Arthritis & Rheumatology
DOI 10.1002/art.39753

MUSCLE BIOPSY IN COMBINATION WITH MYOSITIS-SPECIFIC AUTOANTIBODIES AIDS PREDICTION OF OUTCOMES IN JUVENILE DERMATOMYOSITIS

*Claire T. Deakin,[1] *Shireena A. Yasin,[1] Stefania Simou,[1] Katie A. Arnold,[1] Sarah L. Tansley,[2] Zoe E. Betteridge,[2] Neil J. McHugh,[2] Hemlata Varsani,[1] Janice L. Holton,[3] Thomas S. Jacques,[4] Clarissa A. Pilkington,[5] Kiran Nistala,[5,6,7] and Lucy R. Wedderburn[1,5,7,8] on behalf of the UK Juvenile Dermatomyositis Research Group

Corresponding author: Professor Lucy R. Wedderburn

Address: Infection, Immunity, and Inflammation, UCL Institute of Child Health

30 Guilford Street, London WC1N 1EH, United Kingdom

Telephone: +44 207 905 2391

Email: l.wedderburn@ucl.ac.uk

* These authors contributed equally to this work.

1. Infection, Immunity, and Inflammation Programme, UCL Institute of Child Health, London, UK
2. Department of Pharmacy and Pharmacology, University of Bath, Bath, UK
3. Department of Molecular Neuroscience, MRC Centre for Neuromuscular Diseases, UCL Institute of Neurology, London, UK

Received: Dec 17, 2015; Revised: Mar 03, 2016; Accepted: May 10, 2016

4. Developmental Biology and Cancer Programme, UCL Institute of Child Health, London, UK
5. Rheumatology Unit, Great Ormond Street Hospital for Children, London, UK
6. UCL Division of Medicine, London, UK
7. NIHR Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London
8. Arthritis Research UK Centre for Adolescent Rheumatology at UCL, UCLH and GOSH

Running title: Prediction of outcomes in juvenile myositis

Conflict of interest statement: The authors have declared that no conflict of interest exists.

ABSTRACT

Objective:

Juvenile dermatomyositis (JDM) is a rare and severe autoimmune condition characterized by rash and proximal muscle weakness. While some patients respond to standard treatment, others do not. We investigated whether histopathology and myositis-specific autoantibodies (MSA) have prognostic significance.

Methods:

Muscle biopsy samples (n=101) from the UK JDM Cohort and Biomarker Study were stained, analyzed and scored. Autoantibodies were measured (n=90) and longitudinal clinical data were collected (median follow-up 4.9 years). Long-term treatment status was modelled using generalized estimating equations.

Results:

Muscle biopsy scores differed according to MSA. When effects of MSA were accounted for, increased severity of muscle pathology predicted increased risk of remaining on treatment over time: 1.48-fold higher odds (1.12-1.96, p=0.0058) for the global pathology score (hVAS) and 1.10-fold higher odds (1.01-1.21, p=0.038) for the total biopsy score for the standardized score tool. A protective effect was identified in patients with anti-Mi2 autoantibodies, who had 7.06-fold lower odds (1.41-35.36, p=0.018) of remaining on treatment, despite displaying more severe muscle pathology at biopsy. For patients with anti-NXP-2, anti-TIF1γ autoantibodies or no-detectable autoantibody, increased severity of muscle pathology alone could predict the risk of remaining on treatment, without adjustment for MSA: 1.61-fold higher odds (1.16-2.22, p=0.0040) for hVAS and 1.13-fold higher odds (1.03-1.24, p=0.013) for total biopsy score.

Conclusion:

Muscle pathology, in combination with MSA, predicts the risk of remaining on treatment in JDM and may be useful for discussing probable treatment length with parents and patients. Understanding these associations may identify patients at greater risk of severe disease.

Accurate prediction of outcomes is a common problem in rare diseases. For many rare diseases, including juvenile myositis, patients and clinicians have an unmet need to predict poor outcomes. A further challenge in the study of rare autoimmune diseases is that disease mechanisms may be unknown, which renders biomarker research difficult. Juvenile dermatomyositis (JDM) is an example of a rare disease where the disease pathogenesis is only partially understood. A chronic autoimmune condition of childhood, JDM is typically characterized by proximal muscle weakness, elevated levels of muscle enzymes in serum, and rashes such as heliotrope rash and Gottron's papules (1). Other clinical features which contribute major morbidity include calcinosis, ulceration, treatment-resistant rash and involvement of the gut, lungs and brain. While some patients achieve remission following standard disease management, others fail to respond. In a recent long-term outcome study of 59 adults who had had JDM, with median follow-up of 16 years, 51% still had active disease (2). At present, early biomarkers of disease that are associated with long-term outcomes have not been identified.

To facilitate research into biological mechanisms, biomarkers and disease outcomes, the UK JDM Cohort and Biomarker Study (JDCBS; n=506 at time of writing) was established to collect serial clinical data and biospecimens (3). Such studies open the potential for the classification of rare diseases into sub-types defined by biomarkers and associated with predictable outcomes, and investigation of disease mechanisms that drive these sub-types. Biomarker research may eventually enable development of therapies directed against more relevant targets for particular subtypes, ultimately leading to better clinical outcomes. Myositis-specific autoantibodies (MSA) have been identified in both adult-onset dermatomyositis (DM) and JDM, and include anti-Mi2, anti-melanoma differentiation-associated gene 5 (MDA5), anti-transcriptional intermediary factor 1-γ (TIF1γ; p155/140) and anti-nuclear matrix protein (NXP-2; p140, also identified as the anti-MJ autoantibody).

Associations have been demonstrated between MSA and certain clinical features, suggesting that these autoantibodies may be useful biomarkers (4–11). However, little is known about the biological mechanisms underlying different MSA sub-types or how they relate to long-term prognosis.

We have previously developed and validated a standardized score tool to quantify abnormalities in JDM muscle biopsies (12,13). Use of immunohistochemistry to predict prognosis and inform treatment is well-established in more prevalent diseases such as malignancy. Here, we applied the standardized JDM score tool to a large cohort of biopsy samples (n=101) and tested the hypothesis that early JDM muscle pathology contains information predictive of long-term treatment status.

PATIENTS AND METHODS

Patients, biopsy material and clinical data

Pediatric patients with definite or probable JDM (14) were recruited to the UK JDCBS (n=506). Written informed parental consent and age appropriate assent were obtained from participants prior to inclusion in the study. This research was approved by the Northern & Yorkshire Medical Research and Ethics Committee (MREC), UK. All muscle biopsy samples available from the JDCBS (3) were analyzed where tissue was of sufficient quantity and quality (n=101). All tissue samples were obtained by open quadriceps biopsy under general anesthetic. Most of these patients (94.1%) were treated at Great Ormond Street Hospital for Children (GOSH), a major referral center, where the policy is to perform routine biopsy at time of diagnosis in patients with JDM. Consequently, a wide range of severities from mild to moderate are represented in the biopsied patients (Figure S1, A and B). Although the distribution of disease severity scores at diagnosis was more skewed towards increased

severity in those with a biopsy than cased who did not have a biopsy (Table S1, Figure S1,C and D), the unbiopsied patients were also more likely to have missing data at diagnosis and therefore these data are difficult to interpret.

