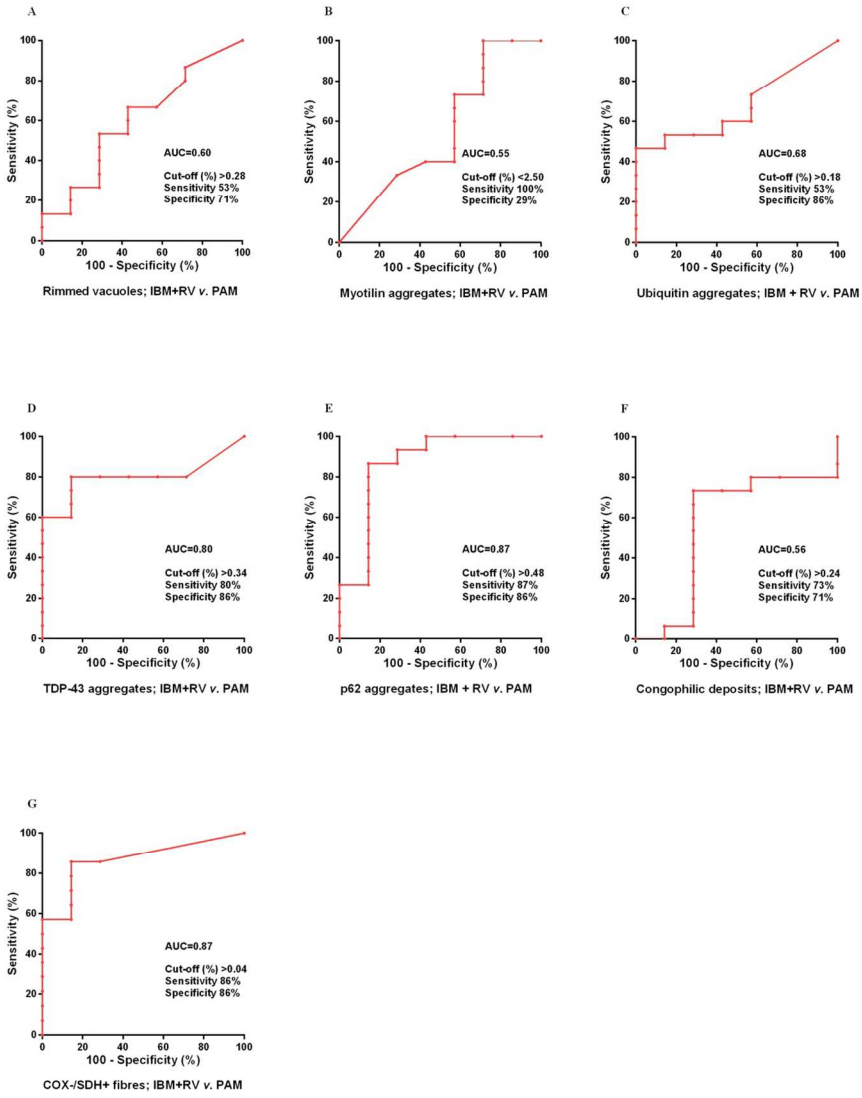


Figure 3 Percentage of fibres containing protein aggregates and COX-/SDH+ fibres in each group
Box and whisker plots illustrating the percentage of fibres in each diagnostic category containing p62 (A), TDP-43 (B), myotilin (C) and ubiquitin (D) immunoreactive aggregates, congophilic deposits (E) and COX-/SDH+ fibres (F). All protein aggregates were present in a greater percentage of fibres in IBM+RV than in IBM–RV. There was no difference in the percentage of COX-/SDH+ muscle fibres between these groups. IBM+RV biopsies had a greater percentage of fibres containing p62 (A) and TDP-43 (B) immunoreactive aggregates and COX-/SDH+ fibres (F) than PAM. Pathological findings were similar in IBM-RV and PM&DM, with no differences in the percentage of fibres containing p62 (A), TDP-43 (B) and ubiquitin (D) immunoreactive aggregates or congophilic deposits (E). However, there was a greater percentage of COX-/SDH+ fibres (F) in IBM–RV than PM&DM and a greater percentage of fibres containing myotilin immunoreactive aggregates (C) in PM&DM than IBM-RV. *Statistically significant results.

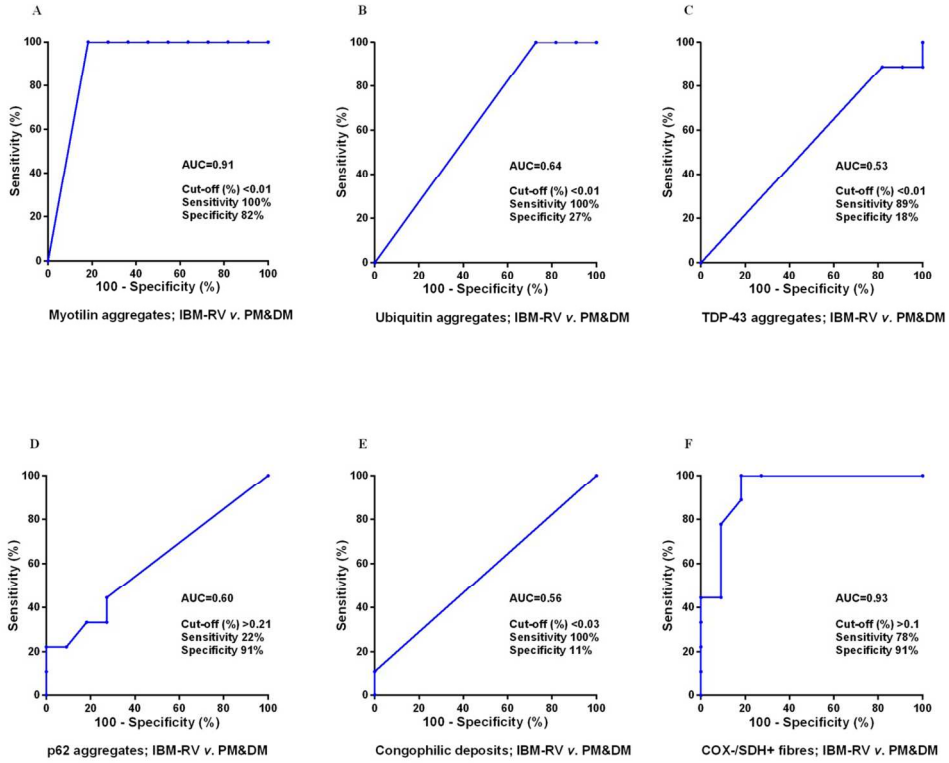
168x226mm (300 x 300 DPI)



Supplementary Figure 2 Sensitivity and specificity of rimmed vacuoles, protein aggregates and mitochondrial changes in IBM+RV compared to PAM

Receiver operating characteristic curves for each test including the area under the curve and optimum cut-off with its associated sensitivity and specificity for rimmed vacuoles (A), myotilin (B), ubiquitin (C), TDP-43 (D), p62 (E) immunoreactive deposits, congophilic deposits (F) and COX-/SDH+ fibres (G). COX/SDH HC staining was the most discriminative test for differentiating IBM+RV and PAM (G). However, there was little difference between COX/SDH HC staining, TDP-43 and p62 IHC staining and none were sufficiently discriminative to be considered diagnostic. AUC = Area under the curve.

165x211mm (300 x 300 DPI)



Supplementary Figure 3 Sensitivity and specificity of protein aggregates and mitochondrial changes in IBM-RV compared to PM&DM

Receiver operating characteristic curves for each test showing the area under the curve and optimum cut-off with its sensitivity and specificity for myotilin (A), ubiquitin (B), TDP-43 (C), p62 (D) immunoreactive deposits, congophilic deposits (E) and COX-/SDH+ fibres (F). COX/SDH histochemical staining (F) and myotilin (G) IHC were the most discriminative tests for differentiating IBM-RV and PM&DM. AUC = Area under the curve.

170x139mm (300 x 300 DPI)

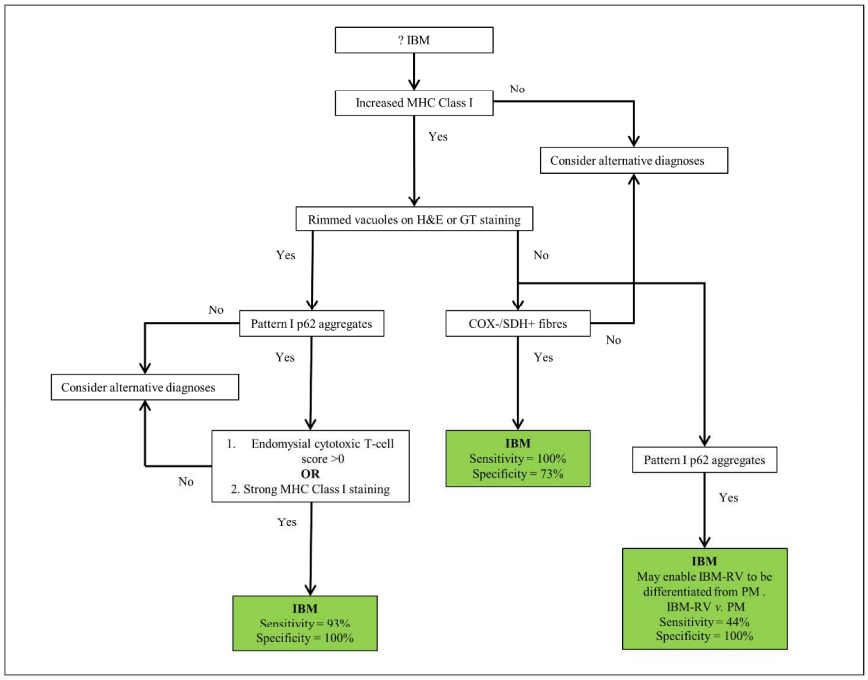


Figure 4 Proposed diagnostic algorithm for IBM based on pathological findings
Flow diagram showing a proposed pathway for diagnosing IBM based on the pathological findings. Increased MHC Class I staining was observed in all cases of IBM and pattern I p62 aggregates in all cases of IBM+RV making them good initial screening tests. Their absence rules-out a diagnosis of IBM and IBM+RV respectively. The presence of an endomysial cytotoxic T-cell score >0 or strong MHC Class I staining in a biopsy with rimmed vacuoles and p62 aggregates secures a diagnosis of IBM+RV. Differentiating IBM-RV and PM&DM pathologically is more challenging. The presence of COX-/SDH+ fibres is not specific to IBM-RV. However, an absence of COX-/SDH+ fibres effectively rules-out a diagnosis of IBM-RV. Pattern I p62 aggregates may enable IBM to be differentiated from PM when present. However, they may lack sensitivity for IBM-RV, therefore their absence does not rule out the diagnosis.

254x190mm (300 x 300 DPI)

STARD checklist for reporting of studies of diagnostic accuracy

(version January 2003)

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	Pg 1,2
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	Pg 2-4
METHODS			
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	Pg 4 and Supplementary Tables 1 and 2
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	Both. Pg 4 and Supplementary Table 2
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	Patients identified from clinics and systematic search of pathological databases
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	Retrospective study
Test methods	7	The reference standard and its rationale.	Clinical features and follow-up
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	Pg 4-6 and Supplementary Table 3
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	Pg 6
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	Two qualified medical doctors. Neuropathologist and Neurologist with an interest and significant experience in muscle pathology.
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	All analyses were blinded and performed in a random order. No clinical information was available at the time of analyses.
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	Pg 6 includes tests for determining diagnostic accuracy including 2x2 tables and ROC curves.
	13	Methods for calculating test reproducibility, if done.	Bland-Altman plots and Cohen's Kappa statistic used.
RESULTS			
Participants	14	When study was performed, including beginning and end dates of recruitment.	2011-2013
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	Included in Supplementary Table 1

	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	Not applicable. Retrospective study.
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	Study performed using tissue taken at the time of the reference standard
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	Diagnosese of control cases included Supplementary Table 2
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	Tables 1 and 2
	20	Any adverse events from performing the index tests or the reference standard.	Not applicable.
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	Included in Tables 1 and 2 and Supplementary Figures 2 and 3
	22	How indeterminate results, missing data and outliers of the index tests were handled.	Only one missing result and this is documented in Table 2. The denominator for calculating the proportion was altered to account for missing case in calculations
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	Included in statistical analysis Pg 6
	24	Estimates of test reproducibility, if done.	Included in statistical analysis Pg 6
DISCUSSION	25	Discuss the clinical applicability of the study findings.	Discussed in discussion Pg 12-15

A retrospective cohort study identifying the principal pathological features useful in the diagnosis of inclusion body myositis

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-004552.R1
Article Type:	Research
Date Submitted by the Author:	18-Feb-2014
Complete List of Authors:	Brady, Stefen; MRC Centre for Neuromuscular Diseases, Division of Neuropathology Squier, Waney Sewry, Caroline Hanna, Mike Hilton-Jones, David Holton, Janice; UCL Institute of Neurology, Department of Molecular Neuroscience and MRC Centre for Neuromuscular Diseases
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Diagnostics
Keywords:	NEUROLOGY, Adult neurology < NEUROLOGY, Neuromuscular disease < NEUROLOGY, Neuropathology < PATHOLOGY

A retrospective cohort study identifying the principal pathological features useful in the diagnosis of inclusion body myositis

Corresponding author:

Dr Janice L Holton

Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

janice.holton@ucl.ac.uk; tel: 00 44 (0)20 3448 4239; fax: 00 44 (0)20 3448 4486.

Authors:

Stefen Brady[1], Waney Squier[2], Caroline Sewry[3,4], Michael Hanna[1], David Hilton-Jones[5], Janice L Holton[6]

[1]MRC Centre for Neuromuscular Diseases, UCL Institute of Neurology and National Hospital for Neurology, Neurosurgery, Queen Square, London, UK.

[2]Department of Neuropathology, University of Oxford, John Radcliffe Hospital, Oxford, UK.

[3]Dubowitz Neuromuscular Centre, Institute of Child Health and Great Ormond Street Hospital for Children, London, UK.

[4]Wolfson Centre of Inherited Neuromuscular Diseases, RJAH Orthopaedic Hospital, Oswestry, UK.

[5]Nuffield Department of Clinical Neurosciences (Clinical Neurology), University of Oxford, John Radcliffe Hospital, Oxford, UK.

[6]Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

Keywords: Neurology, adult neurology, neuromuscular disease, neuropathology

Running title: Pathological criteria for inclusion body myositis

Word count: 2990

ABSTRACT

Objectives

The current pathological diagnostic criteria for sporadic inclusion body myositis (IBM) lack sensitivity. Using immunohistochemical techniques abnormal protein aggregates have been identified in IBM, including some associated with neurodegenerative disorders. Our objective was to investigate the diagnostic utility of a number of markers of protein aggregates together with mitochondrial and inflammatory changes in IBM.

Design

Retrospective cohort study. The sensitivity of pathological features was evaluated in cases of Griggs' definite IBM. The diagnostic potential of the most reliable features was then assessed in clinically-typical IBM with rimmed vacuoles (*n*=15) and clinically-typical IBM without rimmed vacuoles (*n*=9) and IBM mimics - vacuolar myopathies (*n*=7) and steroid-responsive inflammatory myopathies (*n*=11).

Setting

Specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London.

Results

Individual pathological features, in isolation, lacked sensitivity and specificity. However, the morphology and distribution of p62 aggregates in IBM were characteristic and in a myopathy with rimmed vacuoles, the combination of characteristic p62 aggregates and increased sarcolemmal and internal MHC Class I expression or endomysial T-cells were diagnostic for IBM with a sensitivity of 93% and specificity of 100%. In an inflammatory myopathy lacking rimmed vacuoles, the presence of mitochondrial changes was 100% sensitive and 73% specific for IBM; characteristic p62 aggregates were specific (91%), but lacked sensitivity (44%).

Conclusions

We propose an easily applied diagnostic algorithm for the pathological diagnosis of IBM. Additionally our findings support the hypothesis that many of the pathological features considered

typical of IBM develop later in the disease, explaining their poor sensitivity at disease presentation and emphasising the need for revised pathological criteria to supplement the clinical criteria in the diagnosis of IBM.

STRENGTHS AND LIMITATIONS

The present study is a multicentre retrospective evaluation of the diagnostic utility of pathological findings for differentiating IBM from myopathies important in the differential diagnosis – myopathies containing rimmed vacuoles and steroid-responsive inflammatory myopathies.

The main strength of our study was the systematic detailed analysis of well-defined cases. This enabled us to determine the sensitivity and specificity of individual pathological features and produce an easily applied pathological diagnostic algorithm for IBM for use in clinical practice.

Study limitations include the small number of cases and the retrospective design. Further prospective studies are now required in larger cohorts of patients.

INTRODUCTION

Sporadic inclusion body myositis (IBM) is the commonest acquired myopathy in those aged over 50 years.[1] Although classified as an idiopathic inflammatory myopathy, muscle biopsy reveals both degenerative and inflammatory features. The widely used Griggs diagnostic criteria require the presence of several pathological findings,[2] namely rimmed vacuoles, an inflammatory infiltrate with invasion of non-necrotic fibres by mononuclear inflammatory cells (partial invasion), and either amyloid deposits or 15-18 nm tubulofilaments identified by electron microscopy (EM). Although these features in combination are highly specific for IBM, individually they occur in other myopathies, including some important in the differential diagnosis for IBM.[3-7] Moreover, cases of clinically-typical IBM have been reported where the combination of these pathological features is absent causing diagnostic difficulty.[8-11]

Over the last two decades, pathological accumulation of many different proteins has been reported in muscle fibres in IBM.[12] Proteins typically associated with neurodegenerative diseases such as β-amyloid (Aβ), hyperphosphorylated tau and ubiquitin and newer neurodegenerative markers such as p62 and transactivation response DNA binding protein-43 (TDP-43) have been identified, as well as proteins associated with myofibrillar myopathies (MFM), including desmin and αB-crystallin. However, not all observations have been consistently reproduced.[13,14] Mitochondrial changes have also been proposed for inclusion in IBM diagnostic criteria,[15]. Clear guidelines for the incorporation of immunohistochemical findings and mitochondrial changes into diagnostic criteria for IBM have not been established.[16]

Previously, we have shown that the characteristic pattern of weakness associated with IBM is indicative of the diagnosis, even if Griggs pathological features are absent.[11] However, it is not invariably found at presentation. Here we sought to identify which pathological features, other than the Griggs pathological criteria, add further support to the diagnosis of IBM. We systematically investigated which pathological features are present in Griggs pathologically-definite IBM and then

established the diagnostic utility of these features in cases of IBM lacking the Griggs criteria, using myopathies considered in the differential diagnosis of IBM as controls.

MATERIALS AND METHODS

The study received ethical approval from the Departments of Research and Development at Oxford University Hospitals NHS Trust, Oxford and University College London Hospitals NHS Foundation Trust, London.

Cases

All patients were followed by specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London. Biopsies were taken for diagnostic purposes from the deltoid or quadriceps muscles and prior to any treatment.

Methods for demonstrating pathological features in IBM, additional to those defined by the Griggs criteria, were determined in six Griggs pathologically-definite cases of IBM. Cases with no clinical or pathological evidence of neuromuscular disease were used as controls. The diagnostic utility of the pathological features identified was assessed in two groups of clinically-typical IBM; one with rimmed vacuoles on muscle biopsy (IBM+RV; *n*=15), the other without rimmed vacuoles on muscle biopsy (IBM-RV; *n*=9). Disease controls were cases of steroid-responsive inflammatory myopathies [polymyositis and dermatomyositis; (PM&DM); *n*=11] and protein accumulation myopathies with rimmed vacuoles (PAM; *n*=7). Clinical characteristics and inclusion criteria are summarised in Supplementary tables 1 and 2. Tissue from brains donated to the Queen Square Brain Bank for Neurological Disorders was used as positive controls for protein aggregate staining.

Muscle biopsies

Muscles biopsies were snap frozen at the time of surgery in isopentane cooled liquid nitrogen. Until sectioning all samples were stored at -80°C. Serial tissue sections were cut to a thickness of 8 µm, allowed to air dry and stored at -80°C until staining. Prior to staining, tissue sections were allowed to dry at room temperature. Tissue sections were stained with haematoxylin and eosin (H&E), combined cytochrome oxidase (COX) succinate dehydrogenase (SDH) histochemistry and for amyloid using alkalinised Congo red, crystal violet and thioflavin S. Tissue sections for immunohistochemical staining were fixed for 10 minutes, if required, washed for five minutes in running water and incubated in 0.5% hydrogen peroxide to block endogenous peroxidase for 20 minutes. After further washing, tissue sections were incubated in 5% normal goat serum (Vector Laboratories, Burlingame, California) for 30 minutes and then systematically stained for: 1) proteins classically associated with neurodegenerative disease: tau and hyperphosphorylated tau, ubiquitin, Aβ and α-synuclein; 2) proteins more recently reported in neurodegenerative disease: p62, TDP-43, fused in sarcoma protein (FUS) and valosin containing protein (VCP); 3) nuclear membrane proteins: lamin A/C and emerin; 4) proteins associated with MFM: desmin, myotilin and αB-crystallin; and 5) inflammatory cells and major histocompatibility complex class I (MHC Class I): CD3+ T-cells, CD4+ T-cells, CD8+ T-cells, B-cells and macrophages. Primary antibody binding was visualised using Dako REAL™ EnVision™ Detection System which contains horse-radish peroxidase (HRP) labelled goat anti-rabbit/mouse secondary and 3,3'-diaminobenzidine (DAB); following incubation with the relevant primary antibody, tissue sections were washed in phosphate buffered saline (PBS), incubated with HRP labelled goat anti-rabbit/mouse secondary for 30 minutes, washed in PBS and incubated in a 1:50 solution of DAB for three to five minutes. Details of commercial antibodies and conditions used are provided in Supplementary Table 3. IHC for each antibody was performed on all cases simultaneously and including positive and negative controls (Supplementary Figure 1).

Definitions and quantification

The total number of fibres and the number undergoing partial invasion, containing rimmed vacuoles, protein aggregates and COX-negative SDH-positive (COX-/SDH+) fibres were quantified using ImagePro version 6.2 (Media Cybernetics), to ensure that the whole biopsy was systematically analysed. Only transversely-orientated fibres not undergoing necrosis or regeneration were quantified. Tissue sections stained with Congo red were visualised under fluorescent and polarised light. Areas of fluorescence were examined using both rhodamine red (excitation 512-546 nm and emission 600-640 nm) and fluorescein isothiocyanate (excitation 440-480 nm and emission 527-530 nm) filters to exclude auto-fluorescence. Supplementary Table 4 provides definitions of the pathological features assessed. The inflammatory infiltrate and MHC Class I staining were analysed using a modified version of the semi-quantitative juvenile dermatomyositis score-tool (Supplementary Figure 2).[17] Assessments were performed blind to clinical details and diagnosis by a single individual (SB). Ten per cent of slides were re-counted to assess intra-observer reliability and 336 slides were assessed independently by two observers (SB and JLH) to determine inter-observer reliability.

Statistical analysis

Statistical analyses were performed using GraphPad PRISM version 5. Continuous and categorical variables were compared using Mann Whitney *U*-test and chi-squared or Fisher's exact test respectively. Spearman's rank order correlation was used to determine the strength and direction of associations between pathological findings. Linear regression was used to determine relationships between clinical features and pathological findings. Test characteristics were calculated using receiver operating characteristic (ROC) curves and 2x2 contingency tables. A test was considered diagnostic when sensitivity >75% and specificity >95% or sensitivity >95% and specificity >75%. Intra-observer and inter-observer agreement was calculated using Bland-Altman plots and Cohen's kappa statistic (κ). Repeat counts were within 95% confidence intervals using Bland-Altman plots and κ was ≥0.7 indicating good intra-observer and good or excellent inter-observer reliability. Statistical significance was set at $p<0.05$.

RESULTS

Pathological findings in Griggs' pathologically-definite IBM

p62, TDP-43, ubiquitin, myotilin and αB-crystallin immunoreactive aggregates were present in all six IBM cases but not in normal controls (Figures 1A-E). p62 and αB-crystallin immunoreactive aggregates were present in a greater percentage of fibres than the pathological features required in the Griggs criteria ($p<0.05$) (Figure 2). Despite their abundance, αB-crystallin immunoreactive aggregates were difficult to quantify due to a significant variability in their morphology. No immunoreactive deposits were observed in IBM cases or normal controls with antibodies to tau and phosphorylated tau, Aβ, α-synuclein, desmin, emerin, lamin A/C, FUS or VCP. Alkalinised Congo red staining was more sensitive than crystal violet and thioflavin S staining for observing amyloid aggregates (Figure 1F). Tissue sections containing congophilic deposits identified under fluorescence light showed no apple-green birefringence under polarised light. Mitochondrial changes and increased sarcolemmal and sarcoplasmic MHC Class I staining were observed in all six IBM cases, but not in normal controls. The inflammatory infiltrate was predominantly composed of endomysial CD8+ T-cells and macrophages, with relatively few B-cells.

Quantitative analysis of pathological features in IBM and disease controls

Having shown that p62, TDP-43, ubiquitin and myotilin aggregates, congophilic deposits, MHC Class I and inflammatory cells were prevalent in Griggs' pathologically-definite IBM, the presence of these abnormalities, together with mitochondrial changes were assessed in IBM+RV, IBM-RV and disease controls.

The percentage of fibres containing p62, TDP-43, myotilin and ubiquitin aggregates and congophilic deposits were greater in IBM+RV than in IBM-RV; there was no difference in the number of COX-/SDH+ fibres (Figure 3A-F). Protein aggregates were observed in morphologically-normal fibres and in fibres exhibiting Griggs' pathological features. p62 and TDP-43 positive aggregates were present in

a greater percentage of fibres in IBM+RV compared to PAM; however, there were no differences in the percentage of fibres containing myotilin and ubiquitin aggregates or congophilic deposits. The percentage of fibres containing p62, TDP-43 and ubiquitin aggregates or congophilic deposits were similar in IBM-RV and PM&DM; however, myotilin aggregates were present in a greater percentage of fibres in PM&DM and COX-/SDH+ fibres were more abundant in IBM-RV. Analysis of the total inflammatory infiltrate (the sum of the semi-quantitative scores for T-cells, B-cells and macrophages) in the endomysium, perimysium and perivascular areas revealed that there were greater numbers of inflammatory cells in the endomysium and perimysium in IBM+RV than in PAM (p<0.03). The same analysis comparing the sum of the inflammatory cells in IBM-RV and PM&DM revealed that the distribution and intensity of the inflammatory infiltrate was similar.

Diagnostic utility of pathological features in IBM and disease controls

To mimic the diagnostic difficulty encountered in clinical practice, the ability of each test to differentiate between myopathies containing rimmed vacuoles (IBM+RV and PAM) and between inflammatory myopathies (IBM–RV and PM&DM) was assessed.

Diagnostic utility determined using receiver-operating characteristic curves

Individually, the presence of p62 immunoreactive inclusions and COX-/SDH+ fibres had the highest sensitivity and specificity for differentiating IBM+RV from PAM, (Supplementary Figure 3) (Table 1). Differentiating between IBM–RV and PM&DM, myotilin positive inclusions or COX-/SDH+ fibres had the highest sensitivity and specificity for IBM-RV (Supplementary Figure 4) (Table 1). Only the presence of myotilin positive inclusions satisfied criteria to be considered suitable as a diagnostic test (<0.01% of fibres containing myotilin aggregates had a sensitivity of 100% and specificity of 82% for IBM-RV).

Table 1 Test characteristics

Table shows the area under the curve and optimum cut-off for each test with the accompanying sensitivity and specificity. AUC = Area under the curve.

Test feature	IBM+RV *v.* PAM				IBM-RV *v.* PM&DM			
	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity
Rimmed vacuoles	0.60	>0.28	0.53	0.71	-	-	-	-
p62 aggregates	0.87	>0.48	0.87	0.86	0.60	>0.21	0.22	0.91
TDP-43 aggregates	0.80	>0.34	0.80	0.86	0.53	<0.01	0.89	0.18
Ubiquitin aggregates	0.68	>0.18	0.53	0.85	0.64	<0.01	1.00	0.27
Myotilin aggregates	0.55	<0.25	1.00	0.29	0.91	<0.01	1.00	0.82
Congophilic deposits	0.56	>0.24	0.73	0.71	0.56	<0.03	0.11	0.82
COX-/SDH+ fibres	0.87	>0.04	0.86	0.86	0.93	>0.1	0.78	0.91

Diagnostic utility determined by comparing proportion of affected cases in each diagnostic group

In the aforementioned experiments, the number of fibres within each muscle biopsy was quantified. However, this is impractical for routine clinical use. Thus, the proportions of affected cases in each group were compared (Table 2). This revealed that neither staining for protein aggregates nor congophilic deposits could differentiate between IBM+RV and PAM. The pathological findings in IBM-RV and PM&DM were also similar, except that the absence of myotilin immunoreactive aggregates was sensitive and specific for IBM-RV. COX-/SDH+ fibres were also suggestive of IBM-RV; one or more COX-/SDH+ fibres had a sensitivity of 100% and specificity 73% for IBM-RV.

Increased MHC Class I expression lacked specificity. However, strong (diffuse sarcolemmal and sarcoplasmic) MHC Class I up-regulation was diagnostic for IBM+RV, differentiating it from PAM, as were the presence of either endomysial CD3+ T-cell or CD4+ T-cell scores >1 or an endomysial CD8+ T-cell score >0. Partial invasion was specific for IBM+RV, but lacked sensitivity. Although the sum of the inflammatory infiltrate was similar in IBM-RV and PM&DM, analysis of the inflammatory cell sub-types revealed greater numbers of perimysial CD3+ T-cells, CD8+ T-cells and endomysial B-cells [were observed] in PM&DM than in IBM–RV ($p \leq 0.02$), however, this was not diagnostically useful. There was no difference in the proportion of cases with fibres undergoing partial invasion between IBM–RV and PM&DM.

Table 2 Comparison of the proportion of positive cases in each group

Pathological features	IBM+RV	PAM	IBM+RV *v.* PAM		IBM–RV	PM&DM	IBM-RV *v.* PM&DM		IBM+RV *v.* IBM-RV
	n (%)	n (%)	Sensitivity	Specificity	n (%)	n (%)	Sensitivity	Specificity	*p* value
Number of cases	15 (100)	7 (100)			9 (100)	11 (100)			
Aggregated proteins, n (%)									
p62	15 (100)	6 (86)	1.00	0.14	4 (44)	3 (27)‡	0.40	0.72	0.003*
TDP-43	13 (87)	5 (71)	0.87	0.29	1 (11)	2 (18)‡	0.11	0.82	0.001*
Ubiquitin	11 (73)	4 (57)	0.73	0.43	0 (0)	3 (27)‡	0.00	0.73	0.001*
Myotilin	10 (67)	5 (71)	0.67	0.29	0 (0)	9 (82)	0.00	0.18	0.002*
Congophilic deposits	13 (87)	7 (100)	0.87	0.00	1 (11)	0 (0)	0.11	1.00	0.001*
COX-/SDH+ fibres†, n (%)									
Any	12 (86)	2 (29)	0.80	0.71	9 (100)	3 (27)	1.00	0.73	0.5
Inflammatory features, n (%)									
MHC Class I up-regulation	15 (100)	3 (43)	1.00	0.57	9 (100)	11 (100)	1.00	0.00	1.00
Strong MHC Class I up-regulation	14 (93)	0 (0)	0.93	1.00	9 (100)	10 (91)	1.00	0.09	0.53
Partial invasion	10 (67)	0 (0)	0.67	1.00	3 (33)	2 (18)	0.33	0.82	0.11
Endomysial CD3+ T-cell score >1	13 (87)	0 (0)	0.87	1.00	4 (44)	7 (64)	0.44	0.36	0.02*
Endomysial CD4+ T-cell score >1	12 (80)	0 (0)	0.80	1.00	2 (22)	5 (45)	0.22	0.46	0.01*
Endomysial CD8+ T-cell score >0	14 (93)	0 (0)	0.93	1.00	4 (44)	5 (45)	0.44	0.54	0.02*
Endomysial CD68+ macrophage score >1	12 (80)	0 (0)	0.80	1.00	4 (44)	8 (73)	0.44	0.17	0.07

†In IBM with rimmed vacuoles *n*=14. ‡Pathological features present in DM, but not PM cases. *Statistically significant results.

Because IBM-RV is more pathologically akin to PM than DM, analyses were repeated comparing IBM-RV and PM cases (*n*=6). No p62, TDP-43 or ubiquitin immunoreactive aggregates were observed in PM cases and the diagnostic utility of tests for differentiating between IBM-RV and PM yielded similar results to prior analyses between IBM-RV and PM&DM.

Diagnostic utility of categorising the pattern of p62 staining

The pattern of p62 staining could be categorised into four distinct groups (Figure 1G-J). Aggregates observed in IBM were present in vacuolated and non-vacuolated fibres and were strongly stained, discreet and clearly delineated, round or angular and typically located subsarcolemmal, perinuclear and peri-vacuolar (pattern I). This pattern was observed in every IBM case with p62 aggregates, one (9%) case of DM and three (43%) cases of PAM (hereditary IBM, dystrophinopathy and genetically undefined MFM). Defining the pattern of immunoreactivity increased the discriminative value of p62 IHC for differentiating IBM+RV from PAM; pattern I p62 aggregates compared to any p62 aggregates increased the specificity from 14% to 57%, with no loss of sensitivity. Differentiating IBM-RV and PM&DM, pattern I p62 aggregates were highly specific (91%), but lacked sensitivity (44%). Patterns II, III and IV were not observed in any IBM cases. Patterns II and III appeared to be specific for PAM (*n*=2; 26%), both were cases of myotilinopathy (*n*=2; 67%), and DM (*n*=2; 40%) respectively. Pattern IV occurred in a genetically undefined case of MFM. No differences were observed in the morphology of TDP-43, myotilin or ubiquitin aggregates between biopsies.

