Author/s:
Tai, G;Corben, LA;Yiu, EM;Delatycki, MB

Title:
A longitudinal study of the SF-36 version 2 in Friedreich ataxia

Date:
2017-07-01

Citation:
Tai, G., Corben, L. A., Yiu, E. M. & Delatycki, M. B. (2017). A longitudinal study of the SF-36 version 2 in Friedreich ataxia. Acta Neurologica Scandinavica, 136 (1), pp.41-46. https://doi.org/10.1111/ane.12693.

Persistent Link:
https://hdl.handle.net/11343/291780

Received Date : 02-Jun-2016

Revised Date : 04-Sep-2016

Accepted Date : 12-Sep-2016

Article type : Original Article

A longitudinal study of the SF-36 version 2 in Friedreich ataxia.

Geneieve Tai[1], Louise A Corben[1,2,3,4], Eppie M Yiu[1,3,5], Martin B Delatycki[1,2,3,6]

[1]Bruce Lefroy Centre for Genetic Health Research, Murdoch Childrens Research Institute, Parkville, Victoria, 3052, Australia.

[2]School of Psychological Science, Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Victoria, 3168, Australia.

[3]Department of Paediatrics, University of Melbourne, Parkville, Victoria, 3052, Australia.

[4]Department of Occupational Therapy, Monash Health, Clayton, Victoria, 3168, Australia.

[5]Department of Neurology, Royal Children's Hospital, Parkville, Victoria, 3052, Australia.

[6]Department of Clinical Genetics, Austin Health, Heidelberg, Victoria, 3084, Australia.

Corresponding author:

Professor Martin Delatycki

Director, Bruce Lefroy Centre for Genetic Health Research

Murdoch Childrens Research Institute

Flemington Rd, Parkville

Victoria, 3052, Australia

Ph: +61 3 9496 4355

Fax: +61 3 8341 6390

E-mail: martin.delatycki@ghsv.org.au

Running title: SF-36 Version 2 in Friedreich ataxia

Word count: 2873

Abstract

Objectives: The Medical Outcomes Study 36 item Short Form Health Survey (SF-36) is one of the most commonly used patient reported outcome (PRO) measure. This study aimed to examine the relationship between SF-36 version 2 (SF-36V2) summary scores and Friedreich ataxia (FRDA) clinical characteristics, and to investigate the responsiveness of the scale, in comparison to the Friedreich Ataxia Rating Scale (FARS), over one, two and three years.

Materials and Methods: Descriptive statistics were used to examine the characteristics of the cohort at baseline and years 1, 2 and 3. Correlations between FRDA clinical characteristics and SF-36V2 summary scores were reported. Responsiveness was measured using paired t-tests.

Results: We found significant correlations between the Physical Component Summary (PCS) of the SF-36V2 and various FRDA clinical parameters but none for the Mental Component Summary (MCS). No significant changes in the SF-36V2 were seen over one or two years, however PCS scores at Year 3 were significantly lower than at baseline (-3.3, SD (7.6), p=0.01). FARS scores were found to be significantly greater at Years 1, 2 and 3 when compared to baseline.

Conclusions: Our findings suggest that despite physical decline, individuals with FRDA have relatively stable mental wellbeing. This study demonstrates that the SF-36V2 is unlikely to be a useful tool for identifying clinical change in FRDA therapeutic trials.

Keywords: Friedreich ataxia, health status, rating scales, SF-36

Introduction

Friedreich ataxia (FRDA) is a progressive neurodegenerative condition affecting approximately 1 in 29,000 Caucasian individuals [1, 2]. Features of the disorder include progressive ataxia and hypertrophic cardiomyopathy. The majority (~96%) of individuals with FRDA are homozygous for a GAA triplet repeat expansion in intron 1 of *FXN* [3, 4].

There are no treatments proven to delay or halt the progression of FRDA [5]. There are however, a number of pharmaceutical agents in clinical trial that are postulated to slow disease progression in FRDA through various different mechanisms, including the reduction of iron accumulation, diminution of oxidative stress by antioxidants and increasing frataxin expression [6]. An example is the recent study by Seyer and colleagues who tested exogenous interferon gamma-1b, shown to increase frataxin levels in cell lines originating from individuals with FRDA, in an open-label study of 12 children [7]. Measuring disease progression in FRDA is challenging due to its slow progression

and variable phenotype [8]. It is thus vital that tools used to measure disease progression are able to detect clinically significant, albeit small, changes in individuals with FRDA.

