

**ASSESSMENT OF MOOD MEASURES FOR PEOPLE WITH
MULTIPLE SCLEROSIS**

TESSA MARIE HOPKINS, BSc.

**Thesis Submitted to the University of Nottingham
For the degree of Doctor of Clinical Psychology**

DECEMBER 2011

Abstract

There are a lack of valid assessment measures of anxiety and depression available for use with people with multiple sclerosis (MS). As a result of few valid measures individuals with these mood disorders and MS have poor access to treatment; the true prevalence of mood disorders is unknown; research investigating the relationship between anxiety, depression and MS is limited. Some previous attempts to validate measures of anxiety and depression have been conducted in this population, but these have included a number of methodological flaws.

To address the concerns highlighted in the literature the current study attempted to validate three measures commonly used in clinical practice to assess depression and anxiety in people with MS. These were: the Beck Anxiety Inventory (BAI); the Beck Depression Inventory (BDI-II); and the Hospital Anxiety and Depression Scale (HADS). These measures they were compared to a gold standard structured clinical interview (Schedules for Clinical Assessment in Neuropsychiatry, SCAN) in 21 people with MS, in order to establish their criterion validity.

To obtain the optimum cut off scores for each of the measures when used with people with MS, a receiver operating curve was conducted which plotted the sensitivity and specificity of each score on the measure. This new cut off score was transformed using SPSS to ascertain the number of cases of depression or anxiety in the sample using the optimum cut off score for each measure. The number of cases identified for depression and anxiety was compared to the number identified by the gold standard SCAN interview.

Using this methodology the BAI was found not to be valid for use in people with MS. The BDI-II was found to be valid with a cut off score of 18 which yields high sensitivity (89%) and high specificity (92%). The HADS was also found to be valid when a cut off score of 10 was used demonstrating high sensitivity and specificity for both the anxiety subscale (100%, 87%) and the depression

subscale (78%, 92%). The reliability of each of the measures was also assessed and all the measures demonstrated high test-retest reliability.

As a result of the high reliability and validity of the BDI-II and the HADS these measures are recommend for the use of screening for anxiety and depression in people with MS. It is hoped that if routine screening for depression and anxiety occurred in this population then access to treatment would improve. It is acknowledged that screening should occur in conjunction with clinical judgement and support. The measures that have found to be valid could also be used to accurately assess the prevalence of anxiety and depression in people with MS. This would enable the targeting of limited resources to research and services in areas of greatest need. Finally, valid measures would allow further research to be conducted to unpick the complex relationship between anxiety, depression and MS which could ultimately impact on the quality of life of that specific client group.

Statement of contribution

The completion of the thesis involved the input of a number of people. The project design, application for ethical approval, writing of literature review, recruitment of participants, data collection and analysis and write up were all completed by Tessa Hopkins. Dr Roshan das Nair and Professor Nadina Lincoln kindly supervised each stage of the process, in particular the design, ethical application, literature review, analysis and write up. Professor Nadina Lincoln, Faye Yuli and Emma Ford enabled the recruitment of participants. Dr Evangelou contributed to the ethical approval of the study.

Acknowledgements. I would like to say thank you to the participants who gave up their time for the study. I would also like to thank those that kindly supported me throughout the study and provided constructive feedback. These include: my supervisors for the study; members of the Nottingham AWE group, the trainees in my cohort, as well as my partner, close family and friends.

Contents

<i>Title</i>	<i>Subtitle</i>	<i>Page Number</i>
Abstract		ii
Statement of contribution		iv
Contents	Tables	vii
	Figures	lx
	Appendices	x
Journal Article	Abstract	1
	Introduction	2
	Methods	3
	Results	6
	Discussion	14
	References	18
Extended Paper	1. Introduction	22
	2. Methods	67

3. Results	82
4. Discussion	107
References	126
Appendices	166

List of tables

<i>Table</i>	<i>Title</i>	<i>Page Number</i>
1	Cases identified in sample by measures	7
2	Co-ordinates of ROC curve for BAI measure	8
3	Co-ordinates of ROC curve for BDI-II measure	10
4	Co-ordinates of ROC curve for HADS anxiety subscale	11
5	Co-ordinates of ROC curve for HADS depression subscale	12
6	Summary of results	13
7	Criteria for major depressive disorder	25
8	Criteria for generalised anxiety disorder	27
9	Summary of studies assessing prevalence of depression and anxiety in people with MS	39
10	Criteria to assess quality of prevalence studies	48
11	'Cases' of mood disorder found within sample (adapted from Nicholl et al., 2001)	61
12	Descriptive statistics for continuous data	84
13	Tests of normal distribution	88

14	Participant scores on measures being assessed	89
15	Strength of kappa coefficient (adapted from Pallant, 2006)	90
16	Strength of kappa coefficient in comparison with SCAN	90
17	Contingency table	93
18	Metrics for contingency table	93
19	Co-ordinates of ROC curve for BAI	95
20	Contingency table for BAI	96
21	Co-ordinates of ROC curve for BDI-II	98
22	Contingency table for BDI-II	99
23	Co-ordinates of ROC curve for HADS anxiety subscale	101
24	Contingency table for HADS anxiety subscale	102
25	Co-ordinates of ROC curve for HADS depression subscale	104
26	Contingency table for HADS depression subscale	105
27	Epidemiological findings and study sample	114

List of Figures

<i>Figure</i>	<i>Title</i>	<i>Page Number</i>
1	Concept map of depression, anxiety and MS	52
2	Flow chart of recruitment	83
3	Histogram for BAI	85
4	Histogram for BDI-II	85
5	Histogram for HADS-anxiety subscale	86
6	Histogram for HADS-depression subscale	86
7	Receiver operating curve for BAI	94
8	Receiver operative curve for BDI-II	97
9	Receiver operating curve for HADS-anxiety subscale	100
10	Receiver operating curve for HADS depression subscale	103

List of Appendices

<i>Appendix</i>	<i>Title</i>	<i>Page</i>
Appendix A	Conformation of ethical approval	167
Appendix B	Confirmation of ethical amendment	171
Appendix C	Confirmation of ethical approval from NHS trust research and development department	172
Appendix D	Confirmation of sponsorship from university	174
Appendix E	Cover letter of information pack sent to potential participants	175
Appendix F	Consent form	176
Appendix G	Opt-in slip	178
Appendix H	Information sheet	179
Appendix I	Demographic questions sent to potential participants	185

Journal Article

[Written for the Journal of Neurology, Neurosurgery and Psychiatry; see extended paper 1.1 for justification of journal choice]

ABSTRACT

In order to understand the complex nature of the relationship between depression, anxiety and multiple sclerosis (MS) valid assessments are needed. The prevalence of anxiety and depression reported varies widely dependent on the assessment used, although it is often reported as being high in people with MS. Despite the proposed high prevalence, depression and anxiety are often poorly identified in people with MS resulting in poor access to treatment.

To address these issues the current study assessed the validity of three commonly used measures of depression and anxiety for people with MS. The Beck Anxiety Inventory (BAI), Beck Depression Inventory (BDI-II), and the Hospital Anxiety and Depression Scale (HADS) were compared to a gold standard clinical interview in 21 people with MS.

The results found that the BDI-II and HADS were valid measures to detect depression and anxiety in people with MS. An optimum cut off score of 18 for the BDI-II yields high sensitivity (89%) and high specificity (92%). An optimum cut off score of 10 for the HADS demonstrates high sensitivity and specificity for both the anxiety subscale (75%, 100%) and the depression subscale (78%, 92%). The BAI was not found to be valid. It is recommended that the measures BDI-II and HADS are used for screening for anxiety and depression in people with MS. By conducting screening it is hoped that people with MS will have greater access to treatment and future research can be conducted to better understand the relationship between depression, anxiety and MS.

INTRODUCTION

Depression and anxiety are reported to have a severe impact on people with multiple sclerosis (MS) but they continue to be under identified.[1] Together these mood disorders are associated with an increase in mortality[2] and dramatically reduce the quality of life for those suffering from it.[3] As a result of under identification, access to treatment is poor[4] despite evidence that psychological and psychopharmacological treatments can be effective for this client group.[1] Poor access to treatment places individuals at risk of deterioration.[1] Valid measures which can be used for screening are needed within this population to correctly identify potential cases of depression and anxiety in people with MS. *[See extended paper 1.2 for diagnostic criteria and 1.3 for discussion of importance of valid assessments]*

The relationship between depression, anxiety and MS is complex. It is influenced by a number of factors including disability,[5] adjustment to illness[6] and social support.[7] Yet knowledge of the impact of each of these factors and how they may be best managed is limited. Research in this area is limited by the poor validity of the measures that are available to measures anxiety and depression in people with MS. *[See extended paper section 1.4 for discussion of the relationship between depression, anxiety and MS]*

A further limitation of having a lack of valid measures available to clinicians is the difficulties in estimating the prevalence of anxiety and depression in people with MS. When using different measures the prevalence of depression in people with MS reported ranges from 26%[8] to 50%[9] and the prevalence of anxiety from 19% to 90%.[6] Valid measures would enable an accurate estimate of prevalence to be made and as a result, resources for research and services could be targeted. *[See extended paper 1.5 for further discussion of varying prevalence rates in people with MS]*

Assessing the validity of an assessment of depression and anxiety in people with MS is problematic due to the complexity of the illness. For example some of the features of MS such as cognitive impairment[10] may potentially impact on an individual's mood, and therefore assessment of mood disorders. Furthermore, there are commonalities in symptoms of mood disorders and MS,

for example, fatigue and pathological crying. These shared symptoms make differential diagnosis more difficult[11] and may compromise the validity of psychiatric assessment measures. *[See extended paper 1.6 for discussion of complexity of assessment due to MS]*

Some attempts have been made to validate measures of anxiety and depression commonly used in clinical practice.[12] However many of these attempts have been methodologically flawed. One critical flaw is studies not using a gold standard comparison which is reported as necessary to assess criterion related validity.[13] Within psychiatric disorders the gold standard is considered to be a structured diagnostic interview,[14] it is this against which the measures in this study were validated. *[See extended paper 1.7 for discussion of psychometric theory; 1.8 for discussion of validity and reliability]*

Aims

The main aim of the study was to validate measures of anxiety and depression which are commonly used in clinical practice for use in people with MS. This was achieved by comparing the measures to a gold standard diagnostic interview. A secondary aim was to assess test-retest reliability. *[See extended paper 1.9 for further discussion of aims]*

METHODS

This cross-sectional study echoed a methodology used by previous studies with similar aims.[14] and was granted ethical approval through the Integrated Research Application System (IRAS) and the local NHS Trust Research and Development Department. *[See extended paper 2.1 for discussion of design and methodologies used in other studies and 2.2 for discussion of ethical issues. See appendices A-C for ethical approval letters]*

Participants

Participants were recruited from two routes. Firstly from an on-going research projects once their active involvement with those projects was completed and secondly, a local database of patients with MS who had consented to be contacted regarding research. Any participants who had a diagnosis of MS were

eligible to be included in the study. Participants were excluded on the grounds of the diagnosis being recent (within the last three months), their current participation in other studies which also assessed their mood, the ability to give informed consent and being able to use and comprehend English. A sample size of 21 participants was calculated where a Kappa Coefficient was the outcome measure (power =.8; effect size =0.5; $p<.05$). *[See extended paper 2.3 for details of participants]*

Measures

Data was gathered on demographic characteristics from each of the participants, including the disability level due to the MS as measured by the Guy's Neurological Disability Scale.[15] The measures to be validated by the study were those commonly used in adult mental health. These measures were then compared to a gold standard diagnostic interview. *[See extended paper 2.4 for further discussion of demographic measures used]*

Beck Anxiety Inventory (BAI).[16] A 21 item self report inventory to measure the symptoms of anxiety. Participants' rate commonly experienced symptoms of anxiety on a four item scale resulting in a range of scores from 0 to 63, with a score above 10 indicating anxiety.[16] The reliability and validity of the BAI has been widely assessed and shown to be robust.[17] Previous attempts to assess the validity of the measure in people with MS have been made by comparing it to other measures but not to a gold standard clinical interview.[12] *[See extended paper 2.5 for further discussion of BAI]*

Beck Depression Inventory II (BDI-II).[19] A 21 item self-report inventory to measure the severity of depressive symptoms. It includes somatic and cognitive-affective symptoms of depression. Participants choose from one of four statements from each item to describe how they have felt during the previous week. Scores on the measure range from 0 to 63 with a score above 14 indicating depression.[18] The measure has shown to be reliable and valid in an adult population without MS.[19] It has been validated for use in people with MS who have been recently diagnosed,[20] such participants were excluded from this study. *[See extended paper 2.6 for further discussion of BDI-II]*

Hospital Anxiety and Depression Scale (HADS).[21] A14 item self-report inventory where respondents answer multiple choice questions about their feelings in the previous week. Scores on the measure range from 0 to 21, with a score above 8 suggesting possible anxiety or depression.[21] The measure has two subscales for anxiety and depression. It has been shown to both reliable and valid.[22] The HADS has been compared to other measures in an attempt to validate it for use in people with MS[12] and has also been compared to a clinical interview, however, this was not in a UK population.[14] *[See extended paper 2.7 for further discussion of HADS]*

Gold Standard – Structure Clinical Assessment Neuropsychiatry Interview (SCAN).[23] The SCAN is a structured clinical interview which maps onto widely used diagnostic symptoms (International Classification of Diseases 10 (ICD-10) and Diagnostic and Statistical manual for Mental Disorders (DSM-IV-TR)[24, 25]. Although it demonstrates adequate reliability[26] the validity of the measure when compared to other gold standard measures is poor.[27] The poor validity reflects methodological difficulties in validating a gold standard measure rather than an accurate assessment of the validity of the measure.[28] The use of the SCAN as a gold standard in previous studies further adds to the justification of its use in the current study.[29] The evidence suggests that less experienced but trained researchers can apply the SCAN reliably.[30] *[See extended paper 2.8 for further discussion of SCAN]*

Procedure

Potential participants were recruited as described above through their involvement in previous research or the local database. They were sent information about the study and copies of the questionnaires. Participants who consented to the study returned the completed measures to a research associate who then scored the measures. Those who had consented to interview had their details passed onto the researcher; this enabled the interview to remain blind to the questionnaire scores. The researcher completed the interview with the participants who provided a diagnosis of depression and/or anxiety using the ICD-10[24] and DSM-IV-TR[25]. Participants who had completed the interview were then asked to repeat the measures they had

completed and return them to the researcher. The results were then collated and analysed. *[See extended paper 2.9 for further details on the information sent to participants; 2.10 for further discussion of procedure]*

RESULTS

Of the 98 participants contacted for the study 24 opted into the study but only 21 completed the questionnaires and interview. The participants were made up of 6 (25%) male and 18 (75%) female with a mean age of 49.25 years (standard deviation 9.65). The type of MS that the participants were diagnosed with was as follows: 58% had relapsing remitting MS, 21% had secondary progressive MS; 17% had primary progressive MS and 4% of participants were unsure. The sample was representative of the general population of people with MS in terms of age and gender [31] but with slightly less people with secondary progressive MS than would be expected[32]. Participants had received their MS diagnosis between 2 and 34 years prior to their involvement in the study (mean 12.13 years; standard deviation 7.50). The range of the disability, as measured by the GNDS was between 3 and 38 (maximum score on the measure is 60; mean 17.92, standard deviation 9.23). The distribution of scores did not meet the assumptions of a normal distribution. Any missing data was removed using pair wise deletion. *[See extended paper 3.1 for plan of analysis, 3.2 for details of participant recruitment and characteristics]*

Using the original cut off points provided by the manuals for the measures participants identified as potential cases for anxiety were 38-52% and for depression was 43% (see table1; unless otherwise stated all data within tables relates to the baseline measurements). A kappa coefficient was calculated and the agreement between the measures and the gold standard ranged from poor to good. The BAI demonstrated the lowest agreement with the gold standard (.34, $p > .05$). The BDI-II demonstrated the highest agreement (.81, $p < .05$) and the HADS anxiety and depression subscales showed moderate agreements with the gold standard (.70, $p < .05$ for anxiety; .61, $p < .05$ for depression). *[See extended paper for 3.3 for further details of the measures]*

Table 1 Number of cases identified in sample by measures using original cut off scores

Measure	Identified cases (n)	Percentage
BDI-II	9	43
BAI	11	52
HADS-anxiety	9	43
HADS – depression	9	43
SCAN depression	- 9	43
SCAN – anxiety	8	38

Receiver Operating Curves

A receiver operating curve (ROC) was used to plot the sensitivity and specificity of each score on the measure. The co-ordinates from the ROC curve allowed the optimum cut off scores to be determined for each of the measures being assessed. For each measure, the optimum cut off score was chosen on the basis that it was the score that yielded the best balance between high sensitivity and high specificity..[33]

An optimum cut off score was calculated for the BAI of 10 and this yielded adequate sensitivity (75%) and specificity (61%; see table 2). However, when compared to the gold standard interview the agreement was poor and non significant (Kappa coefficient =.34; $p>.05$). Therefore, in the current study, the BAI is not found to be a valid assessment of anxiety in people with MS.

Table 2 Co-ordinates of the ROC curve for BAI

Positive if Less Than or Equal To	Sensitivity	1-Specificity
.5	1.00	.92
2	1.00	.85
3.5	1.00	.77
5	1.00	.62
7	.86	.54
9.5	.75	.39
11.5	.63	.31
12.5	.63	.23
13.5	.63	.15
16	.63	.07
19	.50	.07
22.5	.38	.07
26.0	.38	0
30.5	.25	0
37.5	.14	0
42	0	0

The optimum cut off score for the BDI-II for people with MS in this study was found to be 18, as opposed to the score recommended by the manual of 14. The optimum cut off score of 18 has both high sensitivity (89%) and specificity (92%; see table 3). The area under the curve demonstrates the overall accuracy of a measure.[33] When calculated for the BDI-II the area under the curve was high (.98; confidence interval .93-1.03) implying that the measure is accurate. Using the optimum cut off score 43% of participants were identified by the measure as having depression, in contrast to 38% identified by the lower original cut off score. The agreement between the gold standard diagnosis and the BDI-II with the optimum cut off score was very good (Kappa coefficient =.81, $p<.001$).

Table 3 Co-ordinates of ROC curve for BDI-II measure

Positive if Less Than or Equal To	Sensitivity	1-Specificity
.5	1.00	.83
1.5	1.00	.75
3	1.00	.68
4.5	1.00	.58
5.5	1.00	.33
6.5	1.00	.25
7.5	1.00	.17
10.5	.89	.17
18	.89	.08
24.5	.89	0
26.5	.78	0
28	.67	0
30	.44	0
33.5	.33	0
37	.22	0
40	.11	0
43	0	0

ROC curves were conducted for both the HADS anxiety and depression sub scales separately, yet a score of 10 was found to be the optimum cut off for both. The developers of the HADS suggested that a score of eight or above indicates possible anxiety or depression.[21]

For the HADS anxiety sub scale a cut off score of 10 demonstrated high sensitivity (75%) and perfect specificity (100%; see table 4). Using this cut off

score 29% of the participants were identified as having anxiety from the measure, in contrast to the 43% found when using the cut off recommended by the developers of the measure.[21] The area under the ROC curve was high (.96; confidence interval .89 – 1.04) which gives an overall indication that the measure is accurate. The agreement with the diagnosis from gold standard interview was very good (Kappa coefficient =.90, $p < .01$).

Table 4 Co-ordinates of ROC curve for HADS anxiety subscale measure

Positive if Less Than or Equal To	Sensitivity	1-Specificity
.5	1.00	.69
1.5	1.00	.46
3	1.00	.39
5	1.00	.31
6.5	.86	.23
7.5	.86	.15
8.5	.86	.08
10	.86	0
11.5	.63	0
12.5	.50	0
14	.38	0
16.5	.13	0
19	0	0

For the depression sub scale using the optimum cut off of 10 the sensitivity (78%) and specificity were also high (92%; see table 5). Using this cut off score 33% of participants were identified as having depression from the measure, as oppose to the 43% participants identified using the original cut off score of eight. The area under the ROC curve was also high (.96; confidence interval .88 – 1.03) indicating that the measure is accurate. The agreement with the diagnosis from the gold standard interview was good (Kappa coefficient = .70, $p < .01$).

Table 5 Co-ordinates of ROC curve for HADS depression subscale

Positive if Less Than or Equal To	Sensitivity	1-Specificity
.5	1	.67
1.5	1	.5
2.5	1	.25
4	1	.17
6	.89	.17
7.5	.78	.17
9.5	.78	.08
11.5	.68	0
12.5	.63	0
13.5	.50	0
15	.25	0
16.5	.13	0
18	0	0

Once the optimum cut off scores for each of the measures had been determined the scores from the measures were re-classified to demonstrate the frequency of participants who were indicated as being positive or negative for anxiety and/or depression. This outcome information was placed into a contingency table with the outcome of the gold standard interview. The data in the contingency table allowed sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and the discriminant ability of each of the measures to be calculated. This information has been summarised below (table 6).

The test re-test reliability was completed for each of the measures being assessed. Not all the participants completed the repeat measures and as pair-wise deletion was used the sample size was reduced to 17 participants for this statistical test. As the data was not normally distributed a Spearman's correlation coefficient was conducted. *[See extended paper 3.5 for details of the assessment of reliability of the measures]*

Table 6 Summary of results

	BAI	BDI-II	HADS-anxiety	HADS-depression
Optimum cut off	10	18	10	10
Sensitivity	.75	.89	.87	.78
Specificity	.61	.92	1.00	.92
PPV	.55	.88	1.00	.88
NPV	.80	.92	.93	.85
Area under curve	.81*	.98*	.96*	.96*
Kappa coefficient	.34	.81*	.90*	.70*
Spearman's r	.68*	.93*	.88*	.77*

**Significant at $p < .05$ level*

[See extended paper 3.4 for details of ROC and related analyses for each measure and 3.5 for reliability analyses]

DISCUSSION

Of the measures which were assessed the BDI-II and the HADS were found to have good criterion validity for use in people with MS, although the optimum cut off scores for use in people with MS was slightly higher cut off scores than recommended by their manuals. The BAI was not found to be valid for the population.

BDI-II

The assessment of validity of the BDI-II has been completed before using a similar methodology for people with MS who had been recently diagnosed.[20] The study considering recent diagnosis had suggested a lower cut off score of 13 was needed,[22] as opposed to the higher score of 18 recommended by these results. It may be that as MS is a progressive disease[34] the factors that influence the relationship between depression, anxiety and MS alter over time. As an individual's level of disability increases they may also have reduced social support and an increase in the number of confounding symptoms (e.g. fatigue). An alternative explanation for the different cut off scores for people with MS who have been recently diagnosed[20] is the diagnosis met by the participants. Those with a recent diagnosis may meet the criteria for an adjustment disorder rather than depression.[25]

A more recent attempt to validate the BDI-II did not compare it to a gold standard[12] but to other measures, which at the time, had not been validated. Unsurprisingly, given the lack of validation of the measures involved, poor agreement was found between the measures. This is in contrast to the current study where the measures are being compared to a gold standard clinical interview.

The high reliability of the BDI-II found in the current study (.93) reflects that found in the previous studies[35]. Although it must be noted that there is a paucity of literature assessing the test-retest reliability of the BDI-II, therefore a large scale study would be required to have enough statistical power to accurately assess the test-retest reliability of the measure.

As a result of the good reliability and validity demonstrated by the BDI-II in the current study it is recommended that the BDI-II is used as screening measure for use with people with MS.

HADS

The current study is the second to validate the HADS for use in people with MS, although it is the first in the UK. A previous attempt by a Canadian group at validation concluded that the measure was valid, but that it was the optimum cut off score was 8 rather than 10.[14] This highlights the importance of replication within scientific research[36] and the caution that must be taken in generalising results between countries.

Although the Canadian study did recommend a lower cut off score for use in people with MS than the current study both the cut off scores are within the range recommended by the measures developers (8-10).[21] Therefore it may be that those with MS need to meet the higher end of the range to indicate possible depression or anxiety.

The HADS demonstrated high test-retest reliability in the current study. Although the test-retest reliability of the HADS has not been reported, the high reliability has been found similar diseases such as Parkinson's.[37]

The high reliability and validity of the HADS indicate it is a suitable measure for screening for anxiety and depression in people with MS.

BAI

The BAI was not found to be valid for use in people with MS. It demonstrated poor agreement with the gold standard. This echoes previous studies where the BAI has demonstrated poor agreement with alternative measures of anxiety.[12]

When using the BAI the frequency of participants viewed as potentially having anxiety was much greater than the gold standard (48% vs. 33%). This overestimation may be due to the measures' focus on physical symptoms of anxiety[38] rather than a holistic consideration of all symptoms. By focusing exclusively on the physical manifestation of anxiety there is potentially a greater chance of symptoms present in both MS and anxiety confounding the measure.

The BAI did demonstrate good test-retest reliability, although it was below that of the other measures. The test-retest reliability found in the current study was similar to that found in previous studies[39] although it has not been assessed in people with MS specifically prior to the current study.

[See extended paper section 4.1. for further discussion of the findings in the context of previous research]

Strengths and limitations

The strengths of the study are in the clinical implications and, as a result, the potential contribution of knowledge in understanding the relationship between anxiety, depression and MS. The study also demonstrated a robust methodology which has been shown previously to be of use in assessing the validity of measures.

