Journal of Methods and Measurement in the Social Sciences
Vol. 1, No. 1, 8-18, 2010

Measuring disease severity in Duchenne and Becker muscular dystrophy

Melinda F. Davis, Katalin Scherer,
Timothy M. Miller, and F. John Meaney
University of Arizona

Medical investigations use a wide variety of outcome indicators that are often not comparable. It can be challenging to integrate results across multiple studies that do not share a common metric. Some conditions such as Duchenne and Becker muscular dystrophy have a predictable course of disease progression. Severity can be inferred from a patient's medical history. This paper describes the development of a disease severity measure using common markers of disease progression. Rasch modeling was used to estimate severity using dichotomous events that indicate disease progression. Caregivers of 34 young men with Duchenne or Becker muscular dystrophy completed structured interviews about their care and medical history. Interview questions included surgeries (tendon release, scoliosis, tracheostomy), respiratory equipment (assisted ventilation, cough assist devices), and the use of other medical equipment (e.g., braces, walkers, wheelchairs, transfer boards, hospital beds). The resulting measure had a reliability of .83. The correlation between the severity measure and the Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS) was .68. Preliminary results and item calibrations are provided for the severity measure that can be estimated from caregiver reports or administrative data.

Keywords: Measurement, Method, Item response theory, Muscular dystrophy, Disease progression

Measurement is fundamental to medical research. Evaluating the effect of treatments depends on reliable and valid measures of study outcomes. Progress in medical research builds upon the results of multiple studies. Unfortunately, clinical investigations use a wide variety of outcomes that may, or may not be comparable. Studies may use indicators that are dichotomous (presence/absence) or outcome scales that have never been equated to each other. As a result, a study tells us whether or not a treatment had an effect on a particular outcome, but it can be extremely challenging to integrate the results of multiple investigations.

Some medical conditions have a highly predictable disease progression. Clinicians can readily characterize a patient's severity by examining his or her medical chart. In the same way, these symptoms can be used to create a disease severity scale using Rasch modeling procedures. In this paper, we demonstrate the creation of measure of disease severity for Duchenne and Becker muscular dystrophy (DBMD).

We selected Duchenne and Becker muscular dystrophies because they are medical conditions with a predictable disease progression. Young men with DMD may start using braces and assistive devices between ages 6 and 9; by age 12 most use a wheelchair. In DMD, sentinel events (e.g., loss of ambulation, full time wheelchair use) are often used as outcomes in clinical trials along with clinical scales such as the Barthel Index (Mahoney & Barthel, 1965) and the EK Scale (Steffenson Lyager, Werge, Rahbek, & Mattsson, 2002).

Duchenne and Becker Muscular Dystrophies are the most common childhood forms of muscular dystrophy. Mutations in the dystrophin gene at Xp21 lead to absence or a reduction of functional dystrophin protein in muscle fibers that causes progressive muscle weakness (Koenig et al., 1989). Boys with Duchenne muscular dystrophy (DMD) develop progressive proximal muscle weakness early, and have a predictable clinical course (McDonald, Abresch, Carter, Fowler, Johnson, Kilmer, & Sigford, 1995). Muscle weakness and contractures cause loss of

ambulation in the teens. Wheelchair dependence is followed by worsening contractures, scoliosis, and progression of upper extremity and respiratory muscle weakness. Young men die in their mid 20's from respiratory failure. In the milder form (BMD) ambulation is typically preserved into adulthood, but progression and clinical manifestations are otherwise heterogeneous (McDonald, Abresch, Carter, Fowler, Johnson, & Kilmer, 1995). The clinical spectrum and rate of progression in DBMD are varied, and determined by multiple factors, including the nature of the dystrophin gene mutation, as well as genetic expression of utrophin, a dystrophin like protein (Koenig et al., 1989: Bushby, 1992; Kleopa, Drousiotou, Mavrikiou, Ormiston, & Kyriakides, 2006). Since the early 1990's, treatment in DMD with steroids has prolonged ambulation by several years, but it has not altered the ultimately fatal course.

Outcome studies in DBMD and other neurological conditions employ a variety of indicators to assess disease severity and progression. These include functional motor grading scales (Brooke, Griggs, Mendell, Fenichel, Shumate, & Pellegrino, 1981, Brooke, 1986, Vignos & Archibald, 1960), global functioning scales (Mahoney & Barthel, 1965), and multi-item scales such as the Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS) (Cedarbaum & Stambler, 1997).