Neurological function in FRDA is currently assessed using a variety of neurological rating scales including the Friedreich Ataxia Rating Scale (FARS), International Cooperative Ataxia Rating Scale (ICARS) and, most recently, the Scale for the Assessment and Rating of Ataxia (SARA) [9-11]. These rating scales are administered by trained clinicians, and while they provide a good indicator of disease progression and severity, they do not incorporate the perspectives of individuals with FRDA. Patient reported outcome (PRO) measures encompass aspects of a condition not evaluated by clinician rated tools, taking into account the viewpoints of individuals with the condition. A PRO is described as a measurement of any aspect of an individual's health that comes directly from the individual and has not been interpreted by a healthcare professional in any way [12].

The inclusion of PRO measures in pharmaceutical trials is recommended by the Food and Drug Administration (FDA), as they provide valuable information in determining the impact an intervention or drug has on the perception of an individual's health status [2, 12]. One of the most commonly used PRO measures is the Medical Outcomes Study 36 item Short Form Health Survey (SF-36) [13]. The SF-36 is a generic measure of health status and comprises 36 items that are categorised into eight dimensions measuring physical function, role-physical, bodily pain, general health, vitality, social function, role-emotion and mental health. The SF-36 has been widely studied and an updated version has since been published, the SF-36 version 2 (SF-36V2) [14]. The newer version has improved psychometric qualities in comparison to the original [14]. Ceiling and floor effects, in particular, have been reduced due to improvements in the phrasing of the items as well as the change from dichotomous response categories to five point response categories [14].

Both versions of the SF-36 have previously been examined in FRDA [2, 15, 16]. In one cross-sectional study, individuals with FRDA were found to have worse perception of their health status and quality of life when compared to Australian population norms [15]. Epstein and colleagues reported similar findings when comparing individuals with FRDA to a US cohort [16]. Psychometric properties of the SF-36 was studied by Riazi and colleagues who found high floor and ceiling effects indicating reduced specificity in an FRDA population [2]. No study, however, has examined either version of the SF-36 in FRDA longitudinally.

The objectives of this current study were to examine the relationship between SF-36V2 summary scores and FRDA clinical characteristics, as well as to investigate the responsiveness of the scale, in comparison to the FARS, over one, two and three years.

Material and methods

Participants

Individuals homozygous for a GAA expansion in intron 1 of *FXN* and aged at least 18 years were recruited from a dedicated Friedreich ataxia clinic at Monash Health in Victoria, Australia. The SF-36V2 forms were sent out to potential participants prior to their annual clinic appointment. Participants returned the completed forms via post or in person at their clinic appointment. Individuals could seek assistance from a family member or carer if they had difficulty completing the forms. The FARS was conducted at the same clinic visit – it is scored out of 167, with a higher score indicating more severe disease [10].

Data analysis

Data from the SF-36V2 was scored according to procedures described by Ware et al and Hawthorne et al [17, 18]. Using Australian population data, items were coded, summed and transformed into percentage and T-scores as per Wilson et al [15]. The scale was then summarised into two main components resulting in the physical component summary score (PCS) and the mental component summary score (MCS). A higher score is indicative of better perceived health status.

Descriptive statistics were used to examine the characteristics of the cohort at baseline and years 1, 2 and 3. A summary of the SF-36V2 dimension percentage scores at baseline was provided. Correlations between FRDA clinical characteristics and SF-36V2 summary scores were reported. Data was found to be normally distributed hence Pearson's product-moment correlation coefficients were used. The clinical characteristics that were examined in the correlation analyses were age of disease onset, disease duration, age at review, the smaller (GAA1) and larger (GAA2) intron 1 *FXN* GAA repeat sizes and total FARS score.