Clinicopathological correlation

In IBM+RV, IBM-RV and pathologically-definite IBM, there were no correlations in individual biopsies between pathological features. No relationships were identified between the pathological findings and age at symptom onset, age at biopsy, disease duration or serum creatine kinase. The same results were obtained when the IBM groups were analysed separately and as one.

Proposed diagnostic algorithm

Based on our pathological findings, we propose a diagnostic algorithm for differentiating IBM from its disease mimics (Figure 4).

The algorithm was tested in a further 23 cases that fulfilled the criteria for IBM+RV (*n*=12) and IBM-RV (*n*=11). The algorithm correctly diagnosed 20 (87%) cases: 12 (100%) cases of IBM+RV and eight (73%) cases of IBM-RV. In IBM-RV, COX-/SDH+ fibres were present in 8 (73%) cases, pattern I p62 aggregates in 8 (73%) cases and both in 6 (55%) cases.

DISCUSSION

While Griggs' pathological criteria have been accepted as diagnostic of IBM, many patients who, observed over time undoubtedly have IBM, lack one or more of the Griggs pathological features at presentation, even on repeat biopsy.[8,11] Despite IBM being associated with a characteristic pattern of finger flexor and knee extensor weakness, not all patients have this pattern at disease onset, and muscle biopsy remains an important tool in differentiating IBM from its mimics. We sought to determine which additional pathological features support a diagnosis of IBM, demonstrating that characteristic p62 immunoreactive aggregates, strong MHC Class I upregulation, endomysial CD3+ T-cell score >1, CD8+ T-cell score >0 and COX-/SDH+ fibres are features with sufficient sensitivity and specificity to differentiate IBM from pathologically similar myopathies and we propose an easily applied pathological algorithm for the diagnosis of IBM (Figure 4).

In agreement with previous studies, we observed p62,[18] TDP-43,[19] ubiquitin [13] and αB-crystallin [20] immunoreactive aggregates and a predominantly endomysial inflammatory infiltrate [3] in Griggs pathologically-definite IBM. Diagnostic pathological studies of IBM have concentrated on differentiating IBM from other inflammatory myopathies and two recent quantitative studies have found that TDP-43 and markers of autophagy such as p62 and LC3 may be of diagnostic use.[21,22] However, in these studies only a fraction of each biopsy was analysed i.e. 200 fibres. We have found

this limited quantification does not correlate with the percentage of affected fibres in a biopsy nor does it reflect the way in which a muscle biopsy is assessed. Additionally, studies have lacked vacuolar myopathy control cases as it is believed that the inflammatory changes present in IBM enable it to be easily differentiated from other vacuolar myopathies.[22] However, inflammatory changes are frequently observed in muscular dystrophies and the degree of inflammatory change necessary to confidently diagnose IBM is currently unknown.

To mimic the typical diagnostic conundrums encountered in clinical practice, we evaluated the ability of the pathological findings to differentiate IBM+RV from other vacuolar myopathies and IBM-RV from steroid-responsive inflammatory myopathies. We found that quantitative analysis of protein aggregates, congophilic deposits and COX-/SDH+ fibres was of limited diagnostic use. Analysing the biopsies dichotomously and using a semi-quantitative score-tool revealed that increased MHC Class I labelling was sensitive for IBM making it a good initial screening test, its absence excluding the diagnosis. In agreement with an earlier study, we found p62 aggregates identified the largest number of affected fibres in IBM.[23] Additionally, as a novel finding, the morphology and distribution of p62 aggregates was characteristic in IBM. This characteristic pattern of p62 immunoreactive aggregates was highly sensitive for IBM+RV (100%); their absence from a biopsy containing rimmed vacuoles effectively ruling-out a diagnosis of IBM. We confirmed that the most diagnostically useful pathological findings in IBM+RV were evidence of an immune mediated process; strong MHC Class I staining, endomysial CD3+ T-cell score >1 or an endomysial CD8+ T-cell score >0 were diagnostic. Having identified either of these features in a biopsy containing rimmed vacuoles no extra diagnostic certainty was gained from observing partial invasion, COX-/SDH+ fibres or congophilic deposits.

The most discriminative pathological tests for differentiating between IBM–RV and PM&DM were COX/SDH staining and myotilin IHC. Consistent with a recent study,[9] we found the absence of mitochondrial changes casts doubt on a diagnosis of IBM. There was no difference in the median age between IBM-RV and PM&DM cases to account for the difference observed in COX-/SDH+ fibres.

The presence of myotilin and ubiquitin immunoreactive aggregates appeared to rule out a diagnosis of IBM-RV. However, we believe the presence of these features in IBM+RV indicates that they are unlikely to be diagnostically reliable features for differentiating between IBM-RV and steroid-responsive inflammatory myopathies. Although no pathological feature was able to differentiate IBM-RV from steroid responsive inflammatory myopathies with certainty the presence of characteristic p62 aggregates and the absence of COX-/SDH+ fibres may help in supporting and opposing a diagnosis of IBM-RV respectively. Pattern I p62 immunoreactive aggregates were only present in 44% of the initial IBM-RV cases tested, but they were not observed in PM cases and were very rare in DM. Although pattern I p62 aggregates appear to lack sensitivity their specificity was 91% making their presence highly suggestive of a diagnosis of IBM-RV. However, we identified pattern I p62 in eight out of 11 (73%) further cases of IBM-RV that were assessed indicating a greater sensitivity and that p62 IHC warrants further investigation and validation in a larger, independent series. The diagnostic utility of the other patterns of p62 staining is uncertain. Although pattern II appeared to have some specificity for myotilinopathy the small number of cases makes it drawing any conclusion problematic. In addition to p62 other autophagic proteins have been found in IBM and suggested as diagnostic markers. [22] Autophagy is a cellular mechanism for degrading and recycling cellular proteins and organelles and therefore, altered autophagy could lead to the accumulation of abnormal mitochondria and misfolded aggregation-prone proteins and may also result in altered antigen presentation leading to the widespread increase of MHC Class I and suggests that altered autophagy may play an important role in the pathogenesis of IBM.

Almost all pathological features - protein aggregates, congophilic deposits and inflammation - were more abundant in IBM+RV than IBM-RV. Despite using slightly different inclusion criteria, similar differences have been reported between pathologically-typical and pathologically-atypical IBM.[21] However, we found no differences in the number of COX-/SDH+ fibres, the degree of MHC Class I upregulation, the morphology and distribution of p62 immunoreactive aggregates or the pattern of the inflammation between IBM+RV and IBM-RV, supporting our clinical observations that these are the

same disease. We believe that the pathological differences between IBM+RV and IBM-RV are, in part, due to differences in disease duration. Two studies have shown that rimmed vacuoles are more common in patients who are older at the time of muscle biopsy,[24,11] suggesting that they are associated with chronologically more advanced disease. Therefore, the pathological findings which are more abundant in IBM+RV and thought to be typical of IBM may instead be indicative of chronologically more advanced disease explaining their limited sensitivity at disease presentation. However, possibly due to the number of cases analysed, we were unable to confirm a relationship between pathological features and clinical findings. It could be argued that biopsies from different muscles may have affected the pathological findings observed and differences between IBM groups. However, in a recent review of 59 muscle biopsies from IBM cases in our clinical archive with quadriceps (n=31) and deltoid (n=28) biopsies we found no significant difference in the frequency of pathological findings.

A robust clinicopathological definition of IBM is of paramount importance for diagnosis and for selection and entry of patients into clinical trials. We have shown that certain pathological findings are more abundant than those included in the current pathologically-focussed diagnostic criteria. Moreover, p62 immunoreactive deposits, increased MHC Class I expression, endomysial CD3+ T-cells and CD8+ T-cells and COX-/SDH+ fibres have sufficient sensitivity and specificity to aid in the histological differentiation of IBM from disease mimics, supporting their inclusion in future diagnostic criteria for IBM alongside clinical criteria. Both CD3+ T-cells and CD8+ T-cells are included in the diagnostic algorithm as there was little difference in their sensitivity and specificity for differentiating IBM+RV from PAM. However, IHC staining for CD3+ T-cells is likely to be more widely available and avoids the costs of extra staining to subtype the inflammatory infiltrate enabling diagnostic algorithm to be used by a greater number diagnostic laboratories. Using our diagnostic algorithm, we found there would be little additional diagnostic security in identifying partial invasion, performing EM or staining for amyloid deposits. Finally, mitochondrial changes and MHC Class I up-regulation were the most consistent findings in our IBM cases suggesting that they are central to the

pathogenesis and that further investigation and therapeutic intervention should be directed towards these features.

ACKNOWLEDGEMENTS

SB is funded by the Myositis Support Group. JLH is supported by the Reta Lila Weston Institute for Neurological Studies and the Myositis Support Group. This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme.

CONTRIBUTORSHIP STATEMENT

Dr Stefen Brady - Acquisition of data, analysis and interpretation of data and drafting of manuscript.

Dr Waney Squier - Critical revision of manuscript for important intellectual content.

Prof. Caroline Sewry - Study concept and design and critical revision of manuscript for important intellectual content.

Prof. Mike Hanna - Critical revision of manuscript for important intellectual content.

Dr David Hilton-Jones - Critical revision of manuscript for important intellectual content.

Dr Janice Holton - Study concept and design, critical revision of manuscript for important intellectual content and study supervision.

COMPETING INTERESTS

None

DATA SHARING

All additional data can be found in supplementary tables and figures.

REFERENCES

1. Needham M, James I, Corbett A, Day T, Christiansen F, Phillips B, et al. Sporadic inclusion body myositis: phenotypic variability and influence of HLA-DR3 in a cohort of 57 Australian cases. J Neurol Neurosurg Psychiatr. 2008;79:1056–60.
2. Griggs RC, Askanas V, DiMauro S, Engel A, Karpati G, Mendell JR, et al. Inclusion body myositis and myopathies. Ann Neurol. 1995;38:705–13.
3. Arahata K, Engel AG. Monoclonal antibody analysis of mononuclear cells in myopathies. I: Quantitation of subsets according to diagnosis and sites of accumulation and demonstration and counts of muscle fibers invaded by T cells. Ann Neurol. 1984;16:193–208.
4. Mhiri C, Gherardi R. Inclusion body myositis in French patients. A clinicopathological evaluation. Neuropathol Appl Neurobiol. 1990;16:333–44.
5. Villanova M, Kawai M, Lübke U, Oh SJ, Perry G, Six J, et al. Rimmed vacuoles of inclusion body myositis and oculopharyngeal muscular dystrophy contain amyloid precursor protein and lysosomal markers. Brain Res. 1993;603:343–7.
6. Van der Meulen MF, Hoogendijk JE, Moons KG, Veldman H, Badrising UA, Wokke JH. Rimmed vacuoles and the added value of SMI-31 staining in diagnosing sporadic inclusion body myositis. Neuromuscul Disord. 2001;11:447–51.
7. Ferrer I, Olivé M. Molecular pathology of myofibrillar myopathies. Expert Rev Mol Med. 2008;10:e25.
8. Amato AA, Gronseth GS, Jackson CE, Wolfe GI, Katz JS, Bryan WW, et al. Inclusion body myositis: clinical and pathological boundaries. Ann Neurol. 1996;40:581–6.
9. Chahin N, Engel AG. Correlation of muscle biopsy, clinical course, and outcome in PM and sporadic IBM. Neurology. 2008;70:418–24.
10. Benveniste O, Guiguet M, Freebody J, Dubourg O, Squier W, Maisonobe T, et al. Long-term observational study of sporadic inclusion body myositis. Brain. 2011;134:3176–84.

11. Brady S, Squier W, Hilton-Jones D. Clinical assessment determines the diagnosis of inclusion body myositis independently of pathological features. J Neurol Neurosurg Psychiatr. 2013; 16.

12. Greenberg SA. Theories of the Pathogenesis of Inclusion Body Myositis. Current Rheumatology Reports. 2010;12:221–8.

13. Sherriff FE, Joachim CL, Squier MV, Esiri MM. Ubiquitinated inclusions in inclusion-body myositis patients are immunoreactive for cathepsin D but not β-amyloid. Neuroscience Letters. 1995;194:37–40.

14. Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. BMJ. 2009;339:b2680.

15. Needham M, and Mastaglia FL. (2007). Inclusion body myositis: current pathogenetic concepts and diagnostic and therapeutic approaches. Lancet Neurol. 6, 620–631.

16. Benveniste O, Hilton-Jones D. International Workshop on Inclusion Body Myositis held at the Institute of Myology, Paris, on 29 May 2009. Neuromuscular Disorders. 2010;20:414–21.

17. Wedderburn LR, Varsani H, Li CKC, Newton KR, Amato AA, Banwell B, et al. International consensus on a proposed score system for muscle biopsy evaluation in patients with juvenile dermatomyositis: a tool for potential use in clinical trials. Arthritis Rheum. 2007;57:1192–201.

18. Nogalska A, Terracciano C, D'Agostino C, King Engel W, Askanas V. p62/SQSTM1 is overexpressed and prominently accumulated in inclusions of sporadic inclusion-body myositis muscle fibers, and can help differentiating it from polymyositis and dermatomyositis. Acta Neuropathol. 2009;118:407–413.

19. Weihl CC, Temiz P, Miller SE, Watts G, Smith C, Forman M, Hanson PI, Kimonis V, Pestronk A. TDP-43 accumulation in inclusion body myopathy muscle suggests a common pathogenic mechanism with frontotemporal dementia. J. Neurol. Neurosurg. Psychiatr. 2008;79:1186–1189.

20. Banwell BL, Engel AG. Alpha B-crystallin immunolocalization yields new insights into inclusion body myositis. Neurology. 2000;54:1033–1041.

21. Dubourg O, Wanschitz J, Maisonobe T, Béhin A, Allenbach Y, Herson S, and Benveniste O. Diagnostic value of markers of muscle degeneration in sporadic inclusion body myositis. Acta Myol. 2011. 30, 103–108.

22. Hiniker A, Daniels BH, Lee HS, Margeta M. Comparative utility of LC3, p62 and TDP-43 immunohistochemistry in differentiation of inclusion body myositis from polymyositis and related inflammatory myopathies. Acta Neuropathologica Communications. 2013. 1:29.

23. D'Agostino C, Nogalska A, Engel WK, Askanas V. In sporadic inclusion-body myositis muscle fibres TDP-43-positive inclusions are less frequent and robust than p62-inclusions, and are not associated with paired helical filaments. Neuropathol Appl Neurobiol. 2010; Available from: http://www.ncbi.nlm.nih.gov/pubmed/20626631.

24. Momma K, Noguchi S, Malicdan MCV, Hayashi YK, Minami N, Kamakura K, et al. Rimmed vacuoles in Becker muscular dystrophy have similar features with inclusion myopathies. PLoS ONE. 2012;7:e52002.

Supplementary Figure 1 Control staining in brain and muscle tissue

Positive and negative (no primary) brain control sections and normal muscle stained using immunohistochemistry for: p62 (A-C), TDP-43 (D-F), α B-crystallin (G-I), ubiquitin (J-K) and myotilin (M,N) and alkalinised congo red (O).

(A-C) Negative (A) and positive (B) control sections of AD brain and normal muscle (C) stained for p62. Positive control shows p62 positive neurofibrillary tangles and dystrophic neurites (B). No p62 immunoreactivity is observed in normal muscle (C).

(D-F) Negative (D) and positive (E) control sections of FTLD-TDP brain and normal muscle (F) stained for TDP-43. Positive control shows normal nuclear labelling and mislocalised neuronal cytoplasmic staining with neuropil threads (E). Insert shows a neuron with absent nuclear TDP-43 and a cytoplasmic TDP-43 inclusion (E, red arrow and x100 insert). Nuclear TDP-43 staining is observed in normal muscle.

(G-I) Negative (G) and positive (H) control sections of CBD brain and normal muscle (I) stained for α B-crystallin. Positive control shows neuropil threads and a balloon cell neuron (H; red arrow and x100 insert). No α B-crystallin immunoreactivity is observed in normal muscle (I).

(J-L) Negative (J) and positive (K) control AD brain and normal muscle (L) stained for ubiquitin. Positive control shows dystrophic neurites and neuropil threads (K). No ubiquitin immunoreactivity is observed in normal muscle (L).

(M,N) Negative (M) and positive (N) control muscle stained for myotilin. Mild sarcoplasmic staining is observed in normal muscle (N).

(O) Positive control section of AD brain showing an amyloid plaque (O).

Scale bar represents 100 µm in A-D, F and H-M; and 50 µm in E, N-O.

p62 = Sequestosome 1; AD = Alzheimer's disease; TDP-43 = Transactivation response DNA binding protein 43; FTLD-TDP = Frontotemporal lobar degeneration with TDP-43 positive inclusions; CBD = Corticobasal degeneration.

Supplementary Figure 2 IBM inflammatory score-tool

Score tool modified from the published juvenile dermatomyositis inflammatory (JDM) score tool [17] to specifically assess the type, degree and distribution of inflammation in IBM. The inflammatory domain was augmented to include T-cells, T-cell subtypes, B-cells and macrophages. MHC Class I staining was expanded to include three patterns of labelling. The vascular, muscle fibre and connective tissue domains which are present in the JDM score tool were not included.

Supplementary Figure 3 Sensitivity and specificity of rimmed vacuoles, protein aggregates and mitochondrial changes in IBM+RV compared to PAM

Receiver operating characteristic curves for each test including the area under the curve and optimum cut-off with its associated sensitivity and specificity for rimmed vacuoles (A), myotilin (B), ubiquitin (C), TDP-43 (D), p62 (E) immunoreactive deposits, congophilic deposits (F) and COX-/SDH+ fibres (G). COX/SDH HC staining was the most discriminative test for differentiating IBM+RV and PAM (G). However, there was little difference between COX/SDH HC staining, TDP-43 and p62 IHC staining and none were sufficiently discriminative to be considered diagnostic. AUC = Area under the curve.

Supplementary Figure 4 Sensitivity and specificity of protein aggregates and mitochondrial changes in IBM-RV compared to PM&DM

Receiver operating characteristic curves for each test showing the area under the curve and optimum cut-off with its sensitivity and specificity for myotilin (A), ubiquitin (B), TDP-43 (C), p62 (D) immunoreactive deposits, congophilic deposits (E) and COX-/SDH+ fibres (F). COX/SDH histochemical staining (F) and myotilin (G) IHC were the most discriminative tests for differentiating IBM-RV and PM&DM. AUC = Area under the curve.

Figure 1 Protein aggregates and congophilic deposits in IBM

Stained cryostat sections, showing fibres, often in clusters, containing protein aggregates stained for p62 (A), TDP-43 (B), ubiquitin (C), αB-crystallin (D) and myotilin (E). Protein aggregates were present throughout fibres, and were observed in apparently normal fibres, vacuolated fibres and fibres surrounded by inflammatory infiltrates. In fibres containing TDP-43 aggregates, myonuclear TDP-43 staining was frequently reduced (B). Congophilic deposits were observed in vacuolated fibres using epifluorescence (F). Tissue sections were examined using both the rhodamine red and fluorescein isothiocyanate filters to exclude areas of auto-fluorescence (arrow). Combined fluorescent image is shown. Four patterns of immunoreactivity were observed in IBM and disease controls stained for p62

using IHC (G)(H)(I)(J). Pattern I (G) - strongly stained, discreet and clearly delineated, round or angular aggregates, variable in number and size within a muscle fibre but rarely filling it and predominantly located subsarcolemmal, but also perinuclear and adjacent to vacuoles. Pattern II (H) - large aggregates of variable staining intensity. Pattern III (I) - fine granular aggregates dispersed throughout the fibre. Pattern IV (J) - fine granules and wisps of p62 immunoreactivity set within weakly basophilic inclusions.

Scale bar represents 50 μm in A and D; 25 μm in B, C and E-J.

Figure 2 Percentage of muscle fibres containing protein aggregates and Griggs' pathological features

Box and whisker plot illustrating the percentage of muscle fibres containing pathological abnormalities contained in the Griggs criteria and protein aggregates in Griggs' pathologically-definite IBM. Fibres containing aggregates immunoreactive for p62 and αB-crystallin were more frequent than those containing the current diagnostic pathological features (red bars) ($p<0.05$). Protein aggregates recognised by all antibodies were found in a significantly larger number of fibres than partial invasion ($p<0.02$).

Figure 3 Percentage of fibres containing protein aggregates and COX-/SDH+ fibres in each group

Box and whisker plots illustrating the percentage of fibres in each diagnostic category containing p62 (A), TDP-43 (B), myotilin (C) and ubiquitin (D) immunoreactive aggregates, congophilic deposits (E) and COX-/SDH+ fibres (F). All protein aggregates were present in a greater percentage of fibres in IBM+RV than in IBM–RV. There was no difference in the percentage of COX-/SDH+ muscle fibres between these groups. IBM+RV biopsies had a greater percentage of fibres containing p62 (A) and TDP-43 (B) immunoreactive aggregates and COX-/SDH+ fibres (F) than PAM. Pathological findings

were similar in IBM-RV and PM&DM, with no differences in the percentage of fibres containing p62 (A), TDP-43 (B) and ubiquitin (D) immunoreactive aggregates or congophilic deposits (E). However, there was a greater percentage of COX-/SDH+ fibres (F) in IBM–RV than PM&DM and a greater percentage of fibres containing myotilin immunoreactive aggregates (C) in PM&DM than IBM-RV. *Statistically significant results.

Figure 4 Proposed diagnostic algorithm for IBM based on pathological findings

Flow diagram showing a proposed pathway for diagnosing IBM based on the pathological findings. Increased MHC Class I staining was observed in all cases of IBM and pattern I p62 aggregates in all cases of IBM+RV making them good initial screening tests. Their absence rules-out a diagnosis of IBM and IBM+RV respectively. The presence of endomysial CD3+ T-cell score >1, endomysial CD8+ T-cell score >0 or strong MHC Class I staining in a biopsy with rimmed vacuoles and p62 aggregates secures a diagnosis of IBM+RV. Differentiating IBM-RV and PM&DM pathologically is more challenging. The presence of COX-/SDH+ fibres is not specific to IBM-RV; although COX-/SDH+ fibres were not present in every case of IBM-RV their absence casts doubt on the diagnosis of IBM-RV. Pattern I p62 aggregates may enable IBM to be differentiated from PM when present. However, they may lack sensitivity for IBM-RV, therefore their absence does not rule out the diagnosis.

A retrospective cohort study identifying the principal pathological features useful in the diagnosis of inclusion body myositis

Corresponding author:

Dr Janice L Holton

Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

janice.holton@ucl.ac.uk; tel: 00 44 (0)20 3448 4239; fax: 00 44 (0)20 3448 4486.

Authors:

Stefen Brady[1], Waney Squier[2], Caroline Sewry[3,4], Michael Hanna[1], David Hilton-Jones[5], Janice L Holton[6]

[1]MRC Centre for Neuromuscular Diseases, UCL Institute of Neurology and National Hospital for Neurology, Neurosurgery, Queen Square, London, UK.

[2]Department of Neuropathology, University of Oxford, John Radcliffe Hospital, Oxford, UK.

[3]Dubowitz Neuromuscular Centre, Institute of Child Health and Great Ormond Street Hospital for Children, London, UK.

[4]Wolfson Centre of Inherited Neuromuscular Diseases, RJAH Orthopaedic Hospital, Oswestry, UK.

[5]Nuffield Department of Clinical Neurosciences (Clinical Neurology), University of Oxford, John Radcliffe Hospital, Oxford, UK.

[6]Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

Keywords: Neurology, adult neurology, neuromuscular disease, neuropathology

Running title: Pathological criteria for inclusion body myositis

Word count: 2990

ABSTRACT

Objectives

The current pathological diagnostic criteria for sporadic inclusion body myositis (IBM) lack sensitivity. Using immunohistochemical techniques abnormal protein aggregates have been identified in IBM, including some associated with neurodegenerative disorders. Our objective was to investigate the diagnostic utility of a number of markers of protein aggregates together with mitochondrial and inflammatory changes in IBM.

Design

Retrospective cohort study. The sensitivity of pathological features was evaluated in cases of Griggs' definite IBM. The diagnostic potential of the most reliable features was then assessed in clinically-typical IBM with rimmed vacuoles (*n*=15) and clinically-typical IBM without rimmed vacuoles (*n*=9) and IBM mimics - vacuolar myopathies (*n*=7) and steroid-responsive inflammatory myopathies (*n*=11).

Setting

Specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London.

Results

Individual pathological features, in isolation, lacked sensitivity and specificity. However, the morphology and distribution of p62 aggregates in IBM were characteristic and in a myopathy with rimmed vacuoles, the combination of characteristic p62 aggregates and increased sarcolemmal and internal MHC Class I expression or endomysial T-cells were diagnostic for IBM with a sensitivity of 93% and specificity of 100%. In an inflammatory myopathy lacking rimmed vacuoles, the presence of mitochondrial changes was 100% sensitive and 73% specific for IBM; characteristic p62 aggregates were specific (91%), but lacked sensitivity (44%).

Conclusions

We propose an easily applied diagnostic algorithm for the pathological diagnosis of IBM. Additionally our findings support the hypothesis that many of the pathological features considered

typical of IBM develop later in the disease, explaining their poor sensitivity at disease presentation and emphasising the need for revised pathological criteria to supplement the clinical criteria in the diagnosis of IBM.

STRENGTHS AND LIMITATIONS

The present study is a multicentre retrospective evaluation of the diagnostic utility of pathological findings for differentiating IBM from myopathies important in the differential diagnosis – myopathies containing rimmed vacuoles and steroid-responsive inflammatory myopathies.

The main strength of our study was the systematic detailed analysis of well-defined cases. This enabled us to determine the sensitivity and specificity of individual pathological features and produce an easily applied pathological diagnostic algorithm for IBM for use in clinical practice.

Study limitations include the small number of cases and the retrospective design. Further prospective studies are now required in larger cohorts of patients.

INTRODUCTION

Sporadic inclusion body myositis (IBM) is the commonest acquired myopathy in those aged over 50 years.[1] Although classified as an idiopathic inflammatory myopathy, muscle biopsy reveals both degenerative and inflammatory features. The widely used Griggs diagnostic criteria require the presence of several pathological findings,[2] namely rimmed vacuoles, an inflammatory infiltrate with invasion of non-necrotic fibres by mononuclear inflammatory cells (partial invasion), and either amyloid deposits or 15-18 nm tubulofilaments identified by electron microscopy (EM). Although these features in combination are highly specific for IBM, individually they occur in other myopathies, including some important in the differential diagnosis for IBM.[3-7] Moreover, cases of clinically-typical IBM have been reported where the combination of these pathological features is absent causing diagnostic difficulty.[8-11]

Over the last two decades, pathological accumulation of many different proteins has been reported in muscle fibres in IBM.[12] Proteins typically associated with neurodegenerative diseases such as β-amyloid (Aβ), hyperphosphorylated tau and ubiquitin and newer neurodegenerative markers such as p62 and transactivation response DNA binding protein-43 (TDP-43) have been identified, as well as proteins associated with myofibrillar myopathies (MFM), including desmin and αB-crystallin. However, not all observations have been consistently reproduced.[13,14] Mitochondrial changes have also been proposed for inclusion in IBM diagnostic criteria,[15]. Clear guidelines for the incorporation of immunohistochemical findings and mitochondrial changes into diagnostic criteria for IBM have not been established.[16]

Previously, we have shown that the characteristic pattern of weakness associated with IBM is indicative of the diagnosis, even if Griggs pathological features are absent.[11] However, it is not invariably found at presentation. Here we sought to identify which pathological features, other than the Griggs pathological criteria, add further support to the diagnosis of IBM. We systematically investigated which pathological features are present in Griggs pathologically-definite IBM and then established the diagnostic utility of these features in cases of IBM lacking the Griggs criteria, using myopathies considered in the differential diagnosis of IBM as controls.

MATERIALS AND METHODS

The study received ethical approval from the Departments of Research and Development at Oxford University Hospitals NHS Trust, Oxford and University College London Hospitals NHS Foundation Trust, London.

Cases

All patients were followed by specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London. Biopsies were taken for diagnostic purposes from the deltoid or quadriceps muscles and prior to any treatment.

Methods for demonstrating pathological features in IBM, additional to those defined by the Griggs criteria, were determined in six Griggs pathologically-definite cases of IBM. Cases with no clinical or pathological evidence of neuromuscular disease were used as controls. The diagnostic utility of the pathological features identified was assessed in two groups of clinically-typical IBM; one with rimmed vacuoles on muscle biopsy (IBM+RV; *n*=15), the other without rimmed vacuoles on muscle biopsy (IBM-RV; *n*=9). Disease controls were cases of steroid-responsive inflammatory myopathies [polymyositis and dermatomyositis; (PM&DM); *n*=11] and protein accumulation myopathies with rimmed vacuoles (PAM; *n*=7). Clinical characteristics and inclusion criteria are summarised in Supplementary tables 1 and 2. Tissue from brains donated to the Queen Square Brain Bank for Neurological Disorders was used as positive controls for protein aggregate staining.

Muscle biopsies

Muscles biopsies were snap frozen at the time of surgery in isopentane cooled liquid nitrogen. Until sectioning all samples were stored at -80^{o}C. Serial tissue sections were cut to a thickness of 8 µm, allowed to air dry and stored at -80^{o}C until staining. Prior to staining, tissue sections were allowed to dry at room temperature. Tissue sections were stained with haematoxylin and eosin (H&E), combined cytochrome oxidase (COX) succinate dehydrogenase (SDH) histochemistry and for amyloid using alkalinised Congo red, crystal violet and thioflavin S. Tissue sections for immunohistochemical staining were fixed for 10 minutes, if required, washed for five minutes in running water and incubated in 0.5% hydrogen peroxide to block endogenous peroxidase for 20 minutes. After further washing, tissue sections were incubated in 5% normal goat serum (Vector Laboratories, Burlingame, California) for 30 minutes and then systematically stained for: 1) proteins classically associated with

neurodegenerative disease: tau and hyperphosphorylated tau, ubiquitin, Aβ and α-synuclein; 2) proteins more recently reported in neurodegenerative disease: p62, TDP-43, fused in sarcoma protein (FUS) and valosin containing protein (VCP); 3) nuclear membrane proteins: lamin A/C and emerin; 4) proteins associated with MFM: desmin, myotilin and αB-crystallin; and 5) inflammatory cells and major histocompatibility complex class I (MHC Class I): CD3+ T-cells, CD4+ T-cells, CD8+ T-cells, B-cells and macrophages. Primary antibody binding was visualised using Dako REAL™ EnVision™ Detection System which contains horse-radish peroxidase (HRP) labelled goat anti-rabbit/mouse secondary and 3,3'-diaminobenzidine (DAB); following incubation with the relevant primary antibody, tissue sections were washed in phosphate buffered saline (PBS), incubated with HRP labelled goat anti-rabbit/mouse secondary for 30 minutes, washed in PBS and incubated in a 1:50 solution of DAB for three to five minutes. Details of commercial antibodies and conditions used are provided in Supplementary Table 3. IHC for each antibody was performed on all cases simultaneously and including positive and negative controls (Supplementary Figure 1).

Definitions and quantification

The total number of fibres and the number undergoing partial invasion, containing rimmed vacuoles, protein aggregates and COX-negative SDH-positive (COX-/SDH+) fibres were quantified using ImagePro version 6.2 (Media Cybernetics), to ensure that the whole biopsy was systematically analysed. Only transversely-orientated fibres not undergoing necrosis or regeneration were quantified. Tissue sections stained with Congo red were visualised under fluorescent and polarised light. Areas of fluorescence were examined using both rhodamine red (excitation 512-546 nm and emission 600-640 nm) and fluorescein isothiocyanate (excitation 440-480 nm and emission 527-530 nm) filters to exclude auto-fluorescence. Supplementary Table 4 provides definitions of the pathological features assessed. The inflammatory infiltrate and MHC Class I staining were analysed using a modified version of the semi-quantitative juvenile dermatomyositis score-tool (Supplementary Figure 2).[17] Assessments were performed blind to clinical details and diagnosis by a single individual (SB). Ten

per cent of slides were re-counted to assess intra-observer reliability and 336 slides were assessed independently by two observers (SB and JLH) to determine inter-observer reliability.

Statistical analysis

Statistical analyses were performed using GraphPad PRISM version 5. Continuous and categorical variables were compared using Mann Whitney *U*-test and chi-squared or Fisher's exact test respectively. Spearman's rank order correlation was used to determine the strength and direction of associations between pathological findings. Linear regression was used to determine relationships between clinical features and pathological findings. Test characteristics were calculated using receiver operating characteristic (ROC) curves and 2x2 contingency tables. A test was considered diagnostic when sensitivity >75% and specificity >95% or sensitivity >95% and specificity >75%. Intra-observer and inter-observer agreement was calculated using Bland-Altman plots and Cohen's kappa statistic (κ). Repeat counts were within 95% confidence intervals using Bland-Altman plots and κ was ≥0.7 indicating good intra-observer and good or excellent inter-observer reliability. Statistical significance was set at $p<0.05$.