Reliable disease progression scales are of paramount importance to patients, caregivers, health care providers and health services administration. It allows foreknowledge of problems and special needs to be anticipated as the disease progresses, allows family and patient counseling about the realistic outlook of the future, allows planning for and proper allocation of resources for the special needs of these patients, and may improve the cost of healthcare utilization. The DMD functional grading scales described above are valid and reliable measures of disability. Their usefulness in monitoring disease progression and assessing the effectiveness of therapeutic interventions in clinical trials is yet to be determined. Generation and validation of scales that depend on face-to-face patient examinations, like the MDFRS or the Motor Function Measure (MFM), is a very expensive and time consuming process needing large sample sizes (Lue, Su, Yang, Su, Lu, Lin, & Chen, 2006; Bérard, Payan, Hodgkinson, Fermanian, & the MFM Collaborative Study Group, 2005).

Because disease progression in DBMD is predictable, clinicians can readily assess an individual's level of severity from his medical chart. Severity is relayed as a narrative. The purpose of this investigation is the development and evaluation of a measure of disease severity for DBMD using common markers of disease progression that are readily available to clinicians. Rasch modeling techniques were used to develop the measure. This approach is robust to missing data and provides item calibrations that can be used in scale development. A similar example can be seen for smoking reduction in Davis, Sechrest, and Shapiro (2005).

Method

Participants

The respondents for this study were 34 primary caregivers of young men with DBMD who were recruited from Arizona (22), New York (10), Iowa (1) and Kansas (1). The respondents completed a structured survey by telephone; each respondent received $20 for study participation. Eligibility for the study was limited to caregivers of young men with Duchenne or Becker muscular dystrophy who were born before 1982. The structured interview included questions about family characteristics, use of durable medical goods and services. The study was approved by the University of Arizona Institutional Review Board. Complete survey details and univariate results for palliative care services are presented in Arias, et al. (under review).

The majority of the primary caregivers were women (28/34) and nearly half of the caregivers worked outside of the home (16/34). The average family income was $41,000 (*SD* = 25,000); the average level of education was 12 years (*SD* = 3.6); and the average age was 50.6 (SD = 8.4). Nine of the caregivers were Hispanic, and 7 (20.6%) of the caregivers had relatives (uncles or cousins) with muscular dystrophy. The young men were born between 1965 and 1981, with the majority born after 1975. Their average age 26.2 (*SD* = 4.9). Ninety-one percent of the young men had siblings (31/34), and 25 (74%) had two parents in the home. By the time of the interview, twelve of the young men had passed away and all survey questions were answered in retrospect for their last year.

Survey Instrument

The survey included questions about family characteristics (demographics, language use, family composition), early development (date of diagnosis with DBMD), and service utilization (physician, ancillary health care, home health visits). Data were collected for the presence/absence and date of first use/surgery for the 20 indicators of severity. The severity items included use of medical equipment (braces and splints, wheelchairs, transfer boards, hospital beds), and surgeries (tendon release, scoliosis, tracheostomy). See Table 1 for the complete list indicators included in the disease severity measure. Ten questions adapted from the ALSFRS (Cedarbaum & Stambler, 1997) were also collected. The ALSFRS items included handwriting, cutting food and handling utensils, dressing and hygiene, turning in bed and adjusting bed clothes, walking, climbing stairs and breathing. Descriptive results and detailed survey information are available in Arias et al. (under review).

Statistical Analyses

Descriptive statistics were used to characterize the sample and Rasch modeling was used to develop the measure (Rasch, 1960). This approach is well suited to Functioning data and has been used with the Functional Independence Measure (Heinemann, Linacre, Wright, & Granger, 1993), Activlim (Vandervelde, van den Bergh, Goemans, & Thonnard, 2007), and the International Classification of Function (Farin & Fleitz, 2009). Bigsteps software (Linacre & Wright, 1998) was used for model estimation. Rasch measurement techniques were chosen for several reasons. This approach jointly estimates severity for persons and items at the same time, and provides outcomes in a standard metric. The formula for the Rasch model is

$$P = \exp(b-d)/[1 + \exp(b-d)] ;$$

where b = the ability of the person, and d = the difficulty of the item (Wright & Stone, 1999). Here, b represents the severity of the person and d represents the severity of the item.