Responsiveness is the ability of an instrument to measure clinically relevant change over time [19]. This was measured using paired t-tests to examine the change in SF-36V2 summary scores and FARS between baseline and years 1, 2 and 3. Other indicators of responsiveness were also examined. Effect size, the mean change in score over time divided by the standard deviation of the baseline score, was calculated; a larger effect size indicates a more responsive scale [2, 20]. The standardised response mean (SRM) was calculated by dividing the mean score change by the standard deviation of the change [21-23]. As with effect size, a higher SRM represents a more responsive scale.

All statistical analyses were performed using Stata Statistical Software: Release 13 (StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP).

Ethics Committee approval

Approval for this study was obtained from the Monash Health Human Research Ethics Committee (HREC 02114A). All participants gave informed, written consent in accordance with the Declaration of Helsinki. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Results

One hundred and twenty-two individuals completed the SF-36V2 at baseline. Sixty-two completed the questionnaire at Year 1, approximately 12 months later, 47 at Year 2 (24 months) and 37 at Year 3 (36 months).

The characteristics of the cohort at all four time points are reported in Table 1. Half of the cohort at baseline were male (50 %), and the average age at assessment was 33.7 years (SD 12.5, range 18-82 years). The mean age of disease onset was 16.1 years (SD 8.5, range 2 – 55 years) with a mean disease duration of 17.6 years (SD 10.3, range 2 – 48.3 years). Mean GAA1 repeat size was 627.2 (SD 226.9, range 56 – 1099) and average size of the GAA2 repeat was 864.1 (SD 212.2, range 182 – 1345). The mean FARS score at baseline was 89.9 (SD 30.5, range 19-151). Characteristics for the cohort at Year 1, 2 and 3 were generally similar when compared to baseline.

The summary percentage scores at baseline are shown on Table 2. The mean physical component summary (PCS) score was 33.3 (SD 9.0) and the mental component summary (MCS) score was 48.3 (SD 13.0).

Correlation analyses

Table 3 shows correlations between various clinical characteristics and the PCS and MCS scores of the SF-36V2. The PCS correlated significantly with disease duration (r=-0.40, $p<0.01$), age at review (r=-0.27, $p<0.01$) and total FARS score (r=-0.47, $p<0.01$). There were no significant correlations between the MCS and any of the clinical characteristics of FRDA.

Responsiveness

The responsiveness of the SF-36V2 summary scales, PCS and MCS and FARS was measured over one, two and three years (Table 4).

Responsiveness over one year

There were no significant changes in SF-36V2 summary scores over one year, while the mean increase in FARS score from baseline to Year 1 was 4.3 (SD 6.8, t(61)=5.00, $p<0.01$). As expected,

the largest effect size and standardised response mean (SRM) was seen in the FARS (0.17 and 0.63, respectively).

Responsiveness over two years

As with responsiveness over one year, no significant differences between Year 2 and baseline were found for PCS and MCS. The mean FARS score was significantly greater at Year 2 when compared to baseline with a difference of 4.4 (SD 8.8, t(46)=-3.42, p<0.01).

Responsiveness over three years

Scores from PCS were found to be significantly lower at Year 3 than at baseline (mean difference -3.3 points, SD 7.6, t(36)=-2.63, p=0.01), with a moderate effect size of -0.50. The difference in mean FARS score between baseline and Year 3 was also found to be significant, with a mean difference of 7.5 points (SD 11.2, t(36)=-4.06, p<0.01), and an effect size of 0.31.

Summary scores over three years

Scores for the both SF-36V2 summary scales and FARS for participants who completed both assessments over all time points are presented in Figure 1. PCS scores gradually decreased over time whereas MCS scores are fairly steady over time. FARS scores increased the most from baseline to Year 1 and then further increased from Year 2 to Year 3.

Discussion

In this longitudinal study, we examined the relationship between the SF-36V2 and FRDA clinical characteristics and studied the responsiveness of the SF-36V2 in people with FRDA over one, two and three years.

The physical component summary (PCS) of the SF-36V2 was found to correlate significantly with disease duration, age at review and FARS score, whereas the mental component (MCS) showed no significant correlations with any FRDA disease parameters. This finding suggests that despite physical decline, people have relatively stable mental wellbeing, echoing results reported by Wilson and colleagues [15]. Similar findings were also noted in a longitudinal study of the Friedreich Ataxia Impact Scale (FAIS), an FRDA specific PRO measure, in which no significant correlations between FAIS subscales assessing psychological and social impact and FRDA clinical characteristics were found [24].