RESULTS

Pathological findings in Griggs' pathologically-definite IBM

p62, TDP-43, ubiquitin, myotilin and αB-crystallin immunoreactive aggregates were present in all six IBM cases but not in normal controls (Figures 1A-E). p62 and αB-crystallin immunoreactive aggregates were present in a greater percentage of fibres than the pathological features required in the Griggs criteria ($p<0.05$) (Figure 2). Despite their abundance, αB-crystallin immunoreactive aggregates were difficult to quantify due to a significant variability in their morphology. No immunoreactive deposits were observed in IBM cases or normal controls with antibodies to tau and phosphorylated tau, Aβ, α-synuclein, desmin, emerin, lamin A/C, FUS or VCP. Alkalinised Congo red staining was more sensitive than crystal violet and thioflavin S staining for observing amyloid aggregates (Figure 1F). Tissue sections containing congophilic deposits identified under fluorescence light showed no

apple-green birefringence under polarised light. Mitochondrial changes and increased sarcolemmal and sarcoplasmic MHC Class I staining were observed in all six IBM cases, but not in normal controls. The inflammatory infiltrate was predominantly composed of endomysial CD8+ T-cells and macrophages, with relatively few B-cells.

Quantitative analysis of pathological features in IBM and disease controls

Having shown that p62, TDP-43, ubiquitin and myotilin aggregates, congophilic deposits, MHC Class I and inflammatory cells were prevalent in Griggs' pathologically-definite IBM, the presence of these abnormalities, together with mitochondrial changes were assessed in IBM+RV, IBM-RV and disease controls.

The percentage of fibres containing p62, TDP-43, myotilin and ubiquitin aggregates and congophilic deposits were greater in IBM+RV than in IBM-RV; there was no difference in the number of COX-/SDH+ fibres (Figure 3A-F). Protein aggregates were observed in morphologically-normal fibres and in fibres exhibiting Griggs' pathological features. p62 and TDP-43 positive aggregates were present in a greater percentage of fibres in IBM+RV compared to PAM; however, there were no differences in the percentage of fibres containing myotilin and ubiquitin aggregates or congophilic deposits. The percentage of fibres containing p62, TDP-43 and ubiquitin aggregates or congophilic deposits were similar in IBM-RV and PM&DM; however, myotilin aggregates were present in a greater percentage of fibres in PM&DM and COX-/SDH+ fibres were more abundant in IBM-RV. Analysis of the total inflammatory infiltrate (the sum of the semi-quantitative scores for T-cells, B-cells and macrophages) in the endomysium, perimysium and perivascular areas revealed that there were greater numbers of inflammatory cells in the endomysium and perimysium in IBM+RV than in PAM (p<0.03). The same analysis comparing the sum of the inflammatory cells in IBM-RV and PM&DM revealed that the distribution and intensity of the inflammatory infiltrate was similar.

Diagnostic utility of pathological features in IBM and disease controls

To mimic the diagnostic difficulty encountered in clinical practice, the ability of each test to differentiate between myopathies containing rimmed vacuoles (IBM+RV and PAM) and between inflammatory myopathies (IBM–RV and PM&DM) was assessed.

Diagnostic utility determined using receiver-operating characteristic curves

Individually, the presence of p62 immunoreactive inclusions and COX-/SDH+ fibres had the highest sensitivity and specificity for differentiating IBM+RV from PAM, (Supplementary Figure 3) (Table 1). Differentiating between IBM–RV and PM&DM, myotilin positive inclusions or COX-/SDH+ fibres had the highest sensitivity and specificity for IBM-RV (Supplementary Figure 4) (Table 1). Only the presence of myotilin positive inclusions satisfied criteria to be considered suitable as a diagnostic test (<0.01% of fibres containing myotilin aggregates had a sensitivity of 100% and specificity of 82% for IBM-RV).

Table 1 Test characteristics

Test feature	IBM+RV *v.* PAM				IBM-RV *v.* PM&DM			
	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity
Rimmed vacuoles	0.60	>0.28	0.53	0.71	-	-	-	-
p62 aggregates	0.87	>0.48	0.87	0.86	0.60	>0.21	0.22	0.91
TDP-43 aggregates	0.80	>0.34	0.80	0.86	0.53	<0.01	0.89	0.18
Ubiquitin aggregates	0.68	>0.18	0.53	0.85	0.64	<0.01	1.00	0.27
Myotilin aggregates	0.55	<0.25	1.00	0.29	0.91	<0.01	1.00	0.82
Congophilic deposits	0.56	>0.24	0.73	0.71	0.56	<0.03	0.11	0.82
COX-/SDH+ fibres	0.87	>0.04	0.86	0.86	0.93	>0.1	0.78	0.91

Table shows the area under the curve and optimum cut-off for each test with the accompanying sensitivity and specificity. AUC = Area under the curve.

Diagnostic utility determined by comparing proportion of affected cases in each diagnostic group

In the aforementioned experiments, the number of fibres within each muscle biopsy was quantified. However, this is impractical for routine clinical use. Thus, the proportions of affected cases in each group were compared (Table 2). This revealed that neither staining for protein aggregates nor congophilic deposits could differentiate between IBM+RV and PAM. The pathological findings in IBM-RV and PM&DM were also similar, except that the absence of myotilin immunoreactive aggregates was sensitive and specific for IBM-RV. COX-/SDH+ fibres were also suggestive of IBM-RV; one or more COX-/SDH+ fibres had a sensitivity of 100% and specificity 73% for IBM-RV.

Increased MHC Class I expression lacked specificity. However, strong (diffuse sarcolemmal and sarcoplasmic) MHC Class I up-regulation was diagnostic for IBM+RV, differentiating it from PAM, as were the presence of either endomysial CD3+ T-cell or CD4+ T-cell scores >1 or an endomysial CD8+ T-cell score >0. Partial invasion was specific for IBM+RV, but lacked sensitivity. Although the sum of the inflammatory infiltrate was similar in IBM-RV and PM&DM, analysis of the inflammatory cell sub-types revealed greater numbers of perimysial CD3+ T-cells, CD8+ T-cells and endomysial B-cells [were observed] in PM&DM than in IBM–RV ($p \leq 0.02$), however, this was not diagnostically useful. There was no difference in the proportion of cases with fibres undergoing partial invasion between IBM–RV and PM&DM.

Table 2 Comparison of the proportion of positive cases in each group

Pathological features	IBM+RV	PAM	IBM+RV *v.* PAM		IBM–RV	PM&DM	IBM-RV *v.* PM&DM		IBM+RV *v.* IBM-RV
	n (%)	n (%)	Sensitivity	Specificity	n (%)	n (%)	Sensitivity	Specificity	*p* value
Number of cases	15 (100)	7 (100)			9 (100)	11 (100)			
Aggregated proteins, n (%)									
p62	15 (100)	6 (86)	1.00	0.14	4 (44)	3 (27)‡	0.40	0.72	0.003*
TDP-43	13 (87)	5 (71)	0.87	0.29	1 (11)	2 (18)‡	0.11	0.82	0.001*
Ubiquitin	11 (73)	4 (57)	0.73	0.43	0 (0)	3 (27)‡	0.00	0.73	0.001*
Myotilin	10 (67)	5 (71)	0.67	0.29	0 (0)	9 (82)	0.00	0.18	0.002*
Congophilic deposits	13 (87)	7 (100)	0.87	0.00	1 (11)	0 (0)	0.11	1.00	0.001*
COX-/SDH+ fibres†, n (%)									
Any	12 (86)	2 (29)	0.80	0.71	9 (100)	3 (27)	1.00	0.73	0.5
Inflammatory features, n (%)									
MHC Class I up-regulation	15 (100)	3 (43)	1.00	0.57	9 (100)	11 (100)	1.00	0.00	1.00
Strong MHC Class I up-regulation	14 (93)	0 (0)	0.93	1.00	9 (100)	10 (91)	1.00	0.09	0.53
Partial invasion	10 (67)	0 (0)	0.67	1.00	3 (33)	2 (18)	0.33	0.82	0.11
Endomysial CD3+ T-cell score >1	13 (87)	0 (0)	0.87	1.00	4 (44)	7 (64)	0.44	0.36	0.02*
Endomysial CD4+ T-cell score >1	12 (80)	0 (0)	0.80	1.00	2 (22)	5 (45)	0.22	0.46	0.01*
Endomysial CD8+ T-cell score >0	14 (93)	0 (0)	0.93	1.00	4 (44)	5 (45)	0.44	0.54	0.02*
Endomysial CD68+ macrophage score >1	12 (80)	0 (0)	0.80	1.00	4 (44)	8 (73)	0.44	0.17	0.07

†In IBM with rimmed vacuoles *n*=14. ‡Pathological features present in DM, but not PM cases. *Statistically significant results.

Because IBM-RV is more pathologically akin to PM than DM, analyses were repeated comparing IBM-RV and PM cases (*n*=6). No p62, TDP-43 or ubiquitin immunoreactive aggregates were observed in PM cases and the diagnostic utility of tests for differentiating between IBM-RV and PM yielded similar results to prior analyses between IBM-RV and PM&DM.

Diagnostic utility of categorising the pattern of p62 staining

The pattern of p62 staining could be categorised into four distinct groups (Figure 1G-J). Aggregates observed in IBM were present in vacuolated and non-vacuolated fibres and were strongly stained, discreet and clearly delineated, round or angular and typically located subsarcolemmal, perinuclear and peri-vacuolar (pattern I). This pattern was observed in every IBM case with p62 aggregates, one (9%) case of DM and three (43%) cases of PAM (hereditary IBM, dystrophinopathy and genetically undefined MFM). Defining the pattern of immunoreactivity increased the discriminative value of p62 IHC for differentiating IBM+RV from PAM; pattern I p62 aggregates compared to any p62 aggregates increased the specificity from 14% to 57%, with no loss of sensitivity. Differentiating IBM-RV and PM&DM, pattern I p62 aggregates were highly specific (91%), but lacked sensitivity (44%). Patterns II, III and IV were not observed in any IBM cases. Patterns II and III appeared to be specific for PAM (*n*=2; 26%), both were cases of myotilinopathy (*n*=2; 67%), and DM (*n*=2; 40%) respectively. Pattern IV occurred in a genetically undefined case of MFM. No differences were observed in the morphology of TDP-43, myotilin or ubiquitin aggregates between biopsies.

Clinicopathological correlation

In IBM+RV, IBM-RV and pathologically-definite IBM, there were no correlations in individual biopsies between pathological features. No relationships were identified between the pathological findings and age at symptom onset, age at biopsy, disease duration or serum creatine kinase. The same results were obtained when the IBM groups were analysed separately and as one.

Proposed diagnostic algorithm

Based on our pathological findings, we propose a diagnostic algorithm for differentiating IBM from its disease mimics (Figure 4).

The algorithm was tested in a further 23 cases that fulfilled the criteria for IBM+RV (*n*=12) and IBM-RV (*n*=11). The algorithm correctly diagnosed 20 (87%) cases: 12 (100%) cases of IBM+RV and eight (73%) cases of IBM-RV. In IBM-RV, COX-/SDH+ fibres were present in 8 (73%) cases, pattern I p62 aggregates in 8 (73%) cases and both in 6 (55%) cases.

DISCUSSION

While Griggs' pathological criteria have been accepted as diagnostic of IBM, many patients who, observed over time undoubtedly have IBM, lack one or more of the Griggs pathological features at presentation, even on repeat biopsy.[8,11] Despite IBM being associated with a characteristic pattern of finger flexor and knee extensor weakness, not all patients have this pattern at disease onset, and muscle biopsy remains an important tool in differentiating IBM from its mimics. We sought to determine which additional pathological features support a diagnosis of IBM, demonstrating that characteristic p62 immunoreactive aggregates, strong MHC Class I upregulation, endomysial CD3+ T-cell score >1, CD8+ T-cell score >0 and COX-/SDH+ fibres are features with sufficient sensitivity and specificity to differentiate IBM from pathologically similar myopathies and we propose an easily applied pathological algorithm for the diagnosis of IBM (Figure 4).

In agreement with previous studies, we observed p62,[18] TDP-43,[19] ubiquitin [13] and αB-crystallin [20] immunoreactive aggregates and a predominantly endomysial inflammatory infiltrate [3] in Griggs pathologically-definite IBM. Diagnostic pathological studies of IBM have concentrated on differentiating IBM from other inflammatory myopathies and two recent quantitative studies have found that TDP-43 and markers of autophagy such as p62 and LC3 may be of diagnostic use.[21,22] However, in these studies only a fraction of each biopsy was analysed i.e. 200 fibres. We have found

this limited quantification does not correlate with the percentage of affected fibres in a biopsy nor does it reflect the way in which a muscle biopsy is assessed. Additionally, studies have lacked vacuolar myopathy control cases as it is believed that the inflammatory changes present in IBM enable it to be easily differentiated from other vacuolar myopathies.[22] However, inflammatory changes are frequently observed in muscular dystrophies and the degree of inflammatory change necessary to confidently diagnose IBM is currently unknown.

To mimic the typical diagnostic conundrums encountered in clinical practice, we evaluated the ability of the pathological findings to differentiate IBM+RV from other vacuolar myopathies and IBM-RV from steroid-responsive inflammatory myopathies. We found that quantitative analysis of protein aggregates, congophilic deposits and COX-/SDH+ fibres was of limited diagnostic use. Analysing the biopsies dichotomously and using a semi-quantitative score-tool revealed that increased MHC Class I labelling was sensitive for IBM making it a good initial screening test, its absence excluding the diagnosis. In agreement with an earlier study, we found p62 aggregates identified the largest number of affected fibres in IBM.[23] Additionally, as a novel finding, the morphology and distribution of p62 aggregates was characteristic in IBM. This characteristic pattern of p62 immunoreactive aggregates was highly sensitive for IBM+RV (100%); their absence from a biopsy containing rimmed vacuoles effectively ruling-out a diagnosis of IBM. We confirmed that the most diagnostically useful pathological findings in IBM+RV were evidence of an immune mediated process; strong MHC Class I staining, endomysial CD3+ T-cell score >1 or an endomysial CD8+ T-cell score >0 were diagnostic. Having identified either of these features in a biopsy containing rimmed vacuoles no extra diagnostic certainty was gained from observing partial invasion, COX-/SDH+ fibres or congophilic deposits.

The most discriminative pathological tests for differentiating between IBM–RV and PM&DM were COX/SDH staining and myotilin IHC. Consistent with a recent study,[9] we found the absence of mitochondrial changes casts doubt on a diagnosis of IBM. There was no difference in the median age between IBM-RV and PM&DM cases to account for the difference observed in COX-/SDH+ fibres.

The presence of myotilin and ubiquitin immunoreactive aggregates appeared to rule out a diagnosis of IBM-RV. However, we believe the presence of these features in IBM+RV indicates that they are unlikely to be diagnostically reliable features for differentiating between IBM-RV and steroid-responsive inflammatory myopathies. Although no pathological feature was able to differentiate IBM-RV from steroid responsive inflammatory myopathies with certainty the presence of characteristic p62 aggregates and the absence of COX-/SDH+ fibres may help in supporting and opposing a diagnosis of IBM-RV respectively. Pattern I p62 immunoreactive aggregates were only present in 44% of the initial IBM-RV cases tested, but they were not observed in PM cases and were very rare in DM. Although pattern I p62 aggregates appear to lack sensitivity their specificity was 91% making their presence highly suggestive of a diagnosis of IBM-RV. However, we identified pattern I p62 in eight out of 11 (73%) further cases of IBM-RV that were assessed indicating a greater sensitivity and that p62 IHC warrants further investigation and validation in a larger, independent series. The diagnostic utility of the other patterns of p62 staining is uncertain. Although pattern II appeared to have some specificity for myotilinopathy the small number of cases makes it drawing any conclusion problematic. In addition to p62 other autophagic proteins have been found in IBM and suggested as diagnostic markers. [22] Autophagy is a cellular mechanism for degrading and recycling cellular proteins and organelles and therefore, altered autophagy could lead to the accumulation of abnormal mitochondria and misfolded aggregation-prone proteins and may also result in altered antigen presentation leading to the widespread increase of MHC Class I and suggests that altered autophagy may play an important role in the pathogenesis of IBM.

Almost all pathological features - protein aggregates, congophilic deposits and inflammation - were more abundant in IBM+RV than IBM-RV. Despite using slightly different inclusion criteria, similar differences have been reported between pathologically-typical and pathologically-atypical IBM.[21] However, we found no differences in the number of COX-/SDH+ fibres, the degree of MHC Class I upregulation, the morphology and distribution of p62 immunoreactive aggregates or the pattern of the inflammation between IBM+RV and IBM-RV, supporting our clinical observations that these are the

same disease. We believe that the pathological differences between IBM+RV and IBM-RV are, in part, due to differences in disease duration. Two studies have shown that rimmed vacuoles are more common in patients who are older at the time of muscle biopsy,[24,11] suggesting that they are associated with chronologically more advanced disease. Therefore, the pathological findings which are more abundant in IBM+RV and thought to be typical of IBM may instead be indicative of chronologically more advanced disease explaining their limited sensitivity at disease presentation. However, possibly due to the number of cases analysed, we were unable to confirm a relationship between pathological features and clinical findings. It could be argued that biopsies from different muscles may have affected the pathological findings observed and differences between IBM groups. However, in a recent review of 59 muscle biopsies from IBM cases in our clinical archive with quadriceps (n=31) and deltoid (n=28) biopsies we found no significant difference in the frequency of pathological findings.

A robust clinicopathological definition of IBM is of paramount importance for diagnosis and for selection and entry of patients into clinical trials. We have shown that certain pathological findings are more abundant than those included in the current pathologically-focussed diagnostic criteria. Moreover, p62 immunoreactive deposits, increased MHC Class I expression, endomysial CD3+ T-cells and CD8+ T-cells and COX-/SDH+ fibres have sufficient sensitivity and specificity to aid in the histological differentiation of IBM from disease mimics, supporting their inclusion in future diagnostic criteria for IBM alongside clinical criteria. Both CD3+ T-cells and CD8+ T-cells are included in the diagnostic algorithm as there was little difference in their sensitivity and specificity for differentiating IBM+RV from PAM. However, IHC staining for CD3+ T-cells is likely to be more widely available and avoids the costs of extra staining to subtype the inflammatory infiltrate enabling diagnostic algorithm to be used by a greater number diagnostic laboratories. Using our diagnostic algorithm, we found there would be little additional diagnostic security in identifying partial invasion, performing EM or staining for amyloid deposits. Finally, mitochondrial changes and MHC Class I up-regulation were the most consistent findings in our IBM cases suggesting that they are central to the

pathogenesis and that further investigation and therapeutic intervention should be directed towards these features.

REFERENCES

1. Needham M, James I, Corbett A, Day T, Christiansen F, Phillips B, et al. Sporadic inclusion body myositis: phenotypic variability and influence of HLA-DR3 in a cohort of 57 Australian cases. J Neurol Neurosurg Psychiatr. 2008;79:1056–60.
2. Griggs RC, Askanas V, DiMauro S, Engel A, Karpati G, Mendell JR, et al. Inclusion body myositis and myopathies. Ann Neurol. 1995;38:705–13.
3. Arahata K, Engel AG. Monoclonal antibody analysis of mononuclear cells in myopathies. I: Quantitation of subsets according to diagnosis and sites of accumulation and demonstration and counts of muscle fibers invaded by T cells. Ann Neurol. 1984;16:193–208.
4. Mhiri C, Gherardi R. Inclusion body myositis in French patients. A clinicopathological evaluation. Neuropathol Appl Neurobiol. 1990;16:333–44.
5. Villanova M, Kawai M, Lübke U, Oh SJ, Perry G, Six J, et al. Rimmed vacuoles of inclusion body myositis and oculopharyngeal muscular dystrophy contain amyloid precursor protein and lysosomal markers. Brain Res. 1993;603:343–7.
6. Van der Meulen MF, Hoogendijk JE, Moons KG, Veldman H, Badrising UA, Wokke JH. Rimmed vacuoles and the added value of SMI-31 staining in diagnosing sporadic inclusion body myositis. Neuromuscul Disord. 2001;11:447–51.
7. Ferrer I, Olivé M. Molecular pathology of myofibrillar myopathies. Expert Rev Mol Med. 2008;10:e25.
8. Amato AA, Gronseth GS, Jackson CE, Wolfe GI, Katz JS, Bryan WW, et al. Inclusion body myositis: clinical and pathological boundaries. Ann Neurol. 1996;40:581–6.
9. Chahin N, Engel AG. Correlation of muscle biopsy, clinical course, and outcome in PM and sporadic IBM. Neurology. 2008;70:418–24.

10. Benveniste O, Guiguet M, Freebody J, Dubourg O, Squier W, Maisonobe T, et al. Long-term observational study of sporadic inclusion body myositis. Brain. 2011;134:3176–84.

11. Brady S, Squier W, Hilton-Jones D. Clinical assessment determines the diagnosis of inclusion body myositis independently of pathological features. J Neurol Neurosurg Psychiatr. 2013; 16.

12. Greenberg SA. Theories of the Pathogenesis of Inclusion Body Myositis. Current Rheumatology Reports. 2010;12:221–8.

13. Sherriff FE, Joachim CL, Squier MV, Esiri MM. Ubiquitinated inclusions in inclusion-body myositis patients are immunoreactive for cathepsin D but not β-amyloid. Neuroscience Letters. 1995;194:37–40.

14. Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. BMJ. 2009;339:b2680.

15. Needham M, and Mastaglia FL. (2007). Inclusion body myositis: current pathogenetic concepts and diagnostic and therapeutic approaches. Lancet Neurol. 6, 620–631.

16. Benveniste O, Hilton-Jones D. International Workshop on Inclusion Body Myositis held at the Institute of Myology, Paris, on 29 May 2009. Neuromuscular Disorders. 2010;20:414–21.

17. Wedderburn LR, Varsani H, Li CKC, Newton KR, Amato AA, Banwell B, et al. International consensus on a proposed score system for muscle biopsy evaluation in patients with juvenile dermatomyositis: a tool for potential use in clinical trials. Arthritis Rheum. 2007;57:1192–201.

18. Nogalska A, Terracciano C, D'Agostino C, King Engel W, Askanas V. p62/SQSTM1 is overexpressed and prominently accumulated in inclusions of sporadic inclusion-body myositis muscle fibers, and can help differentiating it from polymyositis and dermatomyositis. Acta Neuropathol. 2009;118:407–413.

19. Weihl CC, Temiz P, Miller SE, Watts G, Smith C, Forman M, Hanson PI, Kimonis V, Pestronk A. TDP-43 accumulation in inclusion body myopathy muscle suggests a common pathogenic mechanism with frontotemporal dementia. J. Neurol. Neurosurg. Psychiatr. 2008;79:1186–1189.

20. Banwell BL, Engel AG. Alpha B-crystallin immunolocalization yields new insights into inclusion body myositis. Neurology. 2000;54:1033–1041.

21. Dubourg O, Wanschitz J, Maisonobe T, Béhin A, Allenbach Y, Herson S, and Benveniste O. Diagnostic value of markers of muscle degeneration in sporadic inclusion body myositis. Acta Myol. 2011. 30, 103–108.

22. Hiniker A, Daniels BH, Lee HS, Margeta M. Comparative utility of LC3, p62 and TDP-43 immunohistochemistry in differentiation of inclusion body myositis from polymyositis and related inflammatory myopathies. Acta Neuropathologica Communications. 2013. 1:29.

23. D'Agostino C, Nogalska A, Engel WK, Askanas V. In sporadic inclusion-body myositis muscle fibres TDP-43-positive inclusions are less frequent and robust than p62-inclusions, and are not associated with paired helical filaments. Neuropathol Appl Neurobiol. 2010; Available from: http://www.ncbi.nlm.nih.gov/pubmed/20626631.

24. Momma K, Noguchi S, Malicdan MCV, Hayashi YK, Minami N, Kamakura K, et al. Rimmed vacuoles in Becker muscular dystrophy have similar features with inclusion myopathies. PLoS ONE. 2012;7:e52002.

ACKNOWLEDGEMENTS

SB is funded by the Myositis Support Group. JLH is supported by the Reta Lila Weston Institute for Neurological Studies and the Myositis Support Group. This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme.

CONTRIBUTORSHIP STATEMENT

Dr Stefen Brady - Acquisition of data, analysis and interpretation of data and drafting of manuscript.

Dr Waney Squier - Critical revision of manuscript for important intellectual content.

Prof. Caroline Sewry - Study concept and design and critical revision of manuscript for important intellectual content.

Prof. Mike Hanna - Critical revision of manuscript for important intellectual content.

Dr David Hilton-Jones - Critical revision of manuscript for important intellectual content.

Dr Janice Holton - Study concept and design, critical revision of manuscript for important intellectual content and study supervision.

DATA SHARING

All additional data can be found in supplementary tables and figures.

Supplementary Figure 1 Control staining in brain and muscle tissue

Positive and negative (no primary) brain control sections and normal muscle stained using immunohistochemistry for: p62 (A-C), TDP-43 (D-F), α B-crystallin (G-I), ubiquitin (J-K) and myotilin (M,N) and alkalinised congo red (O).

(A-C) Negative (A) and positive (B) control sections of AD brain and normal muscle (C) stained for p62. Positive control shows p62 positive neurofibrillary tangles and dystrophic neurites (B). No p62 immunoreactivity is observed in normal muscle (C).

(D-F) Negative (D) and positive (E) control sections of FTLD-TDP brain and normal muscle (F) stained for TDP-43. Positive control shows normal nuclear labelling and mislocalised neuronal cytoplasmic staining with neuropil threads (E). Insert shows a neuron with absent nuclear TDP-43 and a cytoplasmic TDP-43 inclusion (E, red arrow and x100 insert). Nuclear TDP-43 staining is observed in normal muscle.

(G-I) Negative (G) and positive (H) control sections of CBD brain and normal muscle (I) stained for α B-crystallin. Positive control shows neuropil threads and a balloon cell neuron (H; red arrow and x100 insert). No α B-crystallin immunoreactivity is observed in normal muscle (I).

(J-L) Negative (J) and positive (K) control AD brain and normal muscle (L) stained for ubiquitin. Positive control shows dystrophic neurites and neuropil threads (K). No ubiquitin immunoreactivity is observed in normal muscle (L).

(M,N) Negative (M) and positive (N) control muscle stained for myotilin. Mild sarcoplasmic staining is observed in normal muscle (N).

(O) Positive control section of AD brain showing an amyloid plaque (O).

Scale bar represents 100 µm in A-D, F and H-M; and 50 µm in E, N-O.

p62 = Sequestosome 1; AD = Alzheimer's disease; TDP-43 = Transactivation response DNA binding protein 43; FTLD-TDP = Frontotemporal lobar degeneration with TDP-43 positive inclusions; CBD = Corticobasal degeneration.

Supplementary Figure 2 IBM inflammatory score-tool

Score tool modified from the published juvenile dermatomyositis inflammatory (JDM) score tool [17] to specifically assess the type, degree and distribution of inflammation in IBM. The inflammatory domain was augmented to include T-cells, T-cell subtypes, B-cells and macrophages. MHC Class I staining was expanded to include three patterns of labelling. The vascular, muscle fibre and connective tissue domains which are present in the JDM score tool were not included.

Figure 1 Protein aggregates and congophilic deposits in IBM

Stained cryostat sections, showing fibres, often in clusters, containing protein aggregates stained for p62 (A), TDP-43 (B), ubiquitin (C), αB-crystallin (D) and myotilin (E). Protein aggregates were present throughout fibres, and were observed in apparently normal fibres, vacuolated fibres and fibres surrounded by inflammatory infiltrates. In fibres containing TDP-43 aggregates, myonuclear TDP-43 staining was frequently reduced (B). Congophilic deposits were observed in vacuolated fibres using epifluorescence (F). Tissue sections were examined using both the rhodamine red and fluorescein isothiocyanate filters to exclude areas of auto-fluorescence (arrow). Combined fluorescent image is shown. Four patterns of immunoreactivity were observed in IBM and disease controls stained for p62 using IHC (G)(H)(I)(J). Pattern I (G) - strongly stained, discreet and clearly delineated, round or angular aggregates, variable in number and size within a muscle fibre but rarely filling it and predominantly located subsarcolemmal, but also perinuclear and adjacent to vacuoles. Pattern II (H) - large aggregates of variable staining intensity. Pattern III (I) - fine granular aggregates dispersed throughout the fibre. Pattern IV (J) - fine granules and wisps of p62 immunoreactivity set within weakly basophilic inclusions.

Scale bar represents 50 µm in A and D; 25 µm in B, C and E-J.

Figure 2 Percentage of muscle fibres containing protein aggregates and Griggs' pathological features

Box and whisker plot illustrating the percentage of muscle fibres containing pathological abnormalities contained in the Griggs criteria and protein aggregates in Griggs' pathologically-definite IBM. Fibres containing aggregates immunoreactive for p62 and αB-crystallin were more frequent than those containing the current diagnostic pathological features (red bars) ($p<0.05$). Protein aggregates recognised by all antibodies were found in a significantly larger number of fibres than partial invasion ($p<0.02$).

Figure 3 Percentage of fibres containing protein aggregates and COX-/SDH+ fibres in each group

Box and whisker plots illustrating the percentage of fibres in each diagnostic category containing p62 (A), TDP-43 (B), myotilin (C) and ubiquitin (D) immunoreactive aggregates, congophilic deposits (E) and COX-/SDH+ fibres (F). All protein aggregates were present in a greater percentage of fibres in IBM+RV than in IBM–RV. There was no difference in the percentage of COX-/SDH+ muscle fibres between these groups. IBM+RV biopsies had a greater percentage of fibres containing p62 (A) and TDP-43 (B) immunoreactive aggregates and COX-/SDH+ fibres (F) than PAM. Pathological findings were similar in IBM-RV and PM&DM, with no differences in the percentage of fibres containing p62 (A), TDP-43 (B) and ubiquitin (D) immunoreactive aggregates or congophilic deposits (E). However, there was a greater percentage of COX-/SDH+ fibres (F) in IBM–RV than PM&DM and a greater percentage of fibres containing myotilin immunoreactive aggregates (C) in PM&DM than IBM-RV. *Statistically significant results.

Supplementary Figure 3 Sensitivity and specificity of rimmed vacuoles, protein aggregates and mitochondrial changes in IBM+RV compared to PAM

Receiver operating characteristic curves for each test including the area under the curve and optimum cut-off with its associated sensitivity and specificity for rimmed vacuoles (A), myotilin (B), ubiquitin (C), TDP-43 (D), p62 (E) immunoreactive deposits, congophilic deposits (F) and COX-/SDH+ fibres

(G). COX/SDH HC staining was the most discriminative test for differentiating IBM+RV and PAM (G). However, there was little difference between COX/SDH HC staining, TDP-43 and p62 IHC staining and none were sufficiently discriminative to be considered diagnostic. AUC = Area under the curve.

Supplementary Figure 4 Sensitivity and specificity of protein aggregates and mitochondrial changes in IBM-RV compared to PM&DM

Receiver operating characteristic curves for each test showing the area under the curve and optimum cut-off with its sensitivity and specificity for myotilin (A), ubiquitin (B), TDP-43 (C), p62 (D) immunoreactive deposits, congophilic deposits (E) and COX-/SDH+ fibres (F). COX/SDH histochemical staining (F) and myotilin (G) IHC were the most discriminative tests for differentiating IBM-RV and PM&DM. AUC = Area under the curve.

Figure 4 Proposed diagnostic algorithm for IBM based on pathological findings

Flow diagram showing a proposed pathway for diagnosing IBM based on the pathological findings. Increased MHC Class I staining was observed in all cases of IBM and pattern I p62 aggregates in all cases of IBM+RV making them good initial screening tests. Their absence rules-out a diagnosis of IBM and IBM+RV respectively. The presence of endomysial CD3+ T-cell score >1, endomysial CD8+ T-cell score >0 or strong MHC Class I staining in a biopsy with rimmed vacuoles and p62 aggregates secures a diagnosis of IBM+RV. Differentiating IBM-RV and PM&DM pathologically is more challenging. The presence of COX-/SDH+ fibres is not specific to IBM-RV; although COX-/SDH+ fibres were not present in every case of IBM-RV their absence casts doubt on the diagnosis of IBM-RV. Pattern I p62 aggregates may enable IBM to be differentiated from PM when present. However, they may lack sensitivity for IBM-RV, therefore their absence does not rule out the diagnosis.