One of the limitations of the study is the small sample size. It is recommended that any replications of the study aim to increase the sample size and thus statistical power of the study. A second limitation is that although the study used a robust methodology to assess the validity of measures only one type of validity was assessed. The construct validity for the measures in people with MS is still unknown and needs to be addressed in order for the relationship between the constructs to be understood. *[See extended paper section 4.2 for further discussion of the strengths and limitations]*

Clinical implications and future research

This study has three important clinical implications in terms of screening and future research. Firstly, screening is recommended for all those with a chronic physical illness,[40] this study demonstrates that the BDI-II and HADS are valid for use as screening measures in people with MS.

It is hoped that by being better able to identify those who experience anxiety and/or depression and MS they will have be able to access treatment. It has already been established that there are effective treatments for those experiencing depression or anxiety with MS yet access to these is very poor.

Effective treatments for anxiety and depression could be explored further in future research now that valid measures are available. It is hoped that the measures which have been found to be valid in the study can contribute to knowledge by enabling future research to more closely examine the complex relationship between anxiety, depression and MS.

Finally, by having valid measures an accurate assessment of the prevalence of anxiety and depression can be made for people with MS. This would allow the targeting of limited resources towards areas of higher prevalence and need.

A note of caution must be placed in implementing the recommendations from this study. Although it has been recommended that the BDI-II and HADS are used as screening measures for people with MS it should be noted that screening programmes have potential detrimental effects,[41] This is particularly because although the measures did demonstrate high sensitivity and specificity they were not perfect and so some people may still be misidentified. Therefore, any screening should be completed in conjunction with clinical judgement for it to be effective. *[See extended paper section 4.3 for discussion of recommendations for future research; 4.4 for clinical implications; 4.5 for critical reflection and 4.6 for conclusions]*

REFERENCES

- 1 Mohr DC, Goodkin DE. Treatment of depression in multiple sclerosis: Review and meta-analysis. *Clinical Psychology, Science and Practice*.1999;9:1-9.
- 2 Feinstein A, O'Connor P, Gray T, et al. The effects of anxiety on psychiatric morbidity in patients with multiple sclerosis. *Mult Scler*. 1999;5:323-326.
- 3 Lobentanz IS, Asenbaum S, Vass K et al. Factors influencing quality of life in multiple sclerosis patients: disability, depressive mood, fatigue and sleep quality. *Acta Neurol Scand*. 2004;110:6-13.
- 4 Feinstein A. An examination of suicidal intent in patients with multiple sclerosis. *Neurol*. 2002; 59: 674-678.
- 5 Tsivgoulis G, Triantafyllou N, Papageorgiou C, et al. Associations of the Expanded Disability Status Scale with anxiety and depression in multiple sclerosis outpatients. *Acta Neurol Scand*. 2007; 115:67-72.
- 6 Jose Sa M. Psychological aspects of multiple sclerosis. *Clin Neurol Neurosurg*. 2008 ;110 :868-877.
- 7 Beckner V, Howard I, Vella L, et al. Telephone administered psychotherapy for depression in MS patients: moderating role of social support. *J Behav Med*. 2010;33:47-59.
- 8 Mohr DC, Hart SL, Julian L, et al. Screening for depression among patients with multiple sclerosis: two questions may be enough. *Mult Scler*. 2007;13:215-219.
- 9 Sadovnick AD, Remick RA, Allen J, et al. Depression and multiple sclerosis. *Neurology*. 1996;46:628-632..10 Brassington JC, Marsh NV. Neuropsychological aspects of multiple sclerosis. *Neuropsychol Rev*. 1998;8:43-77.
- 11 Siegert RJ, Abernethy DA. Depression in multiple sclerosis: a review. *J Neurol Neurosurg Psychiatry*. 2005;76:469-475.

- 12 Nicholl CR, Lincoln NB, Francis VM, et al. Assessment of emotional problems in people with multiple sclerosis. *Clin Rehabil*. 2001;16:657-688.
- 13 Sullivan LW. *The SAGE glossary of the social and behavioural sciences*. London: Sage Publications Ltd. 2009:533-534.
- 14 Honarmand K, Feinstein A. Validation of the Hospital Anxiety and Depression Scale for use with multiple sclerosis patients. *Mult Scler*. 2009;15:1518-1524.
- 15 Sharrack B, Hughes RA. The Guy's neurological disability scale (GNDS): a new disability measure for multiple sclerosis. *Mult Scler*. 1999;5:223-233.
- 16 Beck, AT, Steer RA. *Beck Anxiety Inventory (BAI) Manual*. Oxford: Pearson, 1990.
- 17 De Ayla RJ, Vonderharr-Carlosn DJ, Kim D. Assessing the reliability of the Beck Anxiety Inventory scores. *Educ Psychol Meas*. 2005;65:742-756.
- 18 Beck AT, Steer RA, Brown GA. *Beck Depression Inventory-II (BDI-II) Manual*. Oxford: Pearson, 1996.
- 19 Joiner TE, Walker RL, Pettit JW et al. Evidence based assessment of depression in adults. *Psychol Assess*. 2005;17:267-277.
- 20 Sullivan MJ, Weinshenker B, Mikail S, et al. Screening for major depression in the early stages of multiple sclerosis. *Can J Neurol Sci*. 1995;22:228-231.
- 21 Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 1983;67:361-370.
- 22 Bjelland I, Dahl AA, Haug TT et al. The validity of the Hospital Anxiety and Depression Scale: an updated literature review. *J Psychosom Res*. 2002;52:69-77.
- 23 Wing JK, Babor T, Brugha T. et al. SCAN: Schedules for clinical assessment in neuropsychiatry. *Arch Gen Psychiatry*. 1990;47:589-593.
- 24 World Health Organisaton. *International Classifcaiton of Diseases (ICD-10)*. Genava: World Helath Organisation . 1992.

- 25 American Psychiatric Association. Diagnosis and statistical manual of the mental disorders: Fourth edition – test revision (DSM-IV-TR). Washington DC: American Psychiatric Association. 2000.
- 26 Rijnders CA, van den Berg JFM, Hodinamont PPG, et al. Psychometric properties of the schedules for clinical assessment in neuropsychiatry (SCAN 2.1). Soc Psychiatry Psychiatri Epidemiol. 2000;25:348-352.
- 27 Brugha TS, Jenkins R, Taub N, et al. A general population comparison of the Composite International Diagnostic Interview (CIDI) and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN). Psychol Med. 2001;31:1001-1013.
- 28 Rosenman SJ, Korten AE, Levings CT. Computerised diagnosis in acute psychiatry: validity of the CIDI-Auto against routine clinical diagnosis. J Psychiatr Res. 1997;31:581-592.
- 29 Lincoln NB, Nicholl CR, Flannaghan T, et al. The validity of questionnaire measures for assessing depression after stroke. Clin Rehabil. 2003;17:842-848.
- 30 Brugha RS, Neinhuis F, Bagghi D, et al. The survey form of SCAN: the feasibility of using experienced lay survey interviewers to administer a semi-structured systematic clinical assessment of psychotic and non-psychotic disorders. Psychol Med. 1999;29:703-711.
- 31 Piccinelli M, Wilkinson G. Gender differences in depression. Br J Psychiatry. 2000;177:486-492.
- 32 Cook SD. Handbook of multiple sclerosis. Third edition. New York: Taylor & Francis. 2005.
- 33 Zou KH, O'Malley AJ, Mauri L. Receiver operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation. 2007;115:654-657.
- 34 Lezak MD, Howieson DB, Loring DW. Neuropsychological assessment (4th edition). New York: Oxford University Press. 2004.

- 35 Sprinkle SD, Lurie DL, Insko SL, et al. Criterion validity, severity cut scores and test retest reliability of the Beck Depression Inventory-II in a university counselling centre sample. *J Couns Psychol.* 2002;49:381-385.
- 36 Reiss HT, Judd CM. *Handbook of research methods in social and personality psychology.* Cambridge:Cambridge University Press. 2000.
- 37 Marinus J, Leentjens AFG, Visser M, et al. Evaluation of the Hospital Anxiety and Depression Scale in patients with Parkinson's disease. *Clin Neuropharmacol.* 2002;25:318-324.
- 38 Wilson KA, Chambless DL, de Beurs E. Beck Anxiety Inventory. In Maurish, ME, ed. *The use of psychological testing for treatment planning and outcomes assessment. Third edition, volume three: instruments for adults.* New Jersey: Lawrence Erlbaum Associates Inc 2006:399-420.
- 39 Fydrich T, Dowdall D, Chambless DL. Reliability and validity of the Beck anxiety inventory. *J Anxiety Disord.* 1992;6:55-61..
- 40 National Institute of Clinical Excellence. *Depression in adults with a chronic physical health problem: treatment and management. Guideline number 91.* London: National Institute for Clinical Excellence 2009.
- 41 Gilbody S, Sheldon T, Wessely S. Health policy: should we screen for depression? *Br Med J.* 2009;332:1027-1030.

Extended Paper

1. INTRODUCTION

1.1 Justification of journal choice

The Journal of Neurology, Neurosurgery and Psychiatry was chosen due to its wide readership (impact factor 4.87) which includes a relevant audience to disseminate the study to, such as neurologists. In addition, the journal publishes research regarding common neurological disorders, including multiple sclerosis (MS), and favours the production of articles that have direct relevance to clinical practice as appropriate to this study current study.

The criteria for submission to this journal have been followed (for submission guidelines see: <http://group.bmj.com/products/journals/instructions-for-authors/formatting> and <http://jnnp.bmj.com/site/about/guidelines.xhtml>).

1.2 Diagnostic criteria

1.2.1 Depression. Depression is a disorder characterised by persistent low mood, negative self concept and changes in activity levels (Beck & Alford, 2009). It is one of the leading causes of disability worldwide (Murray & Lopez, 1997). There are a number of diagnoses within the category of depression, for example, dysthymia is diagnosed when depressive symptomology is present for at least two years (American Psychiatric Association, 2000). Current prevalence rates for depression differ slightly between genders, with estimates of 2- 9% for men and 3 – 14% for women depending on the methodology and sample used (Beck, 2000). More recent reviews have noted a trend in depression being diagnosed more commonly in younger cohorts with the onset decreasing towards late adolescence and early adulthood (e.g. Power, 2004). A number of risk factors for depression are cited in the literature, these include stressful life events, a family history of depression, previous depressive episodes and a poor social network (Carr & McNulty, 2006)

There are two common diagnostic systems for the classification of psychiatric disorders: the Diagnostic and Statistical Manual for Mental Disorders (DSM-IV-TR; American Psychiatric Association, APA, 2000) and the International

Classification of Diseases (ICD-10; World Health Organisation, WHO, 1992). The ICD-10 is the system favoured in the UK by the National Health Service and the National Institute of Clinical Excellence (NICE), although in practice the two systems are used interchangeably (Andrews, Slade & Peters, 1999). There is little difference in the diagnostic criteria for depression between the two systems (see tables eight). Within each diagnostic category of depression there are a number of individual diagnoses outlined in the DSM and ICD manuals. This study included these individual diagnoses in the classification of depression.

Traditionally depression was considered to be an acute illness with episodes that lasted six to nine months. However, more recent studies have demonstrated that the natural course of depression is more complex (e.g. Collaborative Depression Study; Katz & Klerman, 1979) and authors now suggest that relapse should be expected in depression (Carr & McNulty, 2006). Although many patients may recover within the first year (70%), a proportion may still be symptomatic five years after the onset (Boland & Keller, 2002). In addition once people have recovered from depression the illness is likely to reoccur (75% experience at least one additional episode of depression; Boland & Keller, 2002).

Depression can be successfully treated through a range of models. Pharmacological treatments utilise a medical model known as the 'monoamine hypothesis' which focuses on a deficit in monoamine neurotransmitters (Carr & McNulty, 2006). A recent meta-analysis reported that 56% of people with depression responded well to antidepressant medication compared to 42% responding to a placebo (Arroll et al., 2005). Within clinical psychology a number of models have attempted to address depression in different forms. For example, cognitive behavioural therapy (e.g. Butler, Chapman, Forman & Beck, 2006), interpersonal therapy (e.g. Cuijpers, Straten, Andersson & van Oppen, 2008) and systemic therapies (e.g. Barbato & D'Avanzo, 2008) have all shown to be effective. Although recent reviews warn that a publication bias may have overestimated the effects of psychological interventions for depression (Cuijpers, Smit, Bohlmeiger, Hollon & Andersson, 2010). In summary, a range of possible treatments for depression have been shown to be successful and

current advice suggests a combination of pharmacology and psychological therapy is necessary (e.g. NICE, 2009).

Table 7

Criteria for Major Depressive Episode

DSM-IV-TR	ICD - 10
A. Five or more of following symptoms listed are present in same two week period with a change from previous functioning, at least one symptom being depressed mood or loss of pleasure.	A duration of at least two weeks of both typical and specific symptoms
1. Depressed mood most of day nearly every day	A. Typical Symptoms: 1. Depressed mood
2. Loss of pleasure or interest in all, or almost all, activities most of day, nearly every day	2. Loss of interest and enjoyment
3. Significant weight loss or gain or decrease or increase in appetite nearly every day	3. Reduced energy leading to increased fatigue and diminished activity
4. Insomnia or hypersomnia nearly every day	4. Marked tiredness after only slight effort
5. Psychomotor agitation or retardation nearly every day	
6. Fatigue or loss of energy nearly every day	B. Specific Symptoms:
7. Feelings of worthlessness or excessive or inappropriate guilt nearly every day	1. Reduced concentration and attention
8. Diminished ability to think or concentrate or indecisiveness	2. Reduced self-esteem and confidence
9. Recurrent thoughts of death, suicidal ideation, suicide attempt or specific plan	3. Ideas of guilt and unworthiness
B. The symptoms do not meet criteria for a Mixed Episode	4. Bleak and pessimistic views of the future
C. Symptoms cause significant distress or impairment in social, occupational, or other important areas of functioning	5. Ideas or acts of self harm or suicide
D. Symptoms are not due to substance abuse or a general medical condition	6. Disturbed sleep
E. Symptoms cannot be better accounted for by bereavement	7. Diminished appetite

1.2.2 Anxiety. Anxiety is characterised by a feeling of fear when presented with a perceived threat (Rachman, 1998). Similar to depression, there are a range of diagnoses within the category of anxiety disorders. These include panic disorder, social anxiety disorder, specific phobias and obsessive compulsive disorder (DSM-IV-TR, APA, 2000). The criteria for anxiety disorders differ little between the DSM and ICD diagnostic systems (see table nine). Current prevalence rates vary between disorders but the overall range is between 1% and 15% (Carr & McNulty, 2006), with the lifetime risk for anxiety disorders being approximately 15% (Kessler et al., 2005). Furthermore, like depression, lifetime prevalence rates seem to be increasing with higher prevalence rates reported in more studies (Kessler et al., 2005). Again, the risk of relapse is significant and has been shown to differ between genders (43% men; 64% women; Yonkers, Bruce, Dyck, & Keller, 2003).

Anxiety disorders can be effectively treated using both pharmacological and psychological methods. Historically, pharmacological treatments have included the use of benzodiazepines but due to withdrawal and tolerance effects, prescribing practice for this group of drugs is now tightly controlled (Carr & McNulty, 2006). Current pharmacological treatments focus on selective serotonin reuptake inhibitors (SSRIs; NICE, 2004). Psychological methods that have been shown to be effective include cognitive behavioural therapy (e.g. Hofmann & Smits, 2008), acceptance based behaviour therapy (e.g. Roemer, Orsillo & Salters-Pedneault, 2008) and short term psychodynamic psychotherapies (Leichsenring, Raburg & Leibing, 2004). Current guidelines take the efficacy of these into account and a combination of pharmacological and psychological approaches is recommended (NICE, 2004).

Table 8

Criteria for Generalised Anxiety Disorder

DSM-IV-TR	ICD-10
A. Excessive anxiety and worry for more days than not, for at least six months, about a number of events or activities	Primary symptoms of anxiety most days for at least several weeks at a time.
B. Person finds it difficult to control the worry	A. Apprehension e.g. worries about future misfortunes, difficulty in concentrating
C. The anxiety and worry are associated with three (or more) of the following: Restlessness or feeling on edge; being easily fatigued; irritability; muscle tension; difficulty falling or staying asleep; difficulty concentrating	B. Motor tension e.g. inability to relax, trembling
D. Focus of anxiety and worry is not confined to another disorder (e.g. Social Phobia)	
E. The anxiety, worry or physical symptoms cause significant distress or impairment in social, occupational or other important areas of functioning	C. Autonomic over activity e.g. light-headedness, sweating, dizziness, dry mouth
F. Disturbance is not due to mood disorder, psychotic disorder, pervasive developmental disorder, substance abuse or a general medical condition.	D. Must not meet full criteria for depressive episode, phobic anxiety disorder, panic disorder, or obsessive-compulsive disorder

1.2.3 Co-morbidity of anxiety and depression

Together, depression and anxiety are common mental health problems affecting approximately one in six people in the UK (McManus, Meltzer, Brugha, Bebbington & Jenkins, 2009). Although anxiety and depression are distinct disorders there is difficulty distinguishing them empirically (Watson et al., 1995). For example, self report measures often have high correlations between anxiety and depression with coefficients ranging from .45 to .75 (Clark & Watson, 1991). As a result some self report measures may demonstrate overall distress rather than the individual constructs of depression and anxiety (e.g. Hospital Anxiety and Depression Scale; Razavi, Delavaus, Faracques & Robaye, 1990; see section 2.7). Some of the inter-relatedness could be accounted for by an overlap of symptoms in a self report measure. For example, the State-Trait Anxiety Inventory (Spielberger, 1989) contains two items which would classically be asymptomatic of depression rather than anxiety: feelings of failure and unhappiness. Indeed, self report measures were the basis of much of the early research into co-morbidity of depression and anxiety (Watson, 2009). More recent research has utilised other methods to consider depression and anxiety such as clinicians' ratings (Gaynes et al., 2007) and considering the constructs at a diagnostic level (e.g. Maser & Colninger, 1990).

Different models have been developed to explain the high co-morbidity of depression and anxiety. Watson & Kendall (1989) suggested a two factor affective model, with negative affect representing a number of negative mood states, such as anger, sadness, and subjective distress and positive affect representing a number of positive mood states including joy and self confidence. Negative affect is present in both depression and anxiety; positive affect has no relationship with anxiety but is negatively correlated with depression (Watson & Tellegen, 1985). Thus, the absence of positive affect (anhedonia) can be used to distinguish depression from anxiety (Watson et al., 1995).

More recently, an additional factor has been added to the two factor model which relates exclusively to anxiety. Known as the tripartite model (Clark & Watson, 1991), depression and anxiety are grouped into three subtypes. In

addition to the negative affect of general depression experienced in both anxiety and depression, and the absence of positive affect unique to depression, a third factor of physical hyper-arousal and somatic tension is uniquely present in anxiety. This tripartite model has been supported empirically in different populations (e.g. children, Chorpita, 2002; older adults, Cook et al., 2004) and used to inform more specific interventions for individuals in a wide range of areas (e.g. targeting smoking cessation interventions according to where individuals fall within the model; Ameringer & Leventhal, 2010). The tripartite model has implications for clinical use (Buckby et al., 2008), particularly as it reflects the cognitive model of depression and anxiety (Nathan & Langenbucher, 2003). If there is an underlying general distress then treatment of either anxiety or depression should lead to a reduction in the other; this has been demonstrated in the literature (e.g. Norton, Hayes & Hope, 2004). The model also closely links with the current diagnostic classification of anxiety and depression as distinct disorders within the ICD-10 and DSM-IV-TR (Nathan & Langenbucher, 2003; see tables eight and nine).

The tripartite model is has received some criticism, much of which centres on it not fully capturing the complexities of the two disorders. For example, Greaves-Lord et al. (2007) found that hyper-arousal was not exclusively present in anxiety and could be also present in depression. In addition, the hyper-arousal of the anxiety factor may only be related to panic disorder and generalised anxiety disorder rather than other anxiety diagnoses such as social phobia and obsessive compulsive disorder (Watson, Gamez & Simms, 2005). In response, more recent models have included more complexity by taking into account specific symptoms of particular anxiety or depression diagnoses. For example, Brown and Barlow (1992) proposed a hierarchical model for anxiety disorders which described both a unique shared factor of anxiety and depression but also unique components for specific anxiety diagnoses. Other similar models, with both common and unique components for different diagnoses of anxiety and depression, have been developed (e.g. Nienka et al., 1998).

In contrast to these categorical models, which assume depression and anxiety are unique constructs, a continuum model has also been developed (Haslam, 2003). The continuum model suggests there is a single continuum which ranges

from 'pure' anxiety to 'pure' depression. The co-morbidity in the continuum model is explained as being the midpoint between 'pure' anxiety and depression. The adoption of a continuum model has been called for in the development of the DSM-V (e.g. Kessler et al., 2003), this is in light of neurobiological research demonstrating specific neurotransmitters potentially providing a mechanistic link between depression and anxiety (Kasper, 2001). However, current reports suggest that the manual will continue to utilise a categorical approach (Fawcett, 2009).

The current research will utilise the tripartite model of depression and anxiety to explain the co-morbidity. This is based on its close relationship to psychological models of depression and anxiety such as the cognitive model (Nathan & Langenbucher, 2003) and its' reflection of the current diagnostic systems.

1.2.4 Depression, anxiety and physical health. The constructs of depression and anxiety may alter slightly when in different contexts. Of particular relevance for the current study, they may alter in people with physical ill health. Much of the literature considering this is from research on older adults (e.g. Wetherell & Arean, 1997) possibly due to the increased likelihood of physical ill health in this population (Katon, 2003)

Current diagnostic criteria suggest depressive disorder cannot be diagnosed if the symptoms are directly related to a medical condition and that failure to take the physical illness into account can lead to over-diagnosis (DSM-IV; Fiske, Wetherell & Gatz, 2009). Conversely, the presence of physical illness may lead to an assumption of mental health symptoms being due to physical illness and thus leading to under diagnosis. In addition, it may be that an individual's mood is a normal and understandable emotional reaction to physical ill health (MacHale, 2000).

To resolve these complexities, some authors have offered alternative symptoms to distinguish depression from physical illness, this has varied between authors as different measures of depression and statistical analyses have been used. Moffic and Paykel (1975) utilised the Beck Depression Inventory (BDI; Beck & Steer, 1990) and found that there were more symptoms of hopelessness, anxiety psychomotor retardation, agitation and self pity in people with

depression and physical illness. They found no differences in physiological symptoms, somatic anxiety symptoms or feelings of guilt from those with depression and no physical illness. This was not reflected in other research, for example, vanHermet et al. (1993) completed psychiatric interviews with a range of medical and non-medical patients with depression or anxiety from the Netherlands. They found that the symptoms of panic, depressed mood, lack of confidence, sleep delay and social withdrawal distinguished those with physical ill health from those who were regarded as physically healthy. When they applied these key symptoms to a UK sample, as a predictive model of psychiatric disorder in people with a physical illness, they found it had high sensitivity (89%) and specificity (97%). The contrasting results of Moffic and Paykel (1975) and vanHermet et al. (1992) is repeated in a number of studies making similar attempts to find symptoms of mood disorders which distinguish people with physical illness and those without (e.g. Pinquart & Shen, 2011). This pattern of inconsistency continues to occur when considering specific physical illness. For example, one study found that those who have had a stroke and depression had more negative symptoms than those with depression who had no stroke (Paradiso, Vaidya, Tranel, Koser & Robinson, 2008). However, an earlier study using similar methodology found no significant differences between those with and without stroke (Spalletta, Ripa & Caltagirone, 2005).

One solution offered by some authors is for psychiatric criteria to be applied without any modifications so all symptoms will be considered, regardless of cause (MacHale, 2002). However, this approach is likely to increase over-diagnosis of mood disorders in people with physical illness so some authors have offered specific guidelines to identify depression in physical illness (e.g. Hawton, Mayou & Feldman, 1990). Recently, NICE have issued guidance regarding depression in long term physical health conditions (NICE, 2009), but rather than clarifying the discussion it focuses more on treatment as opposed to difficulties with diagnosis.

Although some psychological models can incorporate the onset of physical illness as a stressful life event contributing to depression (e.g. Cognitive Model of Depression; Beck, 2008), Kendler, Gardner & Prescott (2002) developed a

specific model to explain the relationship between physical illness and depression. This model illustrates how an individual may have a genetic vulnerability in addition to childhood adversity and stressful life events. When physical illness occurs, if individuals have unresolved attachment difficulties due to childhood adversity, when physical illness occurs, these people may find it hard to collaborate with medical due to difficulties in developing secure trusting relationships (Ciechanowski, Katon, Russo & Walker, 2001) and as a result, it may take them longer to access support (Druss, Rosenheck, Desai & Perlin, 2002). Finally, if depression or anxiety occurs, individuals may need to make changes to self manage their illness (Katon, 2003).

In summary, depression and anxiety are difficult to diagnose in people with physical illness and a consensus has yet to be reached on the best way to achieve this. In addition, once depression and anxiety occur in individuals with a physical illness, this can have an impact on their ability to access services, to interact with medical staff and to self manage their illness thus, indirectly impact on the illness prognosis.