No significant changes were found in the PCS and MCS over one and two years. Significant change in PCS was found over three years however this was not found for the MCS. One of the main limitations of the original version of the SF-36 was its lack of responsiveness as reported in various

studies [2, 16]. This appears to also be the case for SF-36V2 as demonstrated in this current study, where significant change was only seen after three for the PCS. The responsiveness of an instrument is critical when determining its suitability in detecting change over time. The SF-36V2 is thus unlikely to be useful in measuring change in short term clinical therapeutic trials. The limited responsiveness could also be explained by the heterogeneous nature of the condition, with different FRDA functions being affected at various stages [2].

Both versions of the SF-36 have been studied in FRDA previously [2, 15, 16]. These studies demonstrated a substantial health impact of FRDA on quality of life. Our study confirms this finding. Basic psychometric criteria were fulfilled in the original SF-36 however high floor and ceiling effects were found and small effect sizes were reported [2]. Epstein and colleagues (2008) compared SF-36 scores in individuals with FRDA to US population norms and reported no significant differences in MCS scores between those with FRDA and the general population [16].

Other neurodegenerative conditions in which the SF-36 has been studied include multiple sclerosis (MS) and Parkinson disease (PD). In an MS rehabilitation study, small effect sizes were reported for the eight dimensions, with only two demonstrating significant change between admission and discharge (a period of approximately 20 days). In contrast, other measures used in this study (the Functional Independence Measure (FIM), the London Handicap Scale (LHS) and the General Health Questionnaire (GHQ)) demonstrated significant change in the same period and with larger effect sizes. The authors explain this finding by suggesting the topics measured in the SF-36 are much less specific when compared to the other instruments used in the study [25]. In an assessment of basic assumptions of the SF-36 in another MS cohort, Hobart and colleagues (2001) reported good data quality and variability, however found significant floor and ceiling effects in four of the eight SF-36 dimensions [26].

A study of health related quality of life measures in PD reported greater responsiveness of the SF-36 over PD-specific measures (Parkinson's Disease Questionnaire (PDQ-39) and Parkinson's Disease Quality of Life (PDQUALIF) scale) over 18 months. While PD-specific instruments cover issues that are more relevant to individuals with PD, they were found to have lower validity than the SF-36 [27]. Hobart and colleagues suggested that the SF-36 is better suited to a cross sectional setting rather than in clinical trials or longitudinal studies as the floor and ceiling effects observed may mask the effectiveness of potential therapies or changes in health status over time [26].

While the limitations of using a general health status PRO for a specific condition are known [15], the SF-36V2 remains the most widely used generic health status measure and has been extensively studied. The SF-36V2 enables comparisons with the general population and is relatively

straightforward to administer. It also contains a much lower number of items when compared to the Friedreich Ataxia Impact Scale (FAIS), an FRDA specific PRO measure [28]. The FAIS has been studied in a longitudinal setting to measure its responsiveness over one and two years [24]. Limited responsiveness was found, with only one subscale (speech) demonstrating significant change over one and two years. Considering an FRDA-specific measurement tool was found to be relatively unresponsive to change over time, it is not surprising that we found a similar outcome for a generic health status measurement tool. In addition, FRDA is a slowly progressive disease and studies of change over one year of various clinical measures have revealed either modest or insignificant change [29-32]. This is another factor that makes the minimal change in PCS in the current study unsurprising.

PRO measures encompass individuals' perspectives on disease impact and can provide useful information for therapeutic trials. Nevertheless, this study demonstrates that the SF-36V2 is unlikely to be a useful tool for identifying change in FRDA therapeutic trials.

Acknowledgments

The authors thank the participants for their involvement in this study.

Sources of Funding

Funding was from the Friedreich Ataxia Research Association (Australasia) and the Friedreich Ataxia Research Alliance (USA). EMY is supported by a National Health Medical Research Council (NHMRC) Early Career Fellowship and LAC is supported by a National Health Medical Research Council (NHMRC) Early Career Fellowship. MBD receives grant support from NHMRC. This work was made possible through Victorian State Government Operational Infrastructure Support and Australian Government NHMRC IRIISS.

