

Perceptions of Health Status in Multiple Sclerosis Patients and Their Doctors

Marcelo Kremenchutzky, Len Walt

ABSTRACT: Objective: To compare neurologist and patient perceptions of multiple sclerosis (MS)-related health status. **Methods:** MS patients (n=99) were recruited from six sites in Canada. Following a consultation with their neurologist, patients estimated their relapse frequency, rated their general health and quality of life (QoL), reviewed descriptions of eight health domains and selected the three most important, and completed a utility assessment using the standard gamble (SG). Concurrently, neurologists independently used the same instruments to rate their patients' health status. Assessments were compared on the basis of paired mean values of both groups and the degree of exact agreement quantified by intraclass coefficient (ICC) and kappa analyses, which yield values of 1.0 with 100% agreement. **Results:** There were significant differences ($p < 0.001$) between patient and neurologist ratings for relapses in the last year (0.86 vs. 0.4, respectively), QoL (61.2 vs. 69.7 (maximum score = 100) and utility (0.864 vs. 0.971); ICC analysis revealed moderate to poor levels of agreement (0.56 for QoL to 0.03 for SG). There was little concordance in identification of important health domain and the only significant associations were in bodily pain and social functioning (kappa statistic = 0.24, $p = 0.026$ for both). Neurologists identified physical functioning domains as important, while patients placed more emphasis on mental health domains. **Conclusions:** Discrepancies between neurologist and patient perceptions of MS were observed. The study identifies a need to educate neurologists on the recognition of MS health domains that are important in the definition of patient QoL.

RÉSUMÉ: Perceptions qu'ont les patients atteints de sclérose en plaques et leur médecin de l'état de santé de ces patients. Objectif : Le but de l'étude était de comparer la perception de l'état de santé en relation à la sclérose en plaques (SP) qu'ont les neurologues et les patients. **Méthode :** Des patients atteints de SP (n = 99) ont été recrutés dans six centres au Canada. À la suite d'une consultation avec leur neurologue, les patients estimaient la fréquence de leurs rechutes, évaluaient leur santé générale et leur qualité de vie (QV), révisaient des descriptions de 8 domaines de santé et choisissaient les trois plus importants pour eux et complétaient une évaluation de l'utilité par la méthode du pari standard (PS). Les neurologues utilisaient parallèlement les mêmes instruments pour évaluer l'état de santé de leurs patients. Les évaluations ont été comparées quant aux valeurs moyennes paires des deux groupes et le degré de concordance exacte a été quantifié au moyen du coefficient intraclass (CIC) et de l'analyse kappa, dont la valeur est de 1,0 lorsque la concordance est de 100%. **Résultats :** Il existait des différences significatives ($p < 0,001$) entre l'évaluation faite par le patient et celle faite par le neurologue pour le nombre de rechutes au cours de la dernière année (0,86 contre 0,4 respectivement), la QV (61,2 contre 69,7; score maximum = 100) et l'utilité (0,864 par rapport à 0,971). L'analyse CIC a montré un niveau de modéré à faible pour la concordance (0,56 pour la QV; 0,03 pour le PS). Il y avait peu de concordance dans l'identification des domaines de santé importants et les seules associations significatives étaient pour la douleur physique et le fonctionnement social (statistique kappa = 0,24, $p = 0,026$ pour les deux). Les neurologues ont identifié les domaines de fonctionnement physique comme étant importants alors que les patients mettaient plus d'emphasis sur les domaines de santé mentale. **Conclusions :** Nous avons observé des discordances entre la perception de la SP qu'ont les neurologues par rapport à celle de leurs patients. Cette étude indique un besoin d'éducation chez les neurologues sur la reconnaissance des domaines de santé dans la SP qui sont importants dans la définition de la QV chez les patients.

Can J Neurol Sci. 2013; 40: 210-218

Over the last 20 years there has been increasing interest in shared medical decision making between patients and their physicians.¹⁻⁴ This has been driven in part by the recognition that for some diseases, particularly where a clinical outcome is not always readily quantified by a treating physician, patient self-assessment provides additional input that can be used for diagnosis and treatment decisions.³ In addition, patient input related to quality of life (QoL) can allow physicians to recommend treatment based on a balance between positive clinical benefit and potential QoL improvement and the negative QoL consequences of therapy-related adverse events.⁵⁻⁸ The role of the physician in these situations is to help the patient choose the treatment that will optimize health outcomes and maintain an acceptable QoL, from the patients' perspective. However, effective shared decision making is based on productive patient-

physician communication and this interaction is not always ideal.^{2,5-7,9} Patient perception of illness is a complex multidimensional process involving not only specific symptoms, but the impact of symptoms on physical, social and cognitive functioning and their expectations regarding potential clinical interventions.^{2,10,11} This complex make-up hinders the ability of

From the Schulich School of Medicine (MK), University of Western Ontario, London; Biogen Idec Canada Inc. (LW), Mississauga, Ontario, Canada.