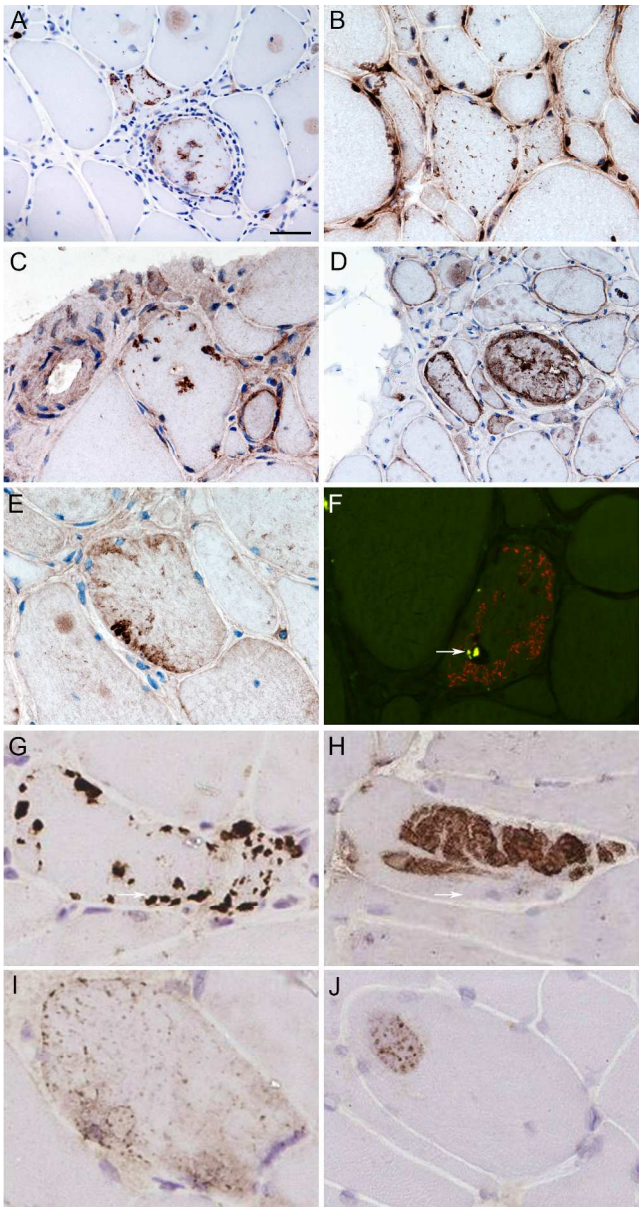


Figure 1 Protein aggregates and congophilic deposits in IBM
Stained cryostat sections, showing fibres, often in clusters, containing protein aggregates stained for p62 (A), TDP-43 (B), ubiquitin (C), αB-crystallin (D) and myotilin (E). Protein aggregates were present throughout fibres, and were observed in apparently normal fibres, vacuolated fibres and fibres surrounded by inflammatory infiltrates. In fibres containing TDP-43 aggregates, myonuclear TDP-43 staining was frequently reduced (B). Congophilic deposits were observed in vacuolated fibres using epifluorescence (F). Tissue sections were examined using both the rhodamine red and fluorescein isothiocyanate filters to exclude areas of auto-fluorescence (arrow). Combined fluorescent image is shown. Four patterns of immunoreactivity were observed in IBM and disease controls stained for p62 using IHC (G)(H)(I)(J). Pattern I (G) - strongly stained, discreet and clearly delineated, round or angular aggregates, variable in number and size within a muscle fibre but rarely filling it and predominantly located subsarcolemmal, but also perinuclear and adjacent to vacuoles. Pattern II (H) - large aggregates of variable staining intensity. Pattern III (I) - fine granular aggregates dispersed throughout the fibre. Pattern IV (J) - fine granules and wisps of

p62 immunoreactivity set within weakly basophilic inclusions.
Scale bar represents 50 μm in A and D; 25 μm in B, C and E-J.

161x305mm (300 x 300 DPI)

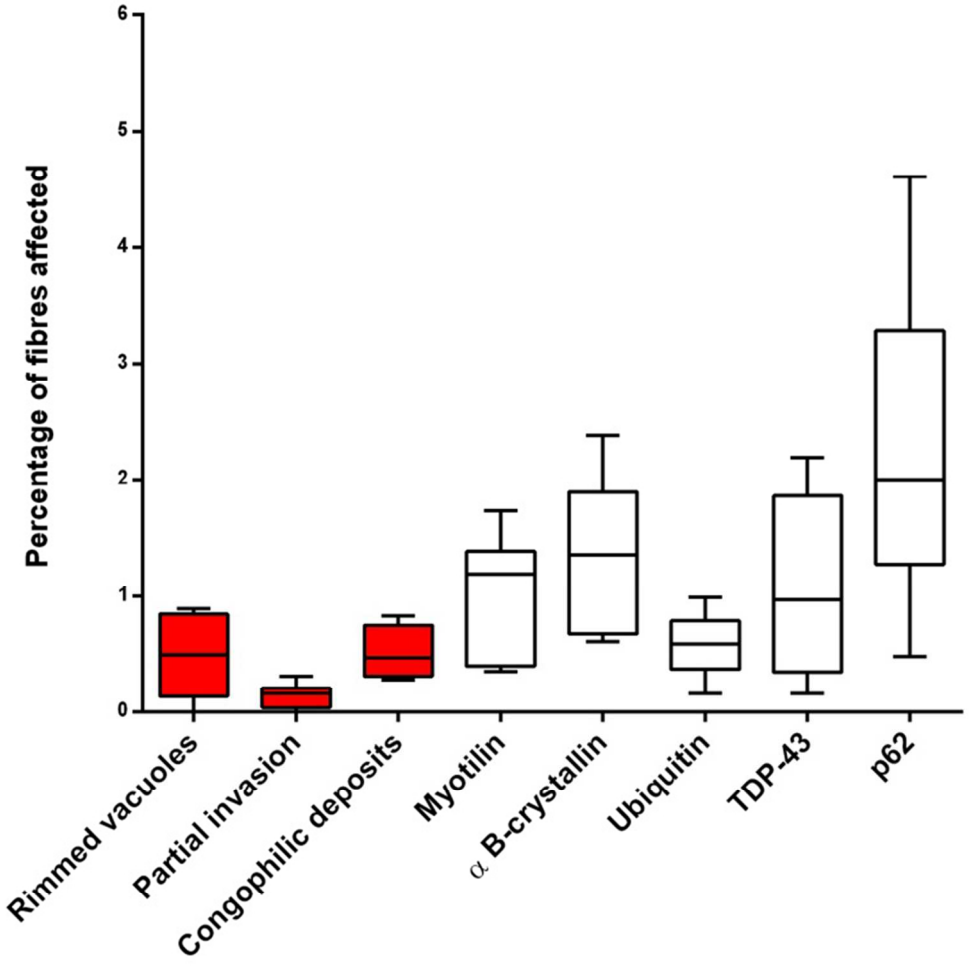


Figure 2 Percentage of muscle fibres containing protein aggregates and Griggs' pathological features
Box and whisker plot illustrating the percentage of muscle fibres containing pathological abnormalities contained in the Griggs criteria and protein aggregates in Griggs' pathologically-definite IBM. Fibres containing aggregates immunoreactive for p62 and αB-crystallin were more frequent than those containing the current diagnostic pathological features (red bars) ($p<0.05$). Protein aggregates recognised by all antibodies were found in a significantly larger number of fibres than partial invasion ($p<0.02$).

97x99mm (300 x 300 DPI)

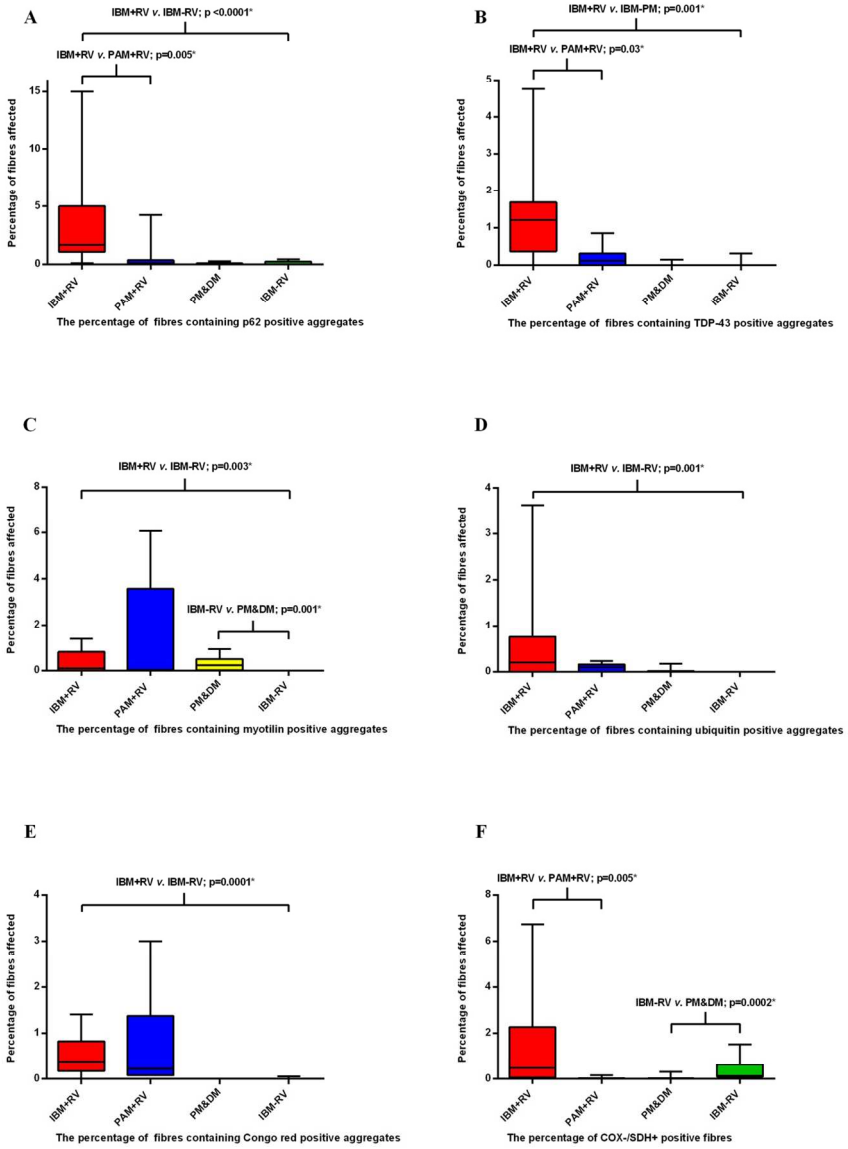


Figure 3 Percentage of fibres containing protein aggregates and COX-/SDH+ fibres in each group
Box and whisker plots illustrating the percentage of fibres in each diagnostic category containing p62 (A), TDP-43 (B), myotilin (C) and ubiquitin (D) immunoreactive aggregates, congophilic deposits (E) and COX-/SDH+ fibres (F). All protein aggregates were present in a greater percentage of fibres in IBM+RV than in IBM–RV. There was no difference in the percentage of COX-/SDH+ muscle fibres between these groups. IBM+RV biopsies had a greater percentage of fibres containing p62 (A) and TDP-43 (B) immunoreactive aggregates and COX-/SDH+ fibres (F) than PAM. Pathological findings were similar in IBM-RV and PM&DM, with no differences in the percentage of fibres containing p62 (A), TDP-43 (B) and ubiquitin (D) immunoreactive aggregates or congophilic deposits (E). However, there was a greater percentage of COX-/SDH+ fibres (F) in IBM–RV than PM&DM and a greater percentage of fibres containing myotilin immunoreactive aggregates (C) in PM&DM than IBM-RV. *Statistically significant results.

168x226mm (300 x 300 DPI)

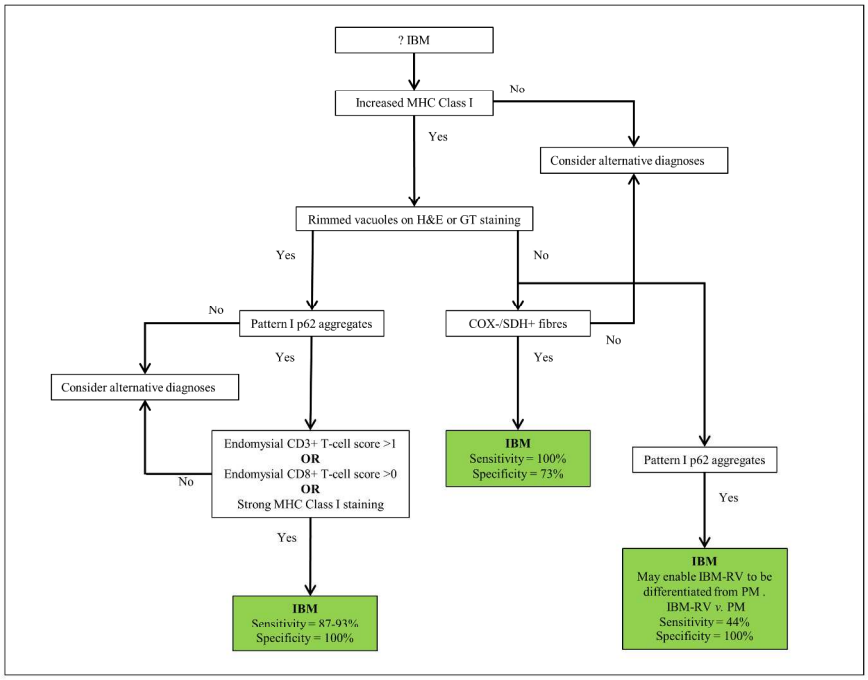


Figure 4 Proposed diagnostic algorithm for IBM based on pathological findings
Flow diagram showing a proposed pathway for diagnosing IBM based on the pathological findings. Increased MHC Class I staining was observed in all cases of IBM and pattern I p62 aggregates in all cases of IBM+RV making them good initial screening tests. Their absence rules-out a diagnosis of IBM and IBM+RV respectively. The presence of endomysial CD3+ T-cell score >1, endomysial CD8+ T-cell score >0 or strong MHC Class I staining in a biopsy with rimmed vacuoles and p62 aggregates secures a diagnosis of IBM+RV. Differentiating IBM-RV and PM&DM pathologically is more challenging. The presence of COX-/SDH+ fibres is not specific to IBM-RV; although COX-/SDH+ fibres were not present in every case of IBM-RV their absence casts doubt on the diagnosis of IBM-RV. Pattern I p62 aggregates may enable IBM to be differentiated from PM when present. However, they may lack sensitivity for IBM-RV, therefore their absence does not rule out the diagnosis.

254x190mm (300 x 300 DPI)

Supplementary Table 1 Clinical characteristics

Characteristic	G-IBM	IBM+RV	IBM-RV	PM&DM	PAM	IBM+RV*	IBM-RV*
Number of cases	6	15	9	11	7	12	11
Male:female	5:1	10:5	4:5	4:3	4:7	10:2	9:2
Median age at symptom onset, years (IQR)	69 (66-70)	54 (49-67)	62 (48-68)	55 (34-65)	46 (24-54)	58 (55-73)	60 (57-72)
Median age at muscle biopsy, years (IQR)	77 (68-78)	64 (59-71)	68 (47-74)	55 (34-65)	54 (29-59)	66 (62-77)	70 (63-74)
Median duration of symptoms, years (IQR)	5 (3-9)	5 (4-7)	3 (2-8)	0 (0-0)	5 (3-9)	5 (4-7)	4 (3-7)
Mean creatine kinase, IU/L, mean (±SD)	377 (±213)	1748 (±1348)	926 (±800)	6744 (±5875)	739 (±320)	662 (±360)	466 (±338)
Mean number of muscle fibres per biopsy	2929 (±1357)	1463 (±954)	1795 (±990)	3534 (±1934)	2749 (±1357)	NA	NA

G-IBM = Griggs' pathologically-definite IBM; IQR = Interquartile range; SD = Standard deviation; NA = Not applicable. *Cases used to test proposed diagnostic flow-chart.

Supplementary Table 2 Clinical inclusion criteria

Diagnostic category	Criteria
G-IBM	Patients fulfilling Griggs' definite criteria (rimmed vacuoles, inflammatory infiltrate with partial invasion of fibres and 15-18 nm tubulofilaments on EM) with prominent finger flexor and knee extensor weakness and CK <12 x ULN.
IBM+RV	Age at symptom onset >45 years, symptoms present for >12 months, finger flexion strength less than shoulder abduction strength and knee extension weakness greater than hip flexion weakness, CK ≤15 x ULN and a muscle biopsy revealing rimmed vacuoles on H&E or GT stained sections without features inconsistent with IBM on a standard diagnostic histological assessment for an inflammatory myopathy*.
IBM-RV	Clinical features and CK as detailed under IBM+RV. Rimmed vacuoles absent on H&E and GT stained sections and without features inconsistent with IBM on a standard diagnostic histological assessment for an inflammatory myopathy*.
PAM	Genetically or clinically and pathologically confirmed cases of PAM with typical rimmed vacuoles present on muscle biopsy and a genetically confirmed dystrophinopathy with typical rimmed vacuoles and protein aggregates present on muscle biopsy. Cases included myotilinopathy (n=2), hIBM with compound heterozygous mutations in GNE (n=1), IBMPFD with mutation in VCP (n=1), genetically unconfirmed cases of myofibrillar myopathy (n=2), and dystrophinopathy with deletion of exons 45-47 (n=1).
PM&DM	Subacute onset of limb girdle weakness, significantly raised CK, inflammatory cell infiltrate present on muscle biopsy and a sustained unequivocal clinical and biochemical response to steroid immunosuppression. DM cases also had to have cutaneous manifestations consistent with the diagnosis.
Normal controls	Patients investigated for cramps or fatigue, normal clinical examination performed by a muscle specialist, normal CK, normal neurophysiological assessment and normal muscle biopsy.

G-IBM = Griggs' pathologically-definite IBM; IBM+RV = Clinically-typical IBM with rimmed vacuoles; IBM-RV = Clinically typical IBM lacking rimmed vacuoles; PAM = Protein accumulation myopathies with rimmed vacuoles; PM&DM = Steroid-responsive inflammatory myopathies; hIBM = Hereditary inclusion body myopathy; IBMPFD = Inclusion body myopathy with Paget's disease and frontotemporal dementia; CK = Creatine kinase; GT = Gomori trichrome; ULN = Upper limit of normal. * Standard histological assessment for inflammatory myopathy includes H&E, GT, Sudan black or oil red O, periodic acid Schiff, nicotinamide adenine dinucleotide dehydrogenase, succinate dehydrogenase, cytochrome c oxidase, combined cytochrome c oxidase and succinate dehydrogenase, phosphorylase, acid and alkaline phosphatase, adenylate deaminase, ATPases at pH 4.2/4.3/9.4 and immunohistochemical staining including neonatal myosin, utrophin, major histocompatibility complex class I, membrane attack complex and a combination of inflammatory cell markers.

Supplementary Table 3 Antibodies and optimum staining conditions

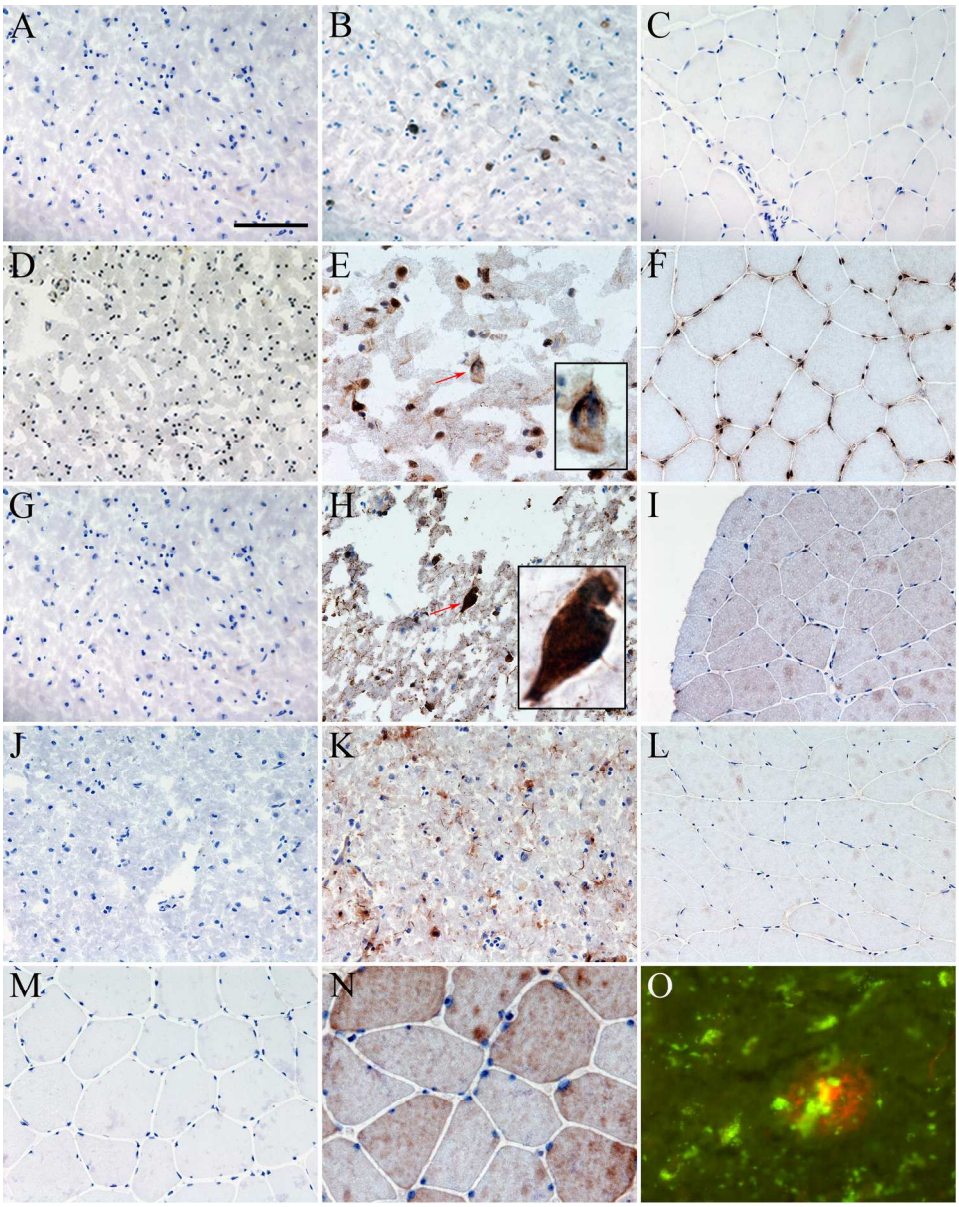
Antibody	Source	Clone	Control tissue	Fixative	Dilution	Primary incubation conditions†
p62	BD Transduction	3/P62	AD brain	A	1:400	1 hour, RT
TDP-43	Proteintech	NA	FTLD-TDP brain	PFA	1:800	24 hours, 4°C
Tau*	Dako	NA	AD brain	A	1:1600	1 hour, RT
Phosphorylated tau**	Autogen Bioclear	AT8	AD brain	A	1:1600	1 hour, RT
Ubiquitin	Dako	NA	AD brain	A	1:100	1 hour, RT
Aβ	Dako	6F/3D	AD brain	PFA and FA	1:100	1 hour, RT
α-synuclein	Abcam	4D6	MSA brain	PBS	1:800	1 hour, RT
FUS	Novus Biologicals	NA	FTLD-FUS brain	A	1:2000	1 hour, RT
Desmin	Dako	D33	Normal muscle	A	1:50	24 hours, 4°C
Myotilin	Novocastra	RSO34	Normal muscle	A	1:500	24 hours, 4°C
αB-crystallin	Novocastra	G2JF	CBD brain	A	1:300	1 hour, RT
VCP	Abcam	5	Normal muscle	A	1:100	1 hour, RT
Lamin A/C	Novocastra	636	Normal muscle	A	1:50	1 hour, RT
Emerin	Novocastra	4G5	Normal muscle	A	1:400	1 hour, RT
MHC Class I	Novocastra	W6/32	Normal muscle	A	1:25	24 hours, 4°C
CD3 (T-cells)	Novocastra	UCHT1	Tonsil	A	1:100	1 hour, RT
CD4 (Helper T-cells)	Novocastra	4B12	Tonsil	A	1:400	1 hour, RT
CD8 (Cytotoxic T-cells)	Novocastra	1A5	Tonsil	A	1:50	1 hour, RT
CD20 (B-cells)	Novocastra	L26	Tonsil	A	1:400	1 hour, RT
CD68 (Macrophages)	Novocastra	KP1	Tonsil	A	1:1600	1 hour, RT

NA = Not applicable; AD = Alzheimer's disease; FTLD-TDP = Frontotemporal lobar degeneration with TDP-43 positive inclusions; MSA = Multiple system atrophy; FTLD-FUS = Frontotemporal lobar degeneration with FUS positive inclusions; CBD = Corticobasal degeneration; A = Acetone; PFA = 4% Paraformaldehyde; FA = Formic acid; PBS = Phosphate buffered saline; RT = Room temperature. Antibodies were directed at * amino acids 243-441 irrespective of phosphorylation and ** phosphorylated Ser202. †Primary antibodies were made up in PBS and primary antibody-antigen binding was visualised with Dako REAL™ EnVision™ Detection System which includes a horseradish-peroxidase labelled goat anti-rabbit/mouse secondary and 1:50 solution of 3,3'-diaminobenzidine as the chromagen.

Supplementary Table 4 Definitions of pathological features

Pathological feature	Definition
Rimmed vacuoles	Irregular vacuole with a granular basophilic rim or containing granular basophilic material when stained with H&E or stained red in the GT. Both H&E and GT stained sections were reviewed before concluding the absence of rimmed vacuoles.
Inflammatory infiltrate and partial invasion	Inflammatory cells must show a nucleus fully circumscribed by a ring of positive staining. T-cells and B-cells must have a lymphoid morphology. Partial invasion was defined as unequivocal invasion of an otherwise structurally normal fibre by one or more inflammatory cells on H&E stained sections or sections stained using IHC.
Protein aggregates	Area of definite staining within a transversely orientated muscle fibre. Diffuse staining affecting the whole of a fibre was not counted nor were protein aggregates in necrotic fibres or regenerating fibres.
Congophilic deposits	Assessed using polarising and fluorescence microscopes. Positive staining using a polarising microscope was defined as congophilic deposits within a muscle fibre that exhibited apple-green birefringence under polarised light. Positive staining with a fluorescence microscope was defined as fluorescent material within a muscle fibre only visible under the rhodamine red filter. Areas of auto-fluorescence were excluded by visualising areas of fluorescence with both rhodamine red and FITC filters.

GT = Gomori trichrome; FITC = Fluorescein isothiocyanate.



Supplementary Figure 1 Control staining in brain and muscle tissue

Positive and negative (no primary) brain control sections and normal muscle stained using immunohistochemistry for: p62 (A-C), TDP-43 (D-F), α B-crystallin (G-I), ubiquitin (J-K) and myotilin (M,N) and alkalinised congo red (O).

(A-C) Negative (A) and positive (B) control sections of AD brain and normal muscle (C) stained for p62. Positive control shows p62 positive neurofibrillary tangles and dystrophic neurites (B). No p62 immunoreactivity is observed in normal muscle (C).

(D-F) Negative (D) and positive (E) control sections of FTLD-TDP brain and normal muscle (F) stained for TDP-43. Positive control shows normal nuclear labelling and mislocalised neuronal cytoplasmic staining with neuropil threads (E). Insert shows a neuron with absent nuclear TDP-43 and a cytoplasmic TDP-43 inclusion (E, red arrow and x100 insert). Nuclear TDP-43 staining is observed in normal muscle.

(G-I) Negative (G) and positive (H) control sections of CBD brain and normal muscle (I) stained for α B-crystallin. Positive control shows neuropil threads and a balloon cell neuron (H; red arrow and x100 insert).

No α B-crystallin immunoreactivity is observed in normal muscle (I).
(J-L) Negative (J) and positive (K) control AD brain and normal muscle (L) stained for ubiquitin. Positive control shows dystrophic neurites and neuropil threads (K). No ubiquitin immunoreactivity is observed in normal muscle (L).
(M,N) Negative (M) and positive (N) control muscle stained for myotilin. Mild sarcoplasmic staining is observed in normal muscle (N).
(O) Positive control section of AD brain showing an amyloid plaque (O).
Scale bar represents 100 μm in A-D, F and H-M; and 50 μm in E, N-O.
p62 = Sequestosome 1; AD = Alzheimer's disease; TDP-43 = Transactivation response DNA binding protein 43; FTLD-TDP = Frontotemporal lobar degeneration with TDP-43 positive inclusions; CBD = Corticobasal degeneration.

171x215mm (300 x 300 DPI)

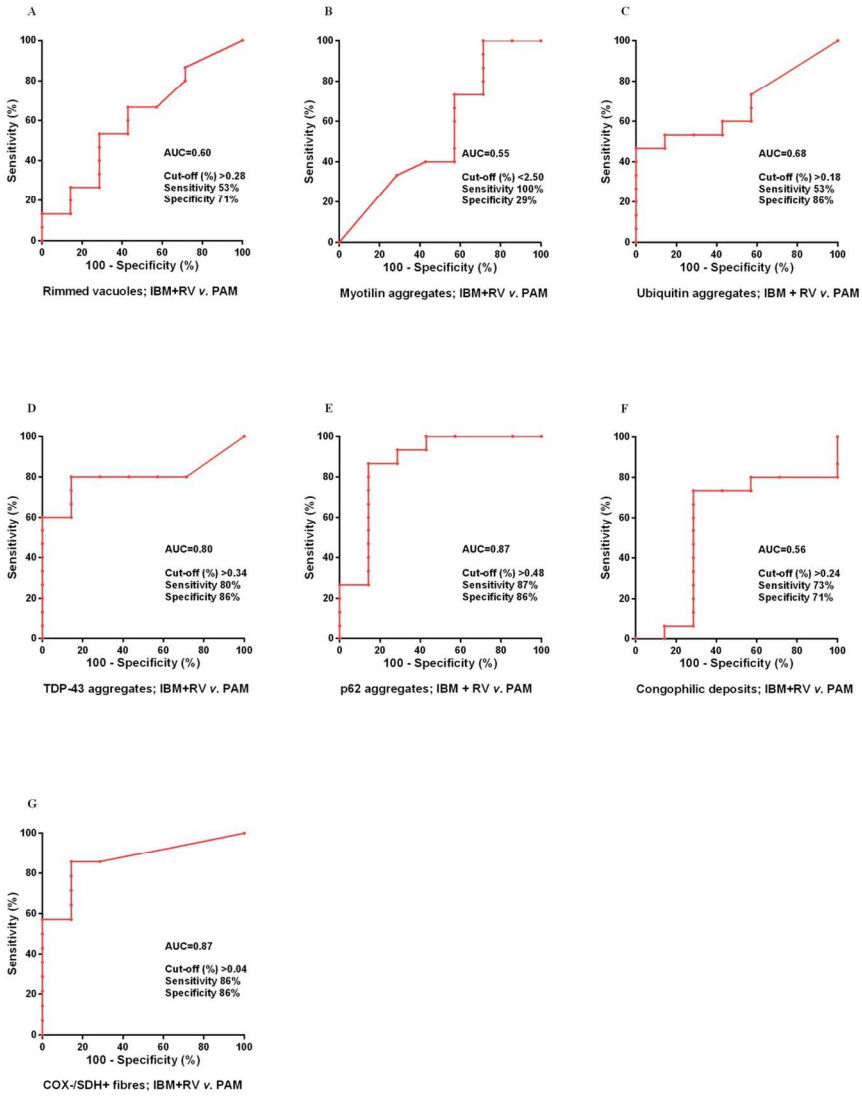
IBM Inflammatory Score Tool

Case Number:_____ Date:_____

	Score	Description
T-cells (CD3)		For each inflammatory cell type in the endomysial, perimysial and perivascular locations score positive infiltrating cells as follows: if none or <4 cells in a x20 field- score 0; if >4 cells in a x20 field and/or 1 cluster (where a cluster is ≥10 cells) - score 1; if >2 clusters in the entire biopsy, and/or diffusely infiltrating cells (i.e.> 20 cells in a x20 field) - score 2.
CD3+ endomysial infiltration	0, 1, 2	
CD3+ perimysial infiltration	0, 1, 2	
CD3+ perivascular infiltration	0, 1, 2	
Helper T-cells (CD4)		
CD4+ endomysial infiltration	0, 1, 2	
CD4+ perimysial infiltration	0, 1, 2	
CD4+ perivascular infiltration	0, 1, 2	
Cytotoxic T-cells (CD8)		
CD8+ endomysial infiltration	0, 1, 2	
CD8+ perimysial infiltration	0, 1, 2	
CD8+ perivascular infiltration	0, 1, 2	
B-cells (CD20)		
CD20+ endomysial infiltration	0, 1, 2	
CD20+ perimysial infiltration	0, 1, 2	
CD20+ perivascular infiltration	0, 1, 2	
Macrophages (CD68)		
CD68+ endomysial infiltration	0, 1, 2	
CD68+ perimysial infiltration	0, 1, 2	
CD68+ perivascular infiltration	0, 1, 2	
MHC Class I	0, 1, 2	For the whole biopsy score as follows: normal (capillary staining only) - score 0; if increased: i) mildly (weak diffuse sarcolemmal staining or scattered positive muscle fibres) - score 1; ii) strongly increased (diffuse definite sarcoplasmic and sarcolemmal increase in staining) score 2.

Supplementary Figure 2 IBM inflammatory score-tool
Score tool modified from the published juvenile dermatomyositis inflammatory (JDM) score tool [17] to specifically assess the type, degree and distribution of inflammation in IBM. The inflammatory domain was augmented to include T-cells, T-cell subtypes, B-cells and macrophages. MHC Class I staining was expanded to include three patterns of labelling. The vascular, muscle fibre and connective tissue domains which are present in the JDM score tool were not included.