Clinical data collected at diagnosis and biopsy included the physician's global assessment (PGA; range 0-10; low scores indicate minimal disease), Manual Muscle Testing and a Subset of Eight Muscles (MMT8; range 0-80; high scores indicate no muscle weakness) (15), Childhood Myositis Assessment Scale (CMAS; range 0-52; high scores indicate no weakness) (16), serum creatine kinase levels (units/L). Treatments received by patients were also recorded at each clinical visit. At diagnosis, all patients received methotrexate and the majority received concomitant steroids in agreement with international protocols (17). Where disease was unresponsive to treatment with methotrexate, other disease-modifying anti-rheumatic drugs (DMARDs) were used, including azathioprine, hydroxychloroquine, intravenous immunoglobulin and cyclophosphamide. For patients who still had refractory disease, anti-tumor necrosis factor biological agents (infliximab or adalimumab) were used. None of the analyzed patients were treated with cyclosporine A.

Histology and biopsy scoring

Histological staining, analysis and scoring of biopsy samples were conducted as described previously, using the validated JDM biopsy score tool to calculate a total biopsy score (12,13). The histopathologist's visual analogue scale score (hVAS) provides a global assessment of the severity of muscle pathology. Values for the total biopsy score (which includes assessment in 4 domains) and hVAS range from 0-27 and 0-10, respectively, with higher scores indicating more severe pathology. All histology and scoring were performed by a single observer (S.A.Y.), blind to the autoantibody status of each JDM case, trained by two highly qualified consultant neuropathologists who are experienced specialists in the field

(T.S.J. and J.L.H.), and who were involved in the development and validation of the JDM biopsy score tool (12,13). For the initial 9 biopsies, scores were firstly cross-compared with those of the two trainers and secondly with those generated by an international panel during the validation of the score tool (13), to ensure reliability. The intraclass correlation coefficient for the hVAS of the observer and those of the international panel for those 9 cases was 0.80 (0.62-0.95), indicating high levels of agreement.

Autoantibody screening

Serum or plasma were screened for autoantibodies as described previously using immunoprecipitation (5,9–11). Specificity for anti-NXP-2 or anti-MDA5 in patients with a 140 kDa band was determined by ELISA as described previously (5,11). Since recent literature has identified important associations between clinical features and MSA (4–11), and relatively fewer MAA cases were present in the biopsy cohort with low numbers in individual groups, we elected to focus on patient groups with sufficient frequency of MSAs to analyze i.e. anti-TIF1γ, anti-NXP-2, anti-MDA5 and anti-Mi2. Patients with MAA or unidentified bands were excluded during the statistical analyses of associations with muscle pathology and associations with muscle pathology and long-term outcomes. Patients with no-detectable autoantibody were included.

Statistics: data analysis and longitudinal modelling

Correlation between total biopsy scores and hVAS was analyzed using Spearman's rank correlation coefficient in R version 3.2.1 (18). A factorial analysis of variance (ANOVA) was conducted using the non-parametric Kruskal-Wallis test in R to identify significant main effects of MSA sub-groups on biopsy scores. Post-hoc comparisons to identify pairs of MSA sub-groups that significantly differed from each other were performed using R package dunn.test, with p-values adjusted using the Bonferroni method (19). For each medication, the

distribution of whether that drug was ever received by patients across each MSA sub-group was analyzed using Fisher's exact test in R, with p-values adjusted using the Bonferroni method.

A longitudinal modelling approach, which could include all available time-points for each patient, was adopted for the analysis of the treatment status outcome in order to make maximal use of the available serial clinical data. A longitudinal approach was preferred over a cross-sectional approach, which is limited to arbitrarily-selected time-points of interest and ignores any other time-points. Recurrent event analysis was preferred over time-to-event analysis, which is limited to time-points up to the first time patients come off treatment and ignores subsequent time-points when patients may come on treatment again. Longitudinal modelling of long-term treatment status was performed using generalized estimating equations (GEE), a longitudinal method for analyzing recurrent events which provides more conservative estimates for modelling binary outcomes than mixed-effects models (20–22). GEE models were fitted using R package geepack and an autoregressive correlation structure (23). Date of diagnosis was the zero time-point. Time from diagnosis was used as the time variable, which ensured that the effects of treatment duration were adjusted for. The no-detectable autoantibody group (n=20) was the reference category for the MSA variable to enable more precise estimates, since this group had the most patients. Although the biopsied patients are predominantly an inception cohort, a mixture of incident and prevalent patients were recruited when the JDCBS was started. For this reason, time from disease onset to diagnosis and time from diagnosis to biopsy were considered as potential confounders in longitudinal modelling. Since time from disease onset to diagnosis and time from diagnosis to biopsy were both found to have significant effects, these confounders were retained as covariates in subsequent analyses. Thus, all parameter estimates are adjusted for the effects of time from disease onset to diagnosis and time from diagnosis to biopsy. Additional

potentially confounding variables were evaluated, and included sex, whether steroids had been received before biopsy, and treatment ever with cyclophosphamide, but none of these had a significant effect in the model, so were not retained. Estimates of odds ratios are presented as odds of being on treatment, with 95% confidence intervals. Since odds ratios below 1 can be difficult to interpret, odds ratios below one are also presented as the odds of being off treatment.

Parameter estimates from the GEE models were used to formulate an equation to calculate the odds and hence predicted probability of being off treatment. To enable predicted probability to be plotted as a function of hVAS or total biopsy score, a fixed time-point of 5 years post-diagnosis was used. Median values for the time from onset to diagnosis and time from diagnosis to biopsy confounding variables were used. Plots were generated using a customized R function and the base plotting system.

Bivariate, univariate and null GEE models were compared using ANOVA for comparison of nested models, and R package MuMIn for calculation of the quasi-Akaike information criterion (QIC) and the proportion of weighting for the preferred model using the function model.sel() (24). ANOVA uses a χ^2 distribution to test the likelihood ratios of the models being compared. QIC is a measure of the relative quality of a GEE model, with lower values indicating an improved fit. For the model comparison analyses, the QIC values, proportion of weighting calculated by model.sel(), χ^2 statistics and p-values are reported.

For longitudinal modelling, p-values <0.05 were considered statistically significant. Summary statistics are presented as median and interquartile range for numeric variables, and as counts and percentages for categorical variables. Ninety-five per cent confidence intervals are presented for all estimated parameters. Figures depicting correlation and distribution of

biopsy scores and forest plot depictions of odds ratios were generated using GraphPad Prism 5.01 (GraphPad Software, San Diego, USA).