RECEIVED AUGUST 2, 2012. FINAL REVISIONS SUBMITTED OCTOBER 19, 2012.
Correspondence to: Marcelo Kremenchutzky, Neurology, Schulich School of Medicine, University of Western Ontario, 339 Windermere Road, London, Ontario, N6A 5A5, Canada. Email: marcelo.kremenchutzky@lhsc.on.ca

the physician to fully understand and act appropriately on what are subjective patient perceptions of disease status and patient opinions about their health do not always correlate with those of professionals taking care of them. Examples of this divergence in views have been identified in cancer,^{7,10} multiple sclerosis,^{3,12} and in musculoskeletal diseases.² For the most part, these studies show that physicians may not always recognize symptoms or outcomes of treatment that are important to patients.² In addition, treatment decisions are in many cases based on a poor understanding of just how much risk patients are willing to accept in return for clinical benefit.^{2,8,12}

Multiple sclerosis is a chronic degenerative neurological condition which has a significant QoL impact and most patients must cope with increasing physical impairment, changing social functioning roles and cognitive decline over time.¹³⁻¹⁵ Although treatment choices have expanded over the last few years, clinical response to therapy is not always positive and new therapies such as natalizumab (Tysabri®) and fingolimod (Gilenya®) are associated with a risk of serious adverse events;^{12,16,17} under these conditions optimal treatment decisions should also include an assessment of the level of risk a patient is willing to tolerate for a possible clinical gain, adding to the complexity of the physician/patient dynamic.^{12,18}

From a physicians' perspective, complex treatment decisions in MS have traditionally been driven in large part by clinically assessable outcomes such as reduction in the number of brain lesions, changes in the frequency of disease flares or progression in patient physical disability.³ However, as in other diseases, additional health domains, beyond the physical disability associated with the disease, are important for patients.^{3,4,12} This makes MS an ideal candidate for shared decision making where the perceptions of the patient and the clinical expertise of the physician combine to make effective clinical decisions.¹⁹⁻²¹ To assess the current status of MS patient/physician communication in Canada, the present pilot study used different tools to quantify the degree of agreement between the two parties in defining health status and to identify QoL domains that are important to patients. In addition, standard gamble techniques were used to quantify differences between physician and patient attitudes regarding health status and the degree of acceptable risk associated with MS therapy. The standard gamble assesses health state utility but also incorporates a measure of risk tolerance by offering individuals a choice between remaining in a defined health state or choosing an alternative medical intervention which has a variable risk of achieving perfect health or death.^{2,22}

METHODS

Study design

This was a prospective, multi-centre, non-interventional study of MS patients and their treating neurologists. Patients meeting inclusion criteria were recruited consecutively when they attended a regularly scheduled neurology clinic visit at one of six participating sites in Ontario and Alberta, Canada. At the conclusion of the consultation, neurologists provided a brief description of the study procedures and patients were asked if they were interested in participating. Interested patients were then asked to read and sign an informed consent form after which study packages were unsealed and the enclosed patient and neurologist questionnaires were completed independently.

Completed questionnaires were then sealed in an envelope and returned to the study coordination centre for data entry and analysis. Both neurologist and patient questionnaires were completed immediately following the consultation but each was blinded to the other's responses.

Any patient with a diagnosis of MS was eligible to participate and there were no restrictions on clinical status or treatment, except for patients receiving natalizumab or weekly intramuscular interferon beta-1a; patients currently treated with either of these drugs were excluded from the study. Natalizumab and interferon beta-1a (Avonex®) are marketed in Canada by Biogen Idec, the sponsor of this study, and patients receiving these products were excluded to avoid any potential bias in neurologist/patient assessments. In addition, natalizumab has a somewhat different benefit/risk profile than current immunomodulatory therapies used in Canada and it was thought that inclusion of data on this drug might skew results to a degree unrelated to its current utilization.¹² The study was approved by the relevant Institutional Review Board for each site.

Questionnaires

All data were collected through neurologist and patient questionnaires. The patient questionnaire collected data on demographics and clinical history of MS including the number of flares in the last year and the number of lifetime flares, as well as MS treatment history. Patients were also asked to assess their MS status (in remission or active disease) and to identify any comorbidities. In addition, the following instruments were included in the questionnaire and used to assess patient health status:

Multiple Sclerosis Quality of Life (MSQOL)-54

The MSQOL-54 is an MS specific health related QoL instrument consisting of the generic Medical Outcomes Study Short Form Health Survey (SF-36),²³ supplemented with additional disease-specific measures.^{24,25} The instrument has been extensively validated in MS.²⁵ The complete MSQOL-54 instrument was included in the patient questionnaire.

Standard gamble

Standard gamble methods were used to assess the value, or utility, patients placed on their current health state. Standard gamble assessment is accomplished by asking subjects to gamble on their health between their current certain health state and an uncertain outcome that has a probability P of full health and a corresponding probability 1 – P of death.^{2,22,26} Since the latter uncertain outcome is based on risk, the standard gamble therefore provides a measure of the risk patients are willing to take for a better outcome and the value they attach to their current health state.²² In the assessment, the probability of the uncertain outcome is varied until the subject has no preference between this alternative and the certain outcome.^{22,26} The result is a utility score between 0 and 1 where the former means death and 1 means perfect health.^{22,26} In the current study, a paper-based standard gamble was used and for patients, the scenario was based on an imaginary medication for MS with a good (perfect health) and a bad (immediate death) outcome; probabilities of immediate death went from 0%, 1%, 2% and 5%,

Table 1: Definitions of SF-36 health domains

SF-36 domain	Domain description*
Physical function	Ability to perform all types of physical activities without limitations due to health
Physical role limitation	Ability to perform all work and other daily activities without limitations due to health
Bodily pain	Having no pain or limitations due to pain
General health	Believing that your personal health is excellent
Vitality	Feeling full of pep and energy all of the time
Social functioning	Ability to perform normal social activities without interference from physical or emotional problems
Emotional role limitation	Ability to perform work and other daily activities without interference from emotional problems
Mental health	Feeling peaceful, happy and calm all of the time

*Descriptions were based on the eight subscales defined by the Short Form Health Survey (SF-36)²³

then increasing in 5% increments to 100% giving a total of 16 different probabilities. Patients were asked to indicate for each scenario if they would take the drug; the utility was based on the point at which patients would not take the imaginary medication.