Conflict of Interest

The authors declare that they have no conflict of interest.

List of abbreviations

SF-36 - Medical Outcomes Study 36 item Short Form Health Survey
PRO – Patient reported outcomes
SF-36V2 - Medical Outcomes Study 36 item Short Form Health Survey Version 2
FRDA – Friedreich ataxia
FARS – Friedreich ataxia rating scale
PCS – Physical Component Summary of the SF-36V2
MCS – Mental Component Summary of the SF-36V2

ICARS - International Cooperative Ataxia Rating Scale

SARA – Scale for the assessment and rating of ataxia

FDA – Food and Drug Administration

GAA1 – the smaller of the intron 1 *FXN* GAA repeat

GAA2 – the larger of the intron 1 *FXN* GAA repeat

MS – Multiple sclerosis

PD – Parkinson Disease

FIM – Functional independence measure

LHS – London Handicap Scale

GHQ – General Health Questionnaire

PDQ-39 – Parkinson's Disease Questionnaire

PDQUALIF – Parkinson's Disease Quality of Life

FAIS – Friedreich ataxia impact scale

References

1. Cossee M, Schmitt M, Campuzano V, Reutenauer L, Moutou C, Mandel JL, *et al.* Evolution of the Friedreich's ataxia trinucleotide repeat expansion: founder effect and premutations. Proceedings of the National Academy of Sciences of the United States of America 1997; 94: 7452-7457.
2. Riazi A, Cano SJ, Cooper M, Bradley JL, Schapira AHV, Hobart JC. Coordinating outcomes measurement in ataxia research: Do some widely used generic rating scales tick the boxes? Movement Disorders 2006; 21: 1396-1403.
3. Durr A, Cossee M, Agid Y, Campuzano V, Mignard C, Penet C, *et al.* Clinical and genetic abnormalities in patients with Friedreich's ataxia. The New England Journal of Medicine 1996; 335: 1169 - 1175.
4. Campuzano V, Montermini L, Molto M, Pianese L, Cossee M, Cavalcanti F, *et al.* Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. Science 1996; 271: 1423 - 1427.
5. Delatycki MB. Evaluating the progression of Friedreich ataxia and its treatment. Journal of Neurology 2009; 256 Suppl 1: 36-41.
6. Perlman SL. A Review of Friedreich Ataxia Clinical Trial Results. Journal of Child Neurology 2012; 27: 1217-1222.
7. Seyer L, Greeley N, Foerster D, Strawser C, Gelbard S, Dong Y, *et al.* Open-label pilot study of interferon gamma-1b in Friedreich ataxia. Acta Neurologica Scandinavica 2015; 132: 7-15.

8. Montermini L, Richter A, Morgan K, Justice CM, Julien D, Castellotti B, *et al.* Phenotypic variability in Friedreich ataxia: role of the associated GAA triplet repeat expansion. Annals of Neurology 1997; 41: 675-682.
9. Trouillas P, Takayanagi T, Hallett M, Currier RD, Subramony SH, Wessel K, *et al.* International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. The Ataxia Neuropharmacology Committee of the World Federation of Neurology. Journal of the Neurological Sciences 1997; 145: 205-211.
10. Subramony SH, May W, Lynch DR, Gomez CM, Fischbeck KH, Hallett M, *et al.* Measuring Friedreich ataxia: Interrater reliability of a neurologic rating scale. Neurology 2005; 64: 1261-1262.
11. Schmitz-Hubsch T, du Montcel ST, Baliko L, Berciano J, Boesch S, Depondt C, *et al.* Scale for the assessment and rating of ataxia: development of a new clinical scale. Neurology 2006; 66: 1717-1720.
12. Health USDo, Evaluation HSFCfD, Research, Health USDo, Evaluation HSFCfB, Research, *et al.* Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health and Quality of Life Outcomes 2006; 4: 79.
13. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Medical Care 1992; 30: 473-483.
14. Jenkinson C, Stewart-Brown S, Petersen S, Paice C. Assessment of the SF-36 version 2 in the United Kingdom. Journal of Epidemiology and Community Health 1999; 53: 46-50.
15. Wilson CL, Fahey MC, Corben LA, Collins VR, Churchyard AJ, Lamont PJ, *et al.* Quality of life in Friedreich ataxia: what clinical, social and demographic factors are important? European Journal of Neurology 2007; 14: 1040-1047.
16. Epstein E, Farmer JM, Tsou A, Perlman S, Subramony SH, Gomez CM, *et al.* Health related quality of life measures in Friedreich Ataxia. Journal of the Neurological Sciences 2008; 272: 123-128.
17. Ware J, Kosinski M, Dewey J. How to Score Version 2 of the SF-36 Health Survey. Lincoln, RI: Quality Metric Inc; 2000.
18. Hawthorne G, Osborne R, Taylor A, Sansoni J. The SF36 Version 2: critical analyses of population weights, scoring algorithms and population norms. Quality of Life Research 2007; 16: 661-673.
19. Lohr Kl. Assessing health status and quality-of-life instruments: Attributes and review criteria. Quality of Life Research 2002; 11: 193-205.
20. Kazis L, Anderson J, Meenan R. Effect sizes for interpreting changes in health status. Medical Care 1989; 27: 178-189.
21. Liang M, Fossel A, Larson M. Comparisons of five health status instruments for orthopedic evaluation. Medical Care 1990; 28.