RESULTS

Demographic, clinical and serological features of the biopsy cohort

The 101 patients in this analysis were predominantly female, white and represented a range of disease severities (Table 1). The MSA analyzed in this study were detected in 58.9% of screened patients, including: anti-TIF1γ (20.0%), anti-NXP-2 (16.7%), anti-MDA5 (12.2%), and anti-Mi2 (5.6%). Myositis-associated autoantibodies (MAA) were detected in 10.0% of patients, unidentified bands were detected in 8.9% of patients and no autoantibodies were detected in 22.2% of patients.

Pathology scores are differentially distributed according to autoantibody status

The biopsy hVAS and total biopsy scores were highly correlated ($R=0.88$, $p<0.0001$; Figure 1A), indicating internal consistency of the tool. Biopsy hVAS and total biopsy scores included low and high scores and were not skewed towards either the more severe or milder ranges (Figure S1). Interestingly, there were clear differences in the distribution of both the hVAS and total biopsy scores between the major MSA sub-groups and the no-detectable autoantibody cases ($p=0.0005$, Figure 1B; and $p=0.0004$, Figure 1C). Anti-Mi2 and anti-MDA5 cases typically displayed severe and mild pathology, respectively. Anti-MDA5 cases had significantly lower hVAS and total biopsy scores than all other groups. Variable levels of severity were observed for the anti-NXP-2, anti-TIF1γ and no-detectable autoantibody groups.

Muscle pathology is associated with long-term treatment status and this effect is influenced by autoantibody status

We next investigated whether MSA and muscle pathology are associated with the long-term outcome of continued medication over time. The treatment status outcome was selected as an outcome of clinical importance to both patients and clinicians. Medications included the immunosuppressive, chemotherapeutic and biological agents detailed in the methods, and were not differently distributed across MSA sub-groups (Table S2). In the GEE models fitted with MSA and either hVAS or total biopsy score as covariates, both hVAS and total biopsy score had a significant effect on long-term treatment status (Figure 2). In the model fitted with hVAS and MSA as covariates, a unit increase in hVAS was associated with 1.48-fold higher odds (1.12-1.96; p=0.0058) of being on treatment over time (Figure 2A).

The overall pattern of the GEE model fitted with both MSA and total biopsy score as covariates was similar to the model fitted with both MSA and hVAS, although the magnitude of the effect sizes and statistical significance were smaller (Figure 2B). A unit increase in total biopsy score was associated with 1.10–fold higher odds (1.01-1.21; p=0.038) of being on treatment over time.

Anti-Mi2 cases have higher odds of coming off treatment

Interestingly, the anti-Mi2 antibody appeared to have a protective effect, with these patients having 7.06-fold lower odds (1.41-35.36, p=0.018) of remaining on treatment over time. This finding was counter-intuitive, since these cases had more severe muscle pathology (Figure 1B and Figure 1C). However, this estimate has wide 95% confidence intervals and warrants cautious interpretation. Anti-Mi2 had a borderline insignificant protective effect in the model fitted with total biopsy score as a covariate. In contrast to the anti-Mi2 cases, patients with the

anti-MDA5 antibody displayed a non-significant trend towards higher odds of remaining on treatment over time, despite having less severe muscle pathology at biopsy.

Muscle pathology or autoantibody status alone do not predict prognosis

To examine the question of whether biopsy hVAS, total biopsy score or MSA alone could predict long-term treatment status, univariate GEE models were fitted (Table 2). In these univariate models, none of hVAS, total biopsy score or MSA sub-groups had significant effects alone, even though these measures had significant effects in the bivariate models that included both muscle pathology and MSA. When the univariate models were compared to the bivariate models, the bivariate models were a better fit for the data (Table 3). Therefore, when all MSA cases were assessed, muscle pathology alone or MSA alone were not predictive of prognosis.

Muscle pathology is a better prognostic indicator than physician's global assessment at diagnosis

We also tested whether substituting muscle pathology score with PGA at diagnosis resulted in better prediction of treatment status. PGA at diagnosis did not have a statistically significant effect (Table 2), and this model was not a better fit than models with MSA and either hVAS or total biopsy score as covariates (Table 3).

For anti-NXP-2, anti-TIF1γ and autoantibody-negative cases, muscle pathology alone predicts long-term treatment status

Given the divergent effects of anti-Mi2 and anti-MDA5 in the GEE models fitted with MSA and muscle pathology scores (Figure 2), we reasoned that taking out these diametrically-opposed groups and any potentially overshadowing effects of those groups would enable further analysis of the anti-NXP-2, anti-TIF1γ and autoantibody-negative cases, which are the

most prevalent MSA groups. In this analysis, muscle pathology alone was associated with long-term treatment status (Figure 3). A unit increase in hVAS was associated with 1.61-fold higher odds (1.16-2.22, p=0.0040) of remaining on treatment over time (Figure 3A), while a unit increase in total biopsy score was associated with 1.13-fold higher odds (1.03-1.24, p=0.013) of remaining on treatment over time (Figure 3B). Inclusion of MSA as a covariate did not improve the fit (Table 3). Furthermore, the estimates for these univariate models where just the anti-NXP-2, anti-TIF1γ and autoantibody-negative cases were considered are similar to those when all MSA sub-groups were considered and the effect of MSA was accounted for (Figure 2). This indicates the equivalence of these sets of models and also the need to account for the effect of MSA when all MSA cases are considered.

Finally, to facilitate interpretation of these models, the predicted probability of being off treatment was plotted as a function of muscle pathology at a given time-point of 5 years post-diagnosis (Figure 3C and Figure 3D). These representations show the predicted probability of being off treatment at 5 years decreases as muscle pathology becomes more severe. For example, a patient with anti-NXP-2, anti-TIF1γ or no-detectable autoantibody and an hVAS score below 2 would have an over 50% probability of being off treatment 5 years after diagnosis. However, if the hVAS score is over 8, the estimated probability of being off treatment at 5 years after diagnosis is below just 6%.

DISCUSSION

Our study demonstrates that MSA are linked to muscle pathology in juvenile myositis. Furthermore, we show MSA influences the relationship between muscle pathology and long-term treatment status in JDM. Such knowledge may assist with identifying patients more likely to respond to treatment, versus those who are less likely to respond and may need more

aggressive treatment early in disease. This is also the first study to identify long-term clinical patterns in JDM patients with anti-Mi2 autoantibodies. It is intriguing that anti-Mi2 was associated with good prognosis despite being linked to severe muscle pathology, and also that there appeared to be an opposite trend for anti-MDA5. It may be that existing immunosuppressive therapies are more effective against the predominant muscle involvement that characterizes anti-Mi2 patients, but not for the anti-MDA5 patients, who have more extra-muscular features. For patients with anti-NXP-2, anti-TIF1γ or no-detectable autoantibodies, the severity of muscle pathology alone predicted the probability of remaining on treatment over time. Our results suggest treatment response is MSA-specific, implying distinct pathophysiology in MSA sub-groups. Therefore, there is a need for further research and consequently therapies targeting specific pathways identified as aberrant in these sub-types.