Health domains

Descriptions of relevant health domains were prepared based on the eight subscales defined by the SF-36 questionnaire, a QoL instrument which has been extensively used to quantify change in QoL in MS patients.^{13,23,27} The descriptive text for each SF-36 defined domain is shown in Table 1 and patients were asked to rank the three domains that were most important to their health.

Immediately after seeing each patient the neurologist completed a questionnaire which recorded their perceptions of the patient's health status. This included an estimate of the number of flares the patient had in the last year and over their lifetime and their impression of the patient's general health and overall QoL recorded on Likert scales ranging from 1 (excellent) to 5 (poor) and from 10 (best possible QoL) to 0 (worst possible QoL), respectively. Both these questions were extracted from the MSQOL-54 instrument (general health (question 1) and QoL (question 53)).²⁴ Neurologists were also asked to rank what they believed were the three most important SF-36 domains for the patient using the same domain descriptions defined for the patient questionnaire (Table 1) and to complete the standard gamble assessment with the assumption that they had the same health status, diagnosis and prognosis as the patient they had just examined. The standard gamble scenarios were identical to those completed by the patient.

Statistical analysis

Descriptive statistics were used to summarize baseline patient demographics and MS clinical status. Patient MSQOL-54 data were scored as previously described.²⁴ Only patient data from

MSQOL-54 questions 1 (general health) and 53 (overall QoL) were used for comparisons with responses to the same questions from neurologists. To aid in overall comparisons, ratings for general health were reversed and scaled to 100 so that possible scores ranged from 0 (poor general health) to 100 (excellent general health). In addition, overall QoL and standard gamble assessments were adjusted to a scale of 0 to 100 with the former defining a poor outcome and 100 the best outcome.

Patient and neurologist ratings on MS clinical status (number of flares), general health, overall QoL and standard gamble were compared using paired t tests. Differences between assessments were expressed as the patient rating – the neurologist rating so that both the magnitude and the direction of differences could be quantified. In addition, agreement between neurologist and patient ratings for all parameters was assessed using intraclass correlation coefficients (ICCs) calculated from a 1-way analysis of variance (ANOVA). Generation of ICCs provides a measure of the degree of agreement between patients and their neurologists at the individual level.^{28,29} The maximum ICC is 1.00, meaning 100% agreement between patient and neurologist ratings; it is generally accepted in assessments of this type that ICC values of < 0.40, 0.41 – 0.60, 0.61 – 0.80 and 0.81 – 1.00 can be interpreted as meaning poor to fair, moderate, good or excellent agreement, respectively.²⁸ An ICC value of 0 is the expected level of agreement based on chance alone and a value of <0 indicates that there is more variation than would be expected from chance alone.^{28,29}

Data on identification of the three most important SF-36 health related QoL domains were expressed as the proportion of patients and neurologists who identified a domain as one of the three most important. The degree of concordance between individual patients and their respective neurologists in the identification of specific domains as one of the three most important was determined using a kappa analysis to assess the level of agreement beyond chance.³⁰ In this analysis when the observed agreement exceeds chance, the kappa statistic is positive with its magnitude reflecting the strength of agreement; a kappa statistic of 1 indicates perfect agreement while a value of < 0.4 defines fair to poor agreement.³¹⁻³³ However, the kappa analysis can be affected by observer bias and the prevalence of responses in observer populations which can lead to an over adjustment for the observed agreement being due to chance.³³ To correct for this the Prevalence-Adjusted Bias-Adjusted Kappa (PABAK), which adjusts for these imbalances, was also calculated to assess the strength of agreement between neurologists and patients.³³

Differences in neurologist and patient ratings were also characterized using box plots. In addition, to identify parameters correlated with differences in ratings, generalized linear models (GLM) were utilized with the difference in ratings between neurologists and patients entered as a dependent variable.² Independent variables entered as fixed effects included patient age, gender, education level, disease duration and patient score for the measure being evaluated.

RESULTS

Six neurologist sites in Ontario and Alberta participated in the study and recruited a total of 99 patients. As shown in Table 2, patients had a mean age of 34.5, the majority were female

Table 2: Baseline demographic and clinical characteristics of the patient population

Variable	Subjects (n=99)
Patient demographics	
Age (Years, mean \pm SD*)	34.5 \pm 11.4
Min-max	17 - 56
Gender (n (% female))	71 (72.4%)
Education (n (%))	
High school graduate	94 (94.9%)
College graduate	35 (35.5%)
University graduate	24 (24.2%)
None/unknown	5 (5.1%)
Age at disease onset (Years, mean \pm SD)	33.8 \pm 12.2
Clinical parameters	
Current MS disease status (n (%))	
In remission	35 (35.4%)
Active disease	
Mild	35 (35.4%)
Moderate	25 (25.3%)
Severe	4 (4.0%)
Current MS therapy (n (%))	
Interferon beta-1b	16 (16.2%)
Interferon beta-1a [†]	20 (20.2%)
Glatiramer acetate	17 (17.2%)
Other therapy	22 (22.2%)
No therapy	27 (27.3%)
Past medical history/concomitant symptoms (n (%))	
Visual problems	24 (24.2%)
Bladder dysfunction	21 (21.2%)
Depression	16 (16.2%)
Spasticity	14 (14.1%)
Asthma	11 (11.1%)
Arthritis	10 (10.1%)
Other	46 (46.5%)
MSQOL-54 score (Mean \pm SD)	
Physical component	60.1 \pm 14.5
Mental component	59.0 \pm 16.2

*Abbreviations are; SD, standard deviation; MSQOL, Multiple Sclerosis Quality of Life; [†]Patients were receiving Rebif®

(72.4%) and age at disease onset was 33.8 years. Based on self-assessed disease status at the time of the study, 35.4% of patients were in remission, while 35.4%, 25.3% and 4.0% classified their disease status as mild, moderate and severe, respectively (Table 2). The majority of patients (53.6%) were currently on immunomodulatory therapy and concomitant MS symptoms (bladder dysfunction, depression, spasticity, and visual impairment) and comorbidities (asthma and arthritis) were common. The physical and mental MSQOL-54 summary scores were similar at 60.1 and 59.0, respectively (Table 2).