1.2.5 Multiple sclerosis (MS) diagnosis. MS is a neurological illness that follows an unpredictable course. It is the most common cause of non-traumatic neurological disability amongst young and middle aged adults (Beiske et al., 2008). Within the UK, the prevalence is estimated to be approximately 107 in every 100,000 people (Robertson, Deans, Fraser, & Compston, 1995), although there is regional variation with higher prevalence rates of MS in Scotland (Forbes, Wilson & Swingler, 1999). The symptoms of MS can be wide ranging, including impaired vision and bladder control, fatigue and cognitive impairment (Lezack, Howeison & Loring, 2004).

The diagnosis of MS based is on the development of clinical symptoms over time (Warren & Warren, 2001). There are a number of diagnostic criteria which are used by different practitioners at different times (e.g. Schumacher et al., 1965). Many of them place the diagnosis into 'possible', 'probable' or 'definite' categories. Despite this, due to the diversity of the disease there is a great potential for misdiagnosis (Burgess, 2003).

MS can be categorised as relapsing remitting, secondary progressive or primary progressive (Burgess, 2003). In relapsing remitting MS there are periods of acute symptoms (relapse) followed by periods of recovery (remission), making the course of the disease unpredictable and uncertain. Over time symptoms increase in severity to a point where the criteria for secondary progressive MS is met and the illness begins an irreversible deterioration (Warren & Warren, 2001). Those with primary progressive MS experience a continuous worsening of their condition from onset, often a gradual process. However, Burgess (2003) does acknowledge that within each type the prognosis and symptoms will still be varied.

1.3 The importance of having valid assessments for mood disorders

Although the literature is limited, treatments for depression and anxiety in MS have been shown to be effective yet individuals with MS are not currently screened for depression or anxiety. Thus access to treatment may be limited..

Research has demonstrated the effectiveness of treatments for mood disorders in people with MS. A meta-analysis by Mohr and Goodkin (1999) concluded that psychotherapy and antidepressant medication are effective in reducing the levels of depression in patients with MS (effect size for psychotherapy $r=.59$, $p<.001$; for antidepressant $r=.71$; $p<.01$). Of the studies using psychotherapies, those which focused on coping skills were significantly more effective than those which were insight-orientated ($z=2.25$, $p<.05$). Mohr and Goodkin (1999) suggest that this may be due to the progressive nature of MS which continually challenges individuals' existing coping skills; possibly this alters as the person adjusts to the disease (see section 1.3.2). Only five studies met the strict criteria to be included in the meta-analysis. This was not sufficient to explore the differences between the treatment options. However, it was enough to conclude that if depression in MS is left untreated, it is likely to worsen. Due to the difficulties with meta-analysis and the strict criteria limiting the studies included, it may be more useful to consider individual studies. For example, cognitive behavioural therapy has been shown to reduce self-injection anxiety in people with MS (ability to self inject post-treatment as measured by Cochran's $Q=12.25$, $p<.05$) (Mohr, Cox, Epstein & Boudewyn, 2002). Although this pilot

study consisted of only eight participants, the clinical impact of reducing self-injection anxiety is clinically significant given that the treatment for the symptoms of MS often requires regular injections (Mohr et al., 2002).

Despite the reported effectiveness of treatments, access is limited. Feinstein (2002) interviewed people with MS and found 31% of patients with a diagnosis of major depression and 35% with suicidal intent had received no psychological help, be it medication or psychotherapy. It is difficult to draw conclusions from a single study but coupled with further research demonstrating that individuals with depression or anxiety rarely access treatment (Layard, 2006) it may be that these results generalise to people with MS. Of those individuals who do not access treatment, complications of assessment may mean those with MS are well represented (see section 1.6).

The consequences of a lack of treatment can be fatal: anxiety and depression in people with MS has been found to correlate with suicidal intent. For example, Feinstein (2002) completed a structured clinical interview and the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983) for a community sample of 140 patients with MS. They found that the lifetime prevalence of suicidal intent was 28.6% (Feinstein, 2002). Suicidal intent was found to be significantly correlated with either major depression or an anxiety disorder ($p < .001$). Unfortunately, the interpretation of the study is limited as suicidal intent was measured as occurring across the individual's lifetime therefore a causal link to a diagnosis of MS cannot be determined. However, an earlier study completed by Feinstein, O'Connor and Feinstein (1999) also demonstrated a link between suicidal intent and depression and/or anxiety. They found people with MS who experienced co-morbid anxiety with depression had increased thoughts of suicide and self harm, somatic complaints and social dysfunction compared to those who experienced anxiety or depression alone. Although these studies have been completed by the same author together they highlight the need for practitioners to be aware of the risk of suicide in individuals with MS and co-morbid mood disorders.

In response to these concerns Mohr & Goodkin (1999) suggest that all MS patients should be routinely screened for depression and that those found to be

depressed should be offered treatment. In light of the literature, this argument should be extended to include patients with anxiety; specifically those with injection phobias, given the severe consequences for their treatment of MS (see section 1.4.3).

1.4 Relationship between depression, anxiety and MS

Although confounding symptoms play a role in the complex relationship between depression, anxiety and MS (e.g. Arnett, Barwick & Beeney, 2008; Bradshaw & Rose, 2008) the focus in the current section is the secondary factors that influence the relationship (see section 1.6 for further discussion of the impact of confounding symptoms). There is a wealth of literature regarding depression and MS but relatively little that considers anxiety (Honarmand & Feinstein, 2009). This disparity is reflected in the present review.

1.4.1 Adjustment to illness. Within health psychology, the concept of psychological adjustment is the process of adjusting to the diagnosis of a chronic illness and the expected impact on an individual's mood (Livneh & Antonak, 2005; see Stanton, Revenson & Tennen, 2007 for review). It has been demonstrated that poor psychological adjustment to illness can lead to mood disorders such as depression and anxiety (e.g. Ramjeet, Koutanji, Barrett, & Scott, 2005). This has been specifically considered in people with MS with a focus on individuals who have been recently diagnosed, with a noted increase in prevalence following diagnosis for both depression (40%, up to two months after diagnosis; Sullivan, Weinshenker, Mikail & Edgley, 2008) and anxiety (34%, up to 24 months after diagnosis; Janssens et al., 2003). As a result of these high prevalence rates, attempts have been made to support people following a diagnosis of MS (e.g. adjustment groups for people with MS; Forman & Lincoln, 2010).

Adjustment in MS is not confined to receiving the diagnosis. The illness is unpredictable in its course, therefore the individual is required to continually adjust, and some authors claim that as a result mood changes are inevitable (Jose Sa, 2008). In support of this, a positive relationship has been found between the perception of uncertainty and depression in people with MS ($r=.559$, $p<.05$; Gold-Spink, Sher & Theodos, 2000).

1.4.2 Disability. As the MS illness progresses it impacts on the individual's level of disability, which may indirectly affect their mood. It has been shown that, in physical illness, higher levels of disability are associated with higher levels of anxiety (Sareen, Cox, Clara & Asmundson, 2005). A similar relationship has been found with depression as disability increases in MS (Tsivgoulis et al., 2007; Chwastiak et al., 2002), however this finding is based on the use of a self report measure of disability: the Expanded Disability Status Scale (Kutzke, 1983). This measure has been criticised for its poor psychometric properties when used in people with MS (Hoogervorst, Kalkers, Uitdehaag & Polman, 2002). Self reports of disability are problematic given the finding that individuals with MS and depression rate themselves as more disabled on self report than is rated by their doctor (Smith & Young, 2000). Although this conclusion was drawn from a case series it does demonstrate the difficulties in using self report measures to assess disability (see section 4.5). In summary, although a positive correlation between severity of disability and mood disorders has been found the exact nature of the relationship remains unclear.

1.4.3 Treatment. A clear association has been found between depression, anxiety and adherence to treatment for MS. Those who are experiencing depression or anxiety may be less likely to comply with medication regimes (Jared, Hancock, Arnett & Lunch, 2010). Furthermore, the medication used as the treatment for MS may have an impact on depression and anxiety, for example Interferon Beta which is used to treat MS has been associated with depression (Jacobs et al., 2000). Treatment of depression and/or anxiety can improve the treatment adherence in people with MS. For example, Mohr et al. (1997a) found that in treating depression participants increased their adherence to their treatment for MS of Interferon Beta. In another study participants with MS who were treated for injection anxiety significantly increased their ability to self inject and thus were able to more easily utilise available medication for MS (Cochrans Q=12.25, $p<.01$; Mohr, Cox, Epstein & Boudewyn, 2002)

1.4.4 Social support. The impact of MS on an individual's mood may be moderated by social support. A chronic illness may lead to strained support and isolation as the disability impacts on the ability to access social activities

(Beckner, Howard, Vella & Mohr, 2010). Social support is often viewed as a buffer to mental health difficulties (Alloway & Bebbington, 1987). Therefore, as the MS progresses the social support may decrease, and thus increasing the risk of mental health problems. Although this has not been researched within MS, it has been demonstrated in comparable chronic illnesses, for example, in patients with rheumatoid arthritis individuals with little social support reported higher levels of symptoms of both depression and arthritis (Revenson, Schiaffon, Magerovitz & Gibofsky, 1991). In addition positive social support has been shown to positively correlate with effective self management in chronic illness across a number of studies (see review completed by Gallant, 2003). However, it must be acknowledged that defining and classifying social support is complex thus any conclusions drawn from such studies should be taken with caution.

1.5 Varying prevalence rates of depression and anxiety in people with MS

Prevalence is the number of existing cases of a disease in a “defined population at a given time” and is presented in terms of percentage (Bonita, Beaglehole & Kjellstrom, 2006, pp.18). The reported prevalence of mood disorders within MS is higher than within a non-clinical population (e.g. Mohr, Hart, Julian & Tasch, 2007). However, as will be shown, the debate continues as to how much higher it is. A crucial factor, when considering prevalence of mood disorders in MS compared to non clinical populations, is difficulty with assessment (Siegert & Abernethy, 2005). The current study aims to help clarify this potential source of confusion.

1.5.1 Review of the literature. To demonstrate the impact of assessments, a number of studies using different assessments have been considered (see table 9). The studies included were found using the databases of MEDLINE and PsycInfo, completing a keyword search using the following terms: ‘prevalence’, ‘multiple sclerosis’ and ‘anxiety’ or ‘depression’ in articles published in the last fifteen years. For articles containing the words ‘prevalence’ ‘multiple sclerosis’ and ‘anxiety’ 40 articles were found. When the term ‘depression’ was substituted for ‘anxiety’ 125 articles were found. The abstracts of these articles were read and many were excluded (e.g. articles published in a

foreign language; those not published in peer reviewed journals). If articles were duplicated in different journals only one copy of the article was included. In total, 15 articles have been included.

Table 9

Summary of studies assessing prevalence of depression and anxiety in people with MS

Authors (Year)	Measures used	Sample	Prevalence found		Comments
			Anxiety (%)	Depression (%)	
Arnett & Randolph (2006)	Beck Depression Inventory (BDI; Beck, Steer & Brown, 1996).	53 patients with MS		49 (time one) 38 (time two)	Assessed longitudinally, three years apart.
Bamer, Cetin, Johnson, Gibbons &	Center for Epidemiologic Studies Depression Scale	530 patients from East Washington		51	CES-D scale is a self-report measure that does not provide a DSM-IV diagnosis (2). Used participants from Chwastiak et al

Ehde (2008)	(CES-D; Radloff, 1977)				(2005) as reference group
Beiske et al., (2008)	Hopkins Symptom Checklist-25 (HSC-25; Derogatis, Lipmann & Covi, 1973).	140 patients with MS in eastern Norway	19.3	31.4	Large reference population (n=1691) but unclear if contained individuals with MS or other physical health conditions.
Chwastiak et al., (2005)	CES-D	739 participants		45.7	Participants recruited through MS association – may have excluded those not part of association.

Dahl, Stordal, Lydersen & Midgard (2009).	HADS	172 pp's with MS, 56,000 controls. Norway.	31.1 (men) 29.7 (women)	26.2 (men) 25.2 (women)	Completed as part of wider population study, large number of 'controls' but not clear if had other health conditions.
Feinstein, O'Connor & Feinstein (2002)	Structured Clinical Interview for DSM-IV (SCID; First, Spitzer, Gibbon & Williams, 1995).	42 patients with relapsing remitting MS		21.4	Nearly half (43%) of participants with depression had a past history of psychiatric illness prior to commencing treatment for MS.
Feinstein, O'Connor, Gray & Feinstein	HADS	152 participants with MS (107	15.8 anxiety alone	4.6 depression alone	Patients recruited from local clinic – excluded if scheduled an

(1999).

female).

additional appointment.

9.2 co-morbid
anxiety and
depression

Korostil &
Feinstein
(2007).

SCID, HADS,
Beck Suicide
Scale (BSS;
Beck, Kovacs &
Weissman,
1979).

140 clinic
attendees with
MS.

Lifetime
prevalence
35.7

20.5 (HADS)

10.7 (HADS)

Data collected from MS clinic
sample so may not generalise to
community sample.

Galeazzi et
al., (2005).

SCID, BDI, State
Trait Anxiety

100 patients –
50 with

46

Found female gender and severity
of disability were risk factors for

	Inventory (Spielberger, 1989).	relapsing remitting MS, 50 matched healthy controls		depression.
Gottberg, Einarsson, Fredikson, von Koch & Holmqvist (2007).	BDI.	166 participants with MS. Stockholm.	19	No reference group was used. Clear description of sample recruitment.
McGuigan & Hutchinson (2006)	BDI.	176 (151 women, 60 men)	35.8	Attempted to find undetected depression in community setting – therefore excluded participants with diagnosis of depression. Also

					gave no clear report of how recruited participants.
Noy et al., (1995)	Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). Hamilton Rating Scale for Anxiety (HRSA; Hamilton, 1959).	20 patients with relapsing remitting MS in Benison and Israel	90	50	Excluded participants with history of any psychiatric illness prior to onset of illness or with previous or concurrent disabling diseases.
Patten, Beck, Williams,	Composite International	136 participants		25.2	No differences found between treatments received. Participants

Barbui & Metz (2003).	Diagnostic Interview Short Form for Major Depression (CIDI; Robins et al., 1989).				recruited from Canadian public funding drug plan thus excluded those with private insurance.
Poder et al., (2009)	HADS, Social Phobia Inventory (Connor, Davidson, Churchill, Sherwood & Weisler, 2000).	251 patients from clinic. USA.	30 social anxiety (SPI). 21 (HADS)	9 (HADS)	Modified DSM criteria for social anxiety to allow it to be considered as diagnosis when symptoms are limited to medical conditions' social impact (secondary social anxiety).

Smith & Young (2000)	BDI, HADS.	88 patients	34 (HADS)	39 (BDI) 17 (HADS)	Reduced the cut off score for BDI from 13 to 10 (see below for discussion).
-------------------------	------------	-------------	-----------	-----------------------	---

To evaluate the quality of the studies that assess prevalence of mood disorders, guidelines for evaluating prevalence studies were considered (Boyle, 1988; see table 10). Using this guidance it is apparent that many of the studies failed on the measurement criteria as they did not use measures that have been found to be both reliable and valid for people with MS. To demonstrate the methodological difficulties in determining the prevalence of anxiety and depression in people with MS, examples of studies using different methodologies will be reported in detail.

Table 10

Criteria to assess quality of prevalence studies

Area	
Sampling	<p>Does the survey design yield a sample of respondents representative of a defined population?</p> <p>Is the target population defined clearly?</p> <p>Was probability sampling used to identify potential respondents?</p> <p>Do the characteristics of respondents match the target population?</p>
Measurement	<p>Do the survey instruments yield reliable and valid measures of psychiatric disorder and other key concepts?</p> <p>Are the data collection methods standardised?</p> <p>Are the survey instruments reliable and valid?</p>
Analysis	<p>Were special features of the sampling design accounted for in the analysis?</p> <p>Do the reports include confidence intervals for statistical means?</p>

Some studies of prevalence have compared current and lifetime prevalence of mood disorders in people with MS. Using the HADS and SCID, Korostil and Feinstein (2007) suggested a lifetime prevalence of 35.7% for any anxiety disorder and, within this, 18.6% is accounted for by generalised anxiety disorder in a sample of 140 MS patients. They argued that anxiety disorders are common within MS but are often

missed and, therefore, not treated. This study used the SCID interview to consider lifetime prevalence and compared this with the HADS which assesses current mood state. Two measures were compared despite the fact they measure different things. This limitation meant that no discrimination was made between patients who had an anxiety disorder prior to the diagnosis of MS and those who had anxiety post diagnosis. As a result no conclusions about a causal relationship between anxiety and MS can be drawn from this study.

A second methodological problem is if cut off scores in the measures used have been generalised from different populations. For example, Smith and Young (2000) considered 88 patients from a MS outpatient rehabilitation clinic. 38% of the patients gave a history of depression and 23% a history of anxiety. Using the standard cut off scores for HADS 17% met the criteria for depression (score of 8) and 34% met the criteria for anxiety (score of 8). When using the BDI, Smith and Young (2000) altered the cut off criteria and they found 39% of the same sample met the criteria for depression. The cut off score was altered on the basis of previous research conducted by Sullivan, Weinshenker, Mikal and Bishop (1995) which suggested a cut off score of 13 should be used in people with MS rather than 10. However, Sullivan et al. (1995) drew this conclusion from considering people with MS who had been newly diagnosed, yet the time since diagnosis within the Smith and Young (2000) study ranged from 4 to 20 years. This highlights a difficulty with classical test theory as results can only be generalised to the population from which the sample was taken, Smith and Young (2000) violated this rule, thus creating difficulty in drawing conclusions from their analysis (see section 1.7 for further discussion of psychometric theory). These shortcomings highlight the need for cut off scores in MS and for mood disorders to be validated in a population that has not been recently diagnosed.

A third methodological difficulty with the studies estimating prevalence is the recruitment of participants. This was demonstrated by Feinstein et al. (1999). They found particularly low prevalence rates compared to

other studies (15.8% anxiety; 4.6% depression; 9.2% co-morbid anxiety and depression) which may be explained by their recruitment of participants; participants were assessed during their annual appointments at the MS clinic. However, participants who had chosen to make an additional appointment were excluded in order to not “contaminate [the] sample with patients whose anxiety may have led them to seek further examination or reassurance” (Feinstein et al., 1999, p.323). By adopting this recruitment process it is likely that many participants with anxiety were excluded. This may explain the lower prevalence rate found in comparison to other studies (e.g. Arnett & Randolph, 2006).

1.5.2 Prevalence of depression and anxiety in physical illness. It is useful to consider how depression and anxiety in MS compare to depression and anxiety in other physical illnesses (see section 1.2.3 for further discussion of the constructs of mood disorders in physical illness). Depression is two to three times more likely in those with a physical health condition than healthy adults (NICE, 2009). The neurological disorder of stroke is often used as a comparison for MS due to its similarities in variety and range of symptoms (e.g. Rickards, 2005). Prevalence rates of depression in stroke are estimated at varying between 10% and 34% (Rickards, 2005). However, as outline above, assessment of prevalence in stroke has similar methodological difficulties as MS (Berg, Lonnqvist, Palomake & Kaste, 2009). Similar problems are present in other physical illness and some attempts have been made to address this with papers reviewing and critiquing the available assessments, for example in Parkinson’s disease (Scharg et al., 2007). Thus, a common theme of a higher prevalence of mood disorders in people with physical illness is reflected across a range of studies. However problems with assessment continue to make accurate prevalence figures hard to obtain (e.g. Strober & Arnett, 2009). The current study is attempting to address this with people with MS.

1.6 Complexity of assessment due to confounding symptoms

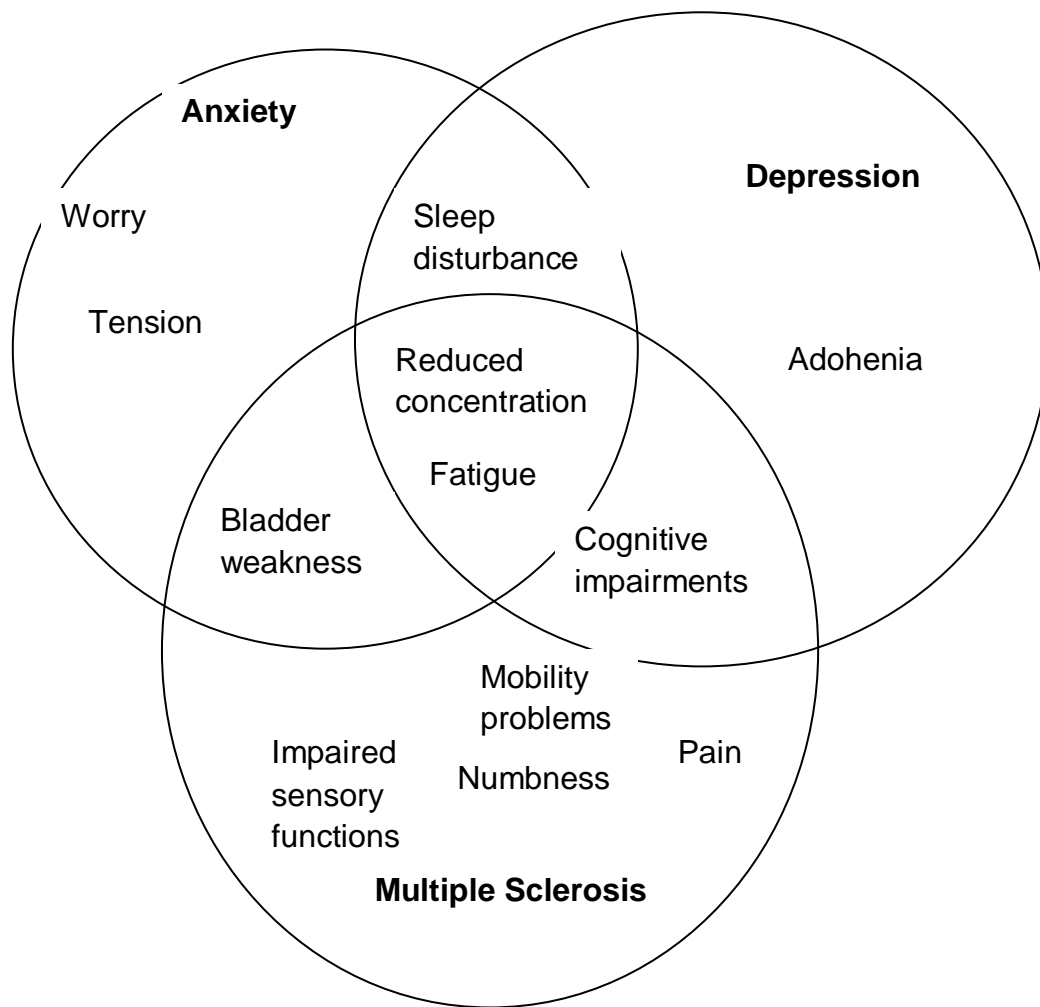
Anxious and depressed people with MS do not always meet the diagnostic criteria for depression or anxiety (see tables eight and nine). Clinicians may attribute confounding symptoms such as concentration problems entirely to MS when a portion of them could be due to depression or anxiety (Mohr & Goodkin, 1999). Conversely, it may be that symptoms of motor tension and autonomic over-activity may be diagnosed as generalised anxiety disorder (ICD-10, WHO, 1992) when they are partially or wholly due to the diagnosis of MS. Potential confounding symptoms will now be considered.

The co-morbidity between depression and anxiety has been discussed previously (see 1.2.3) in the context of a tripartite model. When MS is considered as an additional co-morbidity the similarities between symptoms is further complicated, particularly as many measures of depression and anxiety overestimate somatic symptoms within physical health (see 1.2.4).

In order to further understand the relationship between anxiety, depression and MS, a conceptual map has been developed (see figure one). This is based on the literature and the tripartite model (Clark & Watson, 1991: section 1.2.3) concerning what are likely to be the most common symptoms between depression, anxiety and MS. As can be seen there are some symptoms which are unique to each construct but a range which are present in two or more.

This pattern is reflected in people with MS, where it is the somatic symptoms of mood disorders that are most commonly reported as confounding symptoms as opposed to affective and cognitive symptoms. Specifically, fatigue (e.g. Motl, Suh & Weikert, 2010) is a common symptom in MS and depression and anxiety whereas bladder weakness is less common. A conceptual map demonstrates the different symptoms and how likely it is that they are confounding.

Figure One: Concept map of depression, anxiety and MS



1.6.1 Fatigue. Fatigue is a common compounding symptom of MS and depression as it is present in both. Measures of mood that include an assessment of fatigue may be over-estimating the prevalence of depression within a MS population. However, Mohr, Hart & Goldberg (2003) argue that the commonly reported strong relationship between fatigue and depression in MS is not supported by empirical evidence. They suggest the over-emphasis of the relationship is due to, either fatigue being measured as part of both depression and anxiety, or due to people who are experiencing depression over-estimating self-reported fatigue severity. In assessing this, they treated people with MS and depression for depression and found that the treatment led to an improvement in self-reported fatigue. This confirmed their latter hypothesis (Mohr et al., 2003). This uncontrolled study used three treatments: cognitive behavioural therapy, supportive group therapy and antidepressant medication. It is unclear if these treatments would reduce fatigue in people without depression (Siegert & Abernethy, 2005). The study demonstrates that the relationships between fatigue, depression and MS are more complex and multi-faceted than earlier studies suggest (e.g. Krupp, LaRocca, Muir-Nash & Steinberg, 1989).