In addition to their usefulness for confirming diagnosis, our analysis shows that biopsies contain important information which, in combination with MSA, has prognostic significance. If our findings are replicated using larger patient numbers, performance of muscle biopsy routinely during diagnostic work-up may be justified in JDM. A recent study in adult DM also suggested that histopathology varies with MSA, but that study did not include analysis of MSA or histology against outcome (25). In other fields, such as breast cancer and glomerulonephritis, histology is used to classify heterogeneous disease into sub-types to inform optimal treatment regimens (26,27). Such stratified medicine approaches may have application to rare heterogeneous diseases like JDM.

In our analysis, the effect size and statistical significance of the hVAS were greater than those of the total biopsy score. Although this global pathology assessment correlates well with the standardized biopsy score, these 2 parts of the biopsy tool may measure pathology in different ways. The hVAS has more flexibility and sensitivity to give weight to

features that affect the severity of pathology, and are unaddressed by the specific items within the score tool. Even though the hVAS is based on the individual histopathologist's judgment, the hVAS was found to have high inter- and intra-observer reliability during the development and validation of the score tool (12,13).

Although this study used a large number of JDM biopsy samples (n=101), the relatively low numbers of patients within MSA sub-groups limits the precision of the GEE estimates for the MSA sub-groups, which have wide confidence intervals. For example, the protective effect identified for anti-Mi2 is based on just 5 patients and the estimate has a wide 95% confidence interval, although the statistical significance of the association nonetheless holds. Low numbers also restricted our ability to fit more complex models, such as allowing for interactions between MSA sub-groups and pathology score. Due to the low numbers of individual MSA sub-groups, we consider the most reasonable interpretation of our analysis to be that muscle pathology predicts long-term treatment status and that this effect is influenced by MSA. This finding is based on all MSA patients analyzed (n=69). Ideally, these findings should next be validated in an independent patient cohort, but at present there are few centers that routinely obtain muscle biopsies from JDM patients and to our knowledge, this study represents the largest JDM cohort in which biopsy data are linked to autoantibody status and up to 15 years of clinical data. As other JDM cohorts with biopsy data are built on, it will be important to use these to validate these findings. Low numbers is a challenge for any rare disease study, and the knowledge gained from this study highlights the importance of long-term biospecimen cohort studies in rare diseases. We also recognize that our findings cannot be extrapolated beyond the MSA groups analyzed, and further studies should examine the associations between MAA and pathology using greater numbers of patients.

A second limitation is that the treatment status outcome modelled in this study is linked only indirectly to disease pathology. Treatment status was selected as an outcome that

is meaningful for patients and clinicians, and which could be addressed using our dataset. It also fluctuates less than other outcomes measures we considered, such as "clinically insignificant disease" (28), and thus is more amenable to fitting complex longitudinal models. Defining appropriate outcome measures is still an active area of JDM research, and new measures will facilitate research into biomarkers and outcomes. Since treating clinicians were not all blinded to biopsy and MSA results, it is possible those findings could have influenced treatment practice, although the relationship between histology, MSA and outcomes is still at the research stage, and was not known to clinicians at the time treatment choices were made.

While we sought to include as many biopsy specimens as possible, in practice most of the biopsied patients were treated at one center (GOSH) and displayed a full range of disease severity scores at diagnosis. Since a typical overall range of disease severities is represented, our predictive model does accommodate a range of mild and severe patients. Importantly, a full range of severities of muscle pathologies are represented in the cohort analyzed for biopsy features. However, given that in the UK cohort as a whole those cases who had a muscle biopsy had more severe disease on average than those who did not have a biopsy, we acknowledge that this skew towards greater severity may limit the generalizability of our findings, until more centers generate further samples that represent the typical distribution of disease severity and also have known autoantibody status and longitudinal outcomes data on those cases. Nonetheless, our findings are internally valid with respect to the patients from whom biopsy, MSA status and longitudinal outcomes data are available at present.

In summary, we have shown that muscle pathology and autoantibody status are correlated and that muscle pathology, influenced by MSA status, predicts the probability of remaining on treatment in JDM. Understanding the link between these early biomarkers of disease and long-term outcomes may give further insight into different sub-phenotypes of disease and lead to more tailored therapies. Our biomarker-based modelling may well be

applied to adult forms of inflammatory myositis and may also be a useful approach to the analysis of other rare diseases.

ACKNOWLEDGEMENTS

Funding for the UK JDM Cohort and Biomarker study has been provided by generous grants from the Wellcome Trust UK [085860], Action Medical Research UK [SP4252], The Myositis Support Group UK, Arthritis Research UK [14518, 20164], The Henry Smith Charity and Great Ormond Street Children's Charity [V1268], and the National Institute for Health Research (NIHR) Translational Research Collaboration (TRC) Rare Diseases. This research was supported by the NIHR Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and Institute of Child Health University College London (UCL). The JDM Cohort study is adopted onto the NIHR Comprehensive Research Network. KN is in receipt of a Wellcome Trust Intermediate Clinical Fellowship [097259]. This research has also been supported by The Bath Institute for Rheumatic Diseases and the MRC Confidence in Concepts Scheme (University of Bath). The Arthritis Research UK Centre for Adolescent Rheumatology at UCL, UCL Hospital and GOSH is supported by grants from Arthritis Research UK [20164] and Great Ormond Street Children's Charity. This is a summary of independent research funded by the NIHR's Rare Diseases Translational Research Collaboration. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

The authors thank Dr Maria de Iorio, Department of Statistical Science, UCL and Professor Jugnoo Rahi, Population, Policy and Practice Programme, UCL Institute of Child Health, for statistical guidance.

The Juvenile Dermatomyositis Research Group would like to thank all of the patients and their families who contributed to the UK Juvenile Dermatomyositis Cohort and Biomarker Study. We thank all local research coordinators and principal investigators who have made this research possible. The members who contributed were as follows:

Dr Kate Armon, Mr Joe Ellis-Gage, Ms Holly Roper, Ms Vanja Briggs and Ms Joanna Watts (Norfolk and Norwich University Hospitals), Dr Liza McCann, Mr Ian Roberts, Dr Eileen Baildam, Ms Louise Hanna, Ms Olivia Lloyd and Susan Wadeson (The Royal Liverpool Children's Hospital, Alder Hey, Liverpool), Dr Phil Riley and Ms Ann McGovern (Royal Manchester Children's Hospital, Manchester), Dr Clive Ryder, Mrs. Janis Scott, Mrs. Beverley Thomas, Professor Taunton Southwood, Dr Eslam Al-Abadi (Birmingham Children's Hospital, Birmingham), Dr Sue Wyatt, Mrs Gillian Jackson, Dr Tania Amin, Dr Mark Wood, Dr Vanessa VanRooyen and Ms Deborah Burton (Leeds General Infirmary, Leeds), Dr Joyce Davidson, Dr Janet Gardner-Medwin, Dr Neil Martin, Ms Sue Ferguson, Ms Liz Waxman and Mr Michael Browne (The Royal Hospital for Sick Children, Yorkhill, Glasgow), Dr Mark Friswell, Professor Helen Foster, Mrs Alison Swift, Dr Sharmila Jandial, Ms Vicky Stevenson, Ms Debbie Wade, Dr Ethan Sen, Dr Eve Smith, Ms Lisa Qiao, Mr Stuart Watson and Ms Claire Duong (Great North Children's Hospital, Newcastle), Dr Helen Venning, Dr Rangaraj Satyapal, Mrs Elizabeth Stretton, Ms Mary Jordan, Dr Ellen Mosley, Ms Anna Frost, Ms Lindsay Crate, Dr Kishore Warrier and Stefanie Stafford (Queens Medical Centre, Nottingham), Professor Lucy Wedderburn, Dr Clarissa Pilkington, Dr Nathan Hasson, Mrs Sue Maillard, Ms Elizabeth Halkon, Ms Virginia Brown, Ms Audrey Juggins, Dr Sally Smith, Mrs Sian Lunt, Ms Elli Enayat, Mrs Hemlata Varsani, Miss Laura Kassoumeri, Miss Laura Beard, Miss Katie Arnold, Mrs Yvonne Glackin, Ms Stephanie Simou, Dr Beverley Almeida, Dr Kiran Nistala, Dr Raquel Marques, Dr Shireena Yasin, Dr Claire Deakin, Ms Stefanie Dowle, Ms Charis Papadopoulou (Great Ormond Street Hospital,

London), Dr Kevin Murray (Princess Margaret Hospital, Perth, Western Australia) Dr John Ioannou and Ms Linda Suffield (University College London Hospital, London) Dr Muthana Al-Obaidi, Ms Helen Lee, Ms Sam Leach, Ms Helen Smith, Dr Anne-Marie McMahon, Ms Heather Chisem and Ruth Kingshott (Sheffield's Children's Hospital, Sheffield); Dr Nick Wilkinson, Ms Emma Inness, Ms Eunice Kendall, Mr David Mayers, Ruth Etherton and Dr Kathryn Bailey (Oxford University Hospitals, Oxford); Dr Jacqui Clinch, Ms Natalie Fineman and Ms Helen Pluess-Hall (Bristol Royal Hospital for Children, Bristol); Ms Lindsay Vallance (Royal Aberdeen Children's Hospital); Ms Louise Akeroyd (Bradford Teaching Hospitals); Dr Alice Leahy, Amy Collier, Rebecca Cutts, Dr Hans De Graaf, Dr Brian Davidson, Sarah Hartfree, Danny Pratt (University Hospital Southampton).

The authors declare that they have no competing interests.

REFERENCES

1. Rider LG, Katz JD, Jones OY. Developments in the classification and treatment of the juvenile idiopathic inflammatory myopathies. *Rheum Dis Clin North Am* 2013;39:877–904.

2. Sanner H, Sjaastad I, Flatø B. Disease activity and prognostic factors in juvenile dermatomyositis: a long-term follow-up study applying the Paediatric Rheumatology International Trials Organization criteria for inactive disease and the myositis disease activity assessment tool. *Rheumatology (Oxford)* 2014;53:1578–85.

3. Martin N, Krol P, Smith S, Murray K, Pilkington CA, Davidson JE, et al. A national registry for juvenile dermatomyositis and other paediatric idiopathic inflammatory myopathies: 10 years' experience; the Juvenile Dermatomyositis National (UK and Ireland) Cohort Biomarker Study and Repository for Idiopathic Inflammatory Myopat. *Rheumatology (Oxford)* 2011;50:137–45.

4. Lega J-C, Fabien N, Reynaud Q, Durieu I, Durupt S, Dutertre M, et al. The clinical phenotype associated with myositis-specific and associated autoantibodies: a meta-analysis revisiting the so-called antisynthetase syndrome. *Autoimmun Rev* 2014;13:883–91.

5. Tansley SL, Betteridge ZE, Gunawardena H, Jacques TS, Owens CM, Pilkington C, et al. Anti-MDA5 autoantibodies in juvenile dermatomyositis identify a distinct clinical phenotype: a prospective cohort study. *Arthritis Res Ther* 2014;16:R138.

6. Rider LG, Shah M, Mamyrova G, Huber AM, Rice MM, Targoff IN, et al. The myositis autoantibody phenotypes of the juvenile idiopathic inflammatory myopathies. *Medicine (Baltimore)* 2013;92:223–43.

7. Kobayashi N, Takezaki S, Kobayashi I, Iwata N, Mori M, Nagai K, et al. Clinical and laboratory features of fatal rapidly progressive interstitial lung disease associated with

juvenile dermatomyositis. *Rheumatology (Oxford)* 2015;54:784–91.

8. Hoshino K, Muro Y, Sugiura K, Tomita Y, Nakashima R, Mimori T. Anti-MDA5 and anti-TIF1-gamma antibodies have clinical significance for patients with dermatomyositis. *Rheumatology (Oxford)* 2010;49:1726–33.

9. Gunawardena H, Wedderburn LR, North J, Betteridge Z, Dunphy J, Chinoy H, et al. Clinical associations of autoantibodies to a p155/140 kDa doublet protein in juvenile dermatomyositis. *Rheumatology (Oxford)* 2008;47:324–8.

10. Gunawardena H, Wedderburn LR, Chinoy H, Betteridge ZE, North J, Ollier WER, et al. Autoantibodies to a 140-kd protein in juvenile dermatomyositis are associated with calcinosis. *Arthritis Rheum* 2009;60:1807–14.

11. Tansley SL, Betteridge ZE, Shaddick G, Gunawardena H, Arnold K, Wedderburn LR, et al. Calcinosis in juvenile dermatomyositis is influenced by both anti-NXP2 autoantibody status and age at disease onset. *Rheumatology (Oxford)* 2014;53:2204–8.

12. 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.

13. Varsani H, Charman SC, Li CK, Marie SKN, Amato AA, Banwell B, et al. Validation of a score tool for measurement of histological severity in juvenile dermatomyositis and association with clinical severity of disease. *Ann Rheum Dis* 2015;74:204–10.

14. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med* 1975;292:344–7.

15. Rider LG, Koziol D, Giannini EH, Jain MS, Smith MR, Whitney-Mahoney K, et al. Validation of manual muscle testing and a subset of eight muscles for adult and juvenile idiopathic inflammatory myopathies. *Arthritis Care Res (Hoboken)* 2010;62:465–72.

16. Huber AM, Feldman BM, Rennebohm RM, Hicks JE, Lindsley CB, Perez MD, et al. Validation and clinical significance of the Childhood Myositis Assessment Scale for assessment of muscle function in the juvenile idiopathic inflammatory myopathies. *Arthritis Rheum* 2004;50:1595–603.