Table 3 shows the mean patient and neurologist ratings for MS flares, general health, QoL and standard gamble. There were statistically significant differences ($p < 0.001$) between patient and neurologist ratings for flares in the last year (0.86 by patients vs. 0.40 for neurologists), QoL (61.2 vs. 69.7) and standard gamble (86.4 vs. 97.1), based on paired mean values for both groups. The corresponding ICC values were for the most part low, indicating poor levels of concordance. The highest agreement was for QoL (ICC, 0.56) indicating that although mean scores for this parameter between patients and neurologists were significantly different, on an individual basis, there was a moderate degree of agreement (Table 3). Neurologists rated patient general health higher than patients and underestimated the number of lifetime flares, but the differences were not statistically significant and the ICC analysis for the latter indicated moderate concordance (ICC = 0.41).

Bubble plots of the data for neurologists and patients for flares over the last year, general health, QoL and standard gamble are shown in Figure 1 together with the corresponding correlation coefficients (R^2). In the graphs for general health, QoL and standard gamble, lower scores indicate poorer health; however, for number of flares in the last year, low values mean better clinical outcomes. These data confirm the low level of agreement in all the assessments and show extensive deviation from perfect concordance. To quantify the magnitude of the relationship between neurologist (dependent variable in the analysis) and patient (the predictor in the analysis) ratings, R^2 values were calculated for each plot. These data indicate that the proportion of variation in the neurologists' ratings predicted by the patients' ratings were 10% for flares in the last year, 9% for

Table 3: Comparison of patient and neurologist assessments of disease status, general health, quality of life and standard gamble

Assessment	MS* Flares (n)		General Health [†]	Quality of life [‡]	Standard gamble [¶]
	Over last year	Lifetime			
Patient	0.86	4.45	53.0	61.2	86.4
Neurologist	0.40	4.28	58.1	69.7	97.1
Difference	0.45	0.17	-5.05	-7.98	-10.7
P value	<0.001	0.745	0.096	<0.001	<0.001
ICC	0.21	0.41	0.26	0.56	0.03
(95% CI)	(0.02 – 0.39)	(0.23 – 0.56)	(0.68 – 0.43)	(0.41-0.68)	(-0.22-0.17)

*Abbreviations are; MS, multiple sclerosis, ICC, intraclass coefficient, CI, confidence interval; [†]General health assessed on a 1, excellent to 5, poor Likert scale; for analysis the scoring algorithm was reversed and transformed to a 0-100 scale so that higher scores indicate better general health; [‡]Assessed on a 0 (worst possible quality of life) to 10 (best possible quality of life) scale. For the analysis scores were scaled to 100; [¶] Standard Gamble scores were scaled to 100.

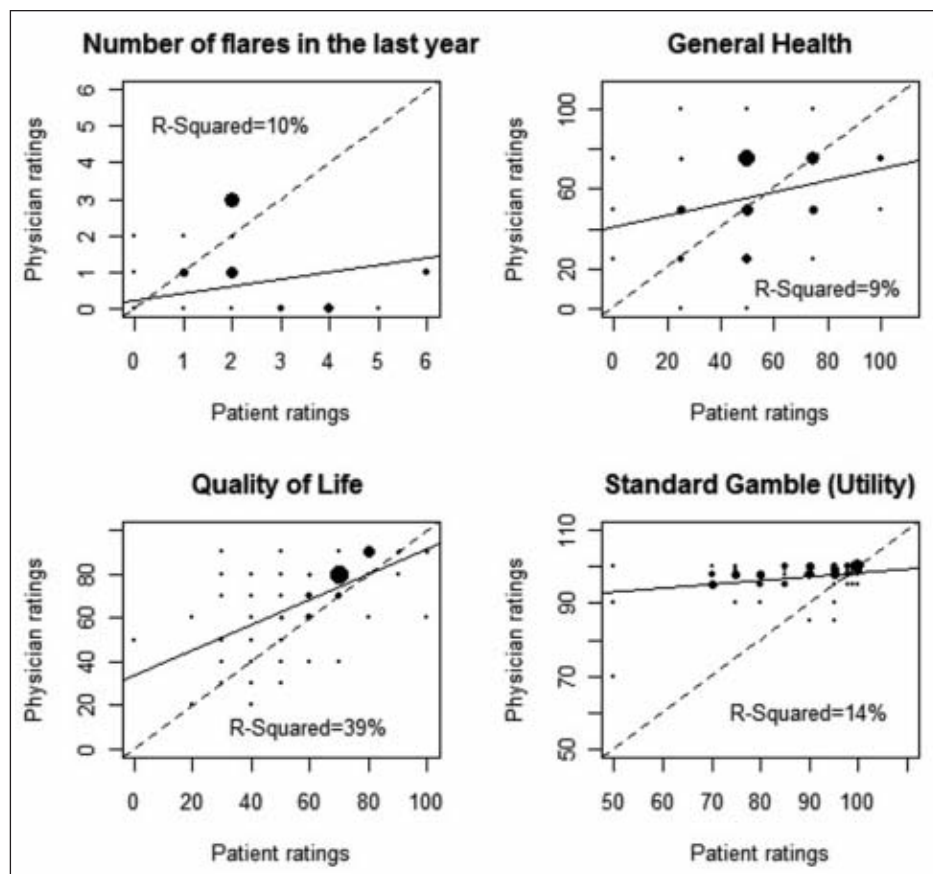


Figure 1: Bubble plots of neurologist and patient ratings for number of flares in the last year, general health, QoL and standard gamble. The R^2 values are the squares of the correlation coefficient for each analysis. The dashed line indicates the expected plot with perfect agreement in scores. Since responses were based on fixed scale responses, the size of data points reflects the number of common responses for each rating. Scores for general health, QoL and standard gamble have been scaled from 0 to 100 with a corresponding range of poor to perfect health.

general health, 39% for QoL and 14% for standard gamble, again confirming highest agreement in QoL assessments.