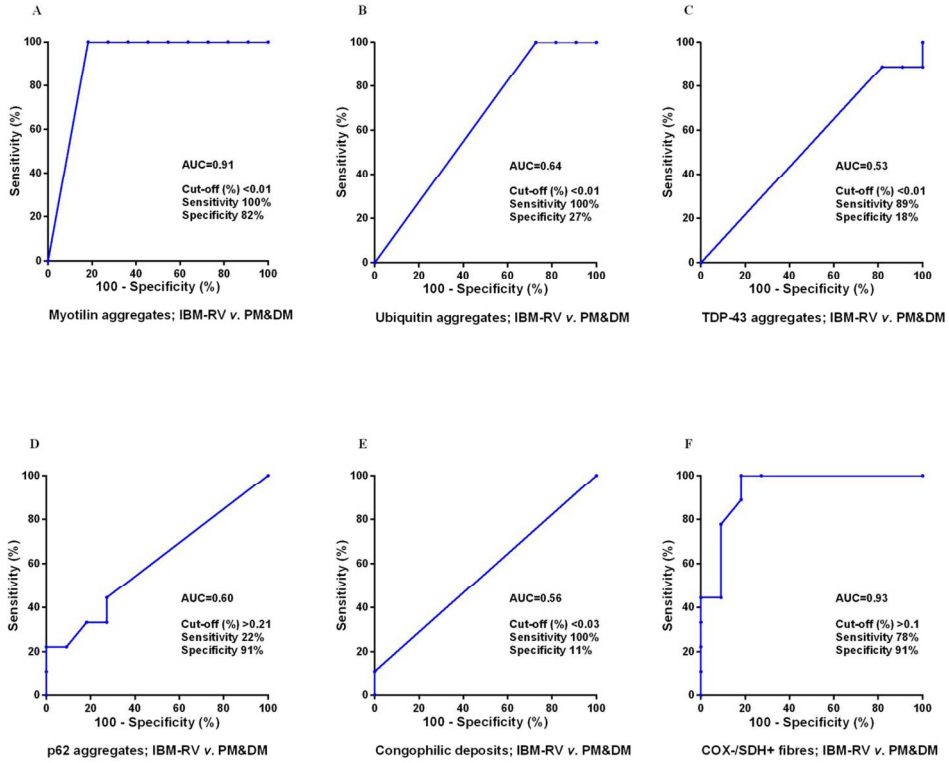
188x255mm (300 x 300 DPI)



Supplementary Figure 3 Sensitivity and specificity of rimmed vacuoles, protein aggregates and mitochondrial changes in IBM+RV compared to PAM

Receiver operating characteristic curves for each test including the area under the curve and optimum cut-off with its associated sensitivity and specificity for rimmed vacuoles (A), myotilin (B), ubiquitin (C), TDP-43 (D), p62 (E) immunoreactive deposits, congophilic deposits (F) and COX-/SDH+ fibres (G). COX/SDH HC staining was the most discriminative test for differentiating IBM+RV and PAM (G). However, there was little difference between COX/SDH HC staining, TDP-43 and p62 IHC staining and none were sufficiently discriminative to be considered diagnostic. AUC = Area under the curve.

165x211mm (300 x 300 DPI)



Supplementary Figure 4 Sensitivity and specificity of protein aggregates and mitochondrial changes in IBM-RV compared to PM&DM

Receiver operating characteristic curves for each test showing the area under the curve and optimum cut-off with its sensitivity and specificity for myotilin (A), ubiquitin (B), TDP-43 (C), p62 (D) immunoreactive deposits, congophilic deposits (E) and COX-/SDH+ fibres (F). COX/SDH histochemical staining (F) and myotilin (G) IHC were the most discriminative tests for differentiating IBM-RV and PM&DM. AUC = Area under the curve.

170x139mm (300 x 300 DPI)

STARD checklist for reporting of studies of diagnostic accuracy

(version January 2003)

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	Pg 1,2
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	Pg 2-4
METHODS			
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	Pg 4 and Supplementary Tables 1 and 2
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	Both. Pg 4 and Supplementary Table 2
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	Patients identified from clinics and systematic search of pathological databases
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	Retrospective study
Test methods	7	The reference standard and its rationale.	Clinical features and follow-up
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	Pg 4-6 and Supplementary Table 3
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	Pg 6
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	Two qualified medical doctors. Neuropathologist and Neurologist with an interest and significant experience in muscle pathology.
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	All analyses were blinded and performed in a random order. No clinical information was available at the time of analyses.
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	Pg 6 includes tests for determining diagnostic accuracy including 2x2 tables and ROC curves.
	13	Methods for calculating test reproducibility, if done.	Bland-Altman plots and Cohen's Kappa statistic used.
RESULTS			
Participants	14	When study was performed, including beginning and end dates of recruitment.	2011-2013
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	Included in Supplementary Table 1

	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	Not applicable. Retrospective study.
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	Study performed using tissue taken at the time of the reference standard
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	Diagnosese of control cases included Supplementary Table 2
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	Tables 1 and 2
	20	Any adverse events from performing the index tests or the reference standard.	Not applicable.
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	Included in Tables 1 and 2 and Supplementary Figures 2 and 3
	22	How indeterminate results, missing data and outliers of the index tests were handled.	Only one missing result and this is documented in Table 2. The denominator for calculating the proportion was altered to account for missing case in calculations
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	Included in statistical analysis Pg 6
	24	Estimates of test reproducibility, if done.	Included in statistical analysis Pg 6
DISCUSSION	25	Discuss the clinical applicability of the study findings.	Discussed in discussion Pg 12-15

BMJ Open

A retrospective cohort study identifying the principal pathological features useful in the diagnosis of inclusion body myositis

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-004552.R2
Article Type:	Research
Date Submitted by the Author:	17-Mar-2014
Complete List of Authors:	Brady, Stefen; MRC Centre for Neuromuscular Diseases, Division of Neuropathology Squier, Waney; University of Oxford, John Radcliffe Hospital, Department of Neuropathology Sewry, Caroline Hanna, Mike; MRC Centre for Neuromuscular Diseases, UCL Institute of Neurology and National Hospital for Neurology, Neurosurgery, Hilton-Jones, David; Nuffield Department of Clinical Neurosciences (Clinical Neurology), University of Oxford, Holton, Janice; UCL Institute of Neurology, Department of Molecular Neuroscience and MRC Centre for Neuromuscular Diseases
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Diagnostics
Keywords:	NEUROLOGY, Adult neurology < NEUROLOGY, Neuromuscular disease < NEUROLOGY, Neuropathology < PATHOLOGY

A retrospective cohort study identifying the principal pathological features useful in the diagnosis of inclusion body myositis

Corresponding author:

Dr Janice L Holton

Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

janice.holton@ucl.ac.uk; tel: 00 44 (0)20 3448 4239; fax: 00 44 (0)20 3448 4486.

Authors:

Stefen Brady[1], Waney Squier[2], Caroline Sewry[3,4], Michael Hanna[1], David Hilton-Jones[5], Janice L Holton[6]

[1]MRC Centre for Neuromuscular Diseases, UCL Institute of Neurology and National Hospital for Neurology, Neurosurgery, Queen Square, London, UK.

[2]Department of Neuropathology, University of Oxford, John Radcliffe Hospital, Oxford, UK.

[3]Dubowitz Neuromuscular Centre, Institute of Child Health and Great Ormond Street Hospital for Children, London, UK.

[4]Wolfson Centre of Inherited Neuromuscular Diseases, RJAH Orthopaedic Hospital, Oswestry, UK.

[5]Nuffield Department of Clinical Neurosciences (Clinical Neurology), University of Oxford, John Radcliffe Hospital, Oxford, UK.

[6]Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

Keywords: Neurology, adult neurology, neuromuscular disease, neuropathology

Running title: Pathological criteria for inclusion body myositis

Word count: 2990

ABSTRACT

Objectives

The current pathological diagnostic criteria for sporadic inclusion body myositis (IBM) lack sensitivity. Using immunohistochemical techniques abnormal protein aggregates have been identified in IBM, including some associated with neurodegenerative disorders. Our objective was to investigate the diagnostic utility of a number of markers of protein aggregates together with mitochondrial and inflammatory changes in IBM.

Design

Retrospective cohort study. The sensitivity of pathological features was evaluated in cases of Griggs' definite IBM. The diagnostic potential of the most reliable features was then assessed in clinically-typical IBM with rimmed vacuoles (*n*=15) and clinically-typical IBM without rimmed vacuoles (*n*=9) and IBM mimics - vacuolar myopathies (*n*=7) and steroid-responsive inflammatory myopathies (*n*=11).

Setting

Specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London.

Results

Individual pathological features, in isolation, lacked sensitivity and specificity. However, the morphology and distribution of p62 aggregates in IBM were characteristic and in a myopathy with rimmed vacuoles, the combination of characteristic p62 aggregates and increased sarcolemmal and internal MHC Class I expression or endomysial T-cells were diagnostic for IBM with a sensitivity of 93% and specificity of 100%. In an inflammatory myopathy lacking rimmed vacuoles, the presence of mitochondrial changes was 100% sensitive and 73% specific for IBM; characteristic p62 aggregates were specific (91%), but lacked sensitivity (44%).

Conclusions

We propose an easily applied diagnostic algorithm for the pathological diagnosis of IBM. Additionally our findings support the hypothesis that many of the pathological features considered

typical of IBM develop later in the disease, explaining their poor sensitivity at disease presentation and emphasising the need for revised pathological criteria to supplement the clinical criteria in the diagnosis of IBM.

STRENGTHS AND LIMITATIONS

The present study is a multicentre retrospective evaluation of the diagnostic utility of pathological findings for differentiating IBM from myopathies important in the differential diagnosis – myopathies containing rimmed vacuoles and steroid-responsive inflammatory myopathies.

The main strength of our study was the systematic detailed analysis of well-defined cases. This enabled us to determine the sensitivity and specificity of individual pathological features and produce an easily applied pathological diagnostic algorithm for IBM for use in clinical practice.

Study limitations include the small number of cases and the retrospective design. Further prospective studies are now required in larger cohorts of patients.

INTRODUCTION

Sporadic inclusion body myositis (IBM) is the commonest acquired myopathy in those aged over 50 years.[1] Although classified as an idiopathic inflammatory myopathy, muscle biopsy reveals both degenerative and inflammatory features. The widely used Griggs diagnostic criteria require the presence of several pathological findings,[2] namely rimmed vacuoles, an inflammatory infiltrate with invasion of non-necrotic fibres by mononuclear inflammatory cells (partial invasion), and either amyloid deposits or 15-18 nm tubulofilaments identified by electron microscopy (EM). Although these features in combination are highly specific for IBM, individually they occur in other myopathies, including some important in the differential diagnosis for IBM.[3-7] Moreover, cases of clinically-typical IBM have been reported where the combination of these pathological features is absent causing diagnostic difficulty.[8-11]

Over the last two decades, pathological accumulation of many different proteins has been reported in muscle fibres in IBM.[12] Proteins typically associated with neurodegenerative diseases such as β-amyloid (Aβ), hyperphosphorylated tau and ubiquitin and newer neurodegenerative markers such as p62 and transactivation response DNA binding protein-43 (TDP-43) have been identified, as well as proteins associated with myofibrillar myopathies (MFM), including desmin and αB-crystallin. However, not all observations have been consistently reproduced.[13,14] Mitochondrial changes have also been proposed for inclusion in IBM diagnostic criteria,[15]. Clear guidelines for the incorporation of immunohistochemical findings and mitochondrial changes into diagnostic criteria for IBM have not been established.[16]

Previously, we have shown that the characteristic pattern of weakness associated with IBM is indicative of the diagnosis, even if Griggs pathological features are absent.[11] However, it is not invariably found at presentation. Here we sought to identify which pathological features, other than the Griggs pathological criteria, add further support to the diagnosis of IBM. We systematically investigated which pathological features are present in Griggs pathologically-definite IBM and then

established the diagnostic utility of these features in cases of IBM lacking the Griggs criteria, using myopathies considered in the differential diagnosis of IBM as controls.

MATERIALS AND METHODS

The study received ethical approval from the Departments of Research and Development at Oxford University Hospitals NHS Trust, Oxford and University College London Hospitals NHS Foundation Trust, London.

Cases

All patients were followed by specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London. Biopsies were taken for diagnostic purposes from the deltoid or quadriceps muscles and prior to any treatment.

Methods for demonstrating pathological features in IBM, additional to those defined by the Griggs criteria, were determined in six Griggs pathologically-definite cases of IBM. Cases with no clinical or pathological evidence of neuromuscular disease were used as controls. The diagnostic utility of the pathological features identified was assessed in two groups of clinically-typical IBM; one with rimmed vacuoles on muscle biopsy (IBM+RV; *n*=15), the other without rimmed vacuoles on muscle biopsy (IBM-RV; *n*=9). Disease controls were cases of steroid-responsive inflammatory myopathies [polymyositis and dermatomyositis; (PM&DM); *n*=11] and protein accumulation myopathies with rimmed vacuoles (PAM; *n*=7). Clinical characteristics and inclusion criteria are summarised in Supplementary tables 1 and 2. Tissue from brains donated to the Queen Square Brain Bank for Neurological Disorders was used as positive controls for protein aggregate staining.

Muscle biopsies

Muscles biopsies were snap frozen at the time of surgery in isopentane cooled liquid nitrogen. Until sectioning all samples were stored at -80°C. Serial tissue sections were cut to a thickness of 8 µm, allowed to air dry and stored at -80°C until staining. Prior to staining, tissue sections were allowed to dry at room temperature. Tissue sections were stained with haematoxylin and eosin (H&E), combined cytochrome oxidase (COX) succinate dehydrogenase (SDH) histochemistry and for amyloid using alkalinised Congo red, crystal violet and thioflavin S. Tissue sections for immunohistochemical staining were fixed for 10 minutes, if required, washed for five minutes in running water and incubated in 0.5% hydrogen peroxide to block endogenous peroxidase for 20 minutes. After further washing, tissue sections were incubated in 5% normal goat serum (Vector Laboratories, Burlingame, California) for 30 minutes and then systematically stained for: 1) proteins classically associated with neurodegenerative disease: tau and hyperphosphorylated tau, ubiquitin, Aβ and α-synuclein; 2) proteins more recently reported in neurodegenerative disease: p62, TDP-43, fused in sarcoma protein (FUS) and valosin containing protein (VCP); 3) nuclear membrane proteins: lamin A/C and emerin; 4) proteins associated with MFM: desmin, myotilin and αB-crystallin; and 5) inflammatory cells and major histocompatibility complex class I (MHC Class I): CD3+ T-cells, CD4+ T-cells, CD8+ T-cells, B-cells and macrophages. Primary antibody binding was visualised using Dako REAL™ EnVision™ Detection System which contains horse-radish peroxidase (HRP) labelled goat anti-rabbit/mouse secondary and 3,3'-diaminobenzidine (DAB); following incubation with the relevant primary antibody, tissue sections were washed in phosphate buffered saline (PBS), incubated with HRP labelled goat anti-rabbit/mouse secondary for 30 minutes, washed in PBS and incubated in a 1:50 solution of DAB for three to five minutes. Details of commercial antibodies and conditions used are provided in Supplementary Table 3. IHC for each antibody was performed on all cases simultaneously and including positive and negative controls (Supplementary Figure 1).

Definitions and quantification

The total number of fibres and the number undergoing partial invasion, containing rimmed vacuoles, protein aggregates and COX-negative SDH-positive (COX-/SDH+) fibres were quantified using ImagePro version 6.2 (Media Cybernetics), to ensure that the whole biopsy was systematically analysed. Only transversely-orientated fibres not undergoing necrosis or regeneration were quantified. Tissue sections stained with Congo red were visualised under fluorescent and polarised light. Areas of fluorescence were examined using both rhodamine red (excitation 512-546 nm and emission 600-640 nm) and fluorescein isothiocyanate (excitation 440-480 nm and emission 527-530 nm) filters to exclude auto-fluorescence. Supplementary Table 4 provides definitions of the pathological features assessed. The inflammatory infiltrate and MHC Class I staining were analysed using a modified version of the semi-quantitative juvenile dermatomyositis score-tool (Supplementary Figure 2).[17] Assessments were performed blind to clinical details and diagnosis by a single individual (SB). Ten per cent of slides were re-counted to assess intra-observer reliability and 336 slides were assessed independently by two observers (SB and JLH) to determine inter-observer reliability.

Statistical analysis

Statistical analyses were performed using GraphPad PRISM version 5. Continuous and categorical variables were compared using Mann Whitney *U*-test and chi-squared or Fisher's exact test respectively. Spearman's rank order correlation was used to determine the strength and direction of associations between pathological findings. Linear regression was used to determine relationships between clinical features and pathological findings. Test characteristics were calculated using receiver operating characteristic (ROC) curves and 2x2 contingency tables. A test was considered diagnostic when sensitivity >75% and specificity >95% or sensitivity >95% and specificity >75%. Intra-observer and inter-observer agreement was calculated using Bland-Altman plots and Cohen's kappa statistic (κ). Repeat counts were within 95% confidence intervals using Bland-Altman plots and κ was ≥0.7 indicating good intra-observer and good or excellent inter-observer reliability. Statistical significance was set at $p<0.05$.

RESULTS

Pathological findings in Griggs' pathologically-definite IBM

p62, TDP-43, ubiquitin, myotilin and αB-crystallin immunoreactive aggregates were present in all six IBM cases but not in normal controls (Figures 1A-E). p62 and αB-crystallin immunoreactive aggregates were present in a greater percentage of fibres than the pathological features required in the Griggs criteria ($p<0.05$) (Figure 2). Despite their abundance, αB-crystallin immunoreactive aggregates were difficult to quantify due to a significant variability in their morphology. No immunoreactive deposits were observed in IBM cases or normal controls with antibodies to tau and phosphorylated tau, Aβ, α-synuclein, desmin, emerin, lamin A/C, FUS or VCP. Alkalinised Congo red staining was more sensitive than crystal violet and thioflavin S staining for observing amyloid aggregates (Figure 1F). Tissue sections containing congophilic deposits identified under fluorescence light showed no apple-green birefringence under polarised light. Mitochondrial changes and increased sarcolemmal and sarcoplasmic MHC Class I staining were observed in all six IBM cases, but not in normal controls. The inflammatory infiltrate was predominantly composed of endomysial CD8+ T-cells and macrophages, with relatively few B-cells.

Quantitative analysis of pathological features in IBM and disease controls

Having shown that p62, TDP-43, ubiquitin and myotilin aggregates, congophilic deposits, MHC Class I and inflammatory cells were prevalent in Griggs' pathologically-definite IBM, the presence of these abnormalities, together with mitochondrial changes were assessed in IBM+RV, IBM-RV and disease controls.

The percentage of fibres containing p62, TDP-43, myotilin and ubiquitin aggregates and congophilic deposits were greater in IBM+RV than in IBM-RV; there was no difference in the number of COX-/SDH+ fibres (Figure 3A-F). Protein aggregates were observed in morphologically-normal fibres and in fibres exhibiting Griggs' pathological features. p62 and TDP-43 positive aggregates were present in

a greater percentage of fibres in IBM+RV compared to PAM; however, there were no differences in the percentage of fibres containing myotilin and ubiquitin aggregates or congophilic deposits. The percentage of fibres containing p62, TDP-43 and ubiquitin aggregates or congophilic deposits were similar in IBM-RV and PM&DM; however, myotilin aggregates were present in a greater percentage of fibres in PM&DM and COX-/SDH+ fibres were more abundant in IBM-RV. Analysis of the total inflammatory infiltrate (the sum of the semi-quantitative scores for T-cells, B-cells and macrophages) in the endomysium, perimysium and perivascular areas revealed that there were greater numbers of inflammatory cells in the endomysium and perimysium in IBM+RV than in PAM (p<0.03). The same analysis comparing the sum of the inflammatory cells in IBM-RV and PM&DM revealed that the distribution and intensity of the inflammatory infiltrate was similar.

Diagnostic utility of pathological features in IBM and disease controls

To mimic the diagnostic difficulty encountered in clinical practice, the ability of each test to differentiate between myopathies containing rimmed vacuoles (IBM+RV and PAM) and between inflammatory myopathies (IBM–RV and PM&DM) was assessed.

Diagnostic utility determined using receiver-operating characteristic curves

Individually, the presence of p62 immunoreactive inclusions and COX-/SDH+ fibres had the highest sensitivity and specificity for differentiating IBM+RV from PAM, (Supplementary Figure 3) (Table 1). Differentiating between IBM–RV and PM&DM, myotilin positive inclusions or COX-/SDH+ fibres had the highest sensitivity and specificity for IBM-RV (Supplementary Figure 4) (Table 1). Only the presence of myotilin positive inclusions satisfied criteria to be considered suitable as a diagnostic test (<0.01% of fibres containing myotilin aggregates had a sensitivity of 100% and specificity of 82% for IBM-RV).

Table 1 Test characteristics

Table shows the area under the curve and optimum cut-off for each test with the accompanying sensitivity and specificity. AUC = Area under the curve.

Test feature	IBM+RV *v.* PAM				IBM-RV *v.* PM&DM			
	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity
Rimmed vacuoles	0.60	>0.28	0.53	0.71	-	-	-	-
p62 aggregates	0.87	>0.48	0.87	0.86	0.60	>0.21	0.22	0.91
TDP-43 aggregates	0.80	>0.34	0.80	0.86	0.53	<0.01	0.89	0.18
Ubiquitin aggregates	0.68	>0.18	0.53	0.85	0.64	<0.01	1.00	0.27
Myotilin aggregates	0.55	<0.25	1.00	0.29	0.91	<0.01	1.00	0.82
Congophilic deposits	0.56	>0.24	0.73	0.71	0.56	<0.03	0.11	0.82
COX-/SDH+ fibres	0.87	>0.04	0.86	0.86	0.93	>0.1	0.78	0.91

Diagnostic utility determined by comparing proportion of affected cases in each diagnostic group

In the aforementioned experiments, the number of fibres within each muscle biopsy was quantified. However, this is impractical for routine clinical use. Thus, the proportions of affected cases in each group were compared (Table 2). This revealed that neither staining for protein aggregates nor congophilic deposits could differentiate between IBM+RV and PAM. The pathological findings in IBM-RV and PM&DM were also similar, except that the absence of myotilin immunoreactive aggregates was sensitive and specific for IBM-RV. COX-/SDH+ fibres were also suggestive of IBM-RV; one or more COX-/SDH+ fibres had a sensitivity of 100% and specificity 73% for IBM-RV.

Increased MHC Class I expression lacked specificity. However, strong (diffuse sarcolemmal and sarcoplasmic) MHC Class I up-regulation was diagnostic for IBM+RV, differentiating it from PAM, as were the presence of either endomysial CD3+ T-cell or CD4+ T-cell scores >1 or an endomysial CD8+ T-cell score >0. Partial invasion was specific for IBM+RV, but lacked sensitivity. Although the sum of the inflammatory infiltrate was similar in IBM-RV and PM&DM, analysis of the inflammatory cell sub-types revealed greater numbers of perimysial CD3+ T-cells, CD8+ T-cells and endomysial B-cells [were observed] in PM&DM than in IBM–RV ($p \leq 0.02$), however, this was not diagnostically useful. There was no difference in the proportion of cases with fibres undergoing partial invasion between IBM–RV and PM&DM.

Table 2 Comparison of the proportion of positive cases in each group

Pathological features	IBM+RV	PAM	IBM+RV *v.* PAM		IBM–RV	PM&DM	IBM-RV *v.* PM&DM		IBM+RV *v.* IBM-RV
	n (%)	n (%)	Sensitivity	Specificity	n (%)	n (%)	Sensitivity	Specificity	*p* value
Number of cases	15 (100)	7 (100)			9 (100)	11 (100)			
Aggregated proteins, n (%)									
p62	15 (100)	6 (86)	1.00	0.14	4 (44)	3 (27)‡	0.40	0.72	0.003*
TDP-43	13 (87)	5 (71)	0.87	0.29	1 (11)	2 (18)‡	0.11	0.82	0.001*
Ubiquitin	11 (73)	4 (57)	0.73	0.43	0 (0)	3 (27)‡	0.00	0.73	0.001*
Myotilin	10 (67)	5 (71)	0.67	0.29	0 (0)	9 (82)	0.00	0.18	0.002*
Congophilic deposits	13 (87)	7 (100)	0.87	0.00	1 (11)	0 (0)	0.11	1.00	0.001*
COX-/SDH+ fibres†, n (%)									
Any	12 (86)	2 (29)	0.80	0.71	9 (100)	3 (27)	1.00	0.73	0.5
Inflammatory features, n (%)									
MHC Class I up-regulation	15 (100)	3 (43)	1.00	0.57	9 (100)	11 (100)	1.00	0.00	1.00
Strong MHC Class I up-regulation	14 (93)	0 (0)	0.93	1.00	9 (100)	10 (91)	1.00	0.09	0.53
Partial invasion	10 (67)	0 (0)	0.67	1.00	3 (33)	2 (18)	0.33	0.82	0.11
Endomysial CD3+ T-cell score >1	13 (87)	0 (0)	0.87	1.00	4 (44)	7 (64)	0.44	0.36	0.02*
Endomysial CD4+ T-cell score >1	12 (80)	0 (0)	0.80	1.00	2 (22)	5 (45)	0.22	0.46	0.01*
Endomysial CD8+ T-cell score >0	14 (93)	0 (0)	0.93	1.00	4 (44)	5 (45)	0.44	0.54	0.02*
Endomysial CD68+ macrophage score >1	12 (80)	0 (0)	0.80	1.00	4 (44)	8 (73)	0.44	0.17	0.07

†In IBM with rimmed vacuoles *n*=14. ‡Pathological features present in DM, but not PM cases. *Statistically significant results.

Because IBM-RV is more pathologically akin to PM than DM, analyses were repeated comparing IBM-RV and PM cases (*n*=6). No p62, TDP-43 or ubiquitin immunoreactive aggregates were observed in PM cases and the diagnostic utility of tests for differentiating between IBM-RV and PM yielded similar results to prior analyses between IBM-RV and PM&DM.

Diagnostic utility of categorising the pattern of p62 staining

The pattern of p62 staining could be categorised into four distinct groups (Figure 1G-J). Aggregates observed in IBM were present in vacuolated and non-vacuolated fibres and were strongly stained, discreet and clearly delineated, round or angular and typically located subsarcolemmal, perinuclear and peri-vacuolar (pattern I). This pattern was observed in every IBM case with p62 aggregates, one (9%) case of DM and three (43%) cases of PAM (hereditary IBM, dystrophinopathy and genetically undefined MFM). Defining the pattern of immunoreactivity increased the discriminative value of p62 IHC for differentiating IBM+RV from PAM; pattern I p62 aggregates compared to any p62 aggregates increased the specificity from 14% to 57%, with no loss of sensitivity. Differentiating IBM-RV and PM&DM, pattern I p62 aggregates were highly specific (91%), but lacked sensitivity (44%). Patterns II, III and IV were not observed in any IBM cases. Patterns II and III appeared to be specific for PAM (*n*=2; 26%), both were cases of myotilinopathy (*n*=2; 67%), and DM (*n*=2; 40%) respectively. Pattern IV occurred in a genetically undefined case of MFM. No differences were observed in the morphology of TDP-43, myotilin or ubiquitin aggregates between biopsies.

Clinicopathological correlation

In IBM+RV, IBM-RV and pathologically-definite IBM, there were no correlations in individual biopsies between pathological features. No relationships were identified between the pathological findings and age at symptom onset, age at biopsy, disease duration or serum creatine kinase. The same results were obtained when the IBM groups were analysed separately and as one.

Proposed diagnostic algorithm

Based on our pathological findings, we propose a diagnostic algorithm for differentiating IBM from its disease mimics (Figure 4).

The algorithm was tested in a further 23 cases that fulfilled the criteria for IBM+RV (*n*=12) and IBM-RV (*n*=11). The algorithm correctly diagnosed 20 (87%) cases: 12 (100%) cases of IBM+RV and eight (73%) cases of IBM-RV. In IBM-RV, COX-/SDH+ fibres were present in 8 (73%) cases, pattern I p62 aggregates in 8 (73%) cases and both in 6 (55%) cases.

DISCUSSION

While Griggs' pathological criteria have been accepted as diagnostic of IBM, many patients who, observed over time undoubtedly have IBM, lack one or more of the Griggs pathological features at presentation, even on repeat biopsy.[8,11] Despite IBM being associated with a characteristic pattern of finger flexor and knee extensor weakness, not all patients have this pattern at disease onset, and muscle biopsy remains an important tool in differentiating IBM from its mimics. We sought to determine which additional pathological features support a diagnosis of IBM, demonstrating that characteristic p62 immunoreactive aggregates, strong MHC Class I upregulation, endomysial CD3+ T-cell score >1, CD8+ T-cell score >0 and COX-/SDH+ fibres are features with sufficient sensitivity and specificity to differentiate IBM from pathologically similar myopathies and we propose an easily applied pathological algorithm for the diagnosis of IBM (Figure 4).

In agreement with previous studies, we observed p62,[18] TDP-43,[19] ubiquitin [13] and αB-crystallin [20] immunoreactive aggregates and a predominantly endomysial inflammatory infiltrate [3] in Griggs pathologically-definite IBM. Diagnostic pathological studies of IBM have concentrated on differentiating IBM from other inflammatory myopathies and two recent quantitative studies have found that TDP-43 and markers of autophagy such as p62 and LC3 may be of diagnostic use.[21,22] However, in these studies only a fraction of each biopsy was analysed i.e. 200 fibres. We have found

this limited quantification does not correlate with the percentage of affected fibres in a biopsy nor does it reflect the way in which a muscle biopsy is assessed. Additionally, studies have lacked vacuolar myopathy control cases as it is believed that the inflammatory changes present in IBM enable it to be easily differentiated from other vacuolar myopathies.[22] However, inflammatory changes are frequently observed in muscular dystrophies and the degree of inflammatory change necessary to confidently diagnose IBM is currently unknown.

To mimic the typical diagnostic conundrums encountered in clinical practice, we evaluated the ability of the pathological findings to differentiate IBM+RV from other vacuolar myopathies and IBM-RV from steroid-responsive inflammatory myopathies. We found that quantitative analysis of protein aggregates, congophilic deposits and COX-/SDH+ fibres was of limited diagnostic use. Analysing the biopsies dichotomously and using a semi-quantitative score-tool revealed that increased MHC Class I labelling was sensitive for IBM making it a good initial screening test, its absence excluding the diagnosis. In agreement with an earlier study, we found p62 aggregates identified the largest number of affected fibres in IBM.[23] Additionally, as a novel finding, the morphology and distribution of p62 aggregates was characteristic in IBM. This characteristic pattern of p62 immunoreactive aggregates was highly sensitive for IBM+RV (100%); their absence from a biopsy containing rimmed vacuoles effectively ruling-out a diagnosis of IBM. We confirmed that the most diagnostically useful pathological findings in IBM+RV were evidence of an immune mediated process; strong MHC Class I staining, endomysial CD3+ T-cell score >1 or an endomysial CD8+ T-cell score >0 were diagnostic. Having identified either of these features in a biopsy containing rimmed vacuoles no extra diagnostic certainty was gained from observing partial invasion, COX-/SDH+ fibres or congophilic deposits.

The most discriminative pathological tests for differentiating between IBM–RV and PM&DM were COX/SDH staining and myotilin IHC. Consistent with a recent study,[9] we found the absence of mitochondrial changes casts doubt on a diagnosis of IBM. There was no difference in the median age between IBM-RV and PM&DM cases to account for the difference observed in COX-/SDH+ fibres.

The presence of myotilin and ubiquitin immunoreactive aggregates appeared to rule out a diagnosis of IBM-RV. However, we believe the presence of these features in IBM+RV indicates that they are unlikely to be diagnostically reliable features for differentiating between IBM-RV and steroid-responsive inflammatory myopathies. Although no pathological feature was able to differentiate IBM-RV from steroid responsive inflammatory myopathies with certainty the presence of characteristic p62 aggregates and the absence of COX-/SDH+ fibres may help in supporting and opposing a diagnosis of IBM-RV respectively. Pattern I p62 immunoreactive aggregates were only present in 44% of the initial IBM-RV cases tested, but they were not observed in PM cases and were very rare in DM. Although pattern I p62 aggregates appear to lack sensitivity their specificity was 91% making their presence highly suggestive of a diagnosis of IBM-RV. However, we identified pattern I p62 in eight out of 11 (73%) further cases of IBM-RV that were assessed indicating a greater sensitivity and that p62 IHC warrants further investigation and validation in a larger, independent series. The diagnostic utility of the other patterns of p62 staining is uncertain. Although pattern II appeared to have some specificity for myotilinopathy the small number of cases makes it drawing any conclusion problematic. In addition to p62 other autophagic proteins have been found in IBM and suggested as diagnostic markers. [22] Autophagy is a cellular mechanism for degrading and recycling cellular proteins and organelles and therefore, altered autophagy could lead to the accumulation of abnormal mitochondria and misfolded aggregation-prone proteins and may also result in altered antigen presentation leading to the widespread increase of MHC Class I and suggests that altered autophagy may play an important role in the pathogenesis of IBM.

Almost all pathological features - protein aggregates, congophilic deposits and inflammation - were more abundant in IBM+RV than IBM-RV. Despite using slightly different inclusion criteria, similar differences have been reported between pathologically-typical and pathologically-atypical IBM.[21] However, we found no differences in the number of COX-/SDH+ fibres, the degree of MHC Class I upregulation, the morphology and distribution of p62 immunoreactive aggregates or the pattern of the inflammation between IBM+RV and IBM-RV, supporting our clinical observations that these are the

same disease. We believe that the pathological differences between IBM+RV and IBM-RV are, in part, due to differences in disease duration. Two studies have shown that rimmed vacuoles are more common in patients who are older at the time of muscle biopsy,[24,11] suggesting that they are associated with chronologically more advanced disease. Therefore, the pathological findings which are more abundant in IBM+RV and thought to be typical of IBM may instead be indicative of chronologically more advanced disease explaining their limited sensitivity at disease presentation. However, possibly due to the number of cases analysed, we were unable to confirm a relationship between pathological features and clinical findings. It could be argued that biopsies from different muscles may have affected the pathological findings observed and differences between IBM groups. However, in a recent review of 59 muscle biopsies from IBM cases in our clinical archive with quadriceps (n=31) and deltoid (n=28) biopsies we found no significant difference in the frequency of pathological findings.