The Rasch model provides infit and outfit statistics which can be used for scale development. Mean square infit and outfit statistics were used to evaluate item fit. The outfit statistic is the average of the squared standard deviations between observed and expected performance and is sensitive to outliers. Items with high outfit statistics have more noise than signal. Infit is an information-weighted fit statistic that is sensitive to overall item performance. Items with high infit statistics are overly predictable from the other items in the measure. Recommended cutoff scores for clinical judgment were used, with lower and upper bounds of 0.5 and 1.7 respectively. Wright and Linacre (1994) note that fit statistics < 0.5 are less productive for measurement, but not degrading. Fit statistics > 2.0 degrade the measurement system. A few indicators were grouped for the final model; hospital bed/special mattress, stander/lift chair and splints/braces.

The items were either highly correlated or 'either/or', as the individual received only one of the two. No items were dropped from the study.

Reliability and the separation indices were used to evaluate the measure. The Separation Index (G) refers to the number of item strata and person strata that can be reliably defined by a particular test. The separation index is estimated as the ratio of the "True" standard deviation to the average measurement error (Wright & Linacre, 1985). Concurrent validity was examined by the correlation between the Rasch measure and the ALSFRS scale.

Data collection included the date of first use, or surgery for 20 severity indicators. The dates were used to create a longitudinal file with multiple years for each individual, starting at age 4 through the most recent year. There were 22.3 annual observations for the 34 young men. Rasch analyses were conducted using the longitudinal file. Typical Rasch analyses include only persons and items from a single test. These analyses incorporate persons, items, and occasions. The use of multiple occasions contributes to the stability of the severity measure.

Chien (2008) provides a justification for Rasch modeling with longitudinal data, noting raw scores provide necessary and sufficient information for estimation. This investigation also demonstrates an application of small sample strategies to Rasch modeling. Van der Linden, Wilson, Wolfe, & Linacre (2002) note that small samples may lack precision, but will still have construct validity. The large number of repeated measures per case compensates for the small number of cases (Figueredo, Petrinovich & Ross, 1992; Figueredo, Cox & Rhine, 1995). Lord (1983) states that "Small n Justifies the Rasch Model" and recommends that Rasch analyses be conducted as soon as there are data.

Results

The model included 17 disease progression indicators (Table 1). Several items were grouped: splints were grouped with braces; standers were grouped with lift chairs, and hospital beds were grouped with special mattresses. DBMD diagnosis was a prerequisite for study enrollment; all

Table 1
Starting age for disease severity indicators (N = 34)

Event	Age *Mean (SD)*	n (%)
Diagnosed with DBMD	5.1 (3.1)	34 (100.0)
Splints/Braces	8.5 (2.1)	26 (76.5)
Tendon release surgery	9.5 (2.2)	11 (32.4)
Walker	10.0 (2.1)	10 (29.4)
Manual wheelchair	10.7 (2.0)	32 (94.1)
Stander/Lift chair	12.3 (2.7)	9 (26.5)
Motorized wheelchair	13.6 (4.2)	29 (85.3)
Shower chair	13.6 (5.2)	25 (73.5)
Bedside commode	13.7 (4.7)	16 (47.1)
Back brace	13.9 (5.7)	8 (23.5)
Scoliosis	14.6 (1.8)	8 (23.5)
Transfer board	15.8 (4.5)	21 (61.8)
Hospital bed/Special mattress	17.4 (4.6)	27 (79.4)
Assisted ventilation	20.5 (4.5)	20 (58.8)
Death	22.4 (5.3)	12 (35.3)
Tracheostomy	23.2 (3.2)	9 (26.5)
Cough assist device	25.4 (2.7)	8 (23.5)

families reported diagnosis with DBMD at an average age of 5.1 (*SD*=3.1). Cough assist devices were used by only eight individuals, who started using the device at an average age of 25.4 years.

Table 2 provides summary results for the Rasch model. Fifty of the 773 annual observations had minimum scores (there were no severity indicators present) and DBMD severity could not be estimated. The separation index for the individuals was 2.21, with a reliability of .83. This indicates that the seventeen items could reliably divide the individuals into two groups. The separation index for the items was 17.39, with a reliability of 1.00. The item separation index indicates that the observations could reliably separate the items into 17 groups. The reliabilities for persons and items were acceptable.