17. Huber AM, Giannini EH, Bowyer SL, Kim S, Lang B, Lindsley CB, et al. Protocols for the initial treatment of moderately severe Juvenile dermatomyositis: Results of a children's Arthritis and Rheumatology Research Alliance consensus conference. *Arthritis Care Res (Hoboken)* 2010;62:NA–NA.

18. Team RDC. R : A Language and Environment for Statistical Computing. Team RDC, ed. *Vienna Austria R Found Stat Comput* 2015:ISBN 3–900051–07–0. Available at: http://www.r-project.org.

19. Dinno A. dunn.test: Dunn's test of multiple comparisons using rank sums. 2015. Available at: http://cran.r-project.org/package=dunn.test.

20. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics* 1988;44:1049–60.

21. Twisk JWR, Smidt N, Vente W de. Applied analysis of recurrent events: a practical overview. *J Epidemiol Community Health* 2005;59:706–10.

22. Twisk JWR. *Applied Longitudinal Analysis for Epidemiology: A Practical Guide*. Second. Cambridge, UK: Cambridge University Press; 2003.

23. Hojsgaard S, Halekoh U, Yan J. The R Package geepack for Generalized Estimating Equations. *J Stat Softw* 2006;15:1–11.

24. Barton K. MuMIn: Multi-model inference. 2015. Available at: http://cran.r-project.org/package=MuMIn.

25. Pinal-Fernandez I, Casciola-Rosen LA, Christopher-Stine L, Corse AM, Mammen AL. The prevalence of individual histopathologic features varies according to autoantibody status in muscle biopsies from patients with dermatomyositis. *J Rheumatol* 2015;42:1448–54.

26. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, et al. Tailoring therapies - improving the management of early breast cancer: St GallenInternational Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol* 2015;26:1533–46.

27. Weening JJ, D'Agati VD, Schwartz MM, Seshan S V, Alpers CE, Appel GB, et al. The classification of glomerulonephritis in systemic lupus erythematosus revisited. *Kidney Int* 2004;65:521–30.

28. Ruperto N, Pistorio A, Ravelli A, Hasija R, Guseinova D, Filocamo G, et al. Criteria to define response to therapy in paediatric rheumatic diseases. *Eur J Clin Pharmacol* 2011;67 Suppl 1:125–31.

Figure 1. Distributions of total biopsy scores and histopathologist's global pathology scores. (A) Correlation of total biopsy score and hVAS, including Spearman R correlation with 95% confidence interval (n=101). Distribution of **(B)** hVAS and **(C)** total biopsy score across MSA groups (n=69). Factorial ANOVA using the Kruskal-Wallis test was performed to analyze the distribution of these scores. There was a significant main effect of MSA on hVAS (χ^2 (4, n=69) = 20.0, p=0.0005), with significant differences in hVAS between the anti-MDA5 and anti-Mi2 (p=0.0001), anti-MDA5 and anti-NXP-2 (p=0.007), anti-MDA5 and anti-TIF1γ (p=0.04), and anti-MDA5 and no-detectable autoantibody (p=0.03) groups. There was also a significant main effect of MSA on total biopsy score (χ^2 (4, n=69) = 20.4, p=0.0004), with significant differences in total biopsy score between the anti-MDA5 and anti-Mi2 (p=0.0009), anti-MDA5 and anti-NXP-2 (p=0.0006), anti-MDA5 and anti-TIF1γ (p=0.01), and anti-MDA5 and no-detectable autoantibody (p=0.04) groups. ANOVA, analysis of variance; hVAS, histopathologist's visual analogue scale global pathology score; MSA, myositis-specific autoantibody.

Figure 2. Longitudinal GEE modelling of treatment status over time according to myositis-specific autoantibody sub-groups and global muscle pathology score or total biopsy score. Forest plots depicting odds ratios with 95% confidence intervals for being on treatment estimated by GEE models fitted with MSA groups and either **(A)** hVAS or **(B)** total biopsy score as predictors. The no-detectable autoantibody group was used as the reference category. GEE, generalized estimating equations; hVAS, histopathologist's visual analogue scale global pathology score; MSA, myositis-specific autoantibody.

Figure 3. Longitudinal GEE models of the association between biopsy score and treatment status over time for selected MSA groups. Forest plots depicting odds ratios with 95% confidence intervals estimated by GEE models fitted with either **(A)** hVAS or **(B)**

total biopsy score as predictors. The no-detectable autoantibody, anti-NXP-2 and anti-TIF1γ MSA groups were included in these analyses. Predicted probability of being off treatment at 5 years post-diagnosis as a function of either **(C)** hVAS or **(D)** total biopsy score for no-detectable autoantibody, anti-NXP-2 and anti-TIF1γ MSA groups, derived from the GEE models. Dotted lines represent 95% confidence intervals. The median values for time from onset to diagnosis (0.214 years) and for time from diagnosis to biopsy (0.0602 years) were used in the calculations of predicted probabilities. GEE, generalized estimating equations; hVAS, histopathologist's visual analogue scale global pathology score; MSA, myositis-specific autoantibody.

Table 1. Demographic, clinical and serological features of the biopsy cohort (n=101)

Feature	Summary statistic
Sex, n (%)	
Male	33 (32.7%)
Female	68 (67.3%)
Ethnicity, n (%)	
White	72 (71.3%)
Black	12 (11.9%)
South Asian	8 (7.9%)
Other	9 (8.9%)
Clinical features at biopsy, median [interquartile range][a]	
Age at onset (years)	6.1 [3.9-9.3]
Physician's global assessment (PGA)	4.1 [2.0-7.0]
MMT8	55.0 [40.0-71.5]
CMAS	29 [18.75-45]
Creatine kinase (units/L)	213 [55-1019]
Clinical features at biopsy, median [interquartile range]	
Time from disease onset to diagnosis, (months)	2.6 [1.5-7.5]
Time from diagnosis to biopsy, (months)	0.72 [0.39-0.92]
Biopsy performed > 1 month after diagnosis, n (%)	17 (16.8%)
On steroids at biopsy, n (%)[b]	12 (12.2%)
Myositis-specific autoantibodies, n (%)[c]	**53 (58.9%)**
Anti-TIF1γ	18 (20.0%)
Anti-NXP-2	15 (16.7%)

Anti-MDA5	11 (12.2%)
Anti-Mi2	5 (5.6%)
Anti-SRP	2 (2.2%)
Anti-PL7	1 (1.1%)
Anti-SAE	1 (1.1%)
Myositis-associated autoantibodies, n (%)	**9 (10.0%)**
Anti-PM-Scl	6 (6.7%)
Anti-U1RNP	2 (2.2%)
Anti-Topo	1 (1.1%)
Unidentified autoantibodies, n (%)	**8 (8.9%)**
No-detectable autoantibodies, n (%)	**20 (22.2%)**

[a]Clinical features were missing for some patients: n=11 (PGA), n=42 (MMT8), n=17 (CMAS), n=30 (creatine kinase).

[b]Steroids not recorded at the biopsy time-point for 3 individuals (3.0%).