1.6.2 Pathological crying. Another symptom which demonstrates the complex nature of mood disorders and MS is pathological crying. Pathological crying occurs when an individual's emotional expression is exaggerated or contradicts the context (Parvizi et al., 2006) and is present in approximately one in ten people with MS (Feinstein, Feinstein, Gray & O'Connor, 1997). It may be that pathological crying is unrecognised or misdiagnosed as a mood disorder by clinicians (Parvizi et al., 2006). Although pathological crying may coexist with depression, the resolution of one does not necessarily follow the resolution of another (Robinson, Parikh, Lipsey, Skarstein & Price, 1993). Therefore, when considering the assessment of mood in MS, pathological crying should be considered separately (Feinstein, Feinstein, Gray & O'Connor, 1997).

1.6.3 Neurological symptoms. The neurological symptoms of MS add to the complexity of assessment of mood disorders. Overall impaired cognitive functioning is present within approximately 54% of people with MS (MacIntosh-Michaelis et al., 1991). The relationship between cognitive impairment and depression is well established (e.g. Kauhanen et al., 1999). As cognitive impairment is present within MS, it may be hypothesised that it would be associated with depression in this population. However, no such link has been widely reported (Brassington & Marsh, 1998). A possible explanation for this is the nature of the cognitive impairment, as much of the impairment within MS patients is in tasks that require attention resources, such as information processing and working memory. It has yet to be shown that improvement in depressive mood correlates with a reduced cognitive impairment in people with MS (Seigert & Abernethy, 2005).

Rather than an overall impairment it may be more pertinent to focus on the specific brain lesions caused by MS and the resultant changes in moods. Brain lesions differ between patients, making the relationship between mood and brain lesions complex to assess. Zorzon et al. (2001) compared patients with MS, chronic rheumatoid diseases and healthy patients. Some moderate positive correlations were found between specific lesions and depression, for example, a right frontal lesion load positively correlated with the Hamilton Rating Scale for Depression (HRSD; $r=.22$, $p<.05$). No significant association between anxiety and Magnetic Resonance Imaging (MRI) abnormalities were found. This implies that anxiety is a reactive response rather than a condition that is linked to brain lesions. Zorzon et al.'s (2001) study was built on by Feinstein et al. (2002), who provided more psychological rigour such as using a structured psychiatric interview to diagnose depression in patients with MS and carefully matching the samples in terms of demographic characteristics. They found that patients with MS experiencing depression, had more lesions in the left inferior medial frontal regions and greater atrophy of the left anterior temporal regions (Feinstein et al., 2002). Unfortunately, without a fuller understanding of

the neuropathology and neuroimaging of depression it is difficult to draw strong conclusions regarding the relationship between the location of brain lesions in people with MS and depression (Siegert & Abernethy, 2005).

1.6.4 Response to confounding symptoms in assessment. In an attempt to acknowledge the potential confounding symptoms with mood disorders and MS some researchers have suggested modifying potential assessment measures. Mohr et al. (1997b) considered the items within the BDI that can be confounded by MS. They compared the BDI scores for a MS group, a group of patients with diagnosis of major depression and a student control group. They found items relating to fatigue, work difficulty and concerns about health contributed to 33% of the total BDI score in patients with MS compared with only 16.7% and 19.2% respectively for patients within the major depression and the control groups. Mohr et al. (1997b) conclude that if the BDI is used in full it may over-estimate the prevalence of depression in individuals with MS (Mohr et al., 1997b). However, an attempt to replicate these results failed (Aikens et al., 1999) and concluded that the full item BDI should be used for routine assessment in people with MS. In conclusion, there is a lack of consensus regarding the BDI and MS and a clear study with a robust methodology such as the current study, is required to further understand how best to use the BDI for a population of people with MS.

1.7 Psychometric theory

Psychometric theory has developed over the past 80 years in line with developments in the philosophy of science (Kline, 2000; see section 4.5). Within psychometrics two theories have developed to manage the inherent difficulty of assessing measures which consider internal states that cannot be easily verified: classical test theory and item response theory (Rust & Golombok, 2009). Classical test theory underpins the current study.

Classical test theory suggests that, due to the imperfection of psychometric measures, the observed score on the measure may not

reflect the individuals true score. It is argued that this is because any observed score on a measure is made up of both the true score and an additional component of random error (Novick, 1966). This holds if one individual completes the same measure an infinite number of times or if there is a single administration of the same measure over a number of individuals (Kline, 2000).

Much of the literature on classical test theory is devoted to managing the random error within observed scores. The focus is to reduce the random error within a measure so the observed score more closely reflects the true score (Kline, 2000). A number of assumptions are made in this process. It is assumed that errors are random, normally distributed and the value of the error is zero, that is, the mean of the distribution of errors over an infinite number of trials (Van der Linden & Hambleton, 2004). As a result of these assumptions classical test theory is unable to deal with systematic errors, such as changes in scores due to learning (Kline, 2000).

Using classical test theory, a standard error of measurement can be calculated. This provides the standard deviation of the distribution of errors around the true score (Kline, 2000), which provides additional assumptions (Emberson & Reise, 2000). The standard error of measurement is considered to be consistent across a population and thus can be generalised to the population from which the sample was drawn (Emberson & Reise, 2000). The standard error is also thought to be the same for each score, regardless of the score. As the tests become longer they become more reliable, as the larger numbers of items (and statistics generated by them) are more stable if based on more items (Kline, 2000).

Classical test theory is widely used to evaluate measures, with a focus on total measure scores. However, critics argue that it creates sample dependent statistics as the statistics only describe measures used with particular populations (Hambleton & Jones, 1993), thus it is difficult to

estimate the true score in samples that have not been tested. In response to this criticism more complex models have been developed.

Item response theory considers individual items and makes more assumptions (Magno, 2009). It assumes an individual item score indicates not only the presence of a latent trait, such as depression, as in classical test theory, but also includes factors about the item itself. This has a number of consequences. Shorter measures can be viewed as more reliable than longer measures; the standard error of measurement will differ between scores but can be generalised across populations, and unbiased estimates of item properties can be obtained from unrepresentative samples (Embretson & Reise, 2000). Despite the advantages of item response theory, it is still considered too complex and technical for many researchers within psychology (Fraley, Waller & Brennan, 2000).

An alternative to the item response theory is generalisability theory (Rust & Golombok, 2009). This utilises analysis of variance models to relate reliability and validity evaluation statistics to test application. It requires sources of error to be identified as test construction and a method of extrapolating this to the eventual use of the test (Rust & Golombok, 2009). Although generalisability theory does increase the conceptual clarity and precision of psychometrics it is complex, time consuming and expensive and thus has not been widely applied (Rust & Golombok, 2009).

Some critics have argued that the concept of a 'true score' used throughout psychometric theory is not justified and was made to fit with the latent trait theory (e.g. Loevinger, 1957). These critics argue that observed scores cannot be split into component parts as one cannot tell from an observed score that anything exists in the brain and thus it is abstract and not of theoretical importance. There are two responses to this critique. First, Carnap (1962) argued that the statistical definition of a true score is that if an infinite number of measures of an observed score were averaged on the same person as the number of observed

scores reaches infinity the errors cancel each other out and a true score is obtained. Carnap (1962) argued that, although this does not occur in practice, it suggests that it is possible to obtain a true score from an observed score. Secondly, Sutcliffe (1965) argued against the critique of the concept of a true score by utilising Plato's theory. This suggests that, just because something is abstract without a physical presence, it does not mean it is not of any use (for example, 'justice').

1.8 Validity and reliability.

The current study is concerned with assessing the concurrent validity of the measures in question, thus it is based within classical test theory. Given this, a brief description is given of the types of validity and reliability which are underpinned by classical test theory.

1.8.1 Validity

Face Validity. The first type of validity to be considered is face validity. Face validity is the "acceptability of test items...for the operation being carried out" (pp.78, Rust & Golombok, 2009). This type of validity is important, as if a participant does not feel that the measure is appropriate for the construct it is measuring, then they may not take the test seriously (Rust & Golombok, 2009).

Content Validity. The second type of validity is content validity (a.k.a. criterion related or domain referenced validity). Content validity reflects the test specification under which the test was constructed and reflects the particular purpose for which the test is being developed. This is important because, if the measure is not reflecting the task specification, then it must be reflecting something else and so is a potential source of bias (Rust & Golombok, 2009). This assessment of validity reflects the functional approach to psychometric testing (as opposed to the trait approach; Rust & Golombok, 2009). It is the basis by which any test construction programme is judged and tends to be assessed more qualitatively than quantitatively, as any deviation from

the validity is more important than the degree of deviation (Rust & Golombok, 2009).

Predictive Validity. Predictive validity is used wherever measures are used to make predictions. The correlation between the test score and a score on the degree of success in a selected field (success on the criterion) is calculated. A difficulty with predictive validity is that not all those selected produce a score on the criterion. For example, if one was attempting to assess the predictive validity of A-level results as a measure of degree success, those who did not attend university would not have a score to compare to. Therefore with a lack of available data the predictive validity would be under estimated. This difficulty is solved by using the available data and extrapolating downward but this is done with a level of uncertainty.

A previous attempt at assessing this type of validity was made by Moran and Mohr (2005) who attempted to assess if a score on a measure of mood predicated the response to an intervention. Utilising the Hamilton Rating Scale for Depression (HRSD) and the BDI as pre and post measures for an intervention, they argued that by assessing people for depression pre and post treatment (when it was assumed they were not depressed) they could determine the validity of the measures (Moran & Mohr, 2005). They found all of the 21 BDI scores showed statistically significant reductions post-treatment (mean 23.7 pre-treatment, mean 10.5 post-treatment, $p < .05$) and 12 of 17 HRSD items showed statistically significant reductions post-treatment (mean 19.3 pre-treatment, mean 10.8 post-treatment, $p < .05$). The items that were not found to produce statistically significant lower scores post-treatment were late insomnia, insight, psychomotor retardation and psychomotor agitation. This suggests that these individual items are not able to predict response to intervention. However Moran and Mohr (2005) only included participants who showed a reduction in two standard errors of measurement on the BDI or HRSD). As a result they may have excluded participants who continued to experience depression but it was not shown by the measure, making the

assumption that the measures were able to capture all cases. Moran and Mohr (2005) acknowledge that this does not demonstrate that the items in the BDI and HRSD are not confounded by MS but they do argue that the BDI and 12 items of the HRSD are effective in capturing change in depression; thus demonstrating predicative validity.

Concurrent Validity. An assessment of the correlation between a new test and existing tests that purport to measure the same construct is known as concurrent validity. A difficulty with concurrent validity is that it does not address the underlying construct so, although two different tests may claim to measure intelligence, the fact they correlate does not mean that they actually measure intelligence, it just means the measures correlate. It also suggests that if the measures do not correlate then the construct validity of the measure may be questioned, therefore in isolation concurrent validity is not sufficient (Rust & Golombok, 2009).

Of concern for the current study are previous attempts at assessing this type of validity. Nicholl, Lincoln, Francis and Stephan (2001) used a sample of 105 participants with MS who completed questionnaires to assess their mood. The questionnaires used were: GHQ-12, GHQ-28, HADS, BDI, BAI, Clinical Outcomes in Routine Evaluations (CORE-OM; Evans et al., 2002), and Brief Symptom Inventory. Each of the mood measures were found to significantly correlate with each other and the measure of disability (Guy's Neurological Disability Scale; Sharrack & Hughes, 1999). However, there was substantial variation in the rates of depression and anxiety depending on the measure used. Within each measurement manual, guidance is given on the cut off required for an individual to meet a 'case' for a mood disorder. The number of 'cases' from the different measurements are shown in table 11.

Table 11

'Cases' of mood disorder found within sample (Nicholl et al., 2001).

Measure	'Cases' identified within sample (%)		
	Anxiety	Depression	Unspecified
GHQ-28			48
HADS	16	18	
BAI	31		
BDI		28	

The results from this study demonstrate the large variation between the different subscales when used with the same population. Although this study compares the measures it does not provide criterion validation for any of the individual scales. For this to occur the scales would need to be compared to a gold standard assessment, such as psychiatric interview as suggested by Nicholl et al. (2001). The most methodologically sound studies aiming to assess measures concurrent validity are those that have compared the measures to a gold standard clinical interview (Sullivan, 2009). The gold standard is assumed to correctly distinguish between those with and without depression and/or anxiety. Establishing a gold standard measurement within psychiatric disorders can be complex. Mental health diagnoses are based on subjective experiences described by the individual, this creates difficulty in using objective tests (Brugha, Bebbington & Jenkins, 1999). The most frequently cited gold standard comparison for psychiatric disorders in the literature is the structured psychiatric interview (e.g. Honarmand & Feinstein, 2009; Patten et al., 2003; Joiner, Walker, Petit, Perez & Cukrowicz, 2005). This is an attempt to add some objectivity to the clinical judgement made within an unstructured psychiatric interview.

Previous studies have utilised a gold standard to assess the concurrent validity of measures in people with MS. Mohr et al. (2007) compared the Patient Health Questionnaire II (PHQ; Kroenke, Spitzer & Williams, 2003) to the structured clinical interview for the DSM-IV (SCID; First et al., 1995). The PHQ consists of just two items and was found to correctly identify 99% of participants who met the diagnosis for major depressive disorder (confidence interval: 91 – 100%). However, the PHQ did have a high false positive rate of 27.3% so it may overestimate those with depression in the MS population.

A further study was conducted by Honarmand & Feinstein (2009). They also used the SCID interview and compared this to the participants' responses to the HADS. They found a score of 8 on this measure provided good sensitivity and specificity for both depression (90% sensitivity, 87% specificity) and anxiety (86% sensitivity, 80% specificity).

Construct Validity. A final type of validity is construct validity which was first proposed by Cronbach and Meehl (1955). This reflects the trait-related approach to psychometrics (as opposed to the function approach, see 1.8.4). It assesses the closeness of the measure to the underlying construct. Smith (2005) argues this is effectively measuring theory as by determining if a measure reflects construct validity one is examining if the measure conforms to a theory of which the construct is a part. Originally construct validity was based within positivist epistemology (see section 4.5) where theories are viewed as “straightforward deviations from observed facts” (pp.396, Smith, 2005). More recent developments in construct validity have acknowledged more recent epistemological understandings such as critical realism (see section 4.5) where theory building is viewed as an ongoing process. Thus, the construct validity of any measure will alter as the construct and theory alters and attempts to consider the construct validity will influence this (Smith, 2005). Therefore theories may not be fully proved or disproved as the debates in the theory continue to be tested over time.

Smith (2005) argues that these developments in construct validity have allowed psychology to move away from single hypotheses to more comprehensive frameworks. This includes the move towards different ways of considering constructs within psychology that are clinically useful, for example a two factor affective model (see section 1.2.2).

1.8.2 Reliability. When considering reliability, the extent to which the test is measured with what is purported to measure (pp.72, Rust & Golombok, 2009) is considered. One needs to make the test as reliable as possible to take into account any variability in interpreting the test results. There are a number of methods to assess reliability.

Test-retest reliability. Test-retest reliability is being assessed within the current study; it is assessed by giving the same measure to the same respondents with an interval between the two administrations. A correlation coefficient is calculated providing a score between zero and one. By basing the coefficient on the mean scores between the first and second administration test, it is assumed only changes in relative ordering or numbering points on scores can affect the result.

Parallel Reliability. Test-retest reliability is not suitable in all cases. For example, if a test of knowledge is being completed people may learn the knowledge in the first administration so the second administration is influenced by memory and motivation rather than the knowledge being assessed. In response to this parallel forms of reliability are used based on classical test theory. Two versions of a test are created. These tests link where each item on one test has an alternative version on a second test. The respondent is then given both versions of the test and a correlation coefficient is calculated for the scores between the two tests.

Although widely praised, parallel forms reliability are rarely used. This may be because once a test is being constructed the aim is to find the best possible items, so, rather than doing this twice, the items are often merged to create one 'super test'.

Split half reliability. An alternative to parallel forms reliability that is more widely used is split half reliability. The test is split in half in a random fashion to create two pseudo parallel forms, so although they are not actually parallel no systemic bias is present in how the two versions are distributed. The two scores from the same respondent for each pseudo parallel form is then used to calculate the correlation coefficient.

However, using split half reliability only provides reliability for half of the test not the whole test. To calculate the reliability of the whole test further calculations must be made. A Spearman-Brown Calculation is conducted. As would be expected, the reliability is always larger for the whole test than the correlation between the two halves because, the more items in the test, the more reliable it is.

When tests are less objective, such as markers of an interview than an additional form of reliability is used. A correlation coefficient is calculated for two measures of the same interview.

1.8.2 Summary. There is a range of ways of assessing the validity and reliability of measures. The ongoing debate regarding the application of psychometric theory demonstrates the variety of possible approaches and the need to ensure the most appropriate measures are utilised depending on the purpose of the assessment (Rust & Golombok, 2009). The studies discussed demonstrate the range of methodologies that have been used to previously validate measures within a population of people with MS, with concurrent validity being the most frequently cited. Despite some authors concluding that there is currently no agreed 'gold standard' in diagnosing mood disorders within MS (Seigert & Abernethy, 2005), it appears that the closest thing would be a structured clinical interview. As such, the most methodologically sound way of assessing validity of measures for people with MS appears to be to use a structured clinical interview as a gold standard.

1.9 Aims of study

The main aim of the study was to validate measures of anxiety and depression for use with people who have MS. To achieve this aim measures commonly used in clinical practice were compared to a diagnostic interview in a sample from an MS population. The measures of mood assessed were the BDI-II, BAI and HADS. Participants' scores on these measures were compared to a 'gold standard' measure: the Structured Clinical Assessment in Neuropsychiatry (SCAN; Win, Babor, Brugha & Burke, 1990).

A secondary aim was to further validate the measures of anxiety and depression using test-retest analysis. This would provide information on the test-retest reliability of the measures.

1.9.1 Hypotheses. There were four hypotheses that were being tested during the study:

1. Participants who met criteria for depression from the clinical interview (SCAN) would have a score of equal to or more than 10 on the Beck Anxiety Inventory.
2. Participants who met criteria for anxiety from the clinical interview (SCAN) would have a score of equal to or more than 14 on the Beck Depression Inventory II.
3. Participants who met criteria for depression from the clinical interview (SCAN) would have a score of equal to or more than eight on the Hospital Anxiety and Depression Scale-Depression.
4. Participants who met criteria for anxiety from the clinical interview (SCAN) would have a score of equal to or more than eight on the Hospital Anxiety and Depression Scale-Anxiety.

The outcome of the study was expressed in terms of a Cohen's Kappa coefficient (1992). The number provided will give an indication of how each of the measures used compares with the diagnosis provided by

the structured clinical interview. The analysis of the data is described within the results section (see section 3).

2. METHODS

2.1 Design

The methodology of the study employed two designs. First: a cross-sectional within subjects design, where all participants completed the same measures, to assess the concurrent validity of the measures and to assess the additional aim of evaluating the test re-test reliability of the measures.

To assess the concurrent validity of the measures they were compared to a 'gold standard' (Sullivan, 2009; see section 1.8). The methodology reflects previous studies that have similar aims (e.g. Honarmand & Feinstein, 2009; De Souza, Jones, & Rickards, 2009). In particular, it involves the use of a diagnostic interview as a 'gold standard' to compare to instruments which have potential use as screening tools (e.g. Lincoln, Nicholl, Flannaghan, Leonard & van der Gucht, 2003; Lloyd-Williams, Friedman & Rudd, 2001; Aben, Verhey, Lousberg, Lodder & Honig, 2002).

2.2 Ethical issues

The research adhered to ethical principles outlined by the British Psychological Society (2009). As with all research there were risks, burdens and benefits for participants, these were discussed by the ethical committee before approval was granted for the study (appendix H).

There was one clear risk to participants in the study, the potential distress that may be caused if they were discussing sensitive issues. This risk was managed by providing participants with information prior to obtaining consent for the study. The researcher who completed the interview is a trainee clinical psychologist who was able to manage distress if and when it arose. In addition, participants were provided with contact details of organisations that could help manage distress as part of the information pack. If a participant was identified as having

very high levels of anxiety or depression they were advised to contact their GP. This occurred with one participant.

A burden was placed on participants in terms of time. This differed between participants. Completion of the questionnaires took approximately 50 minutes. The interview and further questionnaires added up to an additional 130 minutes. This was dependent on participants' responses at interview, due to the nature of the semi-structured interview if participants responded positively to some questions further detail was required and thus the interview took longer. In an attempt to reduce any further burden of time to the participants, the researcher travelled to the participants' homes to complete the interviews.

A potential benefit of participating in the study included the opportunity to discuss the emotional distress they may have been experiencing with a trainee clinical psychologist who is trained to actively listen to and manage such distress. Participants also benefited other patients with MS as the study aimed to provide valid screening tools for depression and anxiety within this population.

Consent was sought from participants involved in the study. It was assumed that all participants had the capacity to consent to the study unless proved otherwise (Mental Capacity Act, 2005). Participants who returned questionnaires were assumed to imply consent to complete the questionnaire measures (see section 2.3.4). Participants who did not return the completed measures excluded themselves from the study. Consent for the interviews was sought by the researcher prior to the interview taking place, a consent form was sent with the information pack and returned before the interview was conducted. Participants had the opportunity to discuss consent and ask the researcher questions both when completing the consent form and returning it by post and immediately prior to the interview taking place. Consent was viewed as a free choice and participants were able to withdraw from the study at any time prior to the data being analysed.

In order to keep the data safe and secure each participant was given a unique identifier code. This code was clearly labelled on any data from that participant; personally identifiable information has been kept separate from other data gathered. Specifically, consent forms have been kept separate and once the interviews were complete the opt-in slips were destroyed. All data will be securely stored at university for the next seven years before being destroyed.

2.3 Participants

2.3.1 Recruitment. Participants were recruited from the Trent Region through two routes. Firstly, participants who had completed their involvement in an ongoing research project evaluating interventions for low mood in people with MS and had given permission to be contacted about future research were contacted. The contact details of these participants were left with a research associate at the University of Nottingham who was involved in the ongoing project. The research associate passed on the contact details of these potential participants to the researcher and they were sent information packs regarding the current study (appendix H).

As not enough potential participants were recruited through the initial route a second route was used. The Trent Region MS clinic routinely asks patients in their outpatient appointments if they would be willing to be contacted for research. These patients' contact details are then held on a database. Potential participants from this database were then sent information packs regarding the current study. Participants were made aware of the inclusion and exclusion criteria prior to them opting in to the study through the information sheet (see appendix H).

2.3.2 Sample Size. A sample size of 21 was calculated for the study. This was based on the outcome being a Kappa coefficient, an effect size of 0.5, power of 0.8 and significance level of 0.05. These figures are based on standard use for psychological studies (Field, 2005). The calculation was confirmed using the programme GPower3 ("GPower3", 2010; Faul, Erdfelder, Lang & Buchner, 2009).

2.3.3 Inclusion criteria There only inclusion criteria for the study was that participants had a diagnosis of MS. The recruitment process ensured all potential participants had a diagnosis of MS. Participants were recruited either through a previous study for which MS was an inclusion criteria or through a database on which inclusion required a diagnosis of MS and attendance to a local MS clinic.

2.3.4 Exclusion criteria. There was four criterion for the exclusion of participants from the study.

Recent diagnosis of MS. Participants were excluded from the study if they had received a diagnosis of MS within the last three months. The justification for this is twofold. First, the impact of the diagnosis and how it is perceived by the individual may mean their mood state is more likely to be unstable at this point (Janssens et al., 2003). Second, it is likely that participants who have recently been diagnosed would have a number of other MS related physical health appointments to attend and participation in a study may pose an additional burden.

The assessment of this exclusion criterion was through the demographic questions which were sent to participants as part of the information pack (see appendix H). Participants were asked to state the time in years since the diagnosis of MS. It was planned that if participants indicated they had been diagnosed within the previous year this would be investigated further by the researcher and those who had been diagnosed in the last three months would be excluded. In reality, no participants required further investigation to establish the time since diagnosis of MS.

Participation in other studies. The second criterion for exclusion of potential participants was if they were currently taking part in other studies which involved the assessment of depression and anxiety. This exclusion criterion was in place based on ethical considerations of not over-assessing participants. The assessment of this criterion was done through recruitment. Participants who actively

were taking part in a parallel study at the same site which involved the assessment of depression and anxiety were removed from the database by the research associate prior to contact details being given to the researcher.

Informed consent. The third criterion was to exclude participants who were unable to give informed consent. Consent was assumed by participants who completed and returned the measures. For the interview, participants completed and returned consent forms prior to the interview being arranged and consent was discussed again and verbally sought immediately prior to the interview taking place (see section 2.2 for further details of consent). When completing the interview it was planned that if the researcher was concerned about the ability of a potential participant to give informed consent for the interview, it would be discussed with the clinical and research supervisors for the study before a decision was made. This situation did not arise.

English language. Participants who were unable to understand or speak English were excluded. The measures in the study, including the interview, were in the English language. This criterion was assumed to have been met by those participants who were able to read the information pack sent to them and return the completed measures that were in English.

2.4 Demographic Measures

Participants were asked a number of demographic questions. Participants provided their age, gender, type of MS and the time since the diagnosis of MS in years. The time since diagnosis confirmed if the participants needed to be excluded from the study (see section 2.3.4).