22. Terwee C, Bot S, de Boer M, van der Windt D, Knol D, Dekker J, *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology 2007; 60: 34-42.
23. Terwee C, Dekker F, Wiersinga W, Prummel M, Bossuyt P. On assessing responsiveness of health-related quality of life instruments: Guidelines for instrument evaluation. Quality of Life Research 2003; 12.
24. Tai G, Yiu EM, Corben LA, Delatycki MB. A longitudinal study of the Friedreich Ataxia Impact Scale. Journal of the Neurological Sciences 2015; 352: 53-57.
25. Freeman JA, Hobart JC, Langdon DW, Thompson AJ. Clinical appropriateness: a key factor in outcome measure selection: the 36 item short form health survey in multiple sclerosis.[see comment]. Journal of Neurology, Neurosurgery & Psychiatry 2000; 68: 150-156.
26. Hobart J, Freeman J, Lamping D, Fitzpatrick R, Thompson A. The SF-36 in multiple sclerosis: why basic assumptions must be tested. Journal of Neurology, Neurosurgery & Psychiatry 2001; 71: 363-370.
27. Brown CA, Cheng EM, Hays RD, Vassar SD, Vickrey BG. SF-36 includes less Parkinson Disease (PD)-targeted content but is more responsive to change than two PD-targeted health-related quality of life measures. Quality of Life Research 2009; 18: 1219-1237.
28. Cano SJ, Riazi A, Schapira AHV, Cooper JM, Hobart JC. Friedreich's ataxia impact scale: A new measure striving to provide the flexibility required by today's studies. Movement Disorders 2009; 24: 984-992.
29. Lynch DR, Farmer JM, Tsou AY, Perlman S, Subramony SH, Gomez CM, *et al.* Measuring Friedreich ataxia: Complementary features of examination and performance measures. . Neurology 2006; 66: 1711-1716.
30. Lynch DR, Farmer JM, Wilson RL, Balcer LJ. Performance measures in Friedreich ataxia: potential utility as clinical outcome tools. Movement Disorders 2005; 20: 777-782.
31. Reetz K, Dogan I, Costa AS, Dafotakis M, Fedosov K, Giunti P, *et al.* Biological and clinical characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS) cohort: a cross-sectional analysis of baseline data. The Lancet Neurology 2015; 14: 174-182.
32. Tai G, Corben LA, Gurrin L, Yiu EM, Churchyard A, Fahey M, *et al.* A study of up to 12 years of follow-up of Friedreich ataxia utilising four measurement tools. Journal of Neurology, Neurosurgery & Psychiatry 2015; 86: 660-666.