Table 2

Summary of 723 measured (Non-extreme) observations

	Raw score	Count	Measure	Model Error	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
Mean	5.8	17	6.63	0.86	1	-0.2	0.92	-0.1
SD	3.4	0	2.13	0.34	0.71	1	1.44	0.4
Maximum	13	17	10.11	1.62	4.27	2.8	9.9	3.8
Minimum	1	17	2.69	0.62	0.16	-2.1	0.03	-0.8
Model RMSE	1.01	Adj. SD	2.23	Separation	2.21	Person	Reliability	.83

Summary of 17 measured items

	Raw score	Count	Measure	Model Error	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
Mean	248.2	723	8.00	0.12	0.99	-0.3	1.27	-0.3
SD	186.5	0	2.28	0.05	0.23	4.3	1.23	2.8
Maximum	695	723	11.78	0.28	1.51	6.9	5.33	5.3
Minimum	12	723	1.80	0.09	0.62	-8.6	0.36	-5.8
Model RMSE	.13	Adj. SD	2.28	Separation	17.39	Item	Reliability	1.00
SE of Item Mean	.57							

The item fit indices are in Table 3. The item with the lowest severity was diagnosis with DMD with a measure of 5.72. The most severe item was death with a measure of 11.78. The majority of the items were productive for measurement. Only two items had outfit statistics that were greater than 2.0; scoliosis surgery, and use of walkers (Table 3). Rasch analyses excluding the two items with high misfit statistics resulted in very similar item calibrations, as did analyses excluding the two young men with BDB. These results are not reported. Several items had outfit statistics that were below 0.50; use of a manual wheelchair, hospital bed/special mattress, cough assist device, and tracheostomy. Low outfit statistics indicate that the items are too predictable based on the other data.

Figure 1 graphically illustrates the disease severity measure. The items are on the right side of the map and the observations are on the left side. Each '#' represents four observations; each '.' represents 1-3 observations. For persons and items, the mean, one and two standard deviations are indicated by "M", "S", and Q" respectively. The item with the lowest severity was diagnosis with DBMD, and the highest severity item was death. As can be seen in Figure 1, the items covered the range of severity with gaps at the extreme ends of the measure.