2.4.1 Assessment of level of disability. In addition to the demographic questions participants were also asked to complete a measure of disability, the Guy's Neurological Disability Scale (GNDS; Sharrack & Hughes, 1999). The GNDS is a 12 item measures designed

specifically for people with MS (Sharrack & Hughes, 1999). Participants are required to answer multiple choice questions that cover different areas of functioning such as memory, fatigue, vision, concentration, speech and communication. The overall score from the measure (maximum 60) provides a summary of the level of impairment the participant experiences due to MS. This summary must be read with caution as it has been demonstrated that people with MS are poor at estimating their own level of disability (Smith & Young, 2000).

The GNDS has demonstrated good psychometrics. It has been shown to have good test-retest reliability ($r = .972$ for the whole scale; $r = .685 - .987$ for the 12 sections; Rossier & Wade, 2002). Although it is shown to be more reliable when used with an interviewer face-to-face ($r = .97$), this is marginal and it continues to be reliable when the individual self reports in writing ($r = .90$; Rossier & Wade, 2002). The validity of the scale was found to be good when compared to the Expanded Disability Scale ($r = .636$) and the Barthel Index ($r = -.757$), a measure of dependence in personal activities in daily living. The sensitivity of the GNDS has yet to be established (Rossier & Waide, 2002).

2.5 Beck Anxiety Inventory (BAI; Beck & Steer, 1990).

The BAI was developed to address high correlations in rating scales of anxiety and depression, arguing that the two disorders needed to be distinguished from each other (Beck, Epstein, Brown & Steer, 1986; see section 1.2.3 for further discussion of the co-morbidity of anxiety and depression). There exist a number of anxiety disorders. Although the BAI is designed to assess each of these it is reported to best distinguish those with a specific panic disorder (e.g. De Beurs, Wilson, Chambless, Goldstein & Feske, 1997; Leyfer, Ruberg & Woodruff-Borden, 2006). This potential bias to a specific disorder may result from the development of the measure to meet the criteria of anxiety disorders in the DSM-III (APA, 1987). Since then the DSM has been revised and the DSM-IV-TR (APA, 2000) has refined the criteria to enable clearer

differential diagnosis between generalised anxiety disorder and panic disorder (Wilson, Chambless & deBeurs, 2006).

Respondents are asked to rate symptoms of anxiety such as '*fear of dying, scared or hands trembling*'. This is completed using a four point scale from them experiencing it '*not at all*' during the last week to '*severely – it bothered me a lot*' (Beck & Steer, 1990). The developers recommend that a cut off score of 10 suggests mild anxiety, 19 suggests a moderate anxiety and 30 suggests a severe anxiety (Beck & Steer, 1990). Although, in practice it is reported that scores of 30 and above are rare (Wilson et al., 2006).

2.5.1 Reliability and validity of BAI. The reliability of the BAI has also been found to be good. It has been reported to demonstrate good internal consistency in a number of studies (e.g. De Ayala, Vonderharr-Carlson and Kim, 2005). The test-retest reliability is also reported as high (e.g. $r=.71$, Osman et al., 2002; $r=.83$, De Beurs et al., 1997).

The construct validity of the BAI has been evaluated through its ability to discriminate between anxiety and other constructs. Although the BAI has been found to correlate highly with measures of depression (e.g. BDI-II, $r=.61$, $p<.001$; Steer, Ranieri, Beck & Clark, 1993) it has been found to clearly discriminate from the depression construct (e.g. De Beurs et al, 1997). This is unsurprising given it was a key aspect of the development of the BAI (Wilson et al., 2006). The ability of the BAI to measure the construct of anxiety has come under criticism due to its' omission of potential cognitive components of the construct, for example, worry (Wilson et al., 2006). However as mentioned previously, this critique does not take into account the cognitive elements present in the scale such as "fear of dying".

Criterion validity, particularly concurrent validity, is high in the BAI. The BAI has been shown to have correlations with other established measures of anxiety, for example, when compared to the Hamilton Anxiety Rating Scale (HARS; $r=.56$, $p<.001$, Beck & Steer, 1991), the

SCL-90-R anxiety subscale ($r=.81$, $p<.001$; Steer et al., 1993) and the State Trait Anxiety Inventory ($r=.58$, $p<.00$; Fydrich, Dowdall & Chambless, 1992). Unsurprisingly, given its development when compared to a structured clinical interview based on the DSM-II (e.g. SCID), the BAI performs well when using the suggested cut off of 10 (sensitivity 76%, specificity 65%; Eack, Singer, & Greeno, 2008).

2.5.2 Use of BAI in people with MS. The BAI has previously been assessed for concurrent validity in people with MS by comparing it with other measures. It was found to correspond poorly with both the HADS (kappa coefficient= .33, $p<.0001$) and the GHQ-12 (kappa coefficient= .30, $p<.005$; Nicholl et al., 2001). The validity of the BAI has not been compared to a gold standard assessment of a clinical interview in people with MS.

As with the other measures being assessed there is a confounding of symptoms between anxiety and MS (see section 1.6). Although this has not been specifically addressed by researchers using the BAI in a sample of people with MS, the BAI has been considered in people with other physical health problems. For example, older adults with medical problems (Wetherell & Arean, 1997) and Parkinson's disease, where concerns have been raised that the BAI overestimates the prevalence of anxiety (Higginson, Fields, Koller & Troster, 2001).

2.6 Beck Depression Inventory-II (BDI-II; Beck, Steer & Brown, 1996).

The BDI-II is the most widely applied clinical and research measure of depressive symptoms (Aikens et al., 1999). It was developed as a revised version of the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The BDI-II is designed to correspond with the DSM-IV-TR criteria for a major depressive episode (APA, 2000) and is recommended as the "premier instrument" for the assessment of severity of depressive symptoms (Joiner et al., 2005, p.274)

To complete the measure participants indicate their agreement with one of four statements across 21 items. For example, participants are asked about their feeling of failure by marking one of four statements: '*I do not feel like I failure; I have failed more than I should have; As I look back I see a lot of failures; I feel a total failure as a person*' (Beck et al., 1996). The measure is noted for its easy administration and scoring, (Hagen, 2007). The manual for the measure suggests that a score of 14 – 19 indicates a mild depression, 20 – 28 moderate and 29 – 63 severe depression (Beck et al., 1996). This scoring pattern has been criticised for being 'bottom heavy' (Hagen, 2007).

2.6.1 Reliability and validity of BDI-II. The reliability of the BDI-II is high (Hagen, 2007). Reviews of the psychometric properties of the BDI-II report high internal consistency ($\alpha > .9$; Joiner et al., 2005) based on studies conducted by the developers of the measure (e.g. Steer, Rissmiller & Beck, 2000).

The construct validity of the BDI-II has been considered through conducting factor analyses. When developed the BDI-II loaded onto two factors, a cognitive factor and a somatic-affective factor (Beck et al., 1996). The finding of these two factors has been replicated with a number of groups including primary care patients (Arnau, Meagher, Norris & Bransom, 2001) and adolescents (Steer, Geetha, Ranieri, & Beck, 1998). Although alternative models (with between one and four factors) have been suggested, the two factor model remains well "suited to the assessment of depression dimensions" (Vanheule, Desmet, Groenvynk, Rosseel & Fontaine, 2008, pp.183).

The BDI-II has demonstrated strong criterion validity. This has been assessed by comparing the BDI-II to established alternative measures of depression. These have included a structured clinical interview ($r = .83$, $p < .05$, Sprinkle et al., 2002), the CES-D ($r = .68$, $p < .001$, Segal, Coolridge, Cahill & O'Riley, 2008) and Montgomery Asberg Depression Rating Scale ($r = .869$, $p < .0001$; Svanborg & Asberg, 2001).

2.6.2 Use of BDI-II in people with MS. A previous attempt at validating the BDI-II has been completed in people with MS using a clinical interview as a gold standard (Sullivan et al., 1995). Following this study an alternative cut off score for the BDI-II of 13 was suggested, demonstrating sensitivity of 71% and specificity of 79% (Sullivan et al., 1995). However, this study only included participants who had been recently diagnosed and there is likely to be a difference in those who have not recently been diagnosed (Janssens et al., 2001). Generalising to a population from which the sample was not originally taken violates classical test theory assumptions and so the results of this study cannot be generalised to all people with MS (see section 1.7)

As described previously (Section 1.8) an attempt at validating the BDI-II for participants with MS who had not recently been diagnosed was made by Nicholl, Lincoln, Francis and Stephan (2001). The BDI-II was not compared to a gold standard and showed poor agreement with the HADS ($\kappa = .12$, $p > .2$; Nicholl et al., 2001).

The psychometric properties of the BDI-II have also been tested in samples which may be comparable to people with MS. This has addressed concerns about the inclusion of somatic items for a sample with physical health problems due to the overlap of symptoms. For example, inclusion of somatic items was supported in participants with chronic pain due to the item total correlations with chronic pain (e.g. loss of energy; $r = .53$; Harris & D'Eon, 2008).

2.7 Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983).

The HADS was developed to assess anxiety and depression in non-psychiatric patients with physical health problems (Zigmond & Snaith, 1983). Despite its' aim of increasing diagnostic accuracy through the omission of somatic symptoms (Snaith, 2003) a recent meta-analysis found alternative measures perform as well as the HADS in medical populations (e.g. BDI; Brennan, Worrall-Davies, McMillan, Gilboody & House, 2010). The HADS has also been criticised for its exclusion of

terms as the severe end of the depression spectrum (e.g. suicidal ideation and psychotic features; Schrag et al., 2007).

Although it has been criticised the HADS continues to be a widely used measure in both research and clinical practice (Bjelland, Dahl, Haug & Neckelmann, 2002). Respondents are required to answer 14 multiple choice questions about how they have been feeling over the previous week. For example, *I feel tense or 'wound up'* scoring in a range from *most of the time* to *not at all*. The measure is split into two subscales of anxiety and depression where the authors suggest a score above eight in both subscales indicates the potential of a disorder and 10 indicating a diagnosis (Zigmond & Snaith, 1983).

2.7.1 Reliability and validity of HADS. The reliability of the HADS has been assessed in a large non-clinical sample of 1792 participants and found to indicate adequate internal consistency with the anxiety and depression subscales being moderately correlated (0.53; Crawford, Henry, Crombie and Taylor, 2001).

The HADS was developed as having a two dimensional construct of depression and anxiety (Zigmond & Smith, 1983). However both factor analyses (Martin, 2005) and the use of a Rasch model (Pallant & Tennant, 2007) have questioned this assumption. One of the items in particular is problematic: item seven, *I can sit at ease and feel relaxed*. This item is labelled as loading onto the anxiety subscale but has been found to correlate higher with the depression subscale (e.g. Mykletun, Stordal & Dahl, 2001). As a result of this confusion some authors suggest the total score of the HADS should be used as a measure of psychological distress rather than two separate scores of depression and anxiety (e.g. Razavi et al., 1990).

Bjelland et al. (2002) reviewed studies that reported the concurrent validity of the HADS. In comparison to other measures the HADS has demonstrated high correlations (e.g. in comparison to BDI, $r = .6 - .8$; Bjelland et al., 2002). Since this review further studies have been published adapting and validating the HADS for specific client groups

where it has continued to demonstrate robust psychometrics (e.g. intellectual disabilities, Dagnan et al., 2008).

2.7.3 Use of HADS in people with MS. The HADS has been used in people with MS previously where it has been found to be less sensitive when compared to other measures of anxiety and depression (e.g. Nicholl et al., 2001), for example, when compared to the BDI-II it the HADS-depression subscale was found to have just 25% sensitivity (Nicholl et al., 2001).

The validity of the HADS has been assessed in people with MS by comparing it to a gold standard clinical interview. It was found that the advised cut off score of 8 provided high sensitivity and specificity for depression (sensitivity: 90%, specificity: 87.3%) and generalised anxiety disorder (sensitivity: 88.5%, specificity: 80.7%) (Honarmand & Feinstein, 2009). This study was conducted on a Canadian population where the prevalence rate for MS (>50 per 100,000, Poppe, Wolfson & Zhu, 2008) is lower than the UK (107 per 100,000; Robertson et al., 1995) and has not been replicated in the UK.

2.8 Gold Standard Measure: Structured Clinical Assessment in Neuropsychiatry (SCAN; Win et al., 1990).

As stated previously the gold standard measure for a psychiatric diagnosis is a structured clinical interview (section 1.8.1). There exist a number of such interviews (e.g. Composite International Diagnostic Interview, CIDI, Robins et al., 1989; Structured Clinical Interview for DSM-IV, SCID, First et al., 1997) which are similar in nature (Brugha, Jenkins, Taub, Meltzer & Babbington, 2001). The SCAN was chosen for the current study.

2.8.1 Justification for choice of measure. In contrast to some alternatives (e.g. CIDI) the SCAN requires that each criterion is being met currently (Brugha et al., 2001). As a result the SCAN is more likely to map onto the measures being assessed which considers the experiences of participants up to a maximum of two weeks prior to the

measure being completed. Although structured, the small amount of flexibility within the SCAN does prevent underreporting of symptoms (Eaton, Neufield Chen & Cai, 2000).

Pragmatically, the training required to complete the interviews reliably was more accessible to the researcher than the possible alternative interviews. The SCAN also maps on to two diagnostic systems, the ICD-10 (WHO, 1992) and the DSM-IV (APA, 2000). Alternatives only map onto the latter. This was felt to be important as the measures being validated against it were to be used in the NHS which favours the ICD-10 system (e.g. NICE, 2009).

Furthermore and importantly for the current study, Rijniers et al., (2000) confirmed research that less experienced but trained interviewers can apply the SCAN reliably (Brugha, Neinhuis, Bagghi, Smith & Meltzer, 1999). The researcher has limited experience of using the SCAN but did complete the training recommended to use the instrument.

2.8.3 Reliability and Validity of SCAN. The SCAN has substantial test-retest reliability for identifying if an individual met a level of diagnostic 'caseness' (.62; Rijniers et al., 2000). For the specific diagnoses being assessed in the current study the test-retest reliability also remained good (.52 for depression and .49 for anxiety disorders; Rijniers et al., 2000).

The complexity of assessing validity in a gold standard measure has caused some authors to question the validity of the structured clinical interview (e.g. Leeman, 1998). The most common way is to assess validity by the measures' "capacity to reach a diagnosis equivalent with the diagnoses reached by other procedures of known validity" (Rosenman, Korton & Levings, 1997, pp.582). This has been achieved by comparing one gold standard with another. For example, the SCAN has been compared to the CIDI (Brugha et al., 2001). The agreement between the two interviews was poor when considering depression and anxiety diagnoses ($\kappa = .43$; Brugha et al., 2001). However, less than half of the participants in the study (44%) completed both

interviews as the majority of participants completed only either the SCAN or the CIDI, making comparisons between the measures in the same participants difficult. This echoes poor agreement found previously in similar comparison studies (e.g. .28 – .62; Andrews, Peters, Guzman & Bird, 1995). The poor agreement reflects both the methodological difficulties in assessing validity for a gold standard measure in psychiatry (Rosenman et al., 1997) and is similar to that found when other measures are compared (e.g. SCID and Personality Disorder Examination yielded kappa of diagnostic agreement of .38; O'Boyle & Self, 1990).

2.8.4 Use of the SCAN as a gold standard. Although the SCAN has not been used to validate measures in participants with MS, it has been used as a gold standard in a number of studies. This has included psychiatric disorders (Bech, Rasmussen, Olsen, Noerholm & Abildgaard, 2001), stroke (Lincoln et al., 2003), Parkinson's disease (Leentjens, Verhey, Lousberg, Spitsbergen, & Wilmink, 2000) and Huntington's disease (De Souza et al., 2009).

2.9 Information packs sent to participants

The information packs sent to the participants provided them with information about the study, measures to complete, a consent form and opt-in slips. A cover letter explained to the participant about the study and why they had been invited to become involved (appendix E). An information sheet, written by the researcher, included information about the aims of the study, the risks and benefits of involvement and contact details for participants (appendix H). Personal contacts of the researcher who have MS were asked to confirm the information sheet was understandable.

A consent form was included for participants to complete if they wished to take part in an interview (appendix F). Opt-in slips were used for participants to provide contact details for either interviews or to request a summary of results once the study was complete (appendix G). Finally, the information packs also contained copies of each of the

measures used in the study as well as a stamped addressed envelope for the participants to return the questionnaires to the research associate.

2.10 Procedure

Participants were recruited as outlined above. Those who consented to participation in the research returned the completed measures and an opt-in slip to a research associate. The research associate scored the measures and passed on contact details of participants who had consented to an interview with the researcher. This allowed the researcher to be blind to the questionnaire scores when conducting the interview.

The researcher then contacted these participants to arrange an interview. The diagnostic interview (SCAN) was completed at a time and place convenient to the participant and recorded with a digital voice recorder. The recording of the interviews allowed the researcher to listen back to the interview in order to confirm the responses to the questions. The researcher remained blind to the participants' questionnaire scores. The interview allows diagnosis of depression and/or anxiety using the ICD-10 and DSM-IV-TR. The recordings of the interviews were copied, clearly labelled and stored securely at a university base. Following the interview, the participants repeated the initial baseline measures and returned them to the researcher. This extra data allowed the analysis of the test-retest reliability of the measures being assessed.

3. RESULTS

3. 1 Plan of analysis

The aim of the study was to validate measures of anxiety and depression for use with people who have MS (see section 1.9 for further discussion of aims). To achieve this, the analysis assessed if the BDI, BAI and HADS would differentiate between those who were seen to have depression/anxiety and those who were not as measured by the SCAN interview. A Receiver Operative Characteristic (ROC) analysis was conducted which allowed the identification of potential new cut off scores for the measures being assessed.

Further analysis was conducted to evaluate the test-retest reliability of the measures being assessed. This was done using Pearson's correlation coefficient (Pallant, 2006). All analysis was conducted using SPSS for Windows version 16.0 (released September 2007).

3.1.2 Missing Data. Where there was missing data in the study pair wise deletion was used, whereby the case is excluded from calculations for which there is no score (Field, 2000). Two alternatives to handle missing data are list wise deletion or replacing the missing values with the mean value for the sample (Pallant, 2006). List wise deletion would have excluded cases with missing data from any analyses and thus reduced the sample size dramatically (Field, 2005). Replacing the missing values with the mean value for the sample is advised against in small samples as it suppresses the true value of the standard deviation (Field, 2005).

All of the participants were asked to repeat the measures following the interview. However, only 17 (81%) participants returned the completed measures to the researcher. Participants who did not repeat the measures were excluded from analyses of the data regarding repeated measures.

3.2 Participant recruitment and characteristics

3.2.1 Recruitment. In total, 98 potential participants were sent information packs regarding the study. Of these participants, 24 opted into the study and all the participants met the inclusion criteria. 21 participants were interviewed, one participant declined an interview when contacted and two participants were not able to be contacted prior to the deadline set for data collection. A flow chart of the recruitment of participants is shown in figure **two**.

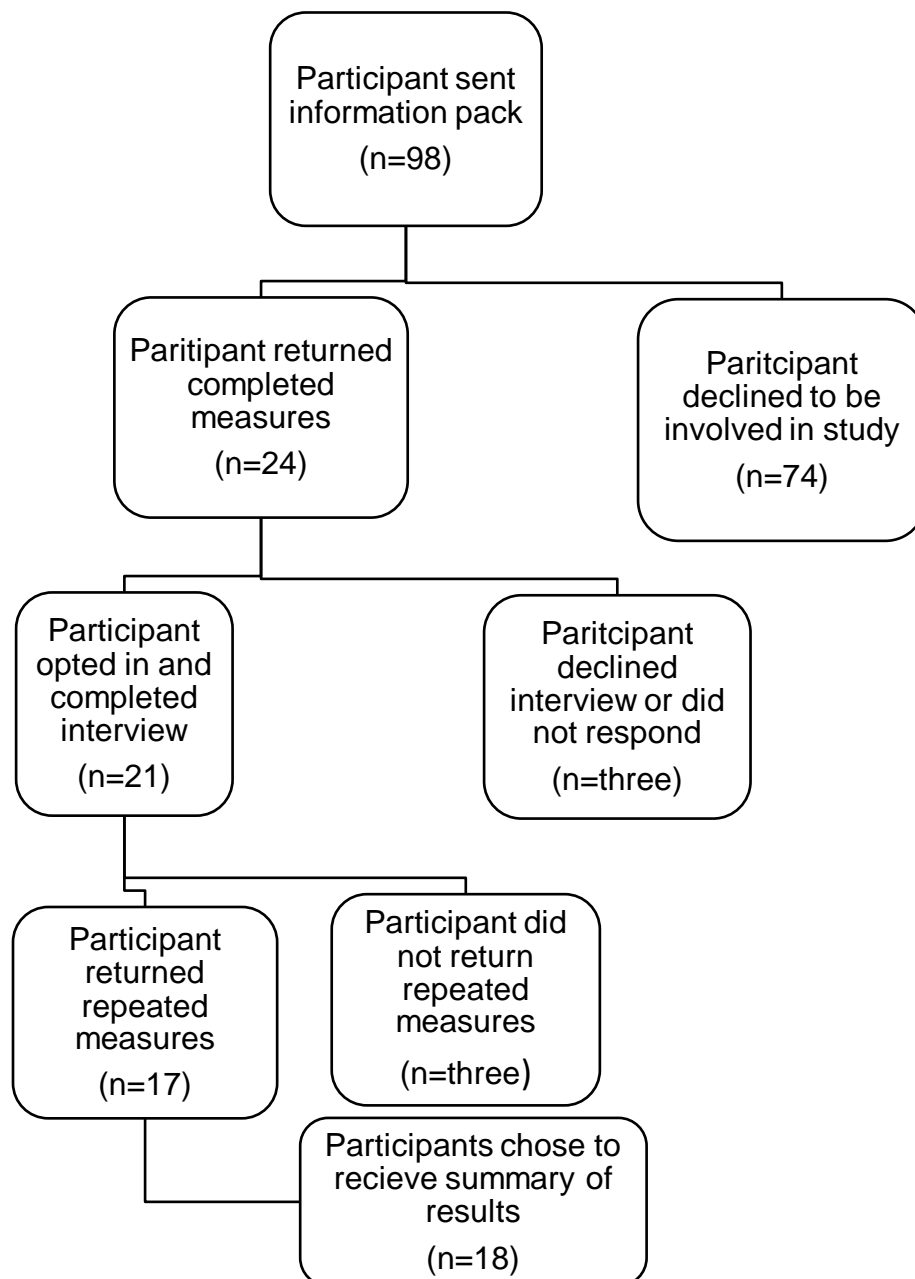


Figure two: Flow chart of recruitment

3.2.2 Participant Characteristics. Demographic information was gathered from participants. The sample consisted of 24 participants: six male (25%) and 18 female (75%). Participants provided details of the type of MS they had been diagnosed with: 14 (58%) of participants had relapsing remitting MS, five (21%) had secondary progressive MS and four (17%) had primary progressive. One (4%) participants was unclear about the type of MS they had.

For the continuous data of age, time since MS diagnosis and level of disability as measured by GNDS score means and standard deviations were calculated. The characteristics of the sample are summarised in the table below (table 12).

Table 12

Descriptive statistics for continuous data

Variable	Mean	Standard Deviation	Range
Age of participant	49.25	9.65	33-67
Years since MS Diagnosis	12.13	7.50	2-34
GNDS Score	17.92	9.23	3 - 38

3.3 Measures

3.3.1 Exploring the distribution of scores. Statistical tests assume that the distribution of data is normal (Pallant, 2006). Therefore, a number of tests were conducted to ascertain if the data was normally distributed and this assumption was not violated. The three assessment measures of BAI, BDI-II and HADS were tested. Initially, the distribution

of the data was assessed visually through a histogram (see figures – three to six).