Table 1. Characteristics of the cohort at Baseline, Year 1, Year 2 and Year 3

Characteristics	Baseline (n=122)	Year 1 (n=62)	Year 2 (n=47)	Year 3 (n=37)
Gender, male, n (%)	61 (50)	34 (54.8)	27 (57.4)	20 (54.1)
Age, years, mean (SD)	33.7 (12.5)	33 (11.3)	34.9 (11.5)	35.3 (11.8)
Range	18 – 82	19 – 59	20 – 58	21 – 59
Onset age, years, mean (SD)	16.1 (8.5)	16 (6.6)	16.1 (6.7)	16.4 (6.7)
Range	2 – 55	3 – 32	3 – 30	6 – 30
Disease duration, years, mean (SD)	17.6 (10.3)	16.6 (8.8)	18.7 (9.5)	18.8 (9.0)
Range	2 – 48.3	3 – 40.9	4 – 40.3	5 – 40.8
GAA1, mean (SD)	627.2 (226.9)	613.0 (209.7)	653.6 (212.3)	642.5 (221.7)
Range	56 – 1099	169 – 1077	182 – 1077	182 – 1050
GAA2, mean (SD)	864.1 (212.2)	859.9 (230.6)	925.7 (175.8)	926.1 (178.5)
Range	182 – 1345	182 – 1345	182 – 1345	182 – 1293
FARS, mean (SD)	89.9 (30.5)	86.5 (25.3)	87.9 (25.4)	88.6 (23.7)
Range	19 – 151	24 – 132	28.2 – 133	30 – 125

Legend: GAA1- smaller FXN intron 1 GAA repeat size, GAA2- larger FXN intron 1 GAA repeat size, FARS- Friedreich Ataxia Rating Scale total score (maximum score is 167).

Table 2. Summary of SF-36V2 percentage scores at baseline

SF-36V2 dimension	N	Mean	SD
Physical functioning	122	23.1	25.2
Role physical	122	59.6	31.5
Bodily pain	122	65.6	20.6
General health	122	50.3	22.2
Vitality	122	49.3	16.2
Social function	122	67.8	26.4
Role emotion	122	78.1	24.6
Mental health	122	69.4	18.5
Physical component summary score (PCS)	122	33.3	9.0
Mental component summary score (MCS)	122	48.3	13.0

Maximum percentage score is 100, with a higher score reflecting better perceived heath status. These scores allow comparison to the Australian population which are standardised to a mean score of 50 (SD 10).

Table 3. Correlations between various clinical parameters and the Physical and Mental Component Summary scores at baseline (n=122)

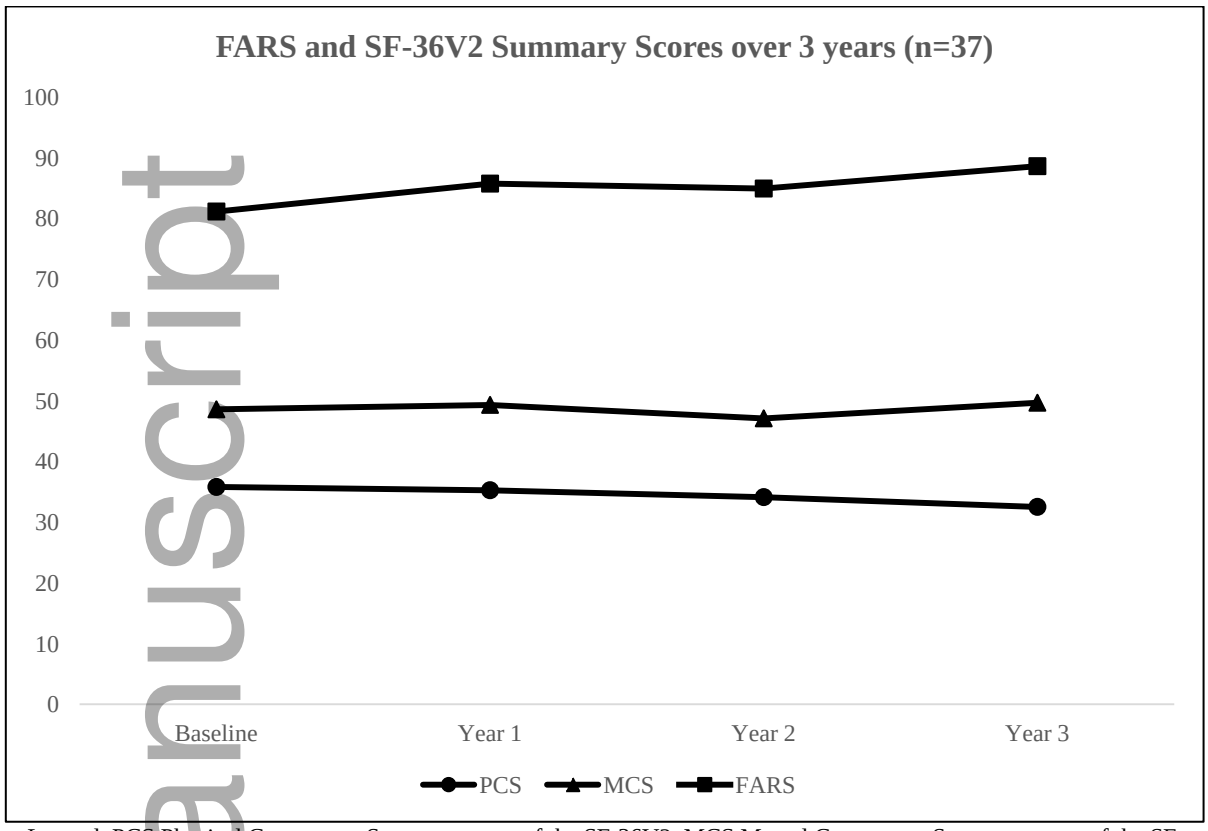
	Physical component summary score (PCS)	**Mental component summary score (MCS)**
Onset age	0.09	-0.13
Disease duration	-0.40**	0.16
Age at review	-0.27**	0.05
GAA1	-0.13	0.11
GAA2	-0.03	0.06
FARS	-0.47**	0.17