A robust clinicopathological definition of IBM is of paramount importance for diagnosis and for selection and entry of patients into clinical trials. We have shown that certain pathological findings are more abundant than those included in the current pathologically-focussed diagnostic criteria. Moreover, p62 immunoreactive deposits, increased MHC Class I expression, endomysial CD3+ T-cells and CD8+ T-cells and COX-/SDH+ fibres have sufficient sensitivity and specificity to aid in the histological differentiation of IBM from disease mimics, supporting their inclusion in future diagnostic criteria for IBM alongside clinical criteria. Both CD3+ T-cells and CD8+ T-cells are included in the diagnostic algorithm as there was little difference in their sensitivity and specificity for differentiating IBM+RV from PAM. However, IHC staining for CD3+ T-cells is likely to be more widely available and avoids the costs of extra staining to subtype the inflammatory infiltrate enabling diagnostic algorithm to be used by a greater number diagnostic laboratories. Using our diagnostic algorithm, we found there would be little additional diagnostic security in identifying partial invasion, performing EM or staining for amyloid deposits. Finally, mitochondrial changes and MHC Class I up-regulation were the most consistent findings in our IBM cases suggesting that they are central to the

pathogenesis and that further investigation and therapeutic intervention should be directed towards these features.

ACKNOWLEDGEMENTS

SB is funded by the Myositis Support Group. JLH is supported by the Reta Lila Weston Institute for Neurological Studies and the Myositis Support Group. This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme.

CONTRIBUTORSHIP STATEMENT

Dr Stefen Brady - Acquisition of data, analysis and interpretation of data and drafting of manuscript.

Dr Waney Squier - Critical revision of manuscript for important intellectual content.

Prof. Caroline Sewry - Study concept and design and critical revision of manuscript for important intellectual content.

Prof. Mike Hanna - Critical revision of manuscript for important intellectual content.

Dr David Hilton-Jones - Critical revision of manuscript for important intellectual content.

Dr Janice Holton - Study concept and design, critical revision of manuscript for important intellectual content and study supervision.

COMPETING INTERESTS

None

DATA SHARING

All additional data can be found in supplementary tables and figures.

REFERENCES

1. 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 Psychiatr. 2008;79:1056–60.
2. Griggs RC, Askanas V, DiMauro S, et al. Inclusion body myositis and myopathies. Ann Neurol. 1995;38:705–13.
3. Arahata K, Engel AG. Monoclonal antibody analysis of mononuclear cells in myopathies. I: Quantitation of subsets according to diagnosis and sites of accumulation and demonstration and counts of muscle fibers invaded by T cells. Ann Neurol. 1984;16:193–208.
4. Mhiri C, Gherardi R. Inclusion body myositis in French patients. A clinicopathological evaluation. Neuropathol Appl Neurobiol. 1990;16:333–44.
5. Villanova M, Kawai M, Lübke U, et al. Rimmed vacuoles of inclusion body myositis and oculopharyngeal muscular dystrophy contain amyloid precursor protein and lysosomal markers. Brain Res. 1993;603:343–7.
6. Van der Meulen MF, Hoogendijk JE, Moons KG, et al. Rimmed vacuoles and the added value of SMI-31 staining in diagnosing sporadic inclusion body myositis. Neuromuscul Disord. 2001;11:447–51.
7. Ferrer I, Olivé M. Molecular pathology of myofibrillar myopathies. Expert Rev Mol Med. 2008;10:e25.
8. Amato AA, Gronseth GS, Jackson CE, et al. Inclusion body myositis: clinical and pathological boundaries. Ann Neurol. 1996;40:581–6.
9. Chahin N, Engel AG. Correlation of muscle biopsy, clinical course, and outcome in PM and sporadic IBM. Neurology. 2008;70:418–24.
10. Benveniste O, Guiguet M, Freebody J, et al. Long-term observational study of sporadic inclusion body myositis. Brain. 2011;134:3176–84.

11. Brady S, Squier W, Hilton-Jones D. Clinical assessment determines the diagnosis of inclusion body myositis independently of pathological features. J Neurol Neurosurg Psychiatr. 2013; 16.

12. Greenberg SA. Theories of the Pathogenesis of Inclusion Body Myositis. Current Rheumatology Reports. 2010;12:221–8.

13. Sherriff FE, Joachim CL, Squier MV, et al. Ubiquitinated inclusions in inclusion-body myositis patients are immunoreactive for cathepsin D but not β-amyloid. Neuroscience Letters. 1995;194:37–40.

14. Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. BMJ. 2009;339:b2680.

15. Needham M, and Mastaglia FL. (2007). Inclusion body myositis: current pathogenetic concepts and diagnostic and therapeutic approaches. Lancet Neurol. 6, 620–631.

16. Benveniste O, Hilton-Jones D. International Workshop on Inclusion Body Myositis held at the Institute of Myology, Paris, on 29 May 2009. Neuromuscular Disorders. 2010;20:414–21.

17. Wedderburn LR, Varsani H, Li CKC, Newton KR, Amato AA, Banwell B, et al. International consensus on a proposed score system for muscle biopsy evaluation in patients with juvenile dermatomyositis: a tool for potential use in clinical trials. Arthritis Rheum. 2007;57:1192–201.

18. Nogalska A, Terracciano C, D'Agostino C, et al. p62/SQSTM1 is overexpressed and prominently accumulated in inclusions of sporadic inclusion-body myositis muscle fibers, and can help differentiating it from polymyositis and dermatomyositis. Acta Neuropathol. 2009;118:407–413.

19. Weihl CC, Temiz P, Miller SE, et al. TDP-43 accumulation in inclusion body myopathy muscle suggests a common pathogenic mechanism with frontotemporal dementia. J. Neurol. Neurosurg. Psychiatr. 2008;79:1186–1189.

20. Banwell BL, Engel AG. Alpha B-crystallin immunolocalization yields new insights into inclusion body myositis. Neurology. 2000;54:1033–1041.

21. Dubourg O, Wanschitz J, Maisonobe T, et al. Diagnostic value of markers of muscle degeneration in sporadic inclusion body myositis. Acta Myol. 2011. 30, 103–108.

22. Hiniker A, Daniels BH, Lee HS, et al. Comparative utility of LC3, p62 and TDP-43 immunohistochemistry in differentiation of inclusion body myositis from polymyositis and related inflammatory myopathies. Acta Neuropathologica Communications. 2013. 1:29.

23. D'Agostino C, Nogalska A, Engel WK, et al. In sporadic inclusion-body myositis muscle fibres TDP-43-positive inclusions are less frequent and robust than p62-inclusions, and are not associated with paired helical filaments. Neuropathol Appl Neurobiol. 2010; Available from: http://www.ncbi.nlm.nih.gov/pubmed/20626631.

24. Momma K, Noguchi S, Malicdan MCV, et al. Rimmed vacuoles in Becker muscular dystrophy have similar features with inclusion myopathies. PLoS ONE. 2012;7:e52002.

Supplementary Figure 1 Control staining in brain and muscle tissue

Positive and negative (no primary) brain control sections and normal muscle stained using immunohistochemistry for: p62 (A-C), TDP-43 (D-F), α B-crystallin (G-I), ubiquitin (J-K) and myotilin (M,N) and alkalinised congo red (O).

(A-C) Negative (A) and positive (B) control sections of AD brain and normal muscle (C) stained for p62. Positive control shows p62 positive neurofibrillary tangles and dystrophic neurites (B). No p62 immunoreactivity is observed in normal muscle (C).
(D-F) Negative (D) and positive (E) control sections of FTLD-TDP brain and normal muscle (F) stained for TDP-43. Positive control shows normal nuclear labelling and mislocalised neuronal cytoplasmic staining with neuropil threads (E). Insert shows a neuron with absent nuclear TDP-43 and a cytoplasmic TDP-43 inclusion (E, red arrow and x100 insert). Nuclear TDP-43 staining is observed in normal muscle.
(G-I) Negative (G) and positive (H) control sections of CBD brain and normal muscle (I) stained for α B-crystallin. Positive control shows neuropil threads and a balloon cell neuron (H; red arrow and x100 insert). No α B-crystallin immunoreactivity is observed in normal muscle (I).
(J-L) Negative (J) and positive (K) control AD brain and normal muscle (L) stained for ubiquitin. Positive control shows dystrophic neurites and neuropil threads (K). No ubiquitin immunoreactivity is observed in normal muscle (L).
(M,N) Negative (M) and positive (N) control muscle stained for myotilin. Mild sarcoplasmic staining is observed in normal muscle (N).
(O) Positive control section of AD brain showing an amyloid plaque (O).
Scale bar represents 100 µm in A-D, F and H-M; and 50 µm in E, N-O.
p62 = Sequestosome 1; AD = Alzheimer's disease; TDP-43 = Transactivation response DNA binding protein 43; FTLD-TDP = Frontotemporal lobar degeneration with TDP-43 positive inclusions; CBD = Corticobasal degeneration.

Supplementary Figure 2 IBM inflammatory score-tool

Score tool modified from the published juvenile dermatomyositis inflammatory (JDM) score tool [17] to specifically assess the type, degree and distribution of inflammation in IBM. The inflammatory domain was augmented to include T-cells, T-cell subtypes, B-cells and macrophages. MHC Class I staining was expanded to include three patterns of labelling. The vascular, muscle fibre and connective tissue domains which are present in the JDM score tool were not included.

Supplementary Figure 3 Sensitivity and specificity of rimmed vacuoles, protein aggregates and mitochondrial changes in IBM+RV compared to PAM

Receiver operating characteristic curves for each test including the area under the curve and optimum cut-off with its associated sensitivity and specificity for rimmed vacuoles (A), myotilin (B), ubiquitin (C), TDP-43 (D), p62 (E) immunoreactive deposits, congophilic deposits (F) and COX-/SDH+ fibres (G). COX/SDH HC staining was the most discriminative test for differentiating IBM+RV and PAM (G). However, there was little difference between COX/SDH HC staining, TDP-43 and p62 IHC staining and none were sufficiently discriminative to be considered diagnostic. AUC = Area under the curve.

Supplementary Figure 4 Sensitivity and specificity of protein aggregates and mitochondrial changes in IBM-RV compared to PM&DM

Receiver operating characteristic curves for each test showing the area under the curve and optimum cut-off with its sensitivity and specificity for myotilin (A), ubiquitin (B), TDP-43 (C), p62 (D) immunoreactive deposits, congophilic deposits (E) and COX-/SDH+ fibres (F). COX/SDH histochemical staining (F) and myotilin (G) IHC were the most discriminative tests for differentiating IBM-RV and PM&DM. AUC = Area under the curve.

Figure 1 Protein aggregates and congophilic deposits in IBM

Stained cryostat sections, showing fibres, often in clusters, containing protein aggregates stained for p62 (A), TDP-43 (B), ubiquitin (C), αB-crystallin (D) and myotilin (E). Protein aggregates were present throughout fibres, and were observed in apparently normal fibres, vacuolated fibres and fibres surrounded by inflammatory infiltrates. In fibres containing TDP-43 aggregates, myonuclear TDP-43 staining was frequently reduced (B). Congophilic deposits were observed in vacuolated fibres using epifluorescence (F). Tissue sections were examined using both the rhodamine red and fluorescein isothiocyanate filters to exclude areas of auto-fluorescence (arrow). Combined fluorescent image is shown. Four patterns of immunoreactivity were observed in IBM and disease controls stained for p62 using IHC (G)(H)(I)(J). Pattern I (G) - strongly stained, discreet and clearly delineated, round or angular aggregates, variable in number and size within a muscle fibre but rarely filling it and

predominantly located subsarcolemmal, but also perinuclear and adjacent to vacuoles. Pattern II (H) - large aggregates of variable staining intensity. Pattern III (I) - fine granular aggregates dispersed throughout the fibre. Pattern IV (J) - fine granules and wisps of p62 immunoreactivity set within weakly basophilic inclusions.

Scale bar represents 50 µm in A and D; 25 µm in B, C and E-J.

Figure 2 Percentage of muscle fibres containing protein aggregates and Griggs' pathological features

Box and whisker plot illustrating the percentage of muscle fibres containing pathological abnormalities contained in the Griggs criteria and protein aggregates in Griggs' pathologically-definite IBM. Fibres containing aggregates immunoreactive for p62 and αB-crystallin were more frequent than those containing the current diagnostic pathological features (red bars) ($p<0.05$). Protein aggregates recognised by all antibodies were found in a significantly larger number of fibres than partial invasion ($p<0.02$).

Figure 3 Percentage of fibres containing protein aggregates and COX-/SDH+ fibres in each group

Box and whisker plots illustrating the percentage of fibres in each diagnostic category containing p62 (A), TDP-43 (B), myotilin (C) and ubiquitin (D) immunoreactive aggregates, congophilic deposits (E) and COX-/SDH+ fibres (F). All protein aggregates were present in a greater percentage of fibres in IBM+RV than in IBM–RV. There was no difference in the percentage of COX-/SDH+ muscle fibres between these groups. IBM+RV biopsies had a greater percentage of fibres containing p62 (A) and TDP-43 (B) immunoreactive aggregates and COX-/SDH+ fibres (F) than PAM. Pathological findings were similar in IBM-RV and PM&DM, with no differences in the percentage of fibres containing p62 (A), TDP-43 (B) and ubiquitin (D) immunoreactive aggregates or congophilic deposits (E). However,

there was a greater percentage of COX-/SDH+ fibres (F) in IBM–RV than PM&DM and a greater percentage of fibres containing myotilin immunoreactive aggregates (C) in PM&DM than IBM-RV. *Statistically significant results.

Figure 4 Proposed diagnostic algorithm for IBM based on pathological findings

Flow diagram showing a proposed pathway for diagnosing IBM based on the pathological findings. Increased MHC Class I staining was observed in all cases of IBM and pattern I p62 aggregates in all cases of IBM+RV making them good initial screening tests. Their absence rules-out a diagnosis of IBM and IBM+RV respectively. The presence of endomysial CD3+ T-cell score >1, endomysial CD8+ T-cell score >0 or strong MHC Class I staining in a biopsy with rimmed vacuoles and p62 aggregates secures a diagnosis of IBM+RV. Differentiating IBM-RV and PM&DM pathologically is more challenging. The presence of COX-/SDH+ fibres is not specific to IBM-RV; although COX-/SDH+ fibres were not present in every case of IBM-RV their absence casts doubt on the diagnosis of IBM-RV. Pattern I p62 aggregates may enable IBM to be differentiated from PM when present. However, they may lack sensitivity for IBM-RV, therefore their absence does not rule out the diagnosis.

A retrospective cohort study identifying the principal pathological features useful in the diagnosis of inclusion body myositis

Corresponding author:

Dr Janice L Holton

Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

janice.holton@ucl.ac.uk; tel: 00 44 (0)20 3448 4239; fax: 00 44 (0)20 3448 4486.

Authors:

Stefen Brady[1], Waney Squier[2], Caroline Sewry[3,4], Michael Hanna[1], David Hilton-Jones[5], Janice L Holton[6]

[1]MRC Centre for Neuromuscular Diseases, UCL Institute of Neurology and National Hospital for Neurology, Neurosurgery, Queen Square, London, UK.

[2]Department of Neuropathology, University of Oxford, John Radcliffe Hospital, Oxford, UK.

[3]Dubowitz Neuromuscular Centre, Institute of Child Health and Great Ormond Street Hospital for Children, London, UK.

[4]Wolfson Centre of Inherited Neuromuscular Diseases, RJAH Orthopaedic Hospital, Oswestry, UK.

[5]Nuffield Department of Clinical Neurosciences (Clinical Neurology), University of Oxford, John Radcliffe Hospital, Oxford, UK.

[6]Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK.

Keywords: Neurology, adult neurology, neuromuscular disease, neuropathology

Running title: Pathological criteria for inclusion body myositis

Word count: 2990

ABSTRACT

Objectives

The current pathological diagnostic criteria for sporadic inclusion body myositis (IBM) lack sensitivity. Using immunohistochemical techniques abnormal protein aggregates have been identified in IBM, including some associated with neurodegenerative disorders. Our objective was to investigate the diagnostic utility of a number of markers of protein aggregates together with mitochondrial and inflammatory changes in IBM.

Design

Retrospective cohort study. The sensitivity of pathological features was evaluated in cases of Griggs' definite IBM. The diagnostic potential of the most reliable features was then assessed in clinically-typical IBM with rimmed vacuoles (*n*=15) and clinically-typical IBM without rimmed vacuoles (*n*=9) and IBM mimics - vacuolar myopathies (*n*=7) and steroid-responsive inflammatory myopathies (*n*=11).

Setting

Specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London.

Results

Individual pathological features, in isolation, lacked sensitivity and specificity. However, the morphology and distribution of p62 aggregates in IBM were characteristic and in a myopathy with rimmed vacuoles, the combination of characteristic p62 aggregates and increased sarcolemmal and internal MHC Class I expression or endomysial T-cells were diagnostic for IBM with a sensitivity of 93% and specificity of 100%. In an inflammatory myopathy lacking rimmed vacuoles, the presence of mitochondrial changes was 100% sensitive and 73% specific for IBM; characteristic p62 aggregates were specific (91%), but lacked sensitivity (44%).

Conclusions

We propose an easily applied diagnostic algorithm for the pathological diagnosis of IBM. Additionally our findings support the hypothesis that many of the pathological features considered

typical of IBM develop later in the disease, explaining their poor sensitivity at disease presentation and emphasising the need for revised pathological criteria to supplement the clinical criteria in the diagnosis of IBM.

STRENGTHS AND LIMITATIONS

The present study is a multicentre retrospective evaluation of the diagnostic utility of pathological findings for differentiating IBM from myopathies important in the differential diagnosis – myopathies containing rimmed vacuoles and steroid-responsive inflammatory myopathies.

The main strength of our study was the systematic detailed analysis of well-defined cases. This enabled us to determine the sensitivity and specificity of individual pathological features and produce an easily applied pathological diagnostic algorithm for IBM for use in clinical practice.

Study limitations include the small number of cases and the retrospective design. Further prospective studies are now required in larger cohorts of patients.

INTRODUCTION

Sporadic inclusion body myositis (IBM) is the commonest acquired myopathy in those aged over 50 years.[1] Although classified as an idiopathic inflammatory myopathy, muscle biopsy reveals both degenerative and inflammatory features. The widely used Griggs diagnostic criteria require the presence of several pathological findings,[2] namely rimmed vacuoles, an inflammatory infiltrate with invasion of non-necrotic fibres by mononuclear inflammatory cells (partial invasion), and either amyloid deposits or 15-18 nm tubulofilaments identified by electron microscopy (EM). Although these features in combination are highly specific for IBM, individually they occur in other myopathies, including some important in the differential diagnosis for IBM.[3-7] Moreover, cases of clinically-typical IBM have been reported where the combination of these pathological features is absent causing diagnostic difficulty.[8-11]

Over the last two decades, pathological accumulation of many different proteins has been reported in muscle fibres in IBM.[12] Proteins typically associated with neurodegenerative diseases such as β-amyloid (Aβ), hyperphosphorylated tau and ubiquitin and newer neurodegenerative markers such as p62 and transactivation response DNA binding protein-43 (TDP-43) have been identified, as well as proteins associated with myofibrillar myopathies (MFM), including desmin and αB-crystallin. However, not all observations have been consistently reproduced.[13,14] Mitochondrial changes have also been proposed for inclusion in IBM diagnostic criteria,[15]. Clear guidelines for the incorporation of immunohistochemical findings and mitochondrial changes into diagnostic criteria for IBM have not been established.[16]

Previously, we have shown that the characteristic pattern of weakness associated with IBM is indicative of the diagnosis, even if Griggs pathological features are absent.[11] However, it is not invariably found at presentation. Here we sought to identify which pathological features, other than the Griggs pathological criteria, add further support to the diagnosis of IBM. We systematically investigated which pathological features are present in Griggs pathologically-definite IBM and then established the diagnostic utility of these features in cases of IBM lacking the Griggs criteria, using myopathies considered in the differential diagnosis of IBM as controls.

MATERIALS AND METHODS

The study received ethical approval from the Departments of Research and Development at Oxford University Hospitals NHS Trust, Oxford and University College London Hospitals NHS Foundation Trust, London.

Cases

All patients were followed by specialist muscle services at the John Radcliffe Hospital, Oxford and the National Hospital for Neurology and Neurosurgery, London. Biopsies were taken for diagnostic purposes from the deltoid or quadriceps muscles and prior to any treatment.

Methods for demonstrating pathological features in IBM, additional to those defined by the Griggs criteria, were determined in six Griggs pathologically-definite cases of IBM. Cases with no clinical or pathological evidence of neuromuscular disease were used as controls. The diagnostic utility of the pathological features identified was assessed in two groups of clinically-typical IBM; one with rimmed vacuoles on muscle biopsy (IBM+RV; *n*=15), the other without rimmed vacuoles on muscle biopsy (IBM-RV; *n*=9). Disease controls were cases of steroid-responsive inflammatory myopathies [polymyositis and dermatomyositis; (PM&DM); *n*=11] and protein accumulation myopathies with rimmed vacuoles (PAM; *n*=7). Clinical characteristics and inclusion criteria are summarised in Supplementary tables 1 and 2. Tissue from brains donated to the Queen Square Brain Bank for Neurological Disorders was used as positive controls for protein aggregate staining.

Muscle biopsies

Muscles biopsies were snap frozen at the time of surgery in isopentane cooled liquid nitrogen. Until sectioning all samples were stored at -80^{o}C. Serial tissue sections were cut to a thickness of 8 µm, allowed to air dry and stored at -80^{o}C until staining. Prior to staining, tissue sections were allowed to dry at room temperature. Tissue sections were stained with haematoxylin and eosin (H&E), combined cytochrome oxidase (COX) succinate dehydrogenase (SDH) histochemistry and for amyloid using alkalinised Congo red, crystal violet and thioflavin S. Tissue sections for immunohistochemical staining were fixed for 10 minutes, if required, washed for five minutes in running water and incubated in 0.5% hydrogen peroxide to block endogenous peroxidase for 20 minutes. After further washing, tissue sections were incubated in 5% normal goat serum (Vector Laboratories, Burlingame, California) for 30 minutes and then systematically stained for: 1) proteins classically associated with

neurodegenerative disease: tau and hyperphosphorylated tau, ubiquitin, Aβ and α-synuclein; 2) proteins more recently reported in neurodegenerative disease: p62, TDP-43, fused in sarcoma protein (FUS) and valosin containing protein (VCP); 3) nuclear membrane proteins: lamin A/C and emerin; 4) proteins associated with MFM: desmin, myotilin and αB-crystallin; and 5) inflammatory cells and major histocompatibility complex class I (MHC Class I): CD3+ T-cells, CD4+ T-cells, CD8+ T-cells, B-cells and macrophages. Primary antibody binding was visualised using Dako REAL™ EnVision™ Detection System which contains horse-radish peroxidase (HRP) labelled goat anti-rabbit/mouse secondary and 3,3'-diaminobenzidine (DAB); following incubation with the relevant primary antibody, tissue sections were washed in phosphate buffered saline (PBS), incubated with HRP labelled goat anti-rabbit/mouse secondary for 30 minutes, washed in PBS and incubated in a 1:50 solution of DAB for three to five minutes. Details of commercial antibodies and conditions used are provided in Supplementary Table 3. IHC for each antibody was performed on all cases simultaneously and including positive and negative controls (Supplementary Figure 1).

Definitions and quantification

The total number of fibres and the number undergoing partial invasion, containing rimmed vacuoles, protein aggregates and COX-negative SDH-positive (COX-/SDH+) fibres were quantified using ImagePro version 6.2 (Media Cybernetics), to ensure that the whole biopsy was systematically analysed. Only transversely-orientated fibres not undergoing necrosis or regeneration were quantified. Tissue sections stained with Congo red were visualised under fluorescent and polarised light. Areas of fluorescence were examined using both rhodamine red (excitation 512-546 nm and emission 600-640 nm) and fluorescein isothiocyanate (excitation 440-480 nm and emission 527-530 nm) filters to exclude auto-fluorescence. Supplementary Table 4 provides definitions of the pathological features assessed. The inflammatory infiltrate and MHC Class I staining were analysed using a modified version of the semi-quantitative juvenile dermatomyositis score-tool (Supplementary Figure 2).[17] Assessments were performed blind to clinical details and diagnosis by a single individual (SB). Ten

per cent of slides were re-counted to assess intra-observer reliability and 336 slides were assessed independently by two observers (SB and JLH) to determine inter-observer reliability.

Statistical analysis

Statistical analyses were performed using GraphPad PRISM version 5. Continuous and categorical variables were compared using Mann Whitney *U*-test and chi-squared or Fisher's exact test respectively. Spearman's rank order correlation was used to determine the strength and direction of associations between pathological findings. Linear regression was used to determine relationships between clinical features and pathological findings. Test characteristics were calculated using receiver operating characteristic (ROC) curves and 2x2 contingency tables. A test was considered diagnostic when sensitivity >75% and specificity >95% or sensitivity >95% and specificity >75%. Intra-observer and inter-observer agreement was calculated using Bland-Altman plots and Cohen's kappa statistic (κ). Repeat counts were within 95% confidence intervals using Bland-Altman plots and κ was ≥0.7 indicating good intra-observer and good or excellent inter-observer reliability. Statistical significance was set at $p<0.05$.

RESULTS

Pathological findings in Griggs' pathologically-definite IBM

p62, TDP-43, ubiquitin, myotilin and αB-crystallin immunoreactive aggregates were present in all six IBM cases but not in normal controls (Figures 1A-E). p62 and αB-crystallin immunoreactive aggregates were present in a greater percentage of fibres than the pathological features required in the Griggs criteria ($p<0.05$) (Figure 2). Despite their abundance, αB-crystallin immunoreactive aggregates were difficult to quantify due to a significant variability in their morphology. No immunoreactive deposits were observed in IBM cases or normal controls with antibodies to tau and phosphorylated tau, Aβ, α-synuclein, desmin, emerin, lamin A/C, FUS or VCP. Alkalinised Congo red staining was more sensitive than crystal violet and thioflavin S staining for observing amyloid aggregates (Figure 1F). Tissue sections containing congophilic deposits identified under fluorescence light showed no

apple-green birefringence under polarised light. Mitochondrial changes and increased sarcolemmal and sarcoplasmic MHC Class I staining were observed in all six IBM cases, but not in normal controls. The inflammatory infiltrate was predominantly composed of endomysial CD8+ T-cells and macrophages, with relatively few B-cells.

Quantitative analysis of pathological features in IBM and disease controls

Having shown that p62, TDP-43, ubiquitin and myotilin aggregates, congophilic deposits, MHC Class I and inflammatory cells were prevalent in Griggs' pathologically-definite IBM, the presence of these abnormalities, together with mitochondrial changes were assessed in IBM+RV, IBM-RV and disease controls.

The percentage of fibres containing p62, TDP-43, myotilin and ubiquitin aggregates and congophilic deposits were greater in IBM+RV than in IBM-RV; there was no difference in the number of COX-/SDH+ fibres (Figure 3A-F). Protein aggregates were observed in morphologically-normal fibres and in fibres exhibiting Griggs' pathological features. p62 and TDP-43 positive aggregates were present in a greater percentage of fibres in IBM+RV compared to PAM; however, there were no differences in the percentage of fibres containing myotilin and ubiquitin aggregates or congophilic deposits. The percentage of fibres containing p62, TDP-43 and ubiquitin aggregates or congophilic deposits were similar in IBM-RV and PM&DM; however, myotilin aggregates were present in a greater percentage of fibres in PM&DM and COX-/SDH+ fibres were more abundant in IBM-RV. Analysis of the total inflammatory infiltrate (the sum of the semi-quantitative scores for T-cells, B-cells and macrophages) in the endomysium, perimysium and perivascular areas revealed that there were greater numbers of inflammatory cells in the endomysium and perimysium in IBM+RV than in PAM (p<0.03). The same analysis comparing the sum of the inflammatory cells in IBM-RV and PM&DM revealed that the distribution and intensity of the inflammatory infiltrate was similar.

Diagnostic utility of pathological features in IBM and disease controls

To mimic the diagnostic difficulty encountered in clinical practice, the ability of each test to differentiate between myopathies containing rimmed vacuoles (IBM+RV and PAM) and between inflammatory myopathies (IBM–RV and PM&DM) was assessed.

Diagnostic utility determined using receiver-operating characteristic curves

Individually, the presence of p62 immunoreactive inclusions and COX-/SDH+ fibres had the highest sensitivity and specificity for differentiating IBM+RV from PAM, (Supplementary Figure 3) (Table 1). Differentiating between IBM–RV and PM&DM, myotilin positive inclusions or COX-/SDH+ fibres had the highest sensitivity and specificity for IBM-RV (Supplementary Figure 4) (Table 1). Only the presence of myotilin positive inclusions satisfied criteria to be considered suitable as a diagnostic test (<0.01% of fibres containing myotilin aggregates had a sensitivity of 100% and specificity of 82% for IBM-RV).

Table 1 Test characteristics

Test feature	IBM+RV *v.* PAM				IBM-RV *v.* PM&DM			
	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity	AUC	Cut-off (% of affected fibres)	Sensitivity	Specificity
Rimmed vacuoles	0.60	>0.28	0.53	0.71	-	-	-	-
p62 aggregates	0.87	>0.48	0.87	0.86	0.60	>0.21	0.22	0.91
TDP-43 aggregates	0.80	>0.34	0.80	0.86	0.53	<0.01	0.89	0.18
Ubiquitin aggregates	0.68	>0.18	0.53	0.85	0.64	<0.01	1.00	0.27
Myotilin aggregates	0.55	<0.25	1.00	0.29	0.91	<0.01	1.00	0.82
Congophilic deposits	0.56	>0.24	0.73	0.71	0.56	<0.03	0.11	0.82
COX-/SDH+ fibres	0.87	>0.04	0.86	0.86	0.93	>0.1	0.78	0.91

Table shows the area under the curve and optimum cut-off for each test with the accompanying sensitivity and specificity. AUC = Area under the curve.

Diagnostic utility determined by comparing proportion of affected cases in each diagnostic group

In the aforementioned experiments, the number of fibres within each muscle biopsy was quantified. However, this is impractical for routine clinical use. Thus, the proportions of affected cases in each group were compared (Table 2). This revealed that neither staining for protein aggregates nor congophilic deposits could differentiate between IBM+RV and PAM. The pathological findings in IBM-RV and PM&DM were also similar, except that the absence of myotilin immunoreactive aggregates was sensitive and specific for IBM-RV. COX-/SDH+ fibres were also suggestive of IBM-RV; one or more COX-/SDH+ fibres had a sensitivity of 100% and specificity 73% for IBM-RV.

Increased MHC Class I expression lacked specificity. However, strong (diffuse sarcolemmal and sarcoplasmic) MHC Class I up-regulation was diagnostic for IBM+RV, differentiating it from PAM, as were the presence of either endomysial CD3+ T-cell or CD4+ T-cell scores >1 or an endomysial CD8+ T-cell score >0. Partial invasion was specific for IBM+RV, but lacked sensitivity. Although the sum of the inflammatory infiltrate was similar in IBM-RV and PM&DM, analysis of the inflammatory cell sub-types revealed greater numbers of perimysial CD3+ T-cells, CD8+ T-cells and endomysial B-cells [were observed] in PM&DM than in IBM–RV ($p \leq 0.02$), however, this was not diagnostically useful. There was no difference in the proportion of cases with fibres undergoing partial invasion between IBM–RV and PM&DM.

Table 2 Comparison of the proportion of positive cases in each group

Pathological features	IBM+RV	PAM	IBM+RV *v.* PAM		IBM–RV	PM&DM	IBM-RV *v.* PM&DM		IBM+RV *v.* IBM-RV
	n (%)	n (%)	Sensitivity	Specificity	n (%)	n (%)	Sensitivity	Specificity	*p* value
Number of cases	15 (100)	7 (100)			9 (100)	11 (100)			
Aggregated proteins, n (%)									
p62	15 (100)	6 (86)	1.00	0.14	4 (44)	3 (27)‡	0.40	0.72	0.003*
TDP-43	13 (87)	5 (71)	0.87	0.29	1 (11)	2 (18)‡	0.11	0.82	0.001*
Ubiquitin	11 (73)	4 (57)	0.73	0.43	0 (0)	3 (27)‡	0.00	0.73	0.001*
Myotilin	10 (67)	5 (71)	0.67	0.29	0 (0)	9 (82)	0.00	0.18	0.002*
Congophilic deposits	13 (87)	7 (100)	0.87	0.00	1 (11)	0 (0)	0.11	1.00	0.001*
COX-/SDH+ fibres†, n (%)									
Any	12 (86)	2 (29)	0.80	0.71	9 (100)	3 (27)	1.00	0.73	0.5
Inflammatory features, n (%)									
MHC Class I up-regulation	15 (100)	3 (43)	1.00	0.57	9 (100)	11 (100)	1.00	0.00	1.00
Strong MHC Class I up-regulation	14 (93)	0 (0)	0.93	1.00	9 (100)	10 (91)	1.00	0.09	0.53
Partial invasion	10 (67)	0 (0)	0.67	1.00	3 (33)	2 (18)	0.33	0.82	0.11
Endomysial CD3+ T-cell score >1	13 (87)	0 (0)	0.87	1.00	4 (44)	7 (64)	0.44	0.36	0.02*
Endomysial CD4+ T-cell score >1	12 (80)	0 (0)	0.80	1.00	2 (22)	5 (45)	0.22	0.46	0.01*
Endomysial CD8+ T-cell score >0	14 (93)	0 (0)	0.93	1.00	4 (44)	5 (45)	0.44	0.54	0.02*
Endomysial CD68+ macrophage score >1	12 (80)	0 (0)	0.80	1.00	4 (44)	8 (73)	0.44	0.17	0.07

†In IBM with rimmed vacuoles *n*=14. ‡Pathological features present in DM, but not PM cases. *Statistically significant results.