Box plots and interquartile range for the differences between neurologist and patient ratings for general health, QoL and standard gamble are shown in Figure 2. These data again show that neurologists tended to rate their patients' health status higher, particularly for QoL and standard gamble, where > 75% of the neurologists gave higher ratings. However, in the case of general health, although 50% of neurologists rated health status higher (a difference of up to 20 points higher on the 0-100 scale), a similar proportion also rated health status worse than patients rated themselves (Figure 2).

To identify patient characteristics related to the possibility of increased divergence between neurologist and patient ratings, a series of GLM analyses were carried out (Figure 3). For health status parameters (general health, QoL, standard gamble) the data indicated that when patients assessed themselves as healthier, their neurologists were more likely to rate the patient as sicker and when patients thought they were in poorer health, neurologists applied a higher health rating. In the case of standard gamble, almost all neurologist ratings were higher than the corresponding patient ratings (Figure 3). For flares in the last year, neurologists consistently underestimated incidence but as the number of reported flares (based on patient report) decreased, the more concordance there was between neurologist/ patient assessments. Generalized linear models plots were repeated for



Figure 2: Box plots of the differences in ratings between patients and neurologists (patient score – neurologist score) for general health, QoL and standard gamble (utility). The dotted line through 0 indicates the expected result if neurologist and patient assessments were identical.

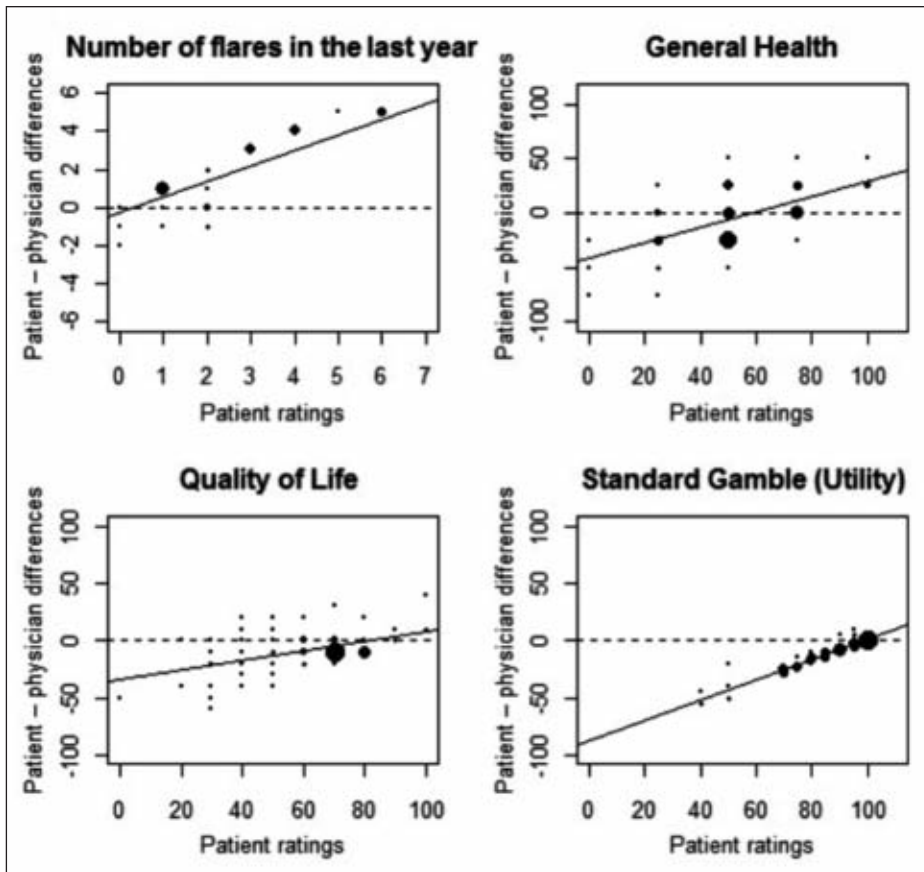


Figure 3: Generalized linear model (GLM) analyses expressed as the difference between patient and neurologist ratings. In the models patient scores were included as an independent variable to control for scale effects. In each case the dashed line through 0 indicates the expected result if neurologist and patient assessments were identical for each parameter. As in Figure 1, the size of data points is proportional to the number of common responses.

various patient baseline characteristics but the only significant predictor for neurologist/patient divergence was the MSQOL-54 physical component score with higher scores leading to decreased divergence in ratings of flares in the last year ($p=0.0239$) and increased divergence in the assessment of lifetime flares ($p=0.0335$), respectively. However, differences in patient-neurologist ratings for all assessed parameters were strongly associated with the patients' scores.