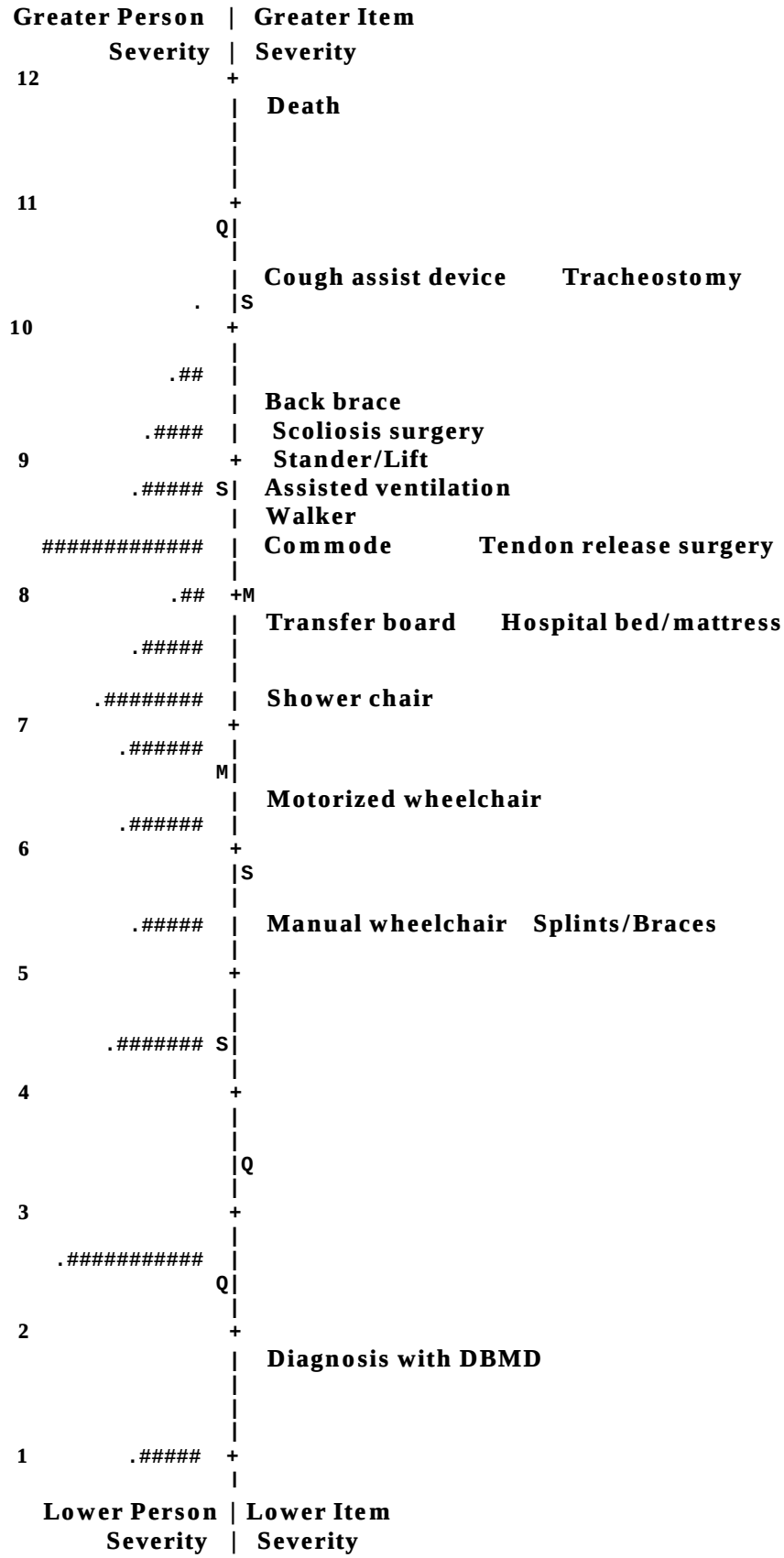

Figure 1: Construct map for disease progression indicators

Table 3
Item fit indices for disease severity indicators (N = 723)

Raw			Infit		Outfit		
Score	Measure	Error	MNSQ	ZSTD	MNSQ	ZSTD	Items
12	11.78	0.28	0.99	0.0	1.56	0.3	Death
43	10.47	0.15	0.89	-1.0	0.36	-1.1	Cough assist device
47	10.37	0.15	0.97	-0.2	0.46	-0.9	Tracheostomy
102	9.43	0.11	1.07	1.1	0.85	-0.3	Back brace
116	9.26	0.11	0.96	-0.7	5.33	5.3	Scoliosis surgery
130	9.09	0.10	0.91	-1.8	0.88	-0.3	Stander/Lift chair
156	8.81	0.10	0.80	-4.5	0.59	-1.4	Assisted ventilation
179	8.59	0.10	1.34	6.9	3.31	5.0	Walker
191	8.47	0.09	1.22	4.8	1.22	0.7	Tendon release surgery
204	8.35	0.09	1.27	5.7	1.26	0.9	Bedside commode
273	7.74	0.09	0.87	-3.0	0.62	-2.4	Transfer board/Hoyer lift
276	7.71	0.09	0.67	-8.6	0.46	-3.7	Hospital bed/mattress
345	7.10	0.09	1.09	1.9	0.92	-0.6	Shower chair
426	6.33	0.10	0.73	-5.1	0.63	-3.7	Motorized wheelchair
510	5.38	0.11	1.51	6.3	1.58	3.4	Splints/braces
515	5.32	0.11	0.62	-6.2	0.36	-5.8	Manual wheelchair
695	1.80	0.20	0.99	0.0	1.15	0.2	Diagnosis with DBMD
248	8.00	0.12	0.99	-0.3	1.27	-0.3	*Mean*
187	2.28	0.05	0.23	4.3	1.23	2.8	*SD*

Figure 2 illustrates the disease progression over time for the young men in this study. Disease progression followed the same general trajectory over time for individuals with Duchenne muscular dystrophy.

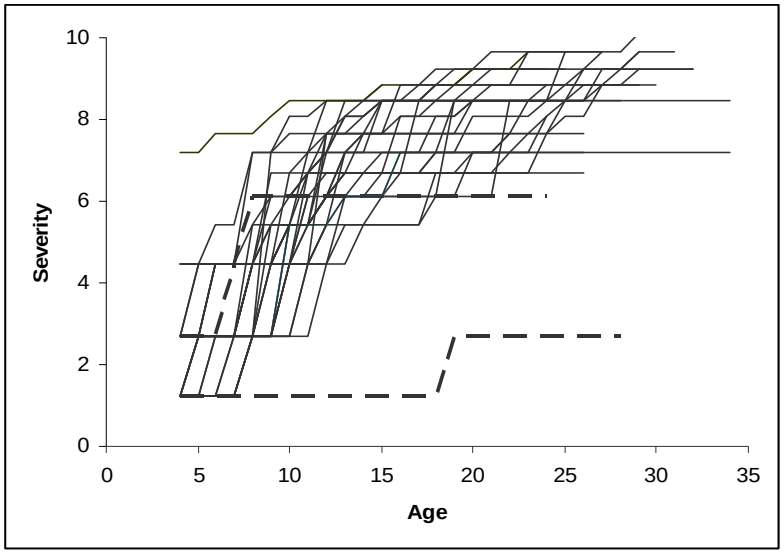

Figure 2: Disease severity by age. Individuals with DMD are indicated by solid lines; those with BMD are indicated by the dashed lines.