Because IBM-RV is more pathologically akin to PM than DM, analyses were repeated comparing IBM-RV and PM cases (*n*=6). No p62, TDP-43 or ubiquitin immunoreactive aggregates were observed in PM cases and the diagnostic utility of tests for differentiating between IBM-RV and PM yielded similar results to prior analyses between IBM-RV and PM&DM.

Diagnostic utility of categorising the pattern of p62 staining

The pattern of p62 staining could be categorised into four distinct groups (Figure 1G-J). Aggregates observed in IBM were present in vacuolated and non-vacuolated fibres and were strongly stained, discreet and clearly delineated, round or angular and typically located subsarcolemmal, perinuclear and peri-vacuolar (pattern I). This pattern was observed in every IBM case with p62 aggregates, one (9%) case of DM and three (43%) cases of PAM (hereditary IBM, dystrophinopathy and genetically undefined MFM). Defining the pattern of immunoreactivity increased the discriminative value of p62 IHC for differentiating IBM+RV from PAM; pattern I p62 aggregates compared to any p62 aggregates increased the specificity from 14% to 57%, with no loss of sensitivity. Differentiating IBM-RV and PM&DM, pattern I p62 aggregates were highly specific (91%), but lacked sensitivity (44%). Patterns II, III and IV were not observed in any IBM cases. Patterns II and III appeared to be specific for PAM (*n*=2; 26%), both were cases of myotilinopathy (*n*=2; 67%), and DM (*n*=2; 40%) respectively. Pattern IV occurred in a genetically undefined case of MFM. No differences were observed in the morphology of TDP-43, myotilin or ubiquitin aggregates between biopsies.

Clinicopathological correlation

In IBM+RV, IBM-RV and pathologically-definite IBM, there were no correlations in individual biopsies between pathological features. No relationships were identified between the pathological findings and age at symptom onset, age at biopsy, disease duration or serum creatine kinase. The same results were obtained when the IBM groups were analysed separately and as one.

Proposed diagnostic algorithm

Based on our pathological findings, we propose a diagnostic algorithm for differentiating IBM from its disease mimics (Figure 4).

The algorithm was tested in a further 23 cases that fulfilled the criteria for IBM+RV (*n*=12) and IBM-RV (*n*=11). The algorithm correctly diagnosed 20 (87%) cases: 12 (100%) cases of IBM+RV and eight (73%) cases of IBM-RV. In IBM-RV, COX-/SDH+ fibres were present in 8 (73%) cases, pattern I p62 aggregates in 8 (73%) cases and both in 6 (55%) cases.

DISCUSSION

While Griggs' pathological criteria have been accepted as diagnostic of IBM, many patients who, observed over time undoubtedly have IBM, lack one or more of the Griggs pathological features at presentation, even on repeat biopsy.[8,11] Despite IBM being associated with a characteristic pattern of finger flexor and knee extensor weakness, not all patients have this pattern at disease onset, and muscle biopsy remains an important tool in differentiating IBM from its mimics. We sought to determine which additional pathological features support a diagnosis of IBM, demonstrating that characteristic p62 immunoreactive aggregates, strong MHC Class I upregulation, endomysial CD3+ T-cell score >1, CD8+ T-cell score >0 and COX-/SDH+ fibres are features with sufficient sensitivity and specificity to differentiate IBM from pathologically similar myopathies and we propose an easily applied pathological algorithm for the diagnosis of IBM (Figure 4).

In agreement with previous studies, we observed p62,[18] TDP-43,[19] ubiquitin [13] and αB-crystallin [20] immunoreactive aggregates and a predominantly endomysial inflammatory infiltrate [3] in Griggs pathologically-definite IBM. Diagnostic pathological studies of IBM have concentrated on differentiating IBM from other inflammatory myopathies and two recent quantitative studies have found that TDP-43 and markers of autophagy such as p62 and LC3 may be of diagnostic use.[21,22] However, in these studies only a fraction of each biopsy was analysed i.e. 200 fibres. We have found

this limited quantification does not correlate with the percentage of affected fibres in a biopsy nor does it reflect the way in which a muscle biopsy is assessed. Additionally, studies have lacked vacuolar myopathy control cases as it is believed that the inflammatory changes present in IBM enable it to be easily differentiated from other vacuolar myopathies.[22] However, inflammatory changes are frequently observed in muscular dystrophies and the degree of inflammatory change necessary to confidently diagnose IBM is currently unknown.

To mimic the typical diagnostic conundrums encountered in clinical practice, we evaluated the ability of the pathological findings to differentiate IBM+RV from other vacuolar myopathies and IBM-RV from steroid-responsive inflammatory myopathies. We found that quantitative analysis of protein aggregates, congophilic deposits and COX-/SDH+ fibres was of limited diagnostic use. Analysing the biopsies dichotomously and using a semi-quantitative score-tool revealed that increased MHC Class I labelling was sensitive for IBM making it a good initial screening test, its absence excluding the diagnosis. In agreement with an earlier study, we found p62 aggregates identified the largest number of affected fibres in IBM.[23] Additionally, as a novel finding, the morphology and distribution of p62 aggregates was characteristic in IBM. This characteristic pattern of p62 immunoreactive aggregates was highly sensitive for IBM+RV (100%); their absence from a biopsy containing rimmed vacuoles effectively ruling-out a diagnosis of IBM. We confirmed that the most diagnostically useful pathological findings in IBM+RV were evidence of an immune mediated process; strong MHC Class I staining, endomysial CD3+ T-cell score >1 or an endomysial CD8+ T-cell score >0 were diagnostic. Having identified either of these features in a biopsy containing rimmed vacuoles no extra diagnostic certainty was gained from observing partial invasion, COX-/SDH+ fibres or congophilic deposits.

The most discriminative pathological tests for differentiating between IBM–RV and PM&DM were COX/SDH staining and myotilin IHC. Consistent with a recent study,[9] we found the absence of mitochondrial changes casts doubt on a diagnosis of IBM. There was no difference in the median age between IBM-RV and PM&DM cases to account for the difference observed in COX-/SDH+ fibres.

The presence of myotilin and ubiquitin immunoreactive aggregates appeared to rule out a diagnosis of IBM-RV. However, we believe the presence of these features in IBM+RV indicates that they are unlikely to be diagnostically reliable features for differentiating between IBM-RV and steroid-responsive inflammatory myopathies. Although no pathological feature was able to differentiate IBM-RV from steroid responsive inflammatory myopathies with certainty the presence of characteristic p62 aggregates and the absence of COX-/SDH+ fibres may help in supporting and opposing a diagnosis of IBM-RV respectively. Pattern I p62 immunoreactive aggregates were only present in 44% of the initial IBM-RV cases tested, but they were not observed in PM cases and were very rare in DM. Although pattern I p62 aggregates appear to lack sensitivity their specificity was 91% making their presence highly suggestive of a diagnosis of IBM-RV. However, we identified pattern I p62 in eight out of 11 (73%) further cases of IBM-RV that were assessed indicating a greater sensitivity and that p62 IHC warrants further investigation and validation in a larger, independent series. The diagnostic utility of the other patterns of p62 staining is uncertain. Although pattern II appeared to have some specificity for myotilinopathy the small number of cases makes it drawing any conclusion problematic. In addition to p62 other autophagic proteins have been found in IBM and suggested as diagnostic markers. [22] Autophagy is a cellular mechanism for degrading and recycling cellular proteins and organelles and therefore, altered autophagy could lead to the accumulation of abnormal mitochondria and misfolded aggregation-prone proteins and may also result in altered antigen presentation leading to the widespread increase of MHC Class I and suggests that altered autophagy may play an important role in the pathogenesis of IBM.

Almost all pathological features - protein aggregates, congophilic deposits and inflammation - were more abundant in IBM+RV than IBM-RV. Despite using slightly different inclusion criteria, similar differences have been reported between pathologically-typical and pathologically-atypical IBM.[21] However, we found no differences in the number of COX-/SDH+ fibres, the degree of MHC Class I upregulation, the morphology and distribution of p62 immunoreactive aggregates or the pattern of the inflammation between IBM+RV and IBM-RV, supporting our clinical observations that these are the

same disease. We believe that the pathological differences between IBM+RV and IBM-RV are, in part, due to differences in disease duration. Two studies have shown that rimmed vacuoles are more common in patients who are older at the time of muscle biopsy,[24,11] suggesting that they are associated with chronologically more advanced disease. Therefore, the pathological findings which are more abundant in IBM+RV and thought to be typical of IBM may instead be indicative of chronologically more advanced disease explaining their limited sensitivity at disease presentation. However, possibly due to the number of cases analysed, we were unable to confirm a relationship between pathological features and clinical findings. It could be argued that biopsies from different muscles may have affected the pathological findings observed and differences between IBM groups. However, in a recent review of 59 muscle biopsies from IBM cases in our clinical archive with quadriceps (n=31) and deltoid (n=28) biopsies we found no significant difference in the frequency of pathological findings.

A robust clinicopathological definition of IBM is of paramount importance for diagnosis and for selection and entry of patients into clinical trials. We have shown that certain pathological findings are more abundant than those included in the current pathologically-focussed diagnostic criteria. Moreover, p62 immunoreactive deposits, increased MHC Class I expression, endomysial CD3+ T-cells and CD8+ T-cells and COX-/SDH+ fibres have sufficient sensitivity and specificity to aid in the histological differentiation of IBM from disease mimics, supporting their inclusion in future diagnostic criteria for IBM alongside clinical criteria. Both CD3+ T-cells and CD8+ T-cells are included in the diagnostic algorithm as there was little difference in their sensitivity and specificity for differentiating IBM+RV from PAM. However, IHC staining for CD3+ T-cells is likely to be more widely available and avoids the costs of extra staining to subtype the inflammatory infiltrate enabling diagnostic algorithm to be used by a greater number diagnostic laboratories. Using our diagnostic algorithm, we found there would be little additional diagnostic security in identifying partial invasion, performing EM or staining for amyloid deposits. Finally, mitochondrial changes and MHC Class I up-regulation were the most consistent findings in our IBM cases suggesting that they are central to the

pathogenesis and that further investigation and therapeutic intervention should be directed towards these features.

REFERENCES

1. Needham M, James I, Corbett A, Day T, Christiansen F, Phillips B, et al. Sporadic inclusion body myositis: phenotypic variability and influence of HLA-DR3 in a cohort of 57 Australian cases. J Neurol Neurosurg Psychiatr. 2008;79:1056–60.
2. Griggs RC, Askanas V, DiMauro S, Engel A, Karpati G, Mendell JR, et al. Inclusion body myositis and myopathies. Ann Neurol. 1995;38:705–13.
3. Arahata K, Engel AG. Monoclonal antibody analysis of mononuclear cells in myopathies. I: Quantitation of subsets according to diagnosis and sites of accumulation and demonstration and counts of muscle fibers invaded by T cells. Ann Neurol. 1984;16:193–208.
4. Mhiri C, Gherardi R. Inclusion body myositis in French patients. A clinicopathological evaluation. Neuropathol Appl Neurobiol. 1990;16:333–44.
5. Villanova M, Kawai M, Lübke U, Oh SJ, Perry G, Six J, et al. Rimmed vacuoles of inclusion body myositis and oculopharyngeal muscular dystrophy contain amyloid precursor protein and lysosomal markers. Brain Res. 1993;603:343–7.
6. Van der Meulen MF, Hoogendijk JE, Moons KG, Veldman H, Badrising UA, Wokke JH. Rimmed vacuoles and the added value of SMI-31 staining in diagnosing sporadic inclusion body myositis. Neuromuscul Disord. 2001;11:447–51.
7. Ferrer I, Olivé M. Molecular pathology of myofibrillar myopathies. Expert Rev Mol Med. 2008;10:e25.
8. Amato AA, Gronseth GS, Jackson CE, Wolfe GI, Katz JS, Bryan WW, et al. Inclusion body myositis: clinical and pathological boundaries. Ann Neurol. 1996;40:581–6.
9. Chahin N, Engel AG. Correlation of muscle biopsy, clinical course, and outcome in PM and sporadic IBM. Neurology. 2008;70:418–24.

10. Benveniste O, Guiguet M, Freebody J, Dubourg O, Squier W, Maisonobe T, et al. Long-term observational study of sporadic inclusion body myositis. Brain. 2011;134:3176–84.

11. Brady S, Squier W, Hilton-Jones D. Clinical assessment determines the diagnosis of inclusion body myositis independently of pathological features. J Neurol Neurosurg Psychiatr. 2013; 16.

12. Greenberg SA. Theories of the Pathogenesis of Inclusion Body Myositis. Current Rheumatology Reports. 2010;12:221–8.

13. Sherriff FE, Joachim CL, Squier MV, Esiri MM. Ubiquitinated inclusions in inclusion-body myositis patients are immunoreactive for cathepsin D but not β-amyloid. Neuroscience Letters. 1995;194:37–40.

14. Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. BMJ. 2009;339:b2680.

15. Needham M, and Mastaglia FL. (2007). Inclusion body myositis: current pathogenetic concepts and diagnostic and therapeutic approaches. Lancet Neurol. 6, 620–631.

16. Benveniste O, Hilton-Jones D. International Workshop on Inclusion Body Myositis held at the Institute of Myology, Paris, on 29 May 2009. Neuromuscular Disorders. 2010;20:414–21.

17. Wedderburn LR, Varsani H, Li CKC, Newton KR, Amato AA, Banwell B, et al. International consensus on a proposed score system for muscle biopsy evaluation in patients with juvenile dermatomyositis: a tool for potential use in clinical trials. Arthritis Rheum. 2007;57:1192–201.

18. Nogalska A, Terracciano C, D'Agostino C, King Engel W, Askanas V. p62/SQSTM1 is overexpressed and prominently accumulated in inclusions of sporadic inclusion-body myositis muscle fibers, and can help differentiating it from polymyositis and dermatomyositis. Acta Neuropathol. 2009;118:407–413.

19. Weihl CC, Temiz P, Miller SE, Watts G, Smith C, Forman M, Hanson PI, Kimonis V, Pestronk A. TDP-43 accumulation in inclusion body myopathy muscle suggests a common pathogenic mechanism with frontotemporal dementia. J. Neurol. Neurosurg. Psychiatr. 2008;79:1186–1189.

20. Banwell BL, Engel AG. Alpha B-crystallin immunolocalization yields new insights into inclusion body myositis. Neurology. 2000;54:1033–1041.

21. Dubourg O, Wanschitz J, Maisonobe T, Béhin A, Allenbach Y, Herson S, and Benveniste O. Diagnostic value of markers of muscle degeneration in sporadic inclusion body myositis. Acta Myol. 2011. 30, 103–108.

22. Hiniker A, Daniels BH, Lee HS, Margeta M. Comparative utility of LC3, p62 and TDP-43 immunohistochemistry in differentiation of inclusion body myositis from polymyositis and related inflammatory myopathies. Acta Neuropathologica Communications. 2013. 1:29.

23. D'Agostino C, Nogalska A, Engel WK, Askanas V. In sporadic inclusion-body myositis muscle fibres TDP-43-positive inclusions are less frequent and robust than p62-inclusions, and are not associated with paired helical filaments. Neuropathol Appl Neurobiol. 2010; Available from: http://www.ncbi.nlm.nih.gov/pubmed/20626631.

24. Momma K, Noguchi S, Malicdan MCV, Hayashi YK, Minami N, Kamakura K, et al. Rimmed vacuoles in Becker muscular dystrophy have similar features with inclusion myopathies. PLoS ONE. 2012;7:e52002.

ACKNOWLEDGEMENTS

SB is funded by the Myositis Support Group. JLH is supported by the Reta Lila Weston Institute for Neurological Studies and the Myositis Support Group. This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme.

CONTRIBUTORSHIP STATEMENT

Dr Stefen Brady - Acquisition of data, analysis and interpretation of data and drafting of manuscript.

Dr Waney Squier - Critical revision of manuscript for important intellectual content.

Prof. Caroline Sewry - Study concept and design and critical revision of manuscript for important intellectual content.

Prof. Mike Hanna - Critical revision of manuscript for important intellectual content.

Dr David Hilton-Jones - Critical revision of manuscript for important intellectual content.

Dr Janice Holton - Study concept and design, critical revision of manuscript for important intellectual content and study supervision.

DATA SHARING

All additional data can be found in supplementary tables and figures.

Supplementary Figure 1 Control staining in brain and muscle tissue

Positive and negative (no primary) brain control sections and normal muscle stained using immunohistochemistry for: p62 (A-C), TDP-43 (D-F), α B-crystallin (G-I), ubiquitin (J-K) and myotilin (M,N) and alkalinised congo red (O).

(A-C) Negative (A) and positive (B) control sections of AD brain and normal muscle (C) stained for p62. Positive control shows p62 positive neurofibrillary tangles and dystrophic neurites (B). No p62 immunoreactivity is observed in normal muscle (C).

(D-F) Negative (D) and positive (E) control sections of FTLD-TDP brain and normal muscle (F) stained for TDP-43. Positive control shows normal nuclear labelling and mislocalised neuronal cytoplasmic staining with neuropil threads (E). Insert shows a neuron with absent nuclear TDP-43 and a cytoplasmic TDP-43 inclusion (E, red arrow and x100 insert). Nuclear TDP-43 staining is observed in normal muscle.

(G-I) Negative (G) and positive (H) control sections of CBD brain and normal muscle (I) stained for α B-crystallin. Positive control shows neuropil threads and a balloon cell neuron (H; red arrow and x100 insert). No α B-crystallin immunoreactivity is observed in normal muscle (I).

(J-L) Negative (J) and positive (K) control AD brain and normal muscle (L) stained for ubiquitin. Positive control shows dystrophic neurites and neuropil threads (K). No ubiquitin immunoreactivity is observed in normal muscle (L).

(M,N) Negative (M) and positive (N) control muscle stained for myotilin. Mild sarcoplasmic staining is observed in normal muscle (N).

(O) Positive control section of AD brain showing an amyloid plaque (O).

Scale bar represents 100 µm in A-D, F and H-M; and 50 µm in E, N-O.

p62 = Sequestosome 1; AD = Alzheimer's disease; TDP-43 = Transactivation response DNA binding protein 43; FTLD-TDP = Frontotemporal lobar degeneration with TDP-43 positive inclusions; CBD = Corticobasal degeneration.

Supplementary Figure 2 IBM inflammatory score-tool

Score tool modified from the published juvenile dermatomyositis inflammatory (JDM) score tool [17] to specifically assess the type, degree and distribution of inflammation in IBM. The inflammatory domain was augmented to include T-cells, T-cell subtypes, B-cells and macrophages. MHC Class I staining was expanded to include three patterns of labelling. The vascular, muscle fibre and connective tissue domains which are present in the JDM score tool were not included.

Figure 1 Protein aggregates and congophilic deposits in IBM

Stained cryostat sections, showing fibres, often in clusters, containing protein aggregates stained for p62 (A), TDP-43 (B), ubiquitin (C), αB-crystallin (D) and myotilin (E). Protein aggregates were present throughout fibres, and were observed in apparently normal fibres, vacuolated fibres and fibres surrounded by inflammatory infiltrates. In fibres containing TDP-43 aggregates, myonuclear TDP-43 staining was frequently reduced (B). Congophilic deposits were observed in vacuolated fibres using epifluorescence (F). Tissue sections were examined using both the rhodamine red and fluorescein isothiocyanate filters to exclude areas of auto-fluorescence (arrow). Combined fluorescent image is shown. Four patterns of immunoreactivity were observed in IBM and disease controls stained for p62 using IHC (G)(H)(I)(J). Pattern I (G) - strongly stained, discreet and clearly delineated, round or angular aggregates, variable in number and size within a muscle fibre but rarely filling it and predominantly located subsarcolemmal, but also perinuclear and adjacent to vacuoles. Pattern II (H) - large aggregates of variable staining intensity. Pattern III (I) - fine granular aggregates dispersed throughout the fibre. Pattern IV (J) - fine granules and wisps of p62 immunoreactivity set within weakly basophilic inclusions.

Scale bar represents 50 µm in A and D; 25 µm in B, C and E-J.

Figure 2 Percentage of muscle fibres containing protein aggregates and Griggs' pathological features

Box and whisker plot illustrating the percentage of muscle fibres containing pathological abnormalities contained in the Griggs criteria and protein aggregates in Griggs' pathologically-definite IBM. Fibres containing aggregates immunoreactive for p62 and αB-crystallin were more frequent than those containing the current diagnostic pathological features (red bars) ($p<0.05$). Protein aggregates recognised by all antibodies were found in a significantly larger number of fibres than partial invasion ($p<0.02$).

Figure 3 Percentage of fibres containing protein aggregates and COX-/SDH+ fibres in each group

Box and whisker plots illustrating the percentage of fibres in each diagnostic category containing p62 (A), TDP-43 (B), myotilin (C) and ubiquitin (D) immunoreactive aggregates, congophilic deposits (E) and COX-/SDH+ fibres (F). All protein aggregates were present in a greater percentage of fibres in IBM+RV than in IBM–RV. There was no difference in the percentage of COX-/SDH+ muscle fibres between these groups. IBM+RV biopsies had a greater percentage of fibres containing p62 (A) and TDP-43 (B) immunoreactive aggregates and COX-/SDH+ fibres (F) than PAM. Pathological findings were similar in IBM-RV and PM&DM, with no differences in the percentage of fibres containing p62 (A), TDP-43 (B) and ubiquitin (D) immunoreactive aggregates or congophilic deposits (E). However, there was a greater percentage of COX-/SDH+ fibres (F) in IBM–RV than PM&DM and a greater percentage of fibres containing myotilin immunoreactive aggregates (C) in PM&DM than IBM-RV. *Statistically significant results.

Supplementary Figure 3 Sensitivity and specificity of rimmed vacuoles, protein aggregates and mitochondrial changes in IBM+RV compared to PAM

Receiver operating characteristic curves for each test including the area under the curve and optimum cut-off with its associated sensitivity and specificity for rimmed vacuoles (A), myotilin (B), ubiquitin (C), TDP-43 (D), p62 (E) immunoreactive deposits, congophilic deposits (F) and COX-/SDH+ fibres

(G). COX/SDH HC staining was the most discriminative test for differentiating IBM+RV and PAM (G). However, there was little difference between COX/SDH HC staining, TDP-43 and p62 IHC staining and none were sufficiently discriminative to be considered diagnostic. AUC = Area under the curve.

Supplementary Figure 4 Sensitivity and specificity of protein aggregates and mitochondrial changes in IBM-RV compared to PM&DM

Receiver operating characteristic curves for each test showing the area under the curve and optimum cut-off with its sensitivity and specificity for myotilin (A), ubiquitin (B), TDP-43 (C), p62 (D) immunoreactive deposits, congophilic deposits (E) and COX-/SDH+ fibres (F). COX/SDH histochemical staining (F) and myotilin (G) IHC were the most discriminative tests for differentiating IBM-RV and PM&DM. AUC = Area under the curve.

Figure 4 Proposed diagnostic algorithm for IBM based on pathological findings

Flow diagram showing a proposed pathway for diagnosing IBM based on the pathological findings. Increased MHC Class I staining was observed in all cases of IBM and pattern I p62 aggregates in all cases of IBM+RV making them good initial screening tests. Their absence rules-out a diagnosis of IBM and IBM+RV respectively. The presence of endomysial CD3+ T-cell score >1, endomysial CD8+ T-cell score >0 or strong MHC Class I staining in a biopsy with rimmed vacuoles and p62 aggregates secures a diagnosis of IBM+RV. Differentiating IBM-RV and PM&DM pathologically is more challenging. The presence of COX-/SDH+ fibres is not specific to IBM-RV; although COX-/SDH+ fibres were not present in every case of IBM-RV their absence casts doubt on the diagnosis of IBM-RV. Pattern I p62 aggregates may enable IBM to be differentiated from PM when present. However, they may lack sensitivity for IBM-RV, therefore their absence does not rule out the diagnosis.

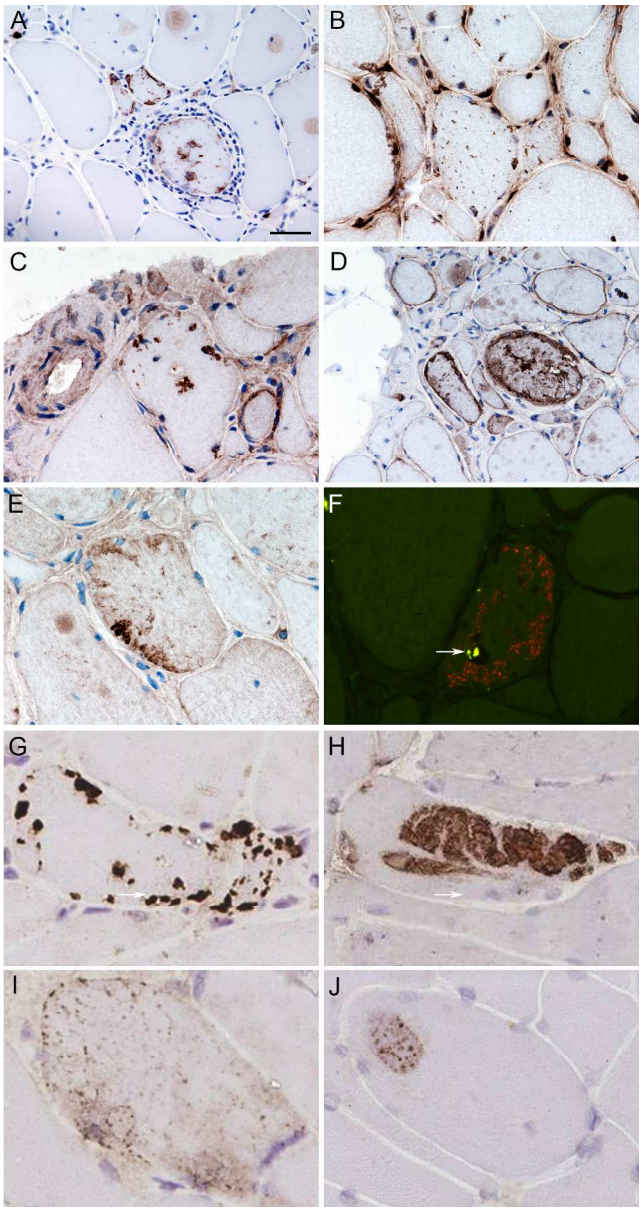


Figure 1 Protein aggregates and congophilic deposits in IBM
Stained cryostat sections, showing fibres, often in clusters, containing protein aggregates stained for p62 (A), TDP-43 (B), ubiquitin (C), αB-crystallin (D) and myotilin (E). Protein aggregates were present throughout fibres, and were observed in apparently normal fibres, vacuolated fibres and fibres surrounded by inflammatory infiltrates. In fibres containing TDP-43 aggregates, myonuclear TDP-43 staining was frequently reduced (B). Congophilic deposits were observed in vacuolated fibres using epifluorescence (F). Tissue sections were examined using both the rhodamine red and fluorescein isothiocyanate filters to exclude areas of auto-fluorescence (arrow). Combined fluorescent image is shown. Four patterns of immunoreactivity were observed in IBM and disease controls stained for p62 using IHC (G)(H)(I)(J). Pattern I (G) - strongly stained, discreet and clearly delineated, round or angular aggregates, variable in number and size within a muscle fibre but rarely filling it and predominantly located subsarcolemmal, but also perinuclear and adjacent to vacuoles. Pattern II (H) - large aggregates of variable staining intensity. Pattern III (I) - fine granular aggregates dispersed throughout the fibre. Pattern IV (J) - fine granules and wisps of

p62 immunoreactivity set within weakly basophilic inclusions.
Scale bar represents 50 μm in A and D; 25 μm in B, C and E-J.

161x305mm (300 x 300 DPI)

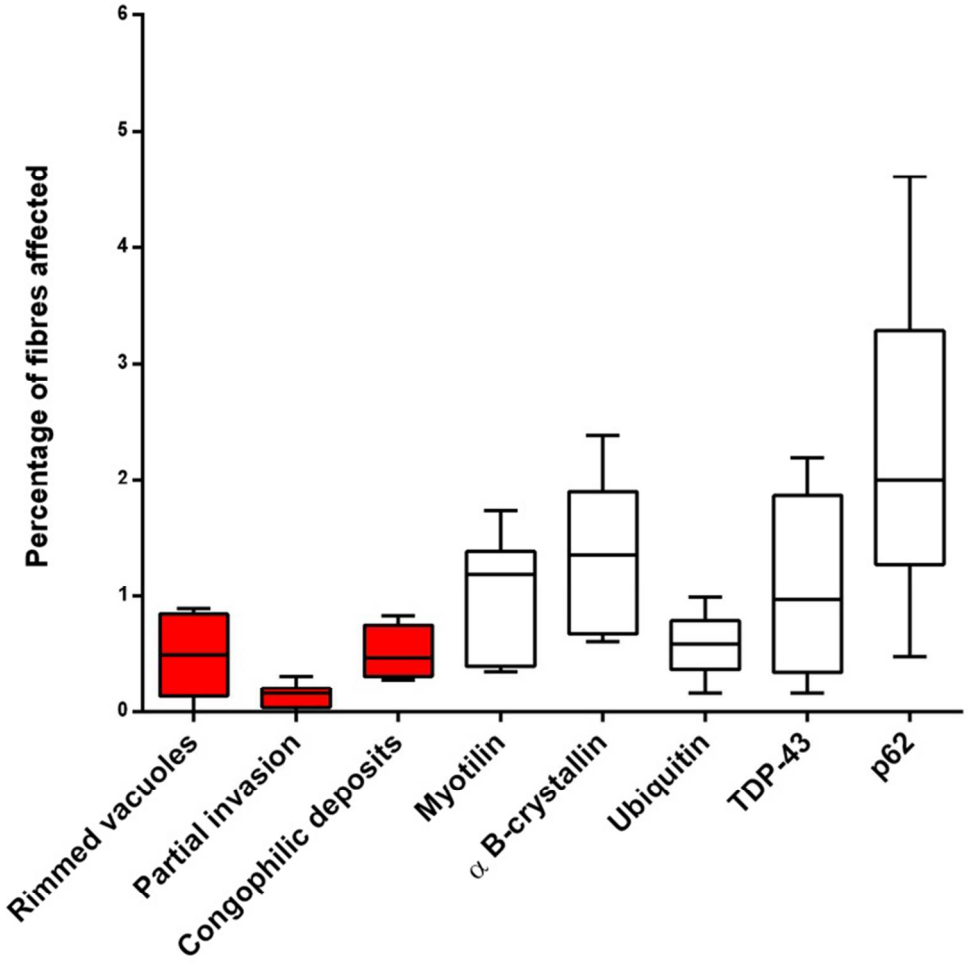


Figure 2 Percentage of muscle fibres containing protein aggregates and Griggs' pathological features
Box and whisker plot illustrating the percentage of muscle fibres containing pathological abnormalities contained in the Griggs criteria and protein aggregates in Griggs' pathologically-definite IBM. Fibres containing aggregates immunoreactive for p62 and αB-crystallin were more frequent than those containing the current diagnostic pathological features (red bars) ($p<0.05$). Protein aggregates recognised by all antibodies were found in a significantly larger number of fibres than partial invasion ($p<0.02$).