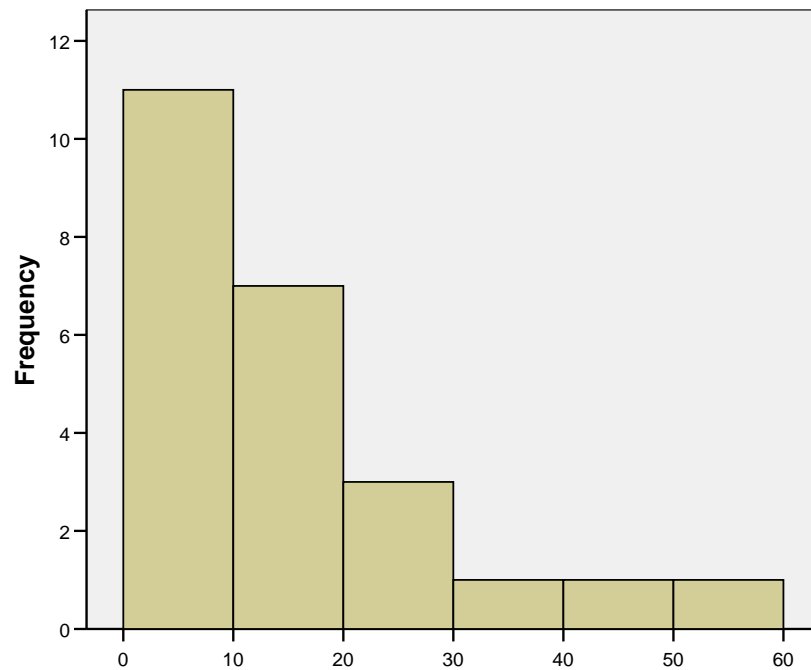


Figure three: Histogram for total scores on BAI

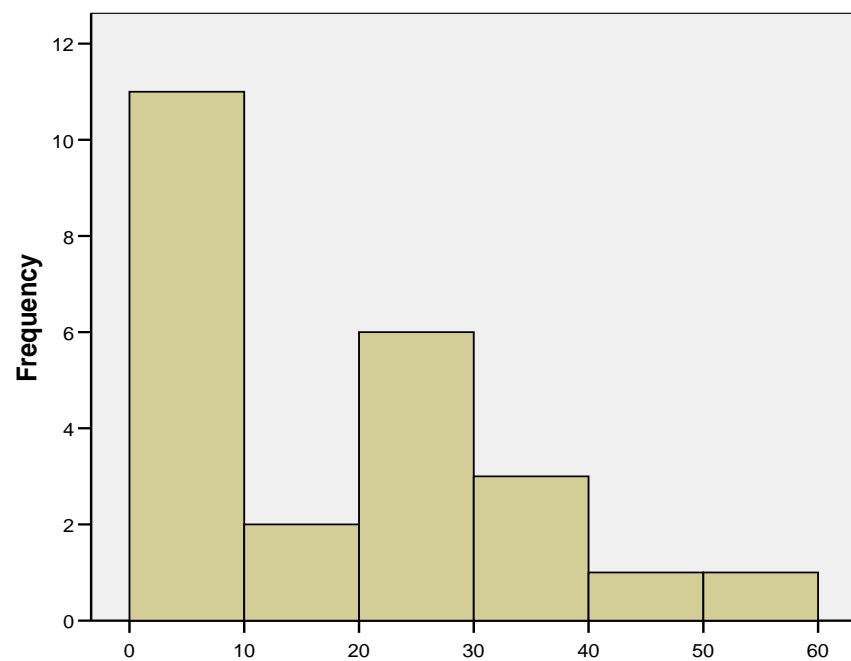


Figure four: Histogram for total scores on BDI-II

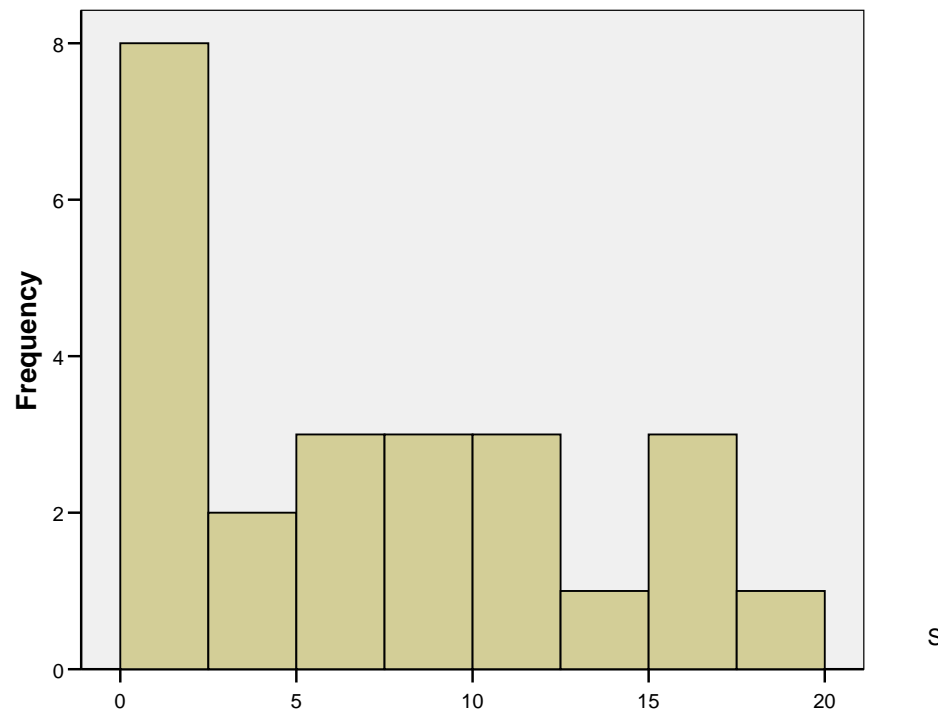


Figure five: Histogram for total scores on HADS – anxiety

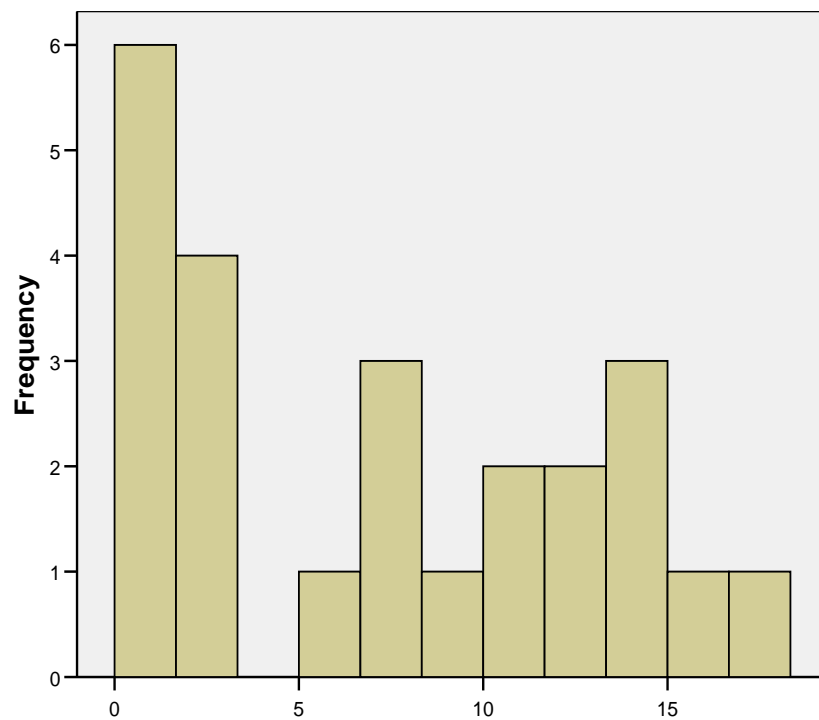


Figure six: Histogram for total scores on HADS - depression

The following criteria were used to assess if the data was normally distributed (adapted from Hinton, Brownlow, McMurray & Cozens, 2004):

- The histogram should be bell shaped
- The tails should meet the x axis at infinity
- The distribution of scores should be symmetrical about the mean

Using these criteria, all the measures were found to be non-normally distributed.

To assess the distribution using a more objective measure the skewness and kurtosis of the scores was computed. If the data is normally distributed then the skewness and kurtosis values should be 0, the further away from this they are the less likely the data is normally distributed (Field, 2005). Field (2005) suggests that skewness and kurtosis are more informative if converted into a z score, this was done by subtracting the mean of the distribution and dividing it by the standard deviation of the distribution. A z score outside of ± 1.96 is considered to be significantly outside of a normal distribution ($p < 0.05$, Field, 2005). Finally the Shapiro-Wilk test was conducted as advised by Pallant (2006), a non significant score indicates the data is distributed normally. Table 13 provides a summary of the tests of normality conducted on the data.

Table 13

Tests of normal distribution

Measure	Skewness	Z Skewness	Kurtosis	Z Kurtosis	Shapiro- Wilk Test	P value for Shapiro-Wilk Test
BAI	1.69	3.57	3.02	3.29	.84	.001
BDI-II	0.76	1.61	-0.21	- 0.23	.90	.019
HADS - anxiety	0.32	0.68	-1.19	-1.30	.92	.046
HADS - depression	0.19	0.40	-1.51	-1.64	.90	.018

Following each of these tests being conducted the overall conclusion can be made that the scores on the BAI, BDI and HADS did not meet the assumptions of a normal distribution within this sample. Therefore, any tests being conducted would be non-parametric (Pallant, 2006).

3.3.2 Descriptive statistics of measures being assessed.

Each of the participants completed a full set of the measures being assessed by the current study. Of the 21 participants included in the study nine (43%) were diagnosed as likely to have depression by the SCAN interview and eight (33%) were diagnosed as likely to have anxiety. The scores from the measures are summarised in table 14, as measures were found to be non-normally distributed the median and inter-quartile range are shown.

Table 14

Participant scores on measures being assessed.

Assessment	Sample size (n)	Inter-quartile Range (25% - 75%)	Median
BAI	24	6 – 20	11.00
BDI-II	24	5 – 29	11.50
HADS – anxiety	24	1 – 12	6.50
HADS – depression	24	1 – 13	7.50
BAI Repeat	17	3 – 14	7.00
BDI-II Repeat	17	3 – 23	7.00
HADS-anxiety Repeat	17	2 – 9	2.00
HADS-depression Repeat	17	2 – 8	2.00

NB: Due to the pair wise deletion of missing data there was a smaller sample size for participants who completed the repeated measure*3.3.3**

Kappa coefficients using advised cut offs. To compare the level of agreement with the advised cut off scores for the measures and the diagnosis provided by the SCAN interview a kappa analysis was conducted. The kappa calculation is preferable to the alternative of a percentage agreement because “it corrects for the probability that raters will agree due to chance alone” (Leech, Barrett & Morgan, 2005). The output of the calculation gives a level of agreement in terms of a Cohen’s Kappa coefficient between zero and one, the level of agreement between measures are labelled according to their strength (see table 15).

Table 15

Strength of Kappa coefficient (adapted from Pallant, 2006)

Kappa coefficient	Level of agreement
0 – 0.4	Poor – no better than chance
0.5 – 0.6	Moderate
0.7	Good
0.8 – 0.9	Very good
1	Perfect

The results of the kappa analysis are shown in table 16. The BDI-II and both of the HADS subscales demonstrated statistically significant moderate to very good agreement with the diagnosis given by the SCAN. The BAI demonstrated non-significant poor agreement.

Table 16

Strength of kappa coefficient of measures in comparison with SCAN

Measure	Cases in meeting criteria for mood disorders in sample (n)	Percentage	Kappa coefficient
BDI-II	9	43	0.81*
BAI	11	52	0.34
HADS-anxiety	9	43	0.70*
HADS – depression	9	43	0.61*
SCAN – depression	9	43	-
SCAN – anxiety	8	38	-

* Statistically significant ($p < .05$)

3.4 ROC analysis

A ROC curve analysis is a commonly used tool for assessing the performance of individual measures in making accurate assessments of diagnosis (Zou, O'Malley & Mauri, 2007). In the current study the ROC curve analysis was used to assess if the BDI, BAI and HADS correctly identified participants with MS who were experiencing depression or anxiety (as measured by the SCAN).

A ROC curve plots sensitivity and specificity over a range of values. For tests examined in the present study, specificity is the probability that the test correctly classifies a person without depression and/or anxiety as negative (Fletcher & Fletcher, 2005). Sensitivity is the probability that the test correctly classifies a person with depression and/or anxiety as positive (Fletcher & Fletcher, 2005). The measures assessed in the current study have a continuous scale; this allows the cut off scores to be altered. Different cut off scores have different levels of sensitivity and specificity (e.g. Honarmand & Feinstein, 2009).

In making the choice for an optimum cut off score for a measure a balance needs to be made between sensitivity and specificity. If the threshold is strict there are fewer false positives but an increased chance of failing to identify true instances of the construct of interest (Swets, Dawes & Monahan, 2000), in the current study the constructs of interest are depression and anxiety. If the threshold is more lenient it will be likely identify there are more true positives but also produce more false positives (Swets et al., 2000). Although a perfect measure would have 100% sensitivity and 100% specificity this is rarely found in practice (Linden, 2006). The standards used in the current study reflect those used in previous research. Previous studies have considered measures with high specificity (>80%) and adequate sensitivity (>60%) to be acceptable (e.g. Eack et al., 2008; Sprinkle et al., 2002).

The ROC curve plots sensitivity on the y axis against 1-specificity on the x axis for different scores on the measure. A 45° diagonal line drawn on the ROC curve corresponds to random chance (connecting

0,0 to 0,1). The gold standard diagnosis is represented as the straight lines forming the upper-left corner of the graph (connecting 0,0 to 0,1 and 0,1 to 1,1; Linden, 2006). The measure being assessed is represented as curve between these two lines, the closer this line is to the gold standard the more accurate the classification from the measure (Swets et al., 2000). If the line is below the diagonal line then the performance on this measure is worse than chance.

The area under the curve (AUC) provides a summary of the diagnostic accuracy of the measure (Zou et al., 2007). If the ROC curve corresponds to random chance then $AUC=0.5$, if it is perfect accuracy then $AUC = 1$. The closer the AUC is to 1, the more accurate the measure is (Zou et al., 2007). For measures to be used as screening tools, as was advocated in the aim of this study an AUC of .8 is thought to be acceptable.

3.4.1 Contingencies Once the optimal cut off score for a measure was identified the new cut off score was transformed within SPSS. Any scores below the new cut off score on the measure are assigned a value of 0 and any scores equal or above the new cut off score are assigned a value of 1. A two-by-two contingency table was used to represent the outcomes with the new cut off score (Fawcett, 2006; see table 17).

The data in the contingency table was used to determine the false positive rate, true positive rate, sensitivity and specificity for the new cut off using the metrics provided in table 18. The metrics also allowed for the positive predictive values (PPV) and negative predictive values (NPV) to be calculated, these allow the probability of the measures giving the correct diagnosis in the population being tested. The PPV is the proportion of individuals who are assessed as having depression or anxiety by the measure that actually have them (Chatburn, 2009). The NPV is the proportion of individuals who are assessed as not having depression or anxiety by the measure that do not have either (Chatburn, 2009).

Table 17

Contingency table (adapted from Zou et al., 2007).

Index Test (BAI/BDI/HADS)	Gold standard (SCAN)		Total
	Positive for depression or anxiety	Negative for depression or anxiety	
Positive for depression or anxiety	A = true positives	B = false positives	A+B = test positives
Negative for depression or anxiety	C = false negatives	D = true neagtives	C + D = test negatives
Total	A+C = diseased	B + D = non- diseased	A + B + C + D = total sample size

Table 18

Metrics for contingency table (adapted from Zou et al., 2007)

	Metrics
False positive rate	$B / (B+D)$
False negative rate	$C / (A+C)$
Sensitivity	$A / (A+C) \times 100$
Specificity	$D / (B + D) \times 100$
Positive predictive Value(PPV)	$A / (A + B)$
Negative Predictive Value (NPV)	$D / (C + D)$
Discriminant Ability	$\text{Sensitivity} + \text{Specificity} / 2$

3.4.2 Analysis for BAI. The ROC curve completed for the BAI demonstrates that it curves to the left but is not quite in the top left hand corner. The AUC is .81 (95% confidence interval = .62 -1), if rounded to one decimal place then the accepted level for this study of .8 is achieved.

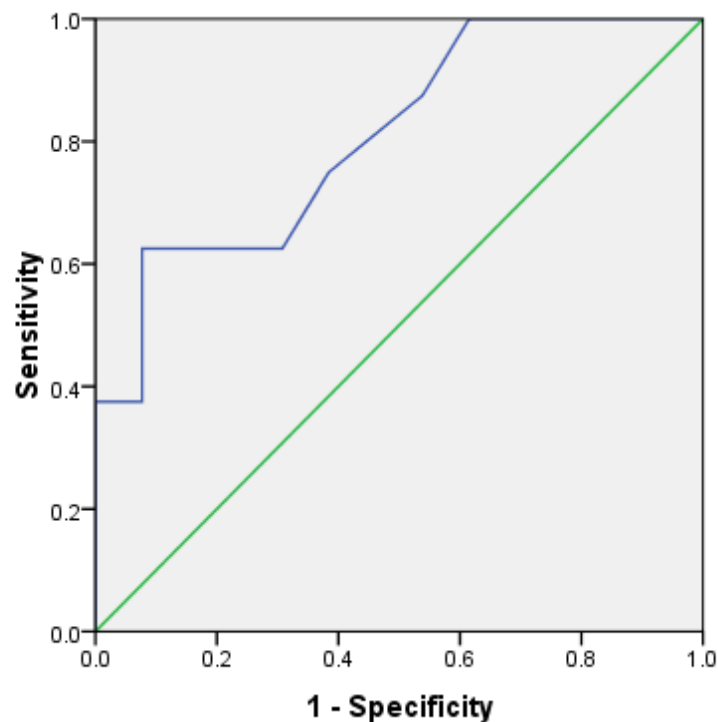


Figure seven: ROC curve for BAI scores

The optimal cut off score for the BAI in people with MS in this sample was calculated using the co-ordinates of the ROC curve (see table 19). The optimum cut off score seems to be 9.5 as it identifies those who have anxiety as diagnosed by the SCAN interview, demonstrating good sensitivity (75%) and adequate specificity (61%). As a score of 9.5 cannot be achieved with the BAI, the cut off will be viewed as those who score 10 or more will be viewed as anxious and those who score nine or less as not anxious.

Table 19

Co-ordinates of ROC curve for BAI

Positive if Less Than or Equal To	Sensitivity	Sensitivity (%)	1-Specificity	Specificity (%)
.5	1.00	100	.92	8
2	1.00	100	.85	15
3.5	1.00	100	.77	23
5	1.00	100	.62	38
7	.86	86	.54	46
9.5	.75	75	.39	61
11.5	.63	63	.31	69
12.5	.63	63	.23	77
13.5	.63	63	.15	85
16	.63	63	.07	93
19	.50	50	.07	93
22.5	.38	38	.07	93
26.0	.38	38	0	100
30.5	.25	25	0	100
37.5	.14	14	0	100
42	0	0	0	100

Contingencies. As it was not appropriate to have a cut off score of 9.5 a score of more than ten was used to indicate anxiety, therefore sensitivity and specificity were calculated for a score of ten. The new optimum cut off score of 10 was redefined in SPSS into binary terms. Where a score of less than nine was given a value of zero and a score of more than or equal to 10 was given a value of one. A contingency table for these scores was then constructed (see table 18).

Table 20

Contingency table for BAI using 10 as a cut off

Index Test - BAI	Gold standard (SCAN)		Total n
	Anxiety	No anxiety	
Anxiety	6	5	11
No anxiety	2	8	10
Total n	8	13	21

This table was then used to calculate the metrics (see table 16). The cut off score of 10 for the BAI identified 11 positive cases and 10 negative cases for possible anxiety. Using 10 as a cut off score the sensitivity of the BAI remains high (75%) as does the specificity (61%). The PPV of the BAI is 55% ($6/11=.55$). The NPV of the BAI is 80% ($8/10=.8$). The discriminant validity of the test provides a useful summary and assumes that the SCAN is perfect (100%), for the BAI the discriminant ability is 68% ($75\%+61\% / 2 =.68$).

Kappa coefficient. The kappa coefficient was calculated to assess the agreement between the BAI and SCAN diagnosis for anxiety when the new cut off score was used. The agreement level was poor (0.34, non-significant).

3.4.3 Analysis for BDI-II. The ROC curve completed for the BDI-II is good; it shows that the measure almost reaches the top left hand of the graph (figure eight). The AUC is also very high (.98, confidence interval .93-1.03) demonstrating the high accuracy of the measure.

The co-ordinates of the curve were used to calculate an optimum cut off for the sample in the study. It appears that the optimum cut off for the BDI-II is 18 (table 21). This provides very high sensitivity (89%) and specificity (92%).

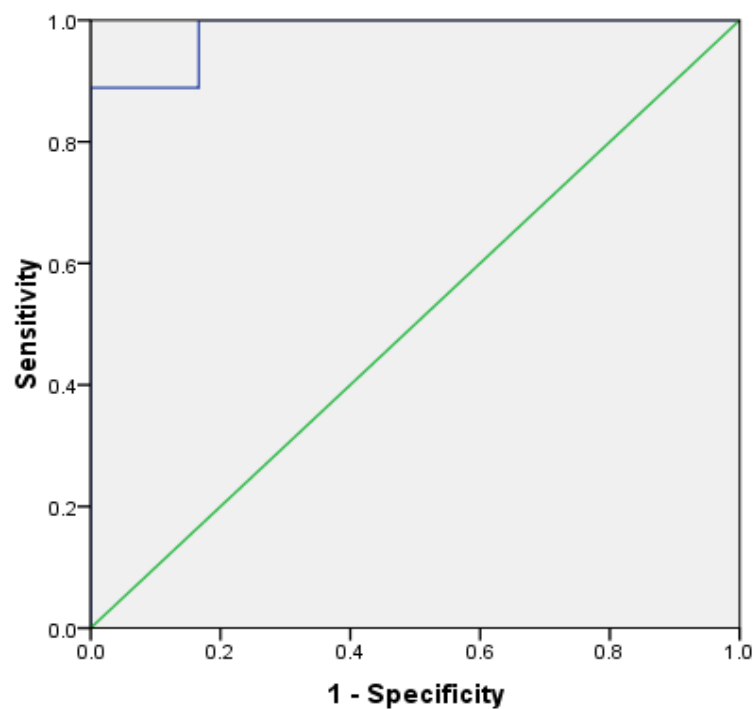


Figure eight: ROC curve for BDI-II

Table 21

Co-ordinates of ROC curve for BDI-II (extension of table 1 in journal article)

Positive if Less Than or Equal To	Sensitivity	Sensitivity (%)	1-Specificity	Specificity (%)
.5	1.00	100	.83	17
1.5	1.00	100	.75	25
3	1.00	100	.68	32
4.5	1.00	100	.58	42
5.5	1.00	100	.33	67
6.5	1.00	100	.25	75
7.5	1.00	100	.17	83
10.5	.89	89	.17	83
18	.89	89	.08	92
24.5	.89	89	0	100
26.5	.78	78	0	100
28	.67	67	0	100
30	.44	44	0	100
33.5	.33	33	0	100
37	.22	22	0	100
40	.11	11	0	100

Contingencies. The new cut off score of 18 was then transformed in SPSS where all scores below 18 were given a value of zero and those equal to or above 18 were given a value of one. This allowed the contingencies to be calculated using the new cut off score (table 23).

Table 22

Contingency table for BDI-II

Index Test	Gold standard (SCAN)		Total (n)
	Positive for depression	Negative for depression	
BDI-II			
Positive for depression	8	1	9
Negative for depression	1	11	12
Total (n)	9	12	21

Using the new cut off score for depression the BDI-II identified nine participants with possible depression and 12 without. As was previously shown for this cut off score the sensitivity (89%) and specificity (92%) are high. The PPV is 88% ($8/9 = .88$) and the NPV is 92% ($11/12 = .92$). The discriminant validity of the test is $(89\% + 92\% / 2)$ 91% which is high.

Kappa coefficient. To assess the agreement between the BDI-II with the new cut off score and the SCAN a kappa coefficient was calculated. This showed very good agreement (0.81, $p < .001$).

3.4.4 Analysis for HADS – anxiety subscale. The ROC curve completed for the HADS-anxiety subscale shows that the measure almost reaches the top left hand of the graph (figure nine), confirmed by the high AUC (.96, confidence interval .89-1.04).

The coordinates of the ROC curve were used to calculate the optimum cut off score for the HADS-anxiety subscale. The optimum score for this sample is 10 which demonstrates both high sensitivity (86%) and perfect specificity (100%; see table 24).

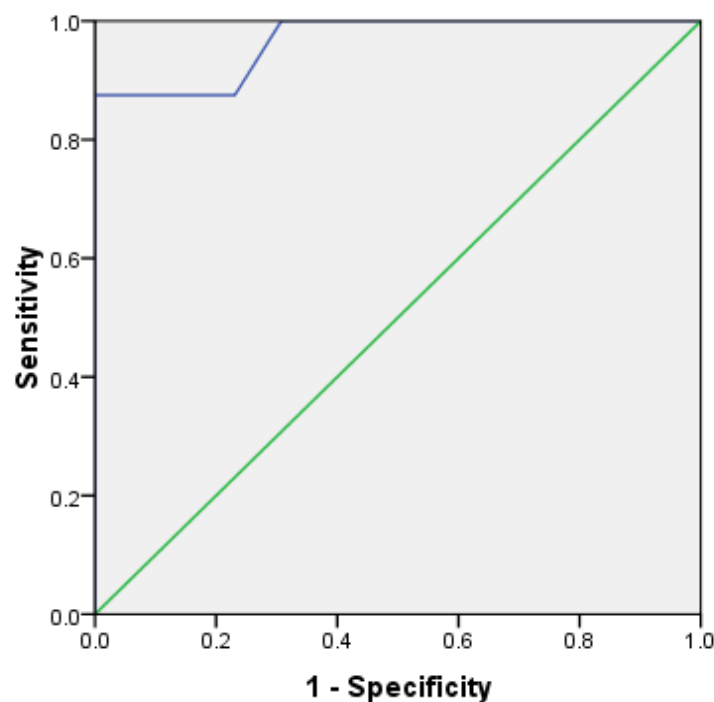


Figure Nine: ROC curve for HADS anxiety subscale

Table 23

Co-ordinates of ROC curve for HADS-anxiety subscale

Positive if Less Than or Equal To	Sensitivity	Sensitivity (%)	1-Specificity	Specificity (%)
.5	1.00	100	.69	31
1.5	1.00	100	.46	54
3	1.00	100	.39	61
5	1.00	100	.31	69
6.5	.86	86	.23	77
7.5	.86	86	.15	85
8.5	.86	86	.08	92
10	.87	87	0	100
11.5	.63	63	0	100
12.5	.50	50	0	100
14	.38	38	0	100
16.5	.13	13	0	100
19	0	0	0	100

Contingencies. The cut off score of 10 was used to transform the data in SPSS, where scores below 10 were given a value of zero and scores equal to or above 10 were given a value of one. This allowed a contingency table to be constructed (table 24).