Concurrent Validity

The caregivers completed ten items adapted from a clinical severity measure (the ALSFRS) for the 22 young men who were still living. The test assesses the individual's functional ability (Cedarbaum & Stambler, 1997). The adapted measure had 4 response options; scores could range from 0 to 30, where 0 is the best score. For this sample, the average score was 2.1 (*SD* = .79). The correlation between the ALSFRS and the Rasch measure for the most recent year was .69 ($p < .05$), which accounted for 47.6% of the variance between the two measures.

Discussion

It can be difficult to compare the results of treatment studies using different outcome measures. This is a common situation in many areas of research from psychology to medicine. Each study reports the effect of a treatment on one or more outcomes, but the outcomes themselves are not comparable unless previous work on scale equating and item calibration has been performed. Research synthesis becomes extraordinarily difficult when a wide variety of outcome measures have been used.

We have demonstrated the development of a disease severity measure using typical markers of disease progression. To illustrate scale development, we chose a medical condition (DBMD) with a very predictable pattern of disease progression. The resulting measure demonstrated adequate reliability and validity coefficients. This measurement approach can be extended to billing and claims data, electronic medical records, chart reviews, and other observational data.

The purpose of this study was to pilot a disease severity measure for muscular dystrophy created from typical markers of disease progression. The results provided substantial evidence for the reliability and validity of the severity measure. We did not anticipate the high levels of reliability and validity that emerged in this investigation. The reliability of .83 was very respectable, particularly for a preliminary investigation. The correlation between ALSFRS and the DBMD severity measure was .69, which provides solid evidence for the concurrent validity of the new measure.

Calculating severity

The item calibrations from Table 3 can be used as weights for scale creation. Table 4 illustrates how severity can be calculated for DBMD. The measure ranges from 0 (no symptoms) to 12 (most severe). The items with the highest severity are used to estimate a person's current level of severity; an individual's lower severity items are not needed. A person with a diagnosis of DBMD and no other indicators would have a severity of 1.8. An individual diagnosed with DBMD, with a shower chair, a cough assist device and a tracheostomy would receive a severity of 10.5. Because the approach is based on the most severe items, the measure is robust to partial or missing data.

DBMD researchers often contend with small samples that are composed of young men of varying ages. Study results provide basic information on treatment effectiveness, but investigators may not be able to compare across studies and ages. Generic DBMD measures that can be estimated from existing markers of disease progression are sorely needed.

The clinical use of these measures may be limited for disease conditions with small sample sizes such as DBMD. A typical regional clinic may only see 20-40 young men, and medical providers will be familiar with each patient's condition. The clinical utility will be greater for progressive conditions with much larger sample sizes such as diabetes or Alzheimer's disease. Severity scores can be estimated from administrative databases. The scores can be used to track

the status of patients in large clinic populations and may identify patients whose condition is rapidly deteriorating.

Table 4
Calculating Severity Scores

Item	Check list	Weight
Diagnosis with DBMD	□ Yes □ No	1.8
Manual wheelchair	□ Yes □ No	5.3
Splints/braces	□ Yes □ No	5.4
Motorized wheelchair	□ Yes □ No	6.3
Shower chair	□ Yes □ No	7.1
Hospital bed/mattress	□ Yes □ No	7.7
Transfer board/Hoyer lift	□ Yes □ No	7.7
Bedside commode	□ Yes □ No	8.4
Tendon release surgery	□ Yes □ No	8.5
Assisted ventilation	□ Yes □ No	8.8
Stander/Lift chair	□ Yes □ No	9.1
Back brace	□ Yes □ No	9.4
Tracheostomy	□ Yes □ No	10.4
Cough assist device	□ Yes □ No	10.5
Severity = the most severe item weight	_____	

Disease severity as an outcome measure. We think the greatest potential for this approach will be in evaluation research. Disease severity measures can be estimated from existing data, retrospectively and prospectively. Such measures can be used to examine the natural history of disease conditions, generate hypothesizes regarding the effects of potential treatments, provide baseline data or supplement outcomes in clinical trials. The most promising use of generic measures may be for cross-clinic comparisons. A remarkable example is reported in Gawande (2004). Patient outcomes were compared across cystic fibrosis clinics; The results revealed subtle differences in treatment approaches and highlighted best practices.