To further quantify the ability of the neurologist to define a patient's health status they were asked to identify what they perceived to be the three most relevant QoL domains for each patient; their assessment was then compared with that of the self-assessment by the patient. As shown in Table 4, these data confirmed that neurologists did not identify the health domains that were important to patients. Comparisons at the group level through mean scores showed that although 63.4% of patients rated mental health as one of the three most important domains, only 11.0% of neurologists provided a similar rating. Instead, 72% of neurologists rated physical function as an important domain compared to only 52.4% of the patients. There was also a difference in the rating of vitality with 30.5% of neurologists identifying this as important compared to only 18.3% of patients.

This analysis was combined with a kappa analysis allowing the patient/neurologist agreement to be quantified and differentiated from agreement based purely on chance. As shown in Table 4, the proportion of observed agreement in identification

of bodily pain (18.3%), emotional role limitation (1.2%), general health (3.7%), mental health (3.6%) and vitality (8.5%) as one of the three most important domains was poor and there was higher observed agreement that these domains were not one of the three most important. The strength of agreement measured by the adjusted Kappa was poor (PABAK between -0.34 to 0.46). In contrast, physical function, physical role limitation and social functioning showed levels of apparent agreement of 37.8%, 58.8% and 25.6%, respectively. However, the corresponding kappa statistics (raw and PABAK) were small (<0.4) indicating poor levels of significant agreement beyond chance and the only significant associations identified were in bodily pain and social functioning both with kappa statistics of 0.24 and p values of 0.026. The kappa analysis indicated that the apparent agreement between neurologists and their patients for physical function (37.8%) was in fact mainly due to chance (corresponding kappa statistics were 0.003 ($p=0.585$) and PABAK=0.02) suggesting that although the majority of participants identified this domain as one of the three most important, there was little evidence that neurologists linked the domains to specific patients. The adjusted kappa analysis indicated that the apparent agreement between neurologists and their patients for physical role limitation (58.5%) was not due to chance, as implied by the raw kappa statistics (0.12 ($p=0.206$)), but in fact reflected a fair level of agreement (PABAK=0.34).

Table 4: Identification of the three most important health related quality of life domains for patients; variation in patient and neurologist assessments

Assessment	Health related quality of life domains*							
	Bodily pain	Emotional role limitation	General health	Mental health	Physical function	Physical role limitation	Social functioning	Vitality
Percent rating domain as important								
Patient	40.2%	14.6%	14.6%	63.4%	52.4%	74.4%	47.6%	18.3%
Neurologist	31.7%	14.6%	23.2%	11.0%	72.0%	75.6%	41.5%	30.5%
Agreement (%)[†]								
One of the 3 domains	18.3%	1.2%	3.7%	3.6%	37.8%	58.5%	25.6%	8.5%
Not one of the 3 domains	46.3%	72.0%	65.9%	29.3%	13.4%	8.5%	36.6%	59.8%
Kappa statistic[‡] (95% CI)	0.24 (0.03–0.45)	-0.07 (-0.25 – 0.10)	0.02 (-0.20 – 0.23)	-0.11 (-0.23 – 0.02)	0.003 (-0.20 – 0.20)	0.12 (-0.11 – 0.35)	0.24 (0.03 – 0.45)	0.16 (-0.06 – 0.38)
P value	0.026	0.442	0.563	0.055	0.585	0.206	0.026	0.117
PABAK[§]	0.29	0.46	0.39	-0.34	0.02	0.34	0.24	0.37

*Domains are described in Table 1; †Agreement between patients and neurologists based on common identification of each domain as one of the three most important domains or not one of the three most important domains; ‡Kappa analysis based on identification of specific domains as one of the three most important; §Prevalence-Adjusted Bias-Adjusted Kappa: adjusted kappa for imbalances caused by differences in prevalence and bias³³

DISCUSSION

In the present study, agreement between neurologist and patient assessments of MS health status was examined using observable clinical outcomes (number of flares) and more subjective measures of general health, QoL and utility. All the measures provided evidence of a lack of congruence between neurologist and patient ratings of health status. From a clinical perspective, neurologists significantly underestimated the number of flares patients had suffered in the last year, although estimates of life-time flares were more accurate. For other outcomes (general health, QoL, utility), neurologists generally rated their patients' health status better than the patients themselves, based on mean scores for both populations. At the individual level, where patient health status assessments were compared with assessments from their specific neurologists using an ICC analysis, results were even more dramatic with little agreement in ratings. The largest correlation (ICC=0.56) was in QoL indicating a moderate degree of congruence between the neurologists and their patients. However, in the case of standard gamble, the correlation was close to zero (ICC=0.03) and the neurologists estimated significantly higher utility values than their patients. In assessing standard gamble the neurologists were asked to imagine that they had the same health status as their patients. However, because the standard gamble incorporates a measure of risk in defining utility,^{2,22} the neurologist utility assessment for a specific patient likely incorporated their own risk attitude together with a component reflecting the patient's health status.² In the study 40% of the patient utility ratings were < 85 whereas all but one of the neurologists reported scores > 85 indicating that patients were considerably more amenable to higher risk, which was not apparent to their neurologists. A higher risk attitude in MS patients compared to neurologists has also been observed in a

study of patients receiving natalizumab which is associated with a rare but severe adverse event (progressive multifocal leukoencephalopathy (PML)).¹² In the study, which assessed risk tolerance using a non-standard gamble-related questionnaire, patients accepted higher risks of PML than neurologists ($p<0.001$) and were more willing to continue therapy with natalizumab, indicating that neurologists had different perceptions of MS severity and risk attitude than their patients.¹² With new emerging therapies in MS, it is likely that future treatment decisions will frequently involve benefit/risk assessments which will require some degree of patient input and a willingness of neurologists to encourage and utilize this input effectively.^{17,18} The data from the present study suggest that current neurologist/patient perceptions of risk are significantly different and that this divergence may impede the effective utilization of new therapies.