Legend: GAA1- smaller FXN intron 1 GAA repeat size, GAA2- larger FXN intron 1 GAA repeat size, FARS- Friedreich Ataxia Rating Scale total score, **p-value<0.01

Table 4. Responsiveness of SF-36V2 summary scores and FARS

Responsiveness of SF-36V2 summary scores and FARS over one year			
	Physical Component Summary (PCS)	**Mental Component Summary (MCS)**	**FARS**
N	62	62	62
Mean score at Baseline (SD)	34.3 (7.8)	47.3 (13.0)	83.0 (26.8)
Mean score at Year 1 (SD)	34.1 (8.5)	48.2 (13.0)	87.4 (25.5)
Difference between scores (SD) (Year 1 - Baseline)	-0.2 (7.1)	0.8 (10.1)	4.3 (6.8)
T test	-0.23	0.63	5.00
P value	0.82	0.53	**<0.01**
Effect size	-0.02	0.06	0.17
Standardised response mean	-0.03	0.08	0.63
Responsiveness of SF-36V2 summary scores and FARS over two years			
N	47	47	47
Mean score at Baseline (SD)	34.4 (7.4)	48.5 (13.4)	83.5 (26.7)
Mean score at Year 2 (SD)	33.3 (8.4)	48.2 (12.1)	87.9 (25.4)
Difference between scores (SD) (Year 2 - Baseline)	-1.0 (6.4)	-0.2 (12.9)	4.4 (8.8)
T test	-1.10	-0.12	3.42
P value	0.28	0.91	**<0.01**
Effect size	-0.12	-0.02	0.17
Standardised response mean	-0.16	-0.02	0.50
Responsiveness of SF-36V2 summary scores and FARS over three years			
N	37	37	37
Mean score at Baseline (SD)	35.8 (7.2)	48.5 (13.5)	81.1 (24.8)
Mean score at Year 3 (SD)	32.5 (6.6)	49.6 (14.2)	88.6 (23.7)
Difference between scores (SD) (Year 3 - Baseline)	-3.3 (7.6)	1.1 (10.9)	7.5 (11.2)
T test	-2.63	0.61	4.06
P value	**0.01**	0.55	**<0.01**
Effect size	-0.50	0.08	0.31
Standardised response mean	-0.43	0.10	0.67

Figure 1. FARS and SF-36V2 Summary scores over 3 years



Legend: PCS Physical Component Summary score of the SF-36V2, MCS Mental Component Summary score of the SF-36V2, FARS Friedreich Ataxia Rating Scale.