Disease severity items. This investigation incorporated typical medical procedures and commonly used medical goods for individuals with DBMD. The results provide preliminary item calibrations for markers of disease progression. The results of the Rasch modeling procedures identified items with high and low misfit statistics. Scoliosis surgery and use of walkers had high misfit statistics and did not fit the scale well. We speculate that changes over time in medical care may be responsible for item misfit. In addition to changes in medical practices, other factors may be associated with the timing and use of durable medical goods. Families often acquire equipment before absolute need. Some items are bulky and may not be acquired by families without space for them. These factors will attenuate the precision of measurement for such items.

Several markers of severity had very low fit statistics (cough assist device, tracheostomy, hospital bed/special mattress, and manual wheelchair). Responses to these items were too predictable based on the other data. However, items with very low misfit can be particularly valuable in scale construction in observational studies, because their timing appears to be tightly linked to disease progression in DBMD. As Chien (2oo8) notes, overly predictable data rarely

cause a problem, because the data correspond to the measure. Highly predictable indicators of disease progression are useful in measures constructed from observational data.

This paper presents preliminary results for a disease severity measure that can be estimated from self report data, patient charts and administrative records. The measure demonstrated solid reliability and validity coefficients in this preliminary investigation. The intent of this investigation is to demonstrate the development of a scale for disease severity using common markers of disease progression. Such measures can be constructed from archival data and can be used to supplement existing impairment and function measures. Better understanding the clinical progression of muscular dystrophy can improve disease management and prevent complications.

Acknowledgements

We would like to express our appreciation to the individuals with DBMD and their families who participated in the study. This investigation was funded by a cooperative agreement between the Centers for Disease Control and Prevention (CDC) and the Association of University Centers on Disabilities (AUCD RTOI # 2004-03-03; PIs: FJM and MFD). The study was conducted in collaboration with the Muscular Dystrophy Surveillance, Tracking and Research Network (MDSTAR*net*) of sites that are funded through cooperative agreements with the CDC. The Palliative Care Group of MDSTAR*net* assisted in the development of the questionnaire for the study and included: Susan Apkon, Melinda F. Davis, Jane Karwoski, Dennis Matthews, F. John Meaney, Timothy Miller, Shree Pandya, and Christina Trout.

References

Arias, R., Andrews, J., Pandya, S., Pettit, K., Trout, C., Apkon, S., Karwosk, J. I, Cunniff, C., Matthews, D., Miller, T., Davis, M. F., & Meaney, F. J. (under review). Palliative care services in families of males with Duchenne muscular dystrophy. *Muscle and Nerve*.

Benson D. *Measuring Outcomes in Ambulatory Care*. (1992) Chicago, American Hospital Publishing, Inc.

Bérard, C., Payan, C., Hodgkinson, I., Fermanian J., & the MFM Collaborative Study Group. (2005). A motor function measure scale for neuromuscular diseases. Construction and validation study. *Neuromuscular Disorders, 15*, 463-470.

Brooke, M. H. (1986). *A clinician's view of neuromuscular diseases* (2nd ed.). Williams & Wilkins.

Brooke, M. H., Griggs, R. C., Mendell, J.R., Fenichel, G. M., Shumate, J. B., & Pellegrino, R. J. (1981). Clinical trial in Duchenne dystrophy. The design of the protocol. *Muscle and Nerve, 4*, 186-197.

Bushby, K. M. (1992). Genetic and clinical correlations of Xp21 muscular dystrophy. *Journal of Inherited Metabolic Disease, 15*, 551–554.

Cedarbaum, J. M., & Stambler, N. (1997). Performance of the Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS) in multicenter clinical trials. *Journal of the Neurological Sciences* ,*152*(Suppl. 1), S1–S9.

Chien, T.-W. (2008). Repeated Measure Designs and Rasch. *Rasch Measurement Transactions, 22*, 1171.

Davis M. F., Sechrest, L., & Shapiro, D. (2005). Measuring Progress toward Smoking Cessation. *Journal of Applied Measurement, 6*, 164-72.

Davis, M. F. (2009). Avoiding Data Disasters and Other Pitfalls. In S. Sidani, and D. L. Streiner, (Eds.) When Research Goes Off the Rails (pp. 320-326). New York: Guilford Publications.

Davis, M. F., Meaney, F. J., & Duncan, B. Factors Influencing the Use of Complementary and Alternative Medicine in Children. (2004). *Journal of Alternative and Complementary Medicine, 10*, 740-742.

Farin, E., & Fleitz, A. (2009). The development of an ICF-oriented, adaptive physician assessment instrument of mobility, self-care, and domestic life. *International Journal of Rehabilitation Research, 32*, 98-107.