[c]Autoantibodies were screened for in 90 biopsied patients. Percentages reflect the number of patients with a given antibody as a proportion of total tested patients.

Table 2. Summary of alternative GEE models

Predictor Variable	Odds Ratio	95% Confidence Interval	P-value
Univariate models for hVAS, total biopsy score and MSA (n=69)			
hVAS	1.10	0.92 – 1.31	0.28
Total biopsy score	1.03	0.96 – 1.10	0.43
MSA			
No-detectable (n=20)	1.00	-	-
Anti-MDA5 (n=11)	1.69	0.38 – 7.60	0.50
Anti-NXP-2 (n=16)	1.61	0.41 – 6.36	0.50
Anti-TIF1γ (n=17)	2.06	0.46 – 9.28	0.35
Anti-Mi2 (n=5)	0.68	0.24 – 1.90	0.46
Bivariate model with PGA and MSA (n=44)[a]			
PGA	1.27	0.92 – 1.76	0.15
MSA			
No-detectable (n=10)	1.00	-	-
Anti-MDA5 (n=9)	1.56	0.22 – 11.00	0.65
Anti-NXP-2 (n=9)	0.44	0.08 – 2.55	0.36
Anti-TIF1γ (n=12)	1.51	0.20 – 11.16	0.69
Anti-Mi2 (n=4)	0.78	0.09 – 7.00	0.83

GEE, generalized estimating equations; hVAS, histopathologist's visual analogue scale global pathology score; MSA, myositis-specific autoantibody; PGA, physician's global assessment.

[a]PGA at diagnosis was available for n=44 patients.

Table 3. Summary of model comparisons

	QIC[a]	Model selection weight[b]	ANOVA[c] χ^2	ANOVA p-value
Comparison of bivariate models (hVAS and MSA or Total biopsy score and MSA) to nested univariate and null models, for models fitted with all MSA patients (n=69)				
Bivariate (hVAS and MSA)	315	-	-	-
Univariate (hVAS only)	349	1	10.2 (4)	* 0.038
Univariate (MSA only)	355	1	7.6 (1)	** 0.0058
Null (Time only)	350	1	10.5 (5)	0.063
Bivariate (Total biopsy score and MSA)	336	-	-	-
Univariate (Total biopsy score only)	351	0.999	8.6 (4)	0.073
Univariate (MSA only)	355	1	4.3 (1)	* 0.038
Null (Time only)	350	0.999	8.6 (5)	0.13
Comparison of bivariate model fitted with MSA and PGA to bivariate models fitted with MSA and either hVAS or total biopsy score (n=44)[d]				
PGA and MSA	316			
hVAS and MSA	263	0	-	-
Total biopsy score and MSA	293	0	-	-
Comparison of univariate models to bivariate and null models, for models fitted with				

anti-NXP-2, anti-TIF1γ and no-detectable MSA patients (n=52)				
Univariate (hVAS)	203	-	-	-
Bivariate (hVAS and MSA)	199	0.85	2.0 (2)	0.36
Null (Time)	247	1	8.3 (1)	** 0.004
Univariate (Total biopsy score)	228	-	-	-
Bivariate (Total biopsy score and MSA)	235	0.96	0.7 (2)	0.71
Null (Time)	247	1	6.2 (1)	* 0.013

ANOVA, analysis of variance; hVAS, histopathologist's visual analogue scale global pathology score; MSA, myositis-specific autoantibody; PGA, physician's global assessment at diagnosis; QIC, quasi-Akaike information criterion.

[a]The quasi-Akaike information criterion is a measurement of the relative quality of the GEE models. Models with lower values indicate a better fit.

[b]Model selection weight representing the proportion of weight to be given to the bivariate models as compared to their respective nested univariate model (rows 2-4, 6-8, 10-11, 13-14), or the model with PGA and MSA as compared to the models with biopsy score and MSA (rows 16-17), on a scale of 0-1, when the bivariate, univariate and null models are compared as indicated. Values of or close to 1 indicate the preferred model.

[c]ANOVA comparison of bivariate model to nested or null models, with degrees of freedom given in parentheses. The ANOVA tests for a reduction in residual sum of squares, with p-values below 0.05 indicating a significantly improved fit for the data.

[d]PGA at biopsy was available for n=44 patients. For the purpose of these model comparisons, GEE models with MSA and hVAS or MSA and total biopsy score were fitted on the equivalent dataset.