As a further measure of neurologist and patient agreement on health status, both groups identified what they thought were the three most important health-related QoL domains. However, here again there was considerable evidence of a disparity between the patients and their neurologists. While both groups identified physical role limitations as an important health-related QoL domain, the patients placed higher value on mental health than neurologists while neurologists placed more emphasis on physical function and vitality. Moreover, with physical function, which to a neurologist is one of the more quantifiable consequences of MS, the kappa statistic, providing a measure of exact agreement beyond chance alone,³⁰ was 0.003 ($p=0.588$) indicating virtually no specific agreement. This suggests that neurologists may have identified physical function as a significant health domain in MS in general but were unable to specifically identify in which of their patients this domain actually influenced QoL. In reality, mental health had a

predominant role in defining QoL for patients but neurologists were also unable to associate this domain with specific patients (Kappa statistic, -0.11).

These results are in general agreement with a similar study carried out in the United Kingdom where neurologists and MS patients identified SF-36 health domains important in defining QoL, although unlike the present study, the neurologist survey was based on MS patients in general rather than specific patients they treated.³ As in the present study, neurologists overestimated the importance of physical functioning health domains but underestimated the impact of mental health.³ Interestingly, the same study found that in MS patients, domains of physical role limitation, vitality, general health and mental health all correlated with overall health-related QoL, as assessed by the EuroQol questionnaire, while physical functioning did not.³ In the present study the kappa analysis revealed little significant agreement between patients and neurologists in identifying these health domains as important, emphasizing the limited ability of neurologists to identify factors that play significant roles in defining MS patient QoL beyond physical functioning, which does have an impact on patient perception of disease and ability to work.³⁴ However, generalizations on parameters that influence MS QoL are difficult because of the many disease-related symptoms that influence this outcome.^{15,35,36} This complexity clearly adds to the difficulties of defining health domains relevant to QoL in specific MS patients.

There are a number of limitations associated with the present study. This was a small study with only six neurologists participating. However, a one way-ANOVA Fisher's test on response data showed that there were no major differences between neurologists, suggesting a relatively homogeneous group. Nevertheless, the small number of neurologists and patients may limit the generalizability of the results. In addition, the study did not collect MS disability information (for example, expanded disability status scores) and only relied on patient assessed status which in itself is likely subject to reporting bias and to patient perceptions of the status of their disease. The absence of these data precludes clinical definition of the severity of disease in the patients and since degree of disability in MS can impact QoL this may have been a relevant factor in defining health status.³⁶ Also, comparison of clinical outcomes was based on the number of MS flares and neurologists and patients may have had different definitions of flares which possibly contributed to the observed lack of congruence in the estimates of flares over the last 12 months. Recruited patients were also relatively young with a short disease duration (0.7 years based on the difference between mean age and age at diagnosis). Since disease duration has been shown to impact both the physical and mental component summary scores of the SF-36³⁶ this also may have had an impact on patient perception of their MS at this stage of the disease and the ability of the neurologist to recognize important health states. Nevertheless, the data represent a cross-sectional assessment of the status of neurologist/patient agreement in MS and support general conclusions on the challenges of neurologist/patient communication. Finally, instruments used in the study to assess health status have not been validated specifically for a physician/patient comparison. In addition, the neurologists assessed patient general health and QoL using Likert scale questions extracted from the MSQOL-54,

while patients completed them in the context of the complete questionnaire; using these specific MSQOL-54 questions in isolation has not been validated as far as we are aware. However, similar Likert scale-based questionnaires have been used to compare physician and patient perceptions of disease status so this is unlikely to have affected the overall study conclusions.^{37,38}

CONCLUSIONS

Effective intervention in MS requires neurologists to assess patient health status so that treatments balance control of physical disability with maintenance of acceptable QoL. The present study suggests that neurologists have an incomplete understanding of patient perceptions of MS, beyond the physical aspects of the disease, and that in an era of increasingly complex and challenging MS therapies, effective neurologist/patient communication needs to be enhanced.

ACKNOWLEDGEMENTS

This study was funded by a grant from Biogen Idec Canada Inc. The authors thank John McCormick, PhD and Merlin Njoya, MSc from McKesson Specialty Canada, for editorial assistance in preparation of the manuscript and for assistance with the statistical analyses, respectively.

REFERENCES

1. Fowler FJ, Levin CA, Sepucha KR. Informing and involving patients to improve the quality of medical decisions. *Health Aff.* 2011;30(4):699-706.
2. Suarez-Almazor ME, Conner-Spady B, Kendall CJ, Russell AS, Skeith K. Lack of congruence in the ratings of patients' health status by patients and their physicians. *Med Decis Making.* 2001;21(2):113-21.
3. Rothwell PM, McDowell Z, Wong CK, Dorman PJ. Doctors and patients don't agree: cross sectional study of patients' and doctors' perceptions and assessments of disability in multiple sclerosis. *BMJ.* 1997;314(7094):1580-3.
4. Heesen C, Solari A, Giordano A, Kasper J, Kopke S. Decisions on multiple sclerosis immunotherapy: new treatment complexities urge patient engagement. *J Neurol Sci.* 2011;306(1-2):192-7.
5. Janse AJ, Gemke RBB, Uiterwaal CSPM, van der Tweel I, Kimpen JLL, Sinnema G. Quality of life: patients and doctors don't always agree: a meta-analysis. *J Clin Epidemiol.* 2004;57(7):653-61.
6. Oliva EN, Nobile F, Alimena G, et al. Quality of life in elderly patients with acute myeloid leukemia: patients may be more accurate than physicians. *Haematologica.* 2011;96(5):696-702.
7. Sonn GA, Sadetsky N, Presti JC, Litwin MS. Differing perceptions of quality of life in patients with prostate cancer and their doctors. *J Urol.* 2009;182(5):2296-302.
8. Johnson FR, Ozdemir S, Mansfield C, et al. Crohn's disease patients' risk-benefit preferences: serious adverse event risks versus treatment efficacy. *Gastroenterology.* 2007;133(3):769-79.
9. Mantyselka P, Kumpusalo E, Ahonen R, Takala J. Patients' versus general practitioners' assessments of pain intensity in primary care patients with non-cancer pain. *Br J Gen Pract.* 2001;51(473):995-7.
10. Wilson KA, Dowling AJ, Abdoell M, Tannock IF. Perception of quality of life by patients, partners and treating physicians. *Qual Life Res.* 2000;9(9):1041-52.
11. Sherbourne CD, Sturm R, Wells KB. What outcomes matter to patients? *J Gen Intern Med.* 1999;14(6):357-63.
12. Heesen C, Kleiter I, Nguyen F, et al. Risk perception in natalizumab-treated multiple sclerosis patients and their neurologists. *Mult Scler* 2010;16(12):1507-12.