Table 24

Contingency table for HADS-anxiety subscale

Index Test	Gold standard (SCAN)		Total (n)
	Positive for anxiety	Negative for anxiety	
HADS- anxiety subscale			
Positive for anxiety	7	0	6
Negative for anxiety	1	13	15
Total (n)	8	13	21

The cut off score of 10 identified six positive cases for anxiety and 15 negative cases. The table allowed for the matrices to be calculated. As stated previously with this cut off score the specificity was perfect (100%), the sensitivity was good at 87% ($6/8=.87$). The PPV was 100% ($6/6=1$) and NPV was high at 93% ($13/14=.93$). The discriminate validity for the test was very good at 94% ($100\% + 87\% / 2 = .94$).

Kappa coefficient. In order to measure the agreement between the HADS-anxiety subscale with a cut off of 10 and the SCAN diagnosis of anxiety kappa was calculated. A good agreement was found (.90, $p<.01$).

3.4.5 Analysis for HADS – depression subscale. The ROC curve which calculated for the HADS – depression subscale was also good. The AUC was high (.96, confidence interval 88-1.03) demonstrating that overall the measure was quite accurate (see figure 10).

The coordinates of the ROC curve were used to calculate the optimum cut off score, of 9.5. This demonstrated high sensitivity (78%) and higher specificity (92%; see table 23). When scoring the measure no decimal places are used, therefore the cut off score was rounded up to 10.

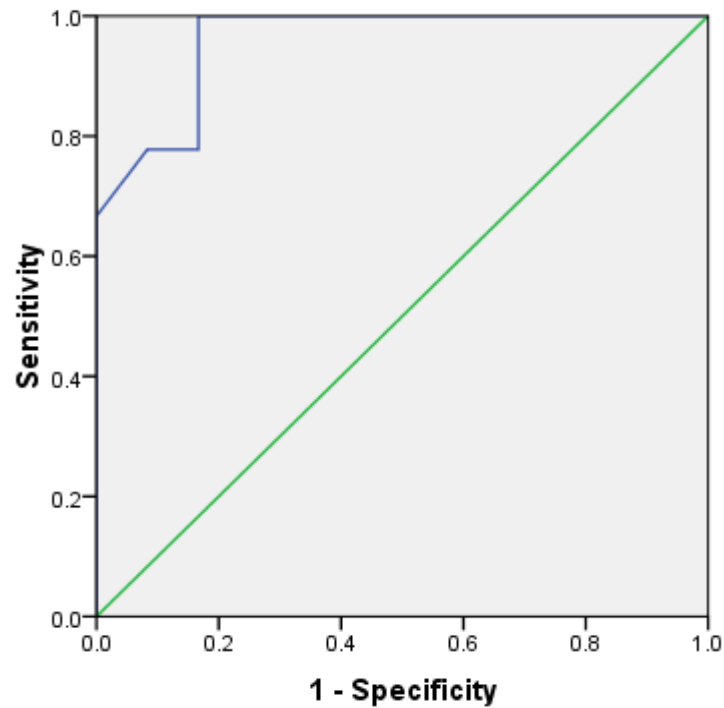


Figure 10: ROC curve for HADS depression subscale

Table 25

Co-ordinates of ROC curve for HADS-depression (extension of table 3 in journal article)

Positive if Less Than or Equal To	Sensitivity	Sensitivity (%)	1-Specificity	Specificity (%)
.5	1	100	.67	33
1.5	1	100	.5	50
2.5	1	100	.25	75
4	1	100	.17	83
6	.89	89	.17	83
7.5	.78	78	.17	83
9.5	.78	78	.08	92
11.5	.68	68	0	100
12.5	.56	63	0	100
13.5	.44	50	0	100
15	.22	25	0	100
16.5	.11	13	0	100
18	0	0	0	100

Contingencies. The optimum cut off score of 9.5 was rounded up to 10, as with the BAI (see section 3.4.2). The new cut off score of 10 was transformed in SPSS so scores of nine or below were given a value of zero and scores of equal to or more than 10 were given a value

of one. This allowed a contingency table to be constructed and the accompanying matrices to be calculated (table 26).

Table 26

Contingency table for HADS-depression (adapted from Zou et al., 2007).

Index Test	Gold standard (SCAN)		Total
	Positive for depression	Negative for depression	
HADS-depression			
Positive for depression	7	1	8
Negative for depression	2	11	13
Total	9	12	21

Using the cut off score of 10 the HADS-depression scale identified seven cases with possible depression and 14 without. The sensitivity of the cut off score is good (78%) and the specificity is high (92%). The PPV value is 88% ($7/8=.88$) and the NPV value is 85% ($11/13=.85$). The discriminant validity of the measure is 85% ($78\% + 92\% / 2$).

Kappa coefficient. The agreement between the HADS-depression subscale with a cut off of 10 and the SCAN diagnosis of depression is good (.70, $p<.01$).

3.5 Reliability

The reliability of a measure is its' ability to consistently reflect the construct being assessed (Field, 2005). Of concern in the current study is test retest reliability (see section 1.8.2 for further discussion of

reliability). If all other things are equal and the measure is reliable participants should achieve the same score at both time points (Field, 2005). The two points of assessment of the measure are then compared using a correlation. Since the current study did not have normally distributed data a non-parametric test was needed and therefore Spearman's correlation coefficient was calculated (Field, 2005).

For the current study the analysis of reliability involved a smaller sample size than the rest of the study (17 participants) as not all participants completed the repeated measures. As for the other measures missing data was excluded pair wise (see section 3.1.2). The results of the analysis are shown in table five (in journal article). The result of the Spearman correlation demonstrate statistically significant correlation in a positive direction, this suggests that each of the measures assessed has good test re-test reliability (Field, 2005).

4. DISCUSSION

Three of the four measures assessed were found to be valid to use with people with MS. The optimum cut off scores for these measures in the sample population was higher than was originally recommended when the scales were developed: the optimum cut off score for the BDI-II was 18 (original cut off score in manual was 14) and the HADS anxiety and depression subscales were 10 (original cut off score recommended by developers was 8). The BAI was not found to be valid for use in people with MS due to its poor agreement with the gold standard SCAN interview. The implications of these results will be discussed in the context of previous and future research as well as their implications for clinical practice. A critical reflection of the study will be given before final conclusions are drawn.

4.1 Placing findings in context of previous research

4.1.1 BAI. The conclusion drawn from the results of this study imply that the cut off score for the BAI does not need to be adjusted when using it in people with MS. However, the poor agreement with the diagnosis provided by the clinical interview suggests that the BAI is not a valid measure to be used in this population.

There has only been one previous study considering the use of the BAI in people with MS (Nicholl et al., 2001). This reflects the paucity of literature in MS regarding anxiety in comparison to depression (e.g. Honarmand & Feinstein, 2009). Nicholl et al. (2001) validated the BAI by comparing it to other measures of anxiety in people with MS and it was found to have similarly poor agreement between measures, consistent with findings of the current study.

The BAI has been validated in people with comparable physical conditions to MS, for example, people with Parkinson's disease. However concerns were expressed that the BAI overestimated the prevalence of anxiety in those with Parkinson's disease (Higginson et al., 2001). Although the BAI did overestimate anxiety in this study (the

BAI suggested 48% of participants had anxiety as opposed to the 33% identified by the SCAN) due to the small sample size it is difficult to draw firm conclusions about this. The overestimation of anxiety in the sample of people with MS may be due to the nature of the measure which focuses on physical symptoms of anxiety rather than cognitive symptoms (Wilson et al., 2006). As a result there may be a strong bias particularly in terms of the confounding symptoms in the measure such as '*numbness or tingling*' which may be present in both MS and anxiety. This may also explain the poor agreement with the SCAN as the SCAN includes a range of symptoms of anxiety rather than exclusively physical symptoms.

The test-retest reliability of the BAI demonstrated that it is a moderately reliable measure (Spearman's $r = .68$). The results from the current study are in the lower end of the range found by previous studies ($r = .67$, Fydrich et al., 1992 to $r = .83$, DeBeurs et al., 1997). The paucity of literature reporting the test-retest reliability of the BAI could be due to the methodological design of many studies. To assess the test-retest reliability of a measure the measure needs to be completed a two time points. This may lead to logistic problems with recruitment. As was demonstrated in this study not all participants are willing to complete the repeated measure resulting in a small sample size. Additionally, some researchers may not ask participants to repeat measures due to ethical concerns about over assessing individuals. A further explanation for the lack of test-retest reliability reporting in the literature is that studies which do repeat measures (e.g. Moran and Mohr, 2005) may do so not to assess reliability of the measure but to detect change following an intervention, therefore although a retest is conducted it is not reported as reflecting the test-retest reliability of the measure.

In summary the BAI was not found to be a valid measure for use in people with MS and is therefore not recommended for screening in people with MS. This reflects the small amount of previous studies which have attempted to validate the BAI in people with MS.

4.1.2 BDI-II. The results of the current study suggest that the BDI-II is a valid measure for use in people with MS when a higher cut off score of 18 is applied (rather than 14). This score demonstrates high sensitivity (89%) and specificity (92%), due to its good agreement with the SCAN interview ($\kappa = .81$, $p < .01$). These results are in contrast with previous research which suggests both a lower cut off score (e.g. 13; Sullivan et al., 1995) and poor agreement with other measures (e.g. $\kappa = .12$ when compared to HADS; Nicholl et al., 2001).

The higher cut off score in the current study in comparison to a previous study may reflect differences in the sample used. Sullivan et al. (1995) used participants who had recently been diagnosed with MS. It may be that the differing recommendations for cut off scores are due to the impact of confounding symptoms. As MS is a progressive disease (Lezack et al., 2004) it may be assumed that those who are recently diagnosed have fewer symptoms than those who have had the disease diagnosed for some time. The higher cut off score in the current study for people with MS may reflect the greater potential number of confounding symptoms, such as fatigue (Mohr et al., 2003), in individuals where the MS has progressed further (see section 1.6) As a result a higher cut off score may be needed for measures assessing depression in people who have not recently been diagnosed with MS to reduce the number of false positives within a sample.

An alternative explanation for the differing cut off scores between Sullivan et al. (1995) and the current study is the possible differential diagnoses given to those recently diagnosed with MS. The gold standard clinical interview to which the results were compared took into account recent life events, which would include the diagnosis of MS. Therefore, participants who have been recently diagnosed (i.e. in the study completed by Sullivan et al., 1995) may have been more likely to receive a diagnosis of adjustment disorder with depressed mood (DSM-IV-TR, APA, 2000) rather than depression, to account for the impact of the diagnosis. The BDI-II would have then been adjusted to match the gold standard in order to validate it, potentially providing a lower cut off

score. In the current study people who had been recently diagnosed were excluded and so any potential difficulties in discriminating between depression and adjustment disorder were reduced. As a result, it is possible that more people were diagnosed with depression as a primary diagnosis (rather than adjustment disorder with depressed mood) and thus the cut off score of the BDI-II was raised to match this. Therefore the results from the current study may be a more accurate and useful reflection of the validity of the BDI-II in MS that that suggested by Sullivan et al. (1995).

Unlike the current study, previous attempts at comparison between the BDI-II and other measures in people with MS have found poor agreement. Nicholl et al. (2001) compared to other measures of depression (e.g. HADS) but did not use a gold standard measure as was used in the current study. Nicholl et al. (2001) used the cut off score of eight for the HADS; however, the current study has demonstrated an optimum cut off score for the HADS measure is 10.

The test-retest reliability of the BDI-II assessed in this study was very good (see section 3.5). This suggests that the construct of depression as measured by the BDI-II in people with MS is stable over time. The good reliability of the BDI-II found in the current study supports similar findings in previous studies: for example, Sprinkle et al. (2002) also found the BDI-II demonstrated a good test-retest reliability of in a sample of university counselling students (Pearson's $r = .96$). There are limited studies which have assessed the test-retest reliability of the BDI-II in comparison to those that have assessed the original BDI (Hagen, 2007; Yin & Fan, 2000 completed a meta-analysis on the BDI and found good test-retest reliability, (Pearson's $r = .69$). In the development of the measure the test-retest reliability is reported as high (Pearson's $r = .93$; Beck et al., 1996). However, this was computed using a small sample size (26 outpatients) and so there continues to be a lack of any large scale assessment of the test-retest reliability of the BDI-II. The paucity of literature considering the test-retest reliability of the BDI-II may reflect its development, as the measure was designed to both measure a

stable construct and be sensitive to change from treatment (Dozois & Covin, 2004). It has been shown with original BDI that scores decrease over time (e.g. Yin & Fan, 2000). The revised BDI-II covers a larger time period for participants to assess (how they feel currently in the BDI versus how they have felt over a two week period in the BDI-II) and as a result it may be expected that the BDI-II will demonstrate greater temporal stability (Dozois & Covin, 2004). However, further research is needed to test this hypothesis.

In summary, although previous studies have suggested a lower cut off score when using the BDI-II in people with MS, these used different samples which lead to difficulties in making comparisons. Other previous studies have not compared the BDI-II to a gold standard. The BDI-II has been found to be both a reliable and valid measure for use in people with MS and therefore it is recommended that it is used as a screening measure.

4.1.3 HADS. The results show both subscales of the HADS can be used as a valid measure of anxiety in people with MS with an increased cut off score of 10 for both subscales (rather than 8 which is recommended by the manual).

The increased cut off score for the HADS is in contrast to that found in previous studies, such as the large scale validation project completed by Honarmand and Feinstein (2009). Given the importance of replication within scientific research (Reiss & Judd, 2000) this suggests concerns may be raised regarding generalising the conclusions from Honarmand and Feinstein (2009). Although the current study had a similar methodology it was completed on a Canadian population rather than in the UK where there is a different healthcare system and a lower prevalence rate of MS (e.g. Poppe et al., 2008).

The higher cut off score suggested in the current study (10) rather than that recommended by Honarmand and Feinstein (2009; 8) may reflect the structure of the HADS measure. In its development it is suggested that a score of eight to 10 on the HADS indicates possible depression

or anxiety and a score of 10 or more suggests probable depression or anxiety (Zigmond & Snaith, 1983). Therefore it may be that those with MS need to meet the higher end of the range to indicate possible depression or anxiety. In addition, the higher cut off score suggested for both anxiety and depression when using the HADS may reflect the confounding symptoms within the disorders, as discussed previously (see section 1.6).

The test-retest reliability of the HADS in this study was good (Spearman's $r = .88$ for depression subscale and $r = .77$ for anxiety subscale). This is similar to findings from previous studies (e.g. Pearson's $r = .79-.80$, Elliott, 1993) and studies considering international versions of the HADS (Herrmann, 1996). Although the test-retest reliability of the HADS has been reported for people with MS, it has been reported for similar groups. For example, when used in people with Parkinson's disease the HADS has demonstrated good test-retest reliability over a two week period (Pearson's $r = .88$; Marinus, Leentjens, Visser, Stiggelbout & van Hilten, 2002)

In summary, the HADS has been shown to be valid and reliable for use in people with MS. Although this study contrasts with previous studies that have validated the HADS in people with MS, this may reflect subtle differences in the samples used. The slightly higher cut off score reflects the higher end of the range recommended by the original manual.

4.1.4. Prevalence of anxiety and depression. The number of people with depression and anxiety found in this study (33-48% depending on the measure) reflected the prevalence rates in some previous studies (e.g. Smith & Young, 2000; Gelazzi et al., 2005) but not all (e.g. Barmer et al., 2008; Feinstein et al., 2002b). This highlights the wide variance in prevalence rates reported in studies as discussed previously. This could possibly reflect the difficulties in assessing these constructs in people with MS (e.g. confounding symptoms, see section 1.6).

In the current study a difference in prevalence was found when using different measures, although this was interpreted as high in comparison to previous research (e.g. Nicholls et al., 2001). The range of variation was greater for anxiety (33-48%) than for depression (all 38%). However if the non-validated measure (BAI) is removed then the range is reduced (33-38%). The small differences between the measures may reflect the small sample size used in the current study.

4.1.5. Participant characteristics. Despite the high numbers of participants invited to take part in the study, only a small number decided to participate (24%). In comparison to other research considering questionnaire data this uptake is low (e.g. 55% expected; Baruch, 1998). Recruitment is cited as the most challenging part of a study (Patel, Doku & Tennakoon, 2003), as was reflected in this study. The low recruitment rate is perhaps surprising given that participants were approached because they had previously agreed to be contacted regarding research in MS. It may have been that participants had already been recruited for a number of studies as the centre in which the study was based produces a high volume of research involving people with MS. The poor recruitment levels may also reflect the study design which involved the participants completing a number of steps. This resulted in a complex information sheet which may have deterred potential participants from opting in to the study. Finally the nature of the study may have deterred some participants. Those with depression may not have felt motivated to be involved, or poor health in potential participants may have adversely affected participation rates in this study (Patel et al., 2003). Despite the difficulties with recruitment there is no proven method of improving participant recruitment to research (e.g. Mapstone, Elbourne & Roberts, 2007). Possible ways to improve recruitment would be to offer participants an incentive or contacting potential participants by telephone rather than simply by letter.

The representativeness of the sample in the current research is outlined in table 27. There is a higher incidence of depression and anxiety in females aged between 25 and 40 years (Piccinelli & Wilkinson, 2000;

Somers, Coldner, Waraich, & Hsu, 2006; Kessler et al., 2007). This is similar to the incidence of MS which is more prominent in females of a similar age (Sadovnick, 2009). Therefore the demographic information gathered as part of the current study demonstrates that it was a representative sample.

Table 27

Epidemiological Findings and Study Sample

	Depression	Anxiety	MS	Study Sample
Age at onset	25 – 45 years	25 – 53 years	20 – 40 years	37 years (mean)
Gender	Female	Female	Female (3:1)	Females (3:1)

The gender difference in prevalence of anxiety and depression cannot be fully explained by the symptoms reported in the different genders (Piccinelli & Wilkinson, 2000), however the different symptoms reported in different genders may be relevant for the current study. When reporting depression, females report symptoms with appetite and sleep change, fatigue and somatic anxiety (Young, Scheftner, Fawcett, & Klerman, 1990). It is these symptoms (according to conceptual map, see figure one) that closely relate to the symptoms of MS. Therefore, having a high proportion of females with a diagnosis of MS, describing symptoms of depression that closely relate to the symptoms of MS, may partially explain the higher prevalence rates of depression in MS.

The type of MS in the sample is more complex. It is estimated that at diagnosis approximately 85% of people with MS are diagnosed as relapsing remitting, 50% of these will then go on to develop secondary progressive after some years and the remaining 15% are diagnosed with primary progressive MS at onset (Cook, 2005). Within the current sample, this was relatively similar with 17% of participants being diagnosed with primary progressive MS. Relapsing remitting MS was

the diagnosis in 58% of participants, a further 21% had secondary progressive MS which will have begun as relapsing remitting MS. Therefore, a total of 79% would have had a diagnosis of relapsing remitting MS at onset but only 36% have gone on to have a diagnosis of secondary progressive MS at the time the current study was conducted. It may be that some of the participants with relapsing remitting MS have yet to develop secondary progressive MS, or this has yet to be formally diagnosed for some participants (Cook, 2005).

Furthermore, the age at which the symptoms of depression begin is the time of the emergence of gender differences in social roles (Wilhem, Parker & Hadzi-Pavlovic, 1997). To understand the impact this has, it is important to note that although MS is diagnosed at a mean age of 30 years (Olek, 2005) as the diagnosis is a process of exclusion (Trojano & Paolicelli, 2001) it can take some time to achieve a diagnosis, which may explain the slightly later age of diagnosis within the current study (37 years). In addition, the similarity of age of diagnosis of multiple sclerosis, anxiety and depression (as summarised in table 28) may explain some of the increased co-morbidity within the disorders. The level of disability experienced by the participants was measured using the GNDS (mean score 17.92; range 3 – 38). Participants' mean level of disability was slightly higher to that reported in previous studies (e.g. 12.0, Stanton, Barnes & Sibler, 2006). The elevated level of disability may reflect both the propensity of people with MS to over-estimate their level of disability (Smith & Young, 2000) and difficulties in completing the GNDS. Although this tool was developed specifically for people with MS in reality it is a complex questionnaire to complete. A review of the completed questionnaires suggested that many participants appeared to have misunderstood the instructions and as such missed out items. As a result of this their scores were elevated as is suggested in the directions for scoring the questionnaires. Therefore the level the measures may not provide an accurate reflection of disability. One way to improve this in future research would be to provide clearer instructions for the measure or for a researcher to complete the

measure with the participant which would allow queries to be answered. Other demographic information gathered from participants suggested that it was a representative sample. For example the type of MS that participants reported was very similar to that reported in studies with much larger sample sizes considering people with MS (e.g. Confavreux & Vukosix, 2005).

In summary, the similarities in the demographic characteristics of the sample in terms of disability and type of MS imply that it is representative of the wider population of people with MS and therefore gives further weight to the generalisability of the findings.

4.2 Strengths and limitations

The results of the study demonstrated that the HADS and BDI-II were valid for use in people with MS whereas the BAI was not. This section will consider the details of the study, examining the strengths and limitations of the methodology and design.

4.2.1 Strengths.

Methodology. As discussed previously (see section 1.8) there are a number of ways in which to validate measures. This study compared measures to a recognised gold standard which is seen as a particularly robust methodology to assess the criterion validity (e.g. Pattern et al., 2003). This was further strengthened by the use of a research associate who scored the measures allowing the researcher to remain blind to the scores when conducting the interviews, reducing interpretation bias.

Contributing to knowledge. The study adds to the wider body of knowledge regarding MS and mood disorders. As is discussed previously this is particularly important because of the paucity of research regarding anxiety in MS (e.g. Honarmand & Feinstein, 2009). In adding to this knowledge base and providing valid measures for use in this population many of the problems outlined within the introduction are addressed. For example, by having valid measures of future research

can accurately assess the prevalence of depression and anxiety in people with MS and further research can be conducted into finding the most effective interventions for this population.

4.2.2 Limitations.

Recruitment and subsequent sample size. The sample size for the study was small, although it allowed enough power for the statistical tests to be meaningful. A larger sample size would have greater statistical power and thus greater confidence in results and generalisability of the findings. The study will be continued once the current thesis is completed by a member of the research team and thus the sample size will be increased. It will be interesting to see if these impacts on the results presented in this thesis.

Measuring validity and reliability. Although the study did find that some of the measures assessed were both valid and reliable for use with an MS population, not all types of reliability and validity were considered. Therefore although it can be concluded that, for example, the BDI-II has good criterion validity, the construct validity for the measures assessed is not known for this population.

Difficulties with sensitivity and specificity. Although the specificity and sensitivity of the measures was very high, for the majority of these it was not 100%. As a result this means that there is still a possibility that people who are experiencing symptoms of depression and anxiety will not be picked up by the measures. This implies that any screening that is conducted using the measures that have been validated should be done with caution (for further discussion of the impact of screening see section 4.5). However, it must be noted that the sensitivity and specificity of the measures was similar to that found when assessed for other client groups (e.g. BDI-II in people with brain injury suggested sensitivity of 87% and specificity of 79%, Homaifar et al., 2009)

4.3 Recommendations for future research

4.3.1 Screening tools. The validation of the measures for depression continues the process of developing them for screening as recommended for all people with long term health conditions (NICE, 2009). However this advice is in conflict with that given by the UK National Screening Committee (Albany, 2010). Anxiety was not recommended for screening by NICE (2009) and has not yet been considered by the National Screening Committee. The strict criteria of the NSC (NSC, 2003) mean that although screening can improve the detection of a disorder such as depression (Allaby, 2010) it also needs to improve the outcome (see section 4.5 for discussion of screening). Future research needs to be conducted assessing if the measures which have been validated in people with MS impact on the outcomes when employed as screening tools.

4.3.2 Replications. This study found that three of the measures were valid for use in people with MS. The cut off scores identified contrasted with previous studies that had used a similar methodology (e.g. Honarmand & Feinstein, 2009). This contrast and variation highlights the importance of replication within scientific research (Reiss & Judd, 2000) which although often recommended is rarely conducted (Reis & Judd, 2000). Therefore it is recommended that future research is conducted to repeat the study to assess reliability of findings and draw conclusions together.

4.4 Clinical implications

4.4.1 Access to treatment. It has been shown that both depression and anxiety within MS are treatable disorders (e.g. Mohr & Goodkin, 1999) and the impact of treatment of depression or anxiety in people with MS on the person's quality of life is significant (e.g. Lobentanz et al., 2004). Despite this access to treatment for depression or anxiety is very poor (e.g. Feinstein, 2002). If participants with MS were given the validated measures by healthcare professionals and researchers it is hoped that access to treatment would be improved.

However, some authors caution against using measures for screening tools that have not been designed for the purpose (e.g. BDI-II; Hagen, 2007) despite the precedent that has been set for using the measures this way (e.g. Lasa, Ayuso-Mateos, Vazquez-Barquero, Diez-Manrique & Dowrick, 2000; Leyfer et al., 2006). Further cause for caution and clinical judgement to be used in conjunction with the measures is the potential misidentification of individuals due to their imperfect levels of sensitivity and specificity (see section 4.2.2).