Figueredo, A. J., Cox, R. L., & Rhine, R. J. (1995). A generalizability analysis of subjective personality assessments in the Stumptail macaque and the Zebra finch. *Multivariate Behavioral Research, 30,* 167-197.

Figueredo, A. J., Ross, D. M., & Petrinovich, L. (1992). The Quantitative Ethology of the Zebra Finch: A Study in Comparative Psychometrics. *Multivariate Behavioral Research, 27,* 435 -458.

Gawande, A. A. (2004) The Bell Curve. *The New Yorker*. [On-line] Available: http://www2.hawaii.edu/~mchlend/website/Gawande_BellCurve.pdf.

Heinemann, A. W., Linacre, J. M., Wright, B. D., & Granger, C. V. (1993). Relationships between impairment and physical disability as measured by the Functional Independence Measure. *Archives of Physical Medicine and Rehabilitation, 74,* 566-573.

Kleopa, K. A., Drousiotou, A., Mavrikiou, E., Ormiston, A., & Kyriakides T. (2006). Naturally occurring utrophin correlates with disease severity in Duchenne muscular dystrophy. *Human Molecular Genetics, 15,* 1623-1628.

Koenig, M., Beggs, A. H., Moyer, M., Scherpf, S., Heindrich, K., Bettecken, T., Meng, G., . . . Kunkel, L. M. (1989). The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. *American Journal of Human Genetics, 45,* 498-506.

Linacre, J. M., & Wright, B. D. (1998). *A User's Guide to BIGSTEPS: Rasch-Model Computer Program*. Chicago: Mesa Press. [On-line] Available: *http:www.winsteps.com /bigsteps.htm* [2009].

Lord, F. M. (1983). Small n Justifies the Rasch Model. In D. J. Weiss, (Ed.), *New Horizons in Testing*. (pp. 51-61). New York: Academic Press.

Lue, Y.-J., Su, C.-Y., Yang, R.-C., Su, W.-L., Lu, Y.-M., Lin, R.-F., & Chen, S.-S. (2006). Development and validation of a muscular dystrophy-specific functional rating scale. *Clinical Rehabilitation, 20,* 804-817.

Mahoney F. I. & Barthel, D. W. (1965). Functional evaluation: The Barthel index. *Maryland State Medical Journal, 14,* 61–65.

McDonald, C. M., Abresch, R. T., Carter, G. T., Fowler, W. M., Jr., Johnson, E. R., Kilmer, D. D., & Sigford, B. J. (1995). Profiles of neuromuscular diseases: Duchenne muscular dystrophy. *American Journal of Physical Medicine and Rehabilitation, 74* (Suppl.), S70–S92.

McDonald, C. M., Abresch, R. T., Carter, G. T., Fowler, W. M., Jr. Johnson, E. R., & Kilmer, D. D. (1995). Profiles of neuromuscular diseases. Becker's muscular dystrophy. *American Journal of Physical Medicine and Rehabilitation, 74* (Suppl. 5), S93-103.

Rasch, G. (1960). *Probabilistic models for some intelligence and achievement tests.* Copenhagen: Danish Institute for Educational Research (Expanded ed., (1980). Chicago: University of Chicago Press).

Steffensen, B., Lyager, S., Werge, B., Rahbek, J., & Mattsson, E. (2002). Physical capacity in non-ambulatory people with Duchenne muscular dystrophy or spinal muscular atrophy. *Developmental Medicine and Child Neurology, 44,* 623–632.

Van der Linden, W. J., Wilson, M., Wolfe, E. & Linacre, J. M. (2002). Review of reviews of Bond & Fox (2001). *Rasch Measurement Transactions, 16,* 871-872.

Vandervelde, L., van den Bergh, P., Goemans, N., & Thonnard, J. (2007). ACTIVLIM: A Rasch-built measure of activity limitations in children and adults with neuromuscular disorders. *Neuromuscular Disorders, 17,* 459-469.

Vignos, P. J, & Archibald, K. C. (1960). Maintenance of ambulation in childhood muscular dystrophy. *Journal of Chronic Diseases, 12,* 273-290.

Wright, B. D., & Linacre, J. M. (1985). Based on: Microscale Manual. Westport, Conn.: Mediax Interactive Technologies, Inc. [On-line] Available: *http://www.rasch.org/rmt/ glossary.htm*.

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*, 370.

Wright, B., & Stone, M. (1999). *Measurement Essentials* (2nd ed.). Wilmington, Delaware: Wide Range, Inc.