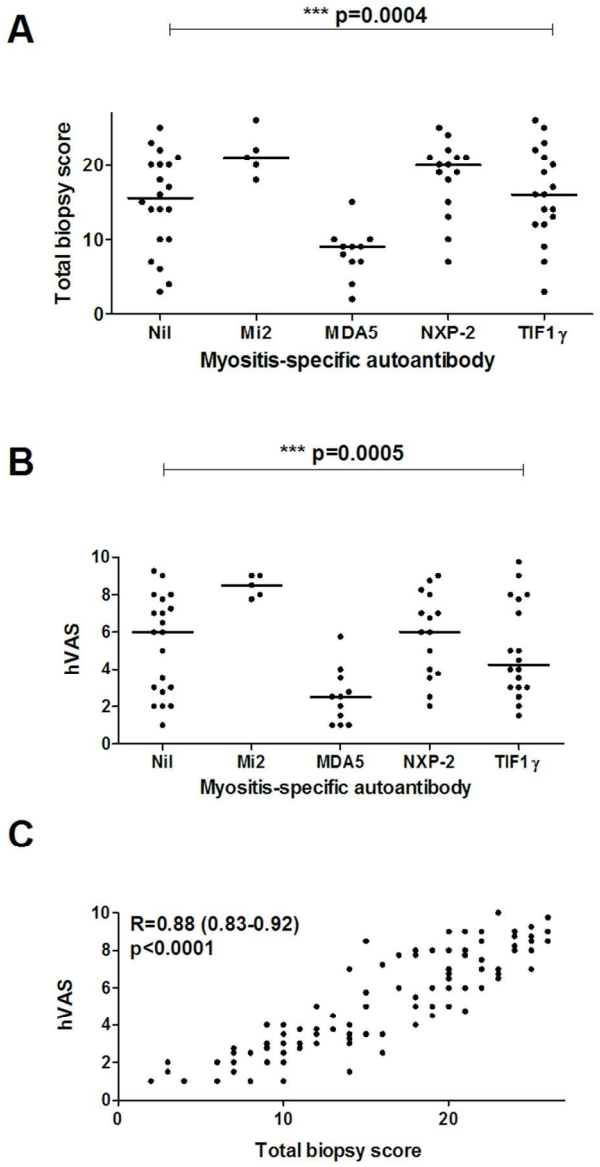

Figure 1. Distributions of total biopsy scores and histopathologist's global pathology scores. (A) Correlation of total biopsy score and hVAS, including Spearman R correlation with 95% confidence interval (n=101). Distribution of (B) hVAS and (C) total biopsy score across MSA groups (n=69). Factorial ANOVA using the Kruskal-Wallis test was performed to analyze the distribution of these scores. There was a significant main effect of MSA on hVAS ($\chi 2$ (4, n=69) = 20.0, p=0.0005), with significant differences in hVAS between the anti-MDA5 and anti-Mi2 (p=0.0001), anti-MDA5 and anti-NXP-2 (p=0.007), anti-MDA5 and anti-TIF1γ (p=0.04), and anti-MDA5 and no-detectable autoantibody (p=0.03) groups. There was also a significant main effect of MSA on total biopsy score ($\chi 2$ (4, n=69) = 20.4, p=0.0004), with significant differences in total biopsy score between the anti-MDA5 and anti-Mi2 (p=0.0009), anti-MDA5 and anti-NXP-2 (p=0.0006), anti-MDA5 and anti-TIF1γ (p=0.01), and anti-MDA5 and no-detectable autoantibody (p=0.04) groups. ANOVA, analysis of variance; hVAS, histopathologist's visual analogue scale global pathology score; MSA, myositis-specific autoantibody.
82x160mm (300 x 300 DPI)

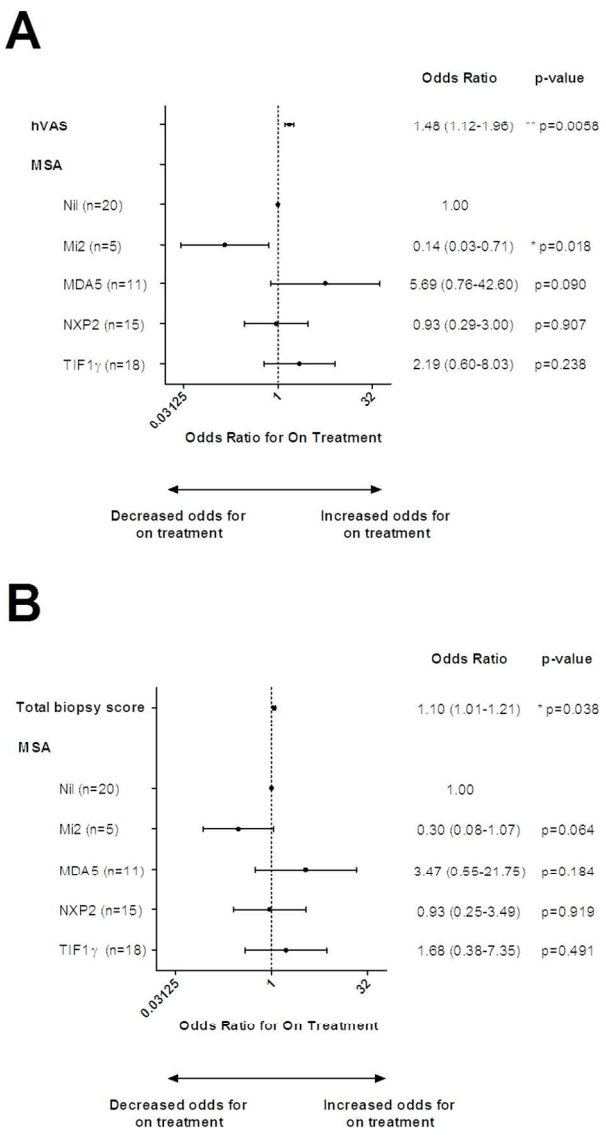

Figure 2. Longitudinal GEE modelling of treatment status over time according to myositis-specific autoantibody sub-groups and global muscle pathology score or total biopsy score. Forest plots depicting odds ratios with 95% confidence intervals for being on treatment estimated by GEE models fitted with MSA groups and either (A) hVAS or (B) total biopsy score as predictors. The no-detectable autoantibody group was used as the reference category. GEE, generalized estimating equations; hVAS, histopathologist's visual analogue scale global pathology score; MSA, myositis-specific autoantibody.
82x150mm (300 x 300 DPI)

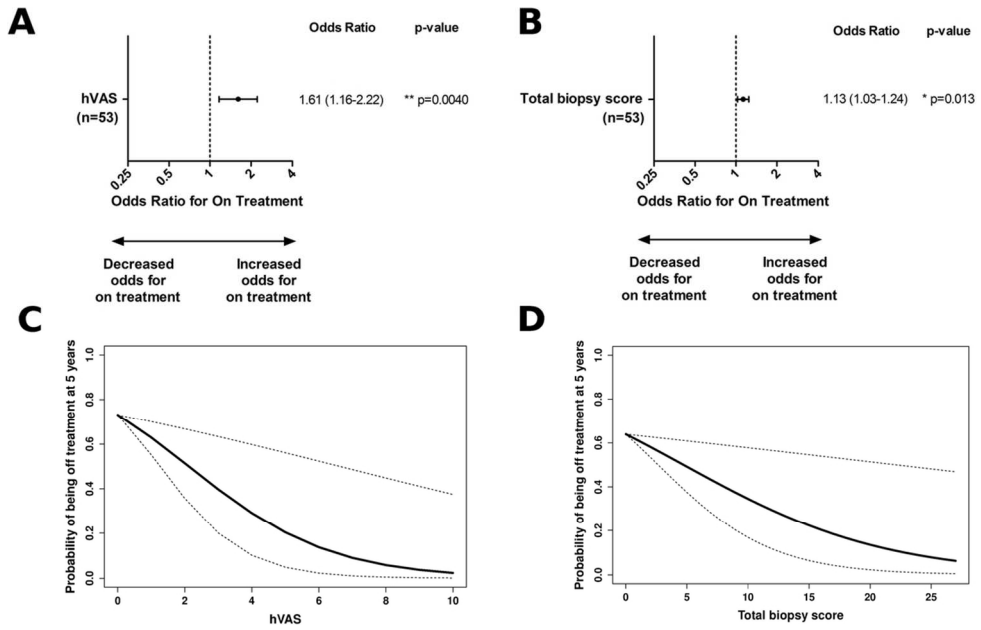

Figure 3. Longitudinal GEE models of the association between biopsy score and treatment status over time for selected MSA groups. Forest plots depicting odds ratios with 95% confidence intervals estimated by GEE models fitted with either (A) hVAS or (B) total biopsy score as predictors. The no-detectable autoantibody, anti-NXP-2 and anti-TIF1γ MSA groups were included in these analyses. Predicted probability of being off treatment at 5 years post-diagnosis as a function of either (C) hVAS or (D) total biopsy score for no-detectable autoantibody, anti-NXP-2 and anti-TIF1γ MSA groups, derived from the GEE models. Dotted lines represent 95% confidence intervals. The median values for time from onset to diagnosis (0.214 years) and for time from diagnosis to biopsy (0.0602 years) were used in the calculations of predicted probabilities. GEE, generalized estimating equations; hVAS, histopathologist's visual analogue scale global pathology score; MSA, myositis-specific autoantibody.
107x66mm (300 x 300 DPI)