13. Spain LA, Tubridy N, Kilpatrick TJ, Adams SJ, Holmes ACN. Illness perception and health-related quality of life in multiple sclerosis. *Acta Neurol Scand*. 2007;116(5):293-9.
14. Forbes A, While A, Taylor M. What people with multiple sclerosis perceive to be important to meeting their needs. *J Adv Nursing*. 2007;58(1):11-22.
15. Miller DM, Allen R. Quality of life in multiple sclerosis: Determinants, measurements, and use in clinical practice. *Curr Neurol Rep*. 2010;10(5):397-406.
16. Yeh EA, Weinstock-Guttman B. Fingolimod: an oral disease-modifying therapy for relapsing multiple sclerosis. *Adv Ther*. 2011;28(4):270-8.
17. Kieseier BC, Stuve O. A critical appraisal of treatment decisions in multiple sclerosis - old versus new. *Nature Rev Neurol*. 2011;7(5):255-62.
18. Johnson FR, Van Houtven G, Ozdemir S, et al. Multiple sclerosis patients' benefit-risk preferences: serious adverse event risks versus treatment efficacy. *J Neurol*. 2009;256(4):554-62.
19. Heesen C, Kasper J, Kopke S, Richter T, Segal J, Muhlhauser I. Informed shared decision making in multiple sclerosis - inevitable or impossible? *J Neurol Sci*. 2007;259(1-2):109-17.
20. Heesen C, Kopke S, Richter T, Kasper J. Shared decision making and self-management in multiple sclerosis - a consequence of evidence. *J Neurol*. 2007;254 (Suppl 2):II116-21.
21. Mendel R, Traut-Mattausch E, Frey D, et al. Do physicians' recommendations pull patients away from their preferred treatment options? *Health Expect*. 2012;15(1):23-31.
22. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the economic evaluation of health care programmes*. 3rd ed. New York: Oxford University Press; 2005.
23. Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36 Health Survey: Manual and interpretation guide*. Boston: The Health Institute, New England Medical Center, 1993.
24. Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A health-related quality of life measure for multiple sclerosis. *Qual Life Res*. 1995;4(3):187-206.
25. Benito-Leon J, Manuel Morales J, Rivera-Navarro J, Mitchell AJ. A review about the impact of multiple sclerosis on health-related quality of life. *Disab Rehab*. 2003;25(23):1291-303.
26. Eisen GM, Locke R, Provenzale D. Health-related quality of life: A primer for gastroenterologists. *Am J Gastroenterol*. 1999;94(8):2017-21.
27. Krokavcova M, Van Dijk JP, Nagyova I, et al. Perceived health status as measured by the SF-36 in patients with multiple sclerosis: a review. *Scand J Caring Sci*. 2009;23(3):529-38.
28. Laugsand EA, Sprangers MAG, Bjordal K, Skorpen F, Kaasa S, Klepstad P. Health care providers underestimate symptom intensities of cancer patients: A multicenter European study. *Health Qual Life Outcomes*. 2010;8:104.
29. Lee J, Koh D, Ong CN. Statistical evaluation of agreement between two methods for measuring a quantitative variable. *Comput Biol Med*. 1989;19(1):61-70.
30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74.
31. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37(5):360-3.
32. Murri R, Ammassari A, Trotta MP, et al. Patient-reported and physician-estimated adherence to HAART. *J Gen Intern Med*. 2004;19(11):1104-10.
33. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46(5):423-9.
34. van Asch P. Impact of mobility impairment in multiple sclerosis 2 - patients' perspectives. *Europ Neurol Rev*. 2011;6(2):115-20.
35. Nortvedt MW, Riise T. The use of quality of life measures in multiple sclerosis research. *Mult Scler*. 2003;9(1):63-72.
36. Fernandez O, Baumstarck-Barrau K, Simeoni M-C, Auquier P and the MusiQoL study group. Patient characteristics and determinants of quality of life in an international population with multiple sclerosis: Assessment using the MusiQoL and SF-36 questionnaires. *Mult Scler*. 2011;17(10):1238-49.
37. Muhlbacher AC, Nubling M. Analysis of physicians' perspectives versus patients' preferences: direct assessment and discrete choice experiments in the therapy of multiple myeloma. *Eur J Health Econ*. 2011;12(3):193-203.
38. Melville JL, Miller EA, Fialkow MF, Lentz GM, Miller JL, Fenner DE. Relationship between patient report and physician assessment of urinary incontinence severity. *Am J Obstet Gynecol*. 2003;189(1):76-80.