97x99mm (300 x 300 DPI)

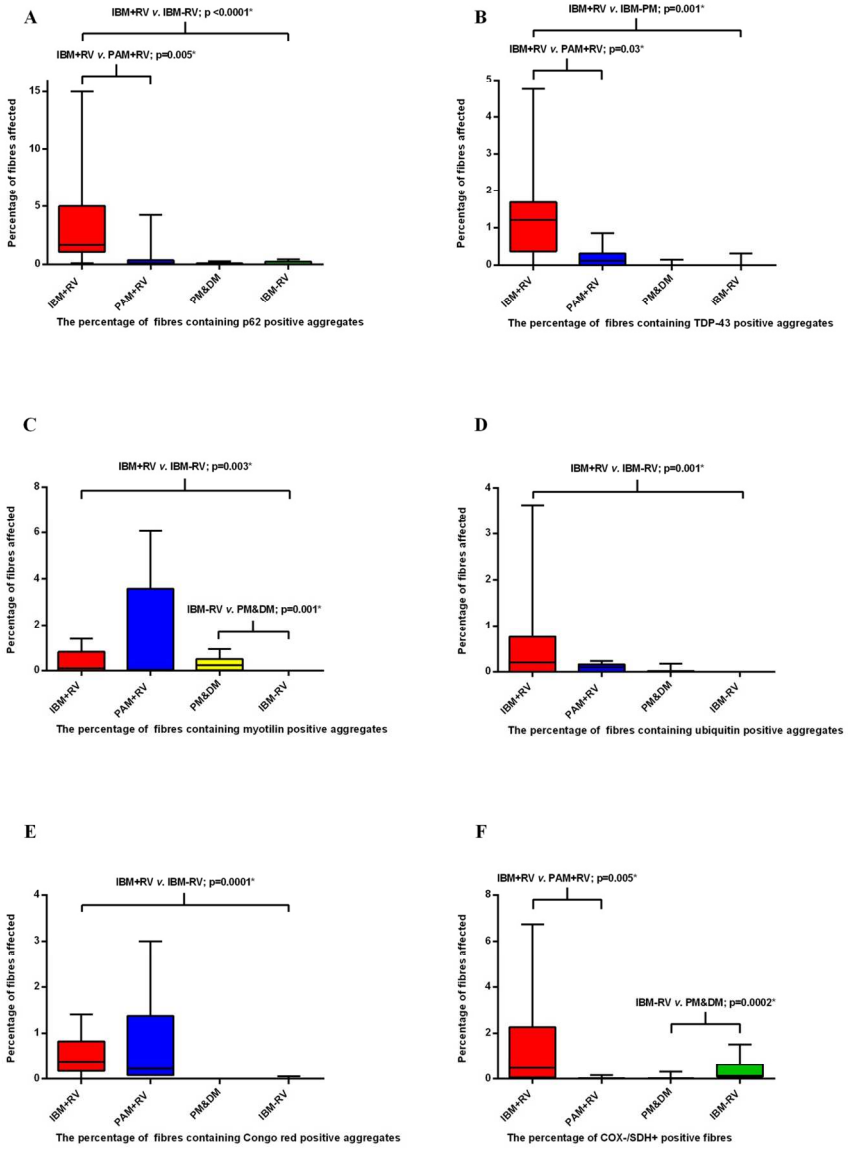


Figure 3 Percentage of fibres containing protein aggregates and COX-/SDH+ fibres in each group
Box and whisker plots illustrating the percentage of fibres in each diagnostic category containing p62 (A), TDP-43 (B), myotilin (C) and ubiquitin (D) immunoreactive aggregates, congophilic deposits (E) and COX-/SDH+ fibres (F). All protein aggregates were present in a greater percentage of fibres in IBM+RV than in IBM–RV. There was no difference in the percentage of COX-/SDH+ muscle fibres between these groups. IBM+RV biopsies had a greater percentage of fibres containing p62 (A) and TDP-43 (B) immunoreactive aggregates and COX-/SDH+ fibres (F) than PAM. Pathological findings were similar in IBM-RV and PM&DM, with no differences in the percentage of fibres containing p62 (A), TDP-43 (B) and ubiquitin (D) immunoreactive aggregates or congophilic deposits (E). However, there was a greater percentage of COX-/SDH+ fibres (F) in IBM–RV than PM&DM and a greater percentage of fibres containing myotilin immunoreactive aggregates (C) in PM&DM than IBM-RV. *Statistically significant results.

168x226mm (300 x 300 DPI)

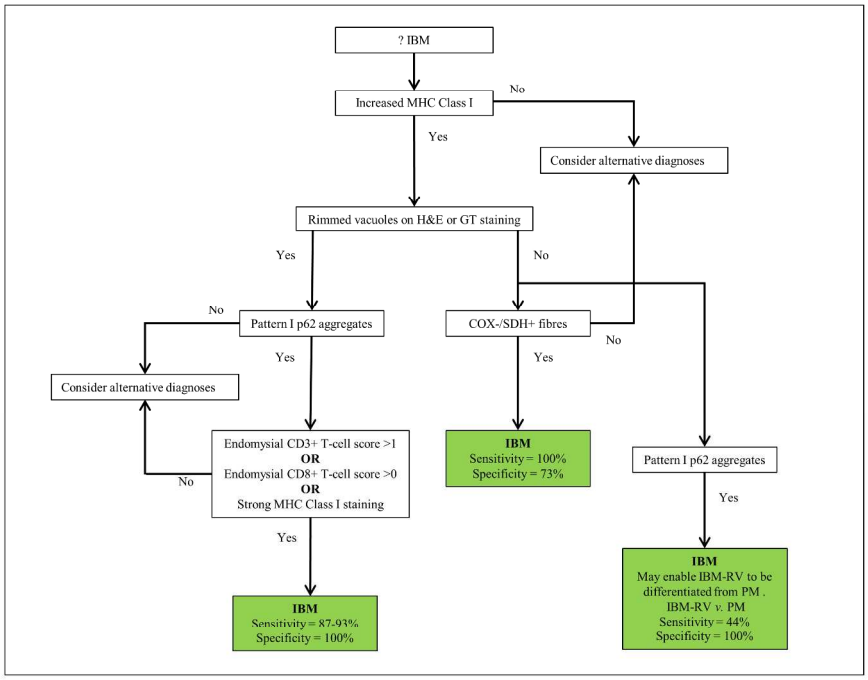


Figure 4 Proposed diagnostic algorithm for IBM based on pathological findings
Flow diagram showing a proposed pathway for diagnosing IBM based on the pathological findings. Increased MHC Class I staining was observed in all cases of IBM and pattern I p62 aggregates in all cases of IBM+RV making them good initial screening tests. Their absence rules-out a diagnosis of IBM and IBM+RV respectively. The presence of endomysial CD3+ T-cell score >1, endomysial CD8+ T-cell score >0 or strong MHC Class I staining in a biopsy with rimmed vacuoles and p62 aggregates secures a diagnosis of IBM+RV. Differentiating IBM-RV and PM&DM pathologically is more challenging. The presence of COX-/SDH+ fibres is not specific to IBM-RV; although COX-/SDH+ fibres were not present in every case of IBM-RV their absence casts doubt on the diagnosis of IBM-RV. Pattern I p62 aggregates may enable IBM to be differentiated from PM when present. However, they may lack sensitivity for IBM-RV, therefore their absence does not rule out the diagnosis.

254x190mm (300 x 300 DPI)

Supplementary Table 1 Clinical characteristics

Characteristic	G-IBM	IBM+RV	IBM-RV	PM&DM	PAM	IBM+RV*	IBM-RV*
Number of cases	6	15	9	11	7	12	11
Male:female	5:1	10:5	4:5	4:3	4:7	10:2	9:2
Median age at symptom onset, years (IQR)	69 (66-70)	54 (49-67)	62 (48-68)	55 (34-65)	46 (24-54)	58 (55-73)	60 (57-72)
Median age at muscle biopsy, years (IQR)	77 (68-78)	64 (59-71)	68 (47-74)	55 (34-65)	54 (29-59)	66 (62-77)	70 (63-74)
Median duration of symptoms, years (IQR)	5 (3-9)	5 (4-7)	3 (2-8)	0 (0-0)	5 (3-9)	5 (4-7)	4 (3-7)
Mean creatine kinase, IU/L, mean (±SD)	377 (±213)	1748 (±1348)	926 (±800)	6744 (±5875)	739 (±320)	662 (±360)	466 (±338)
Mean number of muscle fibres per biopsy	2929 (±1357)	1463 (±954)	1795 (±990)	3534 (±1934)	2749 (±1357)	NA	NA

G-IBM = Griggs' pathologically-definite IBM; IQR = Interquartile range; SD = Standard deviation; NA = Not applicable. *Cases used to test proposed diagnostic flow-chart.

Supplementary Table 2 Clinical inclusion criteria

Diagnostic category	Criteria
G-IBM	Patients fulfilling Griggs' definite criteria (rimmed vacuoles, inflammatory infiltrate with partial invasion of fibres and 15-18 nm tubulofilaments on EM) with prominent finger flexor and knee extensor weakness and CK <12 x ULN.
IBM+RV	Age at symptom onset >45 years, symptoms present for >12 months, finger flexion strength less than shoulder abduction strength and knee extension weakness greater than hip flexion weakness, CK ≤15 x ULN and a muscle biopsy revealing rimmed vacuoles on H&E or GT stained sections without features inconsistent with IBM on a standard diagnostic histological assessment for an inflammatory myopathy*.
IBM-RV	Clinical features and CK as detailed under IBM+RV. Rimmed vacuoles absent on H&E and GT stained sections and without features inconsistent with IBM on a standard diagnostic histological assessment for an inflammatory myopathy*.
PAM	Genetically or clinically and pathologically confirmed cases of PAM with typical rimmed vacuoles present on muscle biopsy and a genetically confirmed dystrophinopathy with typical rimmed vacuoles and protein aggregates present on muscle biopsy. Cases included myotilinopathy (n=2), hIBM with compound heterozygous mutations in GNE (n=1), IBMPFD with mutation in VCP (n=1), genetically unconfirmed cases of myofibrillar myopathy (n=2), and dystrophinopathy with deletion of exons 45-47 (n=1).
PM&DM	Subacute onset of limb girdle weakness, significantly raised CK, inflammatory cell infiltrate present on muscle biopsy and a sustained unequivocal clinical and biochemical response to steroid immunosuppression. DM cases also had to have cutaneous manifestations consistent with the diagnosis.
Normal controls	Patients investigated for cramps or fatigue, normal clinical examination performed by a muscle specialist, normal CK, normal neurophysiological assessment and normal muscle biopsy.

G-IBM = Griggs' pathologically-definite IBM; IBM+RV = Clinically-typical IBM with rimmed vacuoles; IBM-RV = Clinically typical IBM lacking rimmed vacuoles; PAM = Protein accumulation myopathies with rimmed vacuoles; PM&DM = Steroid-responsive inflammatory myopathies; hIBM = Hereditary inclusion body myopathy; IBMPFD = Inclusion body myopathy with Paget's disease and frontotemporal dementia; CK = Creatine kinase; GT = Gomori trichrome; ULN = Upper limit of normal. * Standard histological assessment for inflammatory myopathy includes H&E, GT, Sudan black or oil red O, periodic acid Schiff, nicotinamide adenine dinucleotide dehydrogenase, succinate dehydrogenase, cytochrome c oxidase, combined cytochrome c oxidase and succinate dehydrogenase, phosphorylase, acid and alkaline phosphatase, adenylate deaminase, ATPases at pH 4.2/4.3/9.4 and immunohistochemical staining including neonatal myosin, utrophin, major histocompatibility complex class I, membrane attack complex and a combination of inflammatory cell markers.

Supplementary Table 3 Antibodies and optimum staining conditions

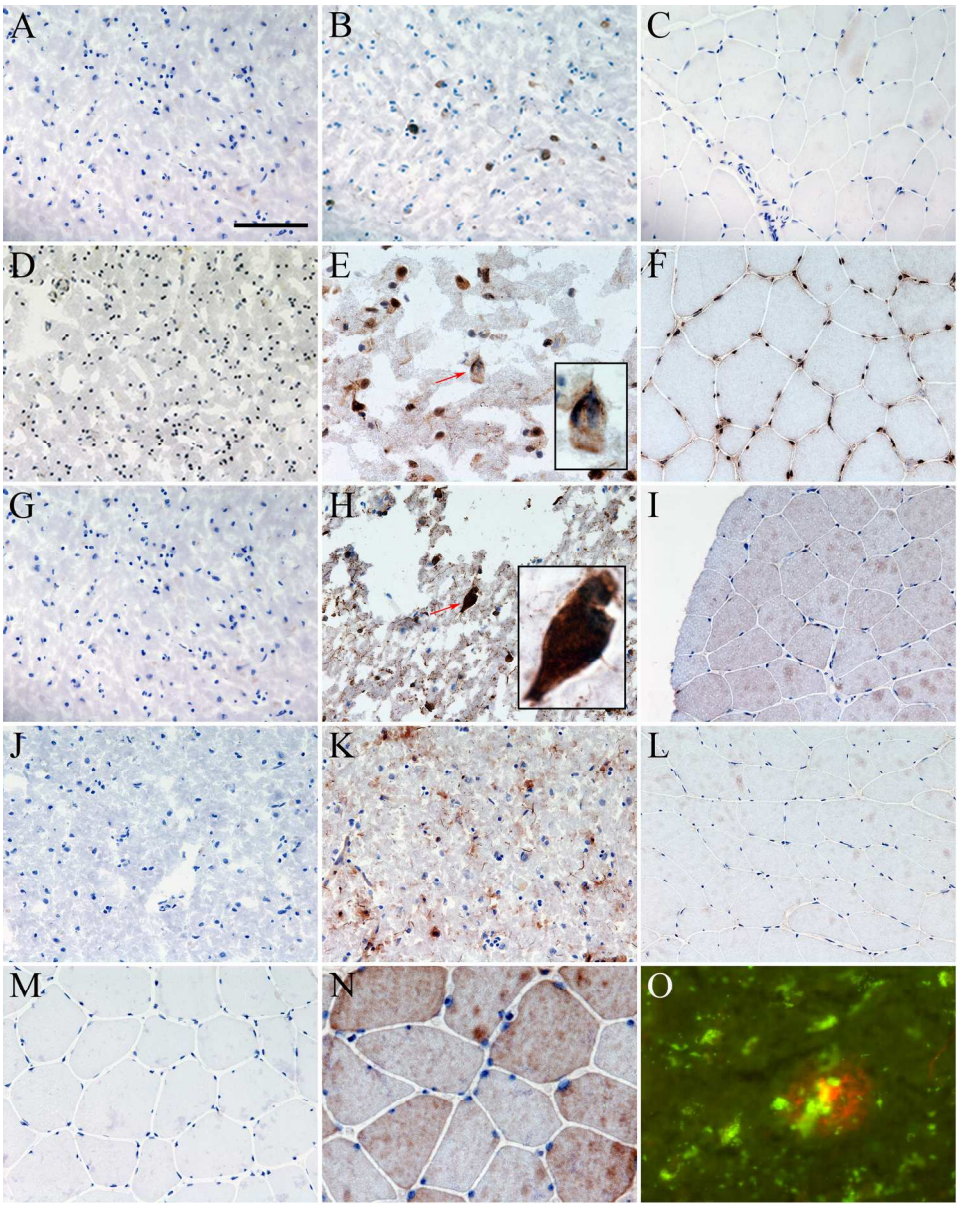
Antibody	Source	Clone	Control tissue	Fixative	Dilution	Primary incubation conditions†
p62	BD Transduction	3/P62	AD brain	A	1:400	1 hour, RT
TDP-43	Proteintech	NA	FTLD-TDP brain	PFA	1:800	24 hours, 4°C
Tau*	Dako	NA	AD brain	A	1:1600	1 hour, RT
Phosphorylated tau**	Autogen Bioclear	AT8	AD brain	A	1:1600	1 hour, RT
Ubiquitin	Dako	NA	AD brain	A	1:100	1 hour, RT
Aβ	Dako	6F/3D	AD brain	PFA and FA	1:100	1 hour, RT
α-synuclein	Abcam	4D6	MSA brain	PBS	1:800	1 hour, RT
FUS	Novus Biologicals	NA	FTLD-FUS brain	A	1:2000	1 hour, RT
Desmin	Dako	D33	Normal muscle	A	1:50	24 hours, 4°C
Myotilin	Novocastra	RSO34	Normal muscle	A	1:500	24 hours, 4°C
αB-crystallin	Novocastra	G2JF	CBD brain	A	1:300	1 hour, RT
VCP	Abcam	5	Normal muscle	A	1:100	1 hour, RT
Lamin A/C	Novocastra	636	Normal muscle	A	1:50	1 hour, RT
Emerin	Novocastra	4G5	Normal muscle	A	1:400	1 hour, RT
MHC Class I	Novocastra	W6/32	Normal muscle	A	1:25	24 hours, 4°C
CD3 (T-cells)	Novocastra	UCHT1	Tonsil	A	1:100	1 hour, RT
CD4 (Helper T-cells)	Novocastra	4B12	Tonsil	A	1:400	1 hour, RT
CD8 (Cytotoxic T-cells)	Novocastra	1A5	Tonsil	A	1:50	1 hour, RT
CD20 (B-cells)	Novocastra	L26	Tonsil	A	1:400	1 hour, RT
CD68 (Macrophages)	Novocastra	KP1	Tonsil	A	1:1600	1 hour, RT

NA = Not applicable; AD = Alzheimer's disease; FTLD-TDP = Frontotemporal lobar degeneration with TDP-43 positive inclusions; MSA = Multiple system atrophy; FTLD-FUS = Frontotemporal lobar degeneration with FUS positive inclusions; CBD = Corticobasal degeneration; A = Acetone; PFA = 4% Paraformaldehyde; FA = Formic acid; PBS = Phosphate buffered saline; RT = Room temperature. Antibodies were directed at * amino acids 243-441 irrespective of phosphorylation and ** phosphorylated Ser202. †Primary antibodies were made up in PBS and primary antibody-antigen binding was visualised with Dako REAL™ EnVision™ Detection System which includes a horseradish-peroxidase labelled goat anti-rabbit/mouse secondary and 1:50 solution of 3,3'-diaminobenzidine as the chromagen.

Supplementary Table 4 Definitions of pathological features

Pathological feature	Definition
Rimmed vacuoles	Irregular vacuole with a granular basophilic rim or containing granular basophilic material when stained with H&E or stained red in the GT. Both H&E and GT stained sections were reviewed before concluding the absence of rimmed vacuoles.
Inflammatory infiltrate and partial invasion	Inflammatory cells must show a nucleus fully circumscribed by a ring of positive staining. T-cells and B-cells must have a lymphoid morphology. Partial invasion was defined as unequivocal invasion of an otherwise structurally normal fibre by one or more inflammatory cells on H&E stained sections or sections stained using IHC.
Protein aggregates	Area of definite staining within a transversely orientated muscle fibre. Diffuse staining affecting the whole of a fibre was not counted nor were protein aggregates in necrotic fibres or regenerating fibres.
Congophilic deposits	Assessed using polarising and fluorescence microscopes. Positive staining using a polarising microscope was defined as congophilic deposits within a muscle fibre that exhibited apple-green birefringence under polarised light. Positive staining with a fluorescence microscope was defined as fluorescent material within a muscle fibre only visible under the rhodamine red filter. Areas of auto-fluorescence were excluded by visualising areas of fluorescence with both rhodamine red and FITC filters.

GT = Gomori trichrome; FITC = Fluorescein isothiocyanate.



Supplementary Figure 1 Control staining in brain and muscle tissue

Positive and negative (no primary) brain control sections and normal muscle stained using immunohistochemistry for: p62 (A-C), TDP-43 (D-F), α B-crystallin (G-I), ubiquitin (J-K) and myotilin (M,N) and alkalinised congo red (O).

(A-C) Negative (A) and positive (B) control sections of AD brain and normal muscle (C) stained for p62. Positive control shows p62 positive neurofibrillary tangles and dystrophic neurites (B). No p62 immunoreactivity is observed in normal muscle (C).

(D-F) Negative (D) and positive (E) control sections of FTLD-TDP brain and normal muscle (F) stained for TDP-43. Positive control shows normal nuclear labelling and mislocalised neuronal cytoplasmic staining with neuropil threads (E). Insert shows a neuron with absent nuclear TDP-43 and a cytoplasmic TDP-43 inclusion (E, red arrow and x100 insert). Nuclear TDP-43 staining is observed in normal muscle.

(G-I) Negative (G) and positive (H) control sections of CBD brain and normal muscle (I) stained for α B-crystallin. Positive control shows neuropil threads and a balloon cell neuron (H; red arrow and x100 insert).

No α B-crystallin immunoreactivity is observed in normal muscle (I).
(J-L) Negative (J) and positive (K) control AD brain and normal muscle (L) stained for ubiquitin. Positive control shows dystrophic neurites and neuropil threads (K). No ubiquitin immunoreactivity is observed in normal muscle (L).
(M,N) Negative (M) and positive (N) control muscle stained for myotilin. Mild sarcoplasmic staining is observed in normal muscle (N).
(O) Positive control section of AD brain showing an amyloid plaque (O).
Scale bar represents 100 μm in A-D, F and H-M; and 50 μm in E, N-O.
p62 = Sequestosome 1; AD = Alzheimer's disease; TDP-43 = Transactivation response DNA binding protein 43; FTLD-TDP = Frontotemporal lobar degeneration with TDP-43 positive inclusions; CBD = Corticobasal degeneration.

171x215mm (300 x 300 DPI)

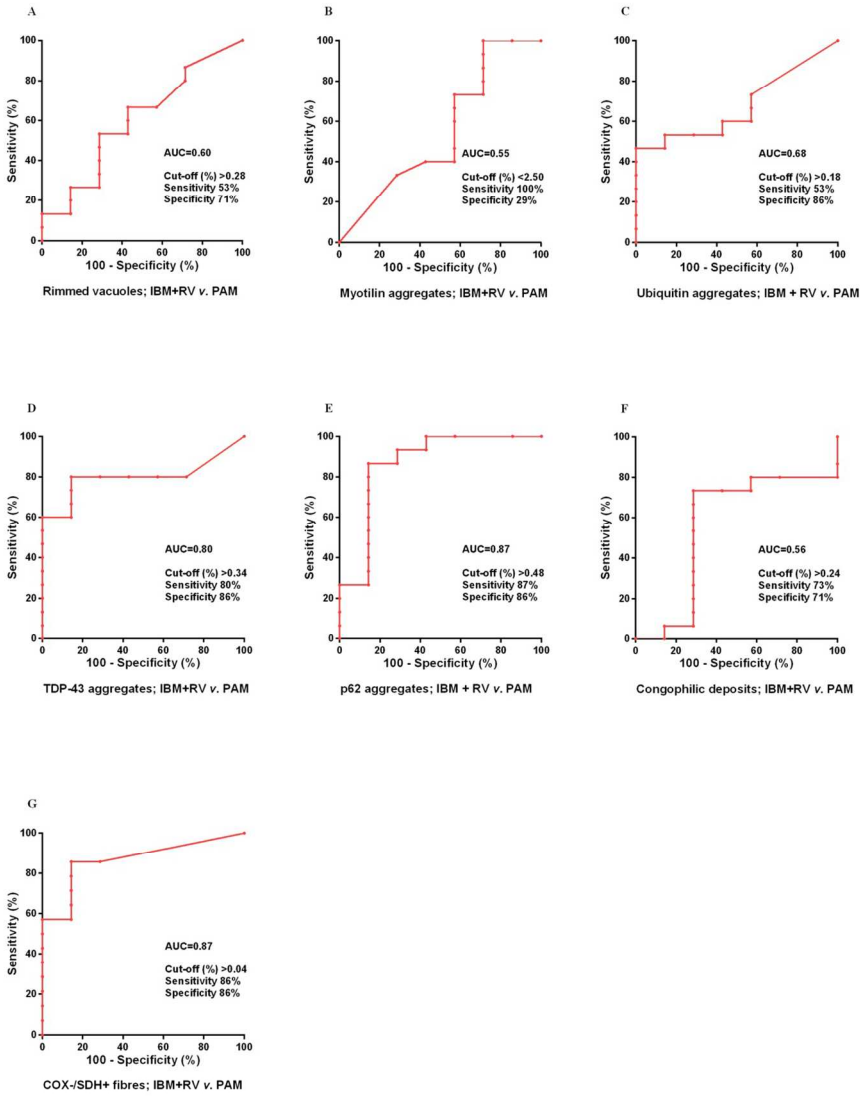
IBM Inflammatory Score Tool

Case Number:_____ Date:_____

	Score	Description
T-cells (CD3)		For each inflammatory cell type in the endomysial, perimysial and perivascular locations score positive infiltrating cells as follows: if none or <4 cells in a x20 field- score 0; if >4 cells in a x20 field and/or 1 cluster (where a cluster is ≥10 cells) - score 1; if >2 clusters in the entire biopsy, and/or diffusely infiltrating cells (i.e.> 20 cells in a x20 field) - score 2.
CD3+ endomysial infiltration	0, 1, 2	
CD3+ perimysial infiltration	0, 1, 2	
CD3+ perivascular infiltration	0, 1, 2	
Helper T-cells (CD4)		
CD4+ endomysial infiltration	0, 1, 2	
CD4+ perimysial infiltration	0, 1, 2	
CD4+ perivascular infiltration	0, 1, 2	
Cytotoxic T-cells (CD8)		
CD8+ endomysial infiltration	0, 1, 2	
CD8+ perimysial infiltration	0, 1, 2	
CD8+ perivascular infiltration	0, 1, 2	
B-cells (CD20)		
CD20+ endomysial infiltration	0, 1, 2	
CD20+ perimysial infiltration	0, 1, 2	
CD20+ perivascular infiltration	0, 1, 2	
Macrophages (CD68)		
CD68+ endomysial infiltration	0, 1, 2	
CD68+ perimysial infiltration	0, 1, 2	
CD68+ perivascular infiltration	0, 1, 2	
MHC Class I	0, 1, 2	For the whole biopsy score as follows: normal (capillary staining only) - score 0; if increased: i) mildly (weak diffuse sarcolemmal staining or scattered positive muscle fibres) - score 1; ii) strongly increased (diffuse definite sarcoplasmic and sarcolemmal increase in staining) score 2.

Supplementary Figure 2 IBM inflammatory score-tool
Score tool modified from the published juvenile dermatomyositis inflammatory (JDM) score tool [17] to specifically assess the type, degree and distribution of inflammation in IBM. The inflammatory domain was augmented to include T-cells, T-cell subtypes, B-cells and macrophages. MHC Class I staining was expanded to include three patterns of labelling. The vascular, muscle fibre and connective tissue domains which are present in the JDM score tool were not included.

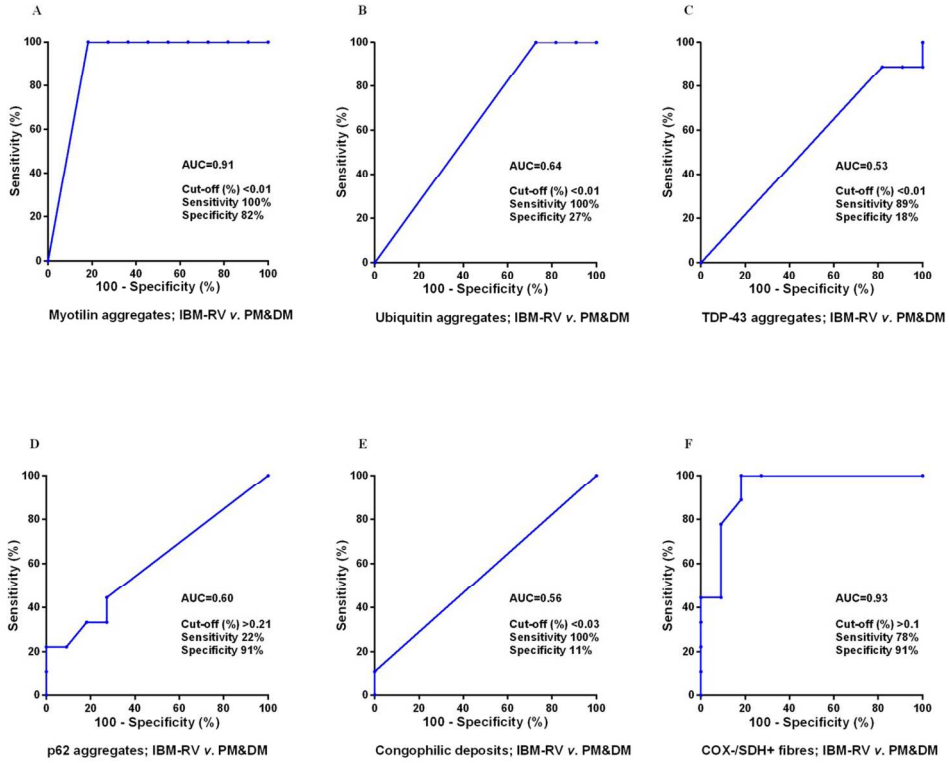
188x255mm (300 x 300 DPI)



Supplementary Figure 3 Sensitivity and specificity of rimmed vacuoles, protein aggregates and mitochondrial changes in IBM+RV compared to PAM

Receiver operating characteristic curves for each test including the area under the curve and optimum cut-off with its associated sensitivity and specificity for rimmed vacuoles (A), myotilin (B), ubiquitin (C), TDP-43 (D), p62 (E) immunoreactive deposits, congophilic deposits (F) and COX-/SDH+ fibres (G). COX/SDH HC staining was the most discriminative test for differentiating IBM+RV and PAM (G). However, there was little difference between COX/SDH HC staining, TDP-43 and p62 IHC staining and none were sufficiently discriminative to be considered diagnostic. AUC = Area under the curve.

165x211mm (300 x 300 DPI)



Supplementary Figure 4 Sensitivity and specificity of protein aggregates and mitochondrial changes in IBM-RV compared to PM&DM

Receiver operating characteristic curves for each test showing the area under the curve and optimum cut-off with its sensitivity and specificity for myotilin (A), ubiquitin (B), TDP-43 (C), p62 (D) immunoreactive deposits, congophilic deposits (E) and COX-/SDH+ fibres (F). COX/SDH histochemical staining (F) and myotilin (G) IHC were the most discriminative tests for differentiating IBM-RV and PM&DM. AUC = Area under the curve.

170x139mm (300 x 300 DPI)

STARD checklist for reporting of studies of diagnostic accuracy

(version January 2003)

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	Pg 1,2
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	Pg 2-4
METHODS			
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	Pg 4 and Supplementary Tables 1 and 2
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	Both. Pg 4 and Supplementary Table 2
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	Patients identified from clinics and systematic search of pathological databases
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	Retrospective study
Test methods	7	The reference standard and its rationale.	Clinical features and follow-up
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	Pg 4-6 and Supplementary Table 3
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	Pg 6
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	Two qualified medical doctors. Neuropathologist and Neurologist with an interest and significant experience in muscle pathology.
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	All analyses were blinded and performed in a random order. No clinical information was available at the time of analyses.
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	Pg 6 includes tests for determining diagnostic accuracy including 2x2 tables and ROC curves.
	13	Methods for calculating test reproducibility, if done.	Bland-Altman plots and Cohen's Kappa statistic used.
RESULTS			
Participants	14	When study was performed, including beginning and end dates of recruitment.	2011-2013
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	Included in Supplementary Table 1

	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	Not applicable. Retrospective study.
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	Study performed using tissue taken at the time of the reference standard
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	Diagnosese of control cases included Supplementary Table 2
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	Tables 1 and 2
	20	Any adverse events from performing the index tests or the reference standard.	Not applicable.
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	Included in Tables 1 and 2 and Supplementary Figures 2 and 3
	22	How indeterminate results, missing data and outliers of the index tests were handled.	Only one missing result and this is documented in Table 2. The denominator for calculating the proportion was altered to account for missing case in calculations
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	Included in statistical analysis Pg 6
	24	Estimates of test reproducibility, if done.	Included in statistical analysis Pg 6
DISCUSSION	25	Discuss the clinical applicability of the study findings.	Discussed in discussion Pg 12-15

Correction

Brady S, Squier W, Sewry C, *et al*. A retrospective cohort study identifying the principal pathological features useful in the diagnosis of inclusion body myositis. *BMJ Open* 2014;4:e004552. During the production process some errors to the values in table 2 were inadvertently introduced, specifically to the values in the rows entitled 'MHC class I upregulation' and 'Strong MHC class I upregulation', and columns 'IBM - RV n (%)', 'IBM - RV vs PM&DM Sensitivity' and 'IBM+RV vs IBM - RV p value.' The correct table is provided below.

BMJ Open 2014;**4**:e004552corr1. doi:10.1136/bmjopen-2013-004552corr1

Table 2 Comparison of the proportion of positive cases in each group

Pathological features	IBM+RV n (%)	PAM n (%)	IBM+RV vs PAM Sensitivity	IBM+RV vs PAM Specificity	IBM - RV n (%)	PM&DM n (%)	IBM - RV vs PM&DM Sensitivity	IBM - RV vs PM&DM Specificity	IBM+RV vs IBM - RV p value
Number of cases	15 (100)	7 (100)			9 (100)	11 (100)			
Aggregated proteins, n (%)									
p62	15 (100)	6 (86)	1.00	0.14	4 (44)	3 (27)*	0.44	0.73	0.003†
TDP-43	13 (87)	5 (71)	0.87	0.29	1 (11)	2 (18)*	0.11	0.82	0.001†
Ubiquitin	11 (73)	4 (57)	0.73	0.43	0 (0)	3 (27)*	0.00	0.73	0.001†
Myotilin	10 (67)	5 (71)	0.67	0.29	0 (0)	9 (82)	0.00	0.18	0.002†
Congophilic deposits	13 (87)	7 (100)	0.87	0.00	1 (11)	0 (0)	0.11	1.00	0.001†
COX−/SDH+ fibres‡, n (%)									
Any	12 (86)	2 (29)	0.86	0.71	9 (100)	3 (27)	1.00	0.73	0.50
Inflammatory features, n (%)									
MHC class I upregulation	15 (100)	3 (43)	1.00	0.57	9 (100)	11 (100)	1.00	0.00	1.00
Strong MHC class I upregulation	14 (93)	0 (0)	0.93	1.00	7 (78)	10 (91)	0.78	0.09	0.53
Partial invasion	10 (67)	0 (0)	0.67	1.00	3 (33)	2 (18)	0.33	0.82	0.21
Endomysial CD3 T-cell score >1	13 (87)	0 (0)	0.87	1.00	4 (44)	7 (64)	0.44	0.36	0.06
Endomysial CD4 T-cell score >1	12 (80)	0 (0)	0.80	1.00	2 (22)	5 (45)	0.22	0.54	0.01†
Endomysial CD8 T-cell score >0	14 (93)	0 (0)	0.93	1.00	7 (78)	9 (82)	0.78	0.18	0.53
Endomysial CD68 macrophage score >1	12 (80)	0 (0)	0.80	1.00	4 (44)	8 (73)	0.44	0.27	0.10

*Pathological features present in DM, but not PM cases.
†Statistically significant results.
‡In IBM with rimmed vacuoles n=14.
COX, cytochrome oxidase; IBM, inclusion body myositis; MHC, major histocompatibility complex; PAM, protein accumulation myopathies with rimmed vacuoles; PM&DM, steroid-responsive inflammatory myopathies; RV, rimmed vacuoles; SDH, succinate dehydrogenase; TDP-43, transactivation response DNA binding protein-43.