4.4.2 Assessment of prevalence. It is hoped that the prevalence of depression and anxiety within this population can be correctly assessed through the identification of appropriate and valid measures contributed by this study. In doing so, commissioners and others can potentially target research and resources if the prevalence of depression and anxiety in MS is found to be particularly high. Although this particular study may not provide an accurate estimate of prevalence due to its small sample size, it allows future studies to do so using measures that have been validated for the population.

4.4.3 Further research. Previous research using the measures validated in the current study has utilised a range of cut off scores even when research has been completed by the same authors. For example, Feinstein et al. (1999) used the HADS with a cut off score of 11 but Korostil & Feinstein (2007) used the same measure with a cut off score of 10 with no justification in either paper for this choice. Therefore, in having the measures validated for use in people with MS the measures can be used in research employing the same valid cut off scores, and thus comparisons can be made between different studies more easily and conclusions generalised. This would allow specific research to be conducted in areas such as comparing treatments for depression and anxiety in people with MS and factors influencing the relationships between depression, anxiety and MS such as disability (Tsivgoulis et al., 2007) and social support (Beckner et al., 2010).

Further research has already begun with a pilot study linked to the current study. Participants from the current study were asked to participate in a study considering the MRI scans depression in people with MS (R, Dineen, personal communication, 9th August 2010).

4.5 Critical reflection

This study aimed to validate mood measures in the MS population. Positivism aims to build up knowledge of phenomena through systematic observation, and then often uses logic to draw inferences or theories about phenomena from these observations. In order to work it assumes phenomena are held stable, so that they can be observed systematically, and that observations are objective and free from bias (Hesse-Biber & Leayy, 2010). The procedures of psychological science have been developed in order to achieve these requirements of holding phenomena stable, so they can be observed (Gliner & Morgan, 2000). Thus, in the current study the constructs of anxiety and depression are assumed to be stable and there are laws governing them, therefore providing appropriate procedures are followed the results of the study can be generalised (Gliner & Morgan, 2000).

A critical realist epistemological position is sometimes contrasted with a positivist position in quantitative research (e.g. Garner, Wagner & Kawulich, 2009), possibly because the two positions share many characteristics. Critical realism assumes that knowledge about phenomena can be generated by observation and making true observations can be technically very difficult (Sayer, 2000). In contrast to positivism, it is sceptical about whether objective observations can ever be made, even if scientific procedure is followed. It is also less sure that phenomena, especially social phenomena can be stable enough to be observed (Parker, 1999, 2000).

These two positions lead to different assumptions being made in the current study. The positivist position suggests that the constructs of anxiety and depression are valid and can be reliably measured by the SCAN diagnostic interview. The critical realist approach would state that

the constructs are valid but cannot be reliably measured by the SCAN alone.

From a positivist stance, the researcher aims to be independent of what is being researched; they are objective observers (Hesse-Biber & Leavy, 2010). To achieve this, in the current study, attempts have been made to control possible observer bias. For example, the researcher was blind to the participants' responses on measures when conducting the SCAN interviews. In addition, the training for the SCAN interview included tests of inter-rater reliability. Thus, it can be assumed, from the positivist position, that should another trained individual complete the interviews undertaken by the researcher, they would achieve similar results. In contrast, the critical realist position would suggest that the inter-rater reliability went some way to demonstrate that the SCAN was successful in measuring the truth of the constructs of depression and anxiety. However, due to its' limitations ,the truth of these constructs is still unknown.

As the current study wishes to find the underlying truth and knowledge it is important that the debate over accessing the truth is resolved. The critical realist position poses an interesting argument but does not allow for the study to draw conclusions that can be generalised. For this to occur, a positivist epistemology was used and as a result classical test theory was applied. This allows the results of the study to be generalised to the population from which the sample was taken (Kline, 2000). In the current study, the results can be applied to individuals with a diagnosis of MS.

The positivist epistemology of the study assumes the evidence for anxiety and depression is interpreted as true descriptions of the construct; there is empirical evidence demonstrating that they are both stable and measurable (using a structured scientific interview). The methodology then allowed different assessment measures of anxiety and depression to be assessed for validity against the structured interview. Furthermore, a robust methodology allows the results of the

study to be generalised to the wider population. However, within the positivist epistemology of the study, some difficulties need to be acknowledged regarding the measurements used; firstly the imperfection of the gold standard interview in mental health and secondly the use of self report measures.

The assumption is made that the constructs of depression and anxiety were valid and could be reliably measured using the SCAN interview as a gold standard. However, it is possible that the gold standard is imperfect (Zou et al., 2007). Although the SCAN has been shown to be reliable (Rijners et al., 2000), its validity in comparison to gold standards utilised in physical health is very poor (e.g. glucose test for diabetes; International Expert Committee, 2009). In addition, its validity in comparison to alternative gold standards measuring anxiety and depression is also poor (Brugha et al., 2001).

The study has also utilised self report measures, which are cited as being the “most widely used measurement tools in psychology [and] also among the most criticised” (Haffel & Howard, 2010 pp.181). An inherent difficulty with self report from a positivist position is the assumption that they suggest an accurate reporting of an underlying truth. When considering internal states measurement becomes difficult as there are no external references and it is assumed that the truth is a fixed point, that is, if someone else interpreted the same data at the same time they would reach the same conclusion.

However, individuals completing self report measures are susceptible to demand characteristics and potential bias as they are unable to accurately observe their own cognitive processes (Nisbett & Wilson, 1977), a problem not found in behavioural or biological measures (Haffel & Howard, 2010). Psychometric theory has attempted to address these inherent difficulties within measurement. Key debates such as classical test theory and item response theory are discussed elsewhere (see section 1.8).

When considering the validity of self report measures there are few studies where both self and clinician ratings are compared to a third source or a biological measure, for example a MRI scan (Joiner et al., 2005). Instead studies may measure both self-report ratings and clinician ratings for the same difficulty in the same participant. Some individual studies have reported agreement between self-report and clinician ratings (e.g. Kaplan et al., 1994; Hopko et al., 2000) but this finding was not replicated in a recent meta-analysis (Cuijpers, Hofmann & Andersson, 2010).

Taking this literature into account it may be possible that using a self report measure led people to overestimate their difficulties, something which is not acknowledged within a positivist epistemology. This has been shown to be the case for people with MS (e.g. Smith & Young, 2000), depression (Corruble, Legrand, Zvenigorowski, Duret & Guelfi, 1999) and anxiety (e.g. Higginson et al., 2001). It is hoped these effects are counterbalanced by the newly suggested cut off scores for the measures.

Within this study, the positivist position would also imply that the results can be generalised (Gliner & Morgan, 2000). Therefore, since the BDI-II, BAI and HADS have been validated they can then be used as screening tools for depression and anxiety in those with a diagnosis of MS, and so the aim of the study was achieved. However, if screening tools are to be used then this should not be done without acknowledgement of the potential ethical issues that may arise.

Much of the criticism of screening measures lies in the danger of identifying false positives (Walker et al., 2007). These concerns demonstrate the stigma that continues to surround mental health difficulties (Gilbody, Sheldon & Wessely, 2009). The impact of misidentifying participants and suggesting someone has a diagnosis of depression and/or anxiety may be more pertinent to individuals with MS. Qualitative studies have shown that communication and information given to people when receiving their MS diagnosis is poor

(e.g. Solari et al., 2007; see Solari et al., 2010 for an extension of this study into developing a questionnaire). For people with MS who may have already received a diagnosis with little support, receiving more diagnoses via screening measures may be inappropriate.

Systematic reviews have been conducted in an attempt to objectively assess the value of screening measures; however, these have resulted in conflicting conclusions. For example when considering screening for depression in primary care Pignone et al. (2002) found that screening can lead to improved patient outcomes but Gilbody, House and Sheldon (2005) found screening had little impact on the management or outcomes of depression. More recent systematic reviews have provided more detail, for example O'Connor, Whitlock, Beil and Gaynes (2009) concluded from their review that screening programmes' impact on improvement of depression outcomes was moderated by the involvement of staff.

A further consideration is the uptake of screening programmes. Although rarely discussed explicitly in the literature, uptakes for screening of mental health problems in healthcare settings are low (30-60%; Gilbody et al., 2005). It may be that patients do not want to be screened. This has been shown in a qualitative study completed by Wittampf et al. (2008) who reported that some patients found screening programmes aversive, particularly if they had acquired a diagnosis of depression through the screening programme they had previously undisclosed.

Screening for depression improves patient outcomes in people with physical illness only when accompanied by effective treatment and follow up. Implementation of wide spread depression screening in medically ill patients would be a costly process that will not benefit patients if sufficient resources are not made available to ensure parity, accessibility, appropriate delivery and correct monitoring of treatment (Evans et al. 2005). Therefore, any screening should be conducted with

caution and as stated by Eaton et al. (2000) should utilise the support of well trained and knowledgeable staff.

4.6 Conclusions

The study found that the BDI-II and the HADS were valid measures to use with people with MS. The strength of the study lies in the clinical implications of this. By having valid measures for the assessment of depression and anxiety it is hoped that difficulties shown in the literature can be addressed. Those with MS can be screened for depression and anxiety using the measures (in conjunction with clinical judgment) and thus may have greater access to treatment. Valid measures will allow the accurate assessment of prevalence of depression and anxiety in MS which in turn may help towards commissioning decisions based around the targeting of resources. Finally, valid measures will allow further research to be conducted on the effectiveness of different treatments for depression and anxiety in people with MS. It is hoped that this will improve the quality of life for people with MS experiencing depression and/or anxiety.

REFERENCES

- Aben, I., Verhey, F., Lousberg, R., Lodder, J., & Honig, A. (2002).
Validity of the Beck depression inventory, Hospital anxiety and
depression scale, SCL-90, and Hamilton depression rating scale
as screening instruments for depression in stroke patients.
Psychosomatics, 43, 386 – 393.
- Aikens, J.E., Reinecke, M.A., Pilskin, N.H., Fischer, J.S., Wiebe, J.S.,
McCracken, L.M., & Taylor, J.L. (1999). Assessing depressive
symptoms in multiple sclerosis: Is it necessary to omit items from
the original Beck depression inventory? *Journal of Behavioural
Medicine*, 22(2), 127 – 142.
- Allaby, M. (2010). *Revised report: Screening for depression. A report for
the UK national screening committee*. London: Public Health
Resources Unit.
- Alloway, R., & Bebbington, P. (1987). The buffer theory of social
support: A review of the literature. *Psychological Medicine*, 17,
91 – 108.
- Ameringer, K.J., & Leventhal, A.M. (2010). Applying the tripartite model
of anxiety and depression to cigarette smoking: An integrative
review. *Nicotine and Tobacco Research*, 12(12), 1183 – 1194.
- Andrews, G., Peters, L., Guzman, A., & Bird, K. (1995). A comparison
of two structured diagnostic interviews: CIDI and SCAN.

Australian and New Zealand Journal of Psychiatry, 29, 124 – 132.

Andrews, G., Slade, T., & Peters, L. (1999). Classification in psychiatry: ICD-10 versus DSM-IV. *British Journal of Psychiatry*, 174, 3 – 5.

American Psychiatric Association (APA). (1987). *Diagnosis and statistical manual of the mental disorders: DSM-III-R*. Washington DC: American Psychiatric Association.

American Psychiatric Association (APA). (2000) *Diagnosis and statistical manual of the mental disorders: Fourth edition – text revision (DSM-IV-TR)*. Washington DC: American Psychiatric Association.

Arnau, R.C., Meagher, M.W., Norris, M.P., & Bramson, R. (2001). Psychometric evaluation of the Beck Depression Inventory with primary care medical patients. *Health Psychology*, 20(2), 112 – 119.

Arnett, P. A., Barwick, F. H., & Beeney, J. E. (2008). Depression in multiple sclerosis: Review and theoretical proposal. *Journal of the International Neuropsychological Society*, 14(5), 691 – 724.

Arnett, P.A., & Randolph, J.J. (2006). Longitudinal course of depression symptoms in multiple sclerosis. *Journal of Neurology, Neurosurgery and Psychiatry*, 77, 606 – 610.

Arroll, B., Macgillivray, S., Ogston, S., Reid, I., Sullivan, F., Williams, B. & Crombie, I. (2005). Efficacy and tolerability of tricyclic

antidepressants and SSRIs compared with placebo treatment for depression in primary care: A meta-analysis. *Annals of Family Medicine*, 2, 449 – 456.

Barbato, A., & D'Avanzo, B. (2008). Efficacy of couple therapy as a treatment for depression: A meta-analysis. *Psychiatric Quarterly*, 79(2), 121 – 132.

Barmer, A.M., Cetin, K., Johnson, K.L., Gibbions, L.E., & Ehde, D.M. (2008). Validation study of prevalence and correlates of depressive symptomatology in multiple sclerosis. *General Hospital Psychiatry*, 30(4), 311 – 317.

Baruch, Y. (1998). Response rate in academic studies: A comparative analysis. *Human Relations*, 53(4), 421 – 438.

Bech, P. (2000). Clinical features of mood disorders and mania. In M.Gelder, J.Lopez-Ibor & N. Andreasen (Eds.), *New Oxford textbook of psychiatry*, pp. 682 – 688). Oxford: Oxford University Press.

Bech, P., Rasmussen, N.A., Olsen, L.R., Noerholm, V., & Abildgaard, W. (2001). The sensitivity and specificity of the Major Depression Inventory using the Present State Examination as the index of diagnostic validity. *Journal of Affective Disorders*, 66, 159 – 164.

Beck, A.T. (2008). The evolution of the cognitive model of depression and its' neurobiological correlates. *American Journal of Psychiatry*, 165, 969 – 977.

- Beck, A.T., & Alford, B.A. (2009). *Depression: Causes and treatment. Second edition*. Pennsylvania: University of Pennsylvania Press.
- Beck, A.T., Epstein, N., Brown, G., & Steer, R.A. (1986). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of consulting and clinical psychology*, 56(6), 893 – 897.
- Beck, A.T., Kovacs, M. & Weissman, A. (1979). The assessment of suicidal intention: The Scale for Suicide Ideation. *Journal of Consulting and Clinical Psychology*, 47(2), 343 – 352.
- Beck, A.T., & Steer, R.A. (1990). *Beck Anxiety Inventory (BAI)*. Oxford: Pearson.
- Beck, A.T., & Steer, R.A. (1991). Relationship between the Beck Anxiety Inventory and the Hamilton Anxiety Rating Scale with anxious outpatients. *Journal of Anxiety Disorders*, 5, 213 – 223.
- Beck, A.T., Steer, R.A., Brown, G.A. (1996). *Beck Depression Inventory – II (BDI-II)*. Oxford: Pearson.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561 – 571.
- Beckner, V., Howard, I., Vella, L., & Mohr, D.C. (2010). Telephone administered psychotherapy for depression in MS patients: moderating role of social support. *Journal of Behavioural Medicine*, 33, 47 – 59.

- Beiske, A.G., Svensson, E., Sandanger, I., Czujiko, B., Pedersen, E.D., Aarseth, J.H., & Myhr, K.M. (2008). Depression and anxiety amongst multiple sclerosis patients. *Journal of Neurology*, 15, 239 – 245.
- Berg, A., Lonnqvist, J., Palmake, H., Kaste, M. (2009). Assessment of depression after stroke: a comparison of different screening instruments. *Stroke*, 40, 523 – 529.
- Bjelland, I., Dahl, A.A., Haug, T.T., & Necklemann, D. (2002). The validity of the hospital anxiety and depression scale: An updated literature review. *Journal of Psychosomatic Research*, 52, 69 – 77.
- Boland, R.J., & Keller, M.B. (2002). The course of depression. In K.L. Davis, D. Charney, J.T. Coyle & C. Nemeroff (Eds.) *Neuropsychopharmacology: The fifth generation of progress* (pp.1009-1015). Washington: Lippincott Williams & Wilkins.
- Bonita, R., Beaglehole, R., & Kjellstrom, T. (2006). *Basic epidemiology. Second edition*. Geneva: World Health Organisation.
- Boyle, M.H. (1998). Guidelines for evaluating prevalence studies. *Evidence-based Mental Health*, 1(2), 37 – 39.
- Bradshaw, J., & Rose, A. (2008). Cognition, depression and fatigue in multiple sclerosis. *Advances in clinical neurosciences and rehabilitation*, 8(4), 15 – 17.

- Brassington, J.C., & Marsh, N.V. (1998). Neuropsychological aspects of multiple sclerosis. *Neuropsychology Review*, 8, 43 – 77.
- Brennan, C., Worrall-Davies, A., McMillan, D., Gilbody, S., & House, A. (2010). The Hospital Anxiety and Depression Scale: A diagnostic meta- analysis of case-finding ability. *Journal of Psychosomatic Research*, 69, 371 – 378.
- British Psychological Society. (2009). *Code of ethics and conduct: Guidance published by the ethics committee of the British Psychological Society*. Leicester: British Psychological Society.
- Brown, T.A., & Barlow, D.H. (1992). Co-morbidity amongst anxiety disorders: implications for treatment and DSM-IV. *Journal of Consulting and Clinical Psychology*, 60, 835 – 844.
- Brugha, T.S., Bebbington, P.E., & Jenkins, R. (1999). A difference that matters: comparisons of structured and semi-structured psychiatric diagnostic interviews in the general population. *Psychological Medicine*, 29, 1013 – 1020.
- Brugha, T.S., Jenkins, R., Taub, N., Meltzer, H., & Bebbington, P.E. (2001). A general population comparison of the Composite International Diagnostic Interview (CIDI) and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN). *Psychological Medicine*, 31, 1001 – 1013.
- Brugha, T.S., Neinhuis, F., Bagghi, D., Smith, J., & Meltzer, H. (1999). The survey form of SCAN: the feasibility of using experienced

lay survey interviewers to administer a semi-structured systematic clinical assessment of psychotic and non-psychotic disorders. *Psychological Medicine*, 29, 703 – 711.

Buckby, J.A., Cotton, S.M., Cosgrave, E.M., Killackey, E.J., & Yung, A.R. (2008). A factor analytic investigation of the tripartite model of affect in a clinical sample of young Australians. *BMC Psychiatry*, 8, 79 – 83.

Burgess, M. (2003). *Multiple Sclerosis: Theory and practice for nurses*. London: Whurr Publishers.

Butler, A.C., Chapman, J.E., Forman, E.M., & Beck, A.T. (2006). The empirical status of cognitive behavioural therapy: A review of meta-analyses. *Clinical Psychology Review*, 26(1), 17 – 31.

Carnap, R. (1962). *Logical foundations of probability*. Chicago: University of Chicago Press.

Carr, A., & McNulty, M. (2006). *The handbook of adult clinical psychology*. East Sussex: Routledge.

Chatburn, R.L. (2009). *Handbook for health care research. Second edition*. London: Jones and Barlett Publishers.

Chorpita, B.F. (2002). The tripartite model and dimensions of anxiety and depression: An examination of structure in a large school sample. *Journal of Abnormal Clinical Psychology*, 30(2), 177 – 190.

- Chwastiak, L., Ehde, D.M., Gibbons, L.E., Sullivan, M., Bowen, J.D., & Kraft, G.H. (2002). Depressive symptoms and severity of illness in multiple sclerosis: Epidemiologic study of a large community sample. *American Journal of Psychiatry*, 159, 1862 – 1868.
- Chwastiak, L.A., Gibbons, L.E., Ehde, D.M., Sullivan, M., Bowen, J.D., Bombardier, C.H., & Kraft, G.H. (2005). Fatigue and psychiatric illness in a large community sample of people with multiple sclerosis. *Journal of Psychosomatic Research*, 59, 291 – 298.
- Ciechanowski, P., Katon, W., Russo, J., & Walker, E. (2001). The patient provider relationship: Attachment theory and adherence to treatment in diabetes. *American Journal for Psychiatry*, 158, 29 – 35.
- Clark, L.A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, 100, 316 - 336.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155 – 159.
- Confavreuz, C, & Vukusic, S. (2005). Natural history of multiple sclerosis: A unifying concept. *Brain*, 129(3), 606 – 619.
- Connor, K.M., Davidson, J.R.T., Churchill, K.E., Sherwod, A., & Wisler, R.H. (2000). Psychometric properties of the Social Phobia

Inventory: New self rating scale. *British Journal of Psychiatry*, 176, 379 – 386.

Cook, J.M., Orvaschel, H., Simco, E., Hersen, M., & Joiner, T. (2004). A test of the tripartite model of depression and anxiety in older adult psychiatric outpatients. *Psychology and Aging*, 19(3), 444 – 451.

Cook, S.D. (2005). *Handbook of multiple sclerosis. Third edition*. New York: Taylor & Fancis.

Corruble, E., Legrand, J.M., Zvenigorowski, H., Duret, C., & Guelfi, J.D. (1999). Concordance between self-report and clinician's assessment of depression. *Journal of Psychiatric Research*, 33(5), 457 – 465.

Crawford, J.R., Henry, J.D., Crombie, C., & Taylor, E.P. (2001). Brief report normative data for the HADS from a large non-clinical sample. *British Journal of Clinical Psychology*, 40, 429 – 434.

Cronbach, L.E., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281 – 302.

Cuijpers, P., Hofmann, L.J., & Andersson, G. (2010). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. *Clinical Psychology Review*, 30(6), 768 – 778.

Cuijpers, P., Smit, F., Bohlmeijer, E., Hollon, S.D., & Andersson, G. (2010). Efficacy of cognitive behavioural therapy and other

psychological treatments for adult depression: A meta-analytic study of publication bias. *British Journal of Psychiatry*, 196(3), 173 – 178.

Cuijpers, P., Straten, A., Andersson, G., & van Oppen, P. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology*, 76(6), 909 – 922.

Dagnan, D., Jahoda, A., McDowell, K., Masson, J., Banks, P., & Hare, D. (2008). The psychometric properties of the hospital anxiety and depression scale adapted for use with people with intellectual disabilities. *Journal of Intellectual Disability Research*, 52(11), 942 – 949.

Dahl, O.P., Stordal, E., Lydersen, S., & Midgard, R. (2009). Anxiety and depression in multiple sclerosis: A comparative population-based study in Nord-Trøndelag County, Norway. *Multiple Sclerosis*, 15(12), 1495 – 1501.

De Ayala, R.J., Vonderharr-Carlson, D.J., & Kim, D. (2005). Assessing the reliability of the Beck anxiety inventory scores. *Educational and psychological measurement*, 65, 742 – 756.

De Beurs, E., Wilson, K.A., Chambless, D.L., Goldstein, A.J., & Feske, U. (1997). Convergent and divergent validity of the Beck Anxiety Inventory for patients with panic disorder and agoraphobia. *Depression and Anxiety*, 6(4), 140 – 146.

- De Souza, J., Jones, L.A., & Rickards, H. (2009). Validation of self-report depression rating scales in Huntington's disease. *Movement disorders*, 25(1), 91 – 96.
- Derogatis, L.R., Lipman, R.S., & Covi, L. (1973) SCL-90: an outpatient psychiatric rating scale: Preliminary report. *Psychopharmacology Bulletin*, 9, 13 – 28.
- Dozois, D.J.A., & Covin, R. (2004). The beck depression inventory-II (BDI-II), Beck hopelessness scale (BHS), and Beck scale for suicide ideation (BSS). In *Comprehensive handbook of psychological assessment. Volume 2: Personality assessment*. M.J. Hilsenroth & D.L. Segal (Eds.) Pp.50 – 69. New Jersey: John Wiley and Sons Inc.
- Durss. B., Rosenheck, A., Desai, M., & Perlin, J. (2002). Quality of preventive medical care for patients with mental disorders. *Medical Care*, 40, 129 – 136.
- Eack, S.M., Singer, J.B., & Greeno, C.G. (2008). Screening for anxiety and depression in community mental health: The Beck Anxiety and Depression Inventories. *Community Mental Health*, 44, 465 – 474.
- Eaton, W.W., Nuefield, K., Chen, L., & Cai, G. (2000). A comparison of self-report and clinical diagnostic interviews for depression. *Archives of General Psychiatry*, 57, 217 – 222.

- Elliott, D. (1993) Comparison of three instruments for measuring patient anxiety in a coronary care unit. *Intensive and Critical Care Nursing*, 9, 195 – 200.
- Embretson, S.W., & Reise, S.P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates Inc.
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J. & Audin, K. (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51 – 60.
- Evans, D.L., Charney, D.S., Lewis, L., Golden, R.N., Gorman, J.M., Krishnan, K.R.R. et al. (2005). Mood disorders in the medically ill: Scientific review and recommendations. *Biological Psychiatry*, 58(3), 175 – 189.
- Faul, F., Erdfelder, E., Lang, A-G., & Buchner, A. (2000). Statistical power analyses using G*Power 3.1: Tests for correlations and regression analyses. *Behaviour Research Methods*, 41, 1149 – 1160.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861 – 874.
- Fawcett, J. (2009). Reports of the DSM-V mood disorders work group. *American Psychiatric Association*. Arlington, Virginia: APA.
- Feinstein, A. (2002). An examination of suicidal intent in patients with multiple sclerosis. *Neurology*, 59, 674 – 678.

- Feinstein, A., Feinstein, K., Gray, T., & O'Connor, P. (1997). Prevalence and neurobehavioral correlates of pathological laughing and crying in multiple sclerosis. *Archives of neurology*, 54(9), 1116 – 1121.
- Feinstein, A., O'Connor, P., & Feinstein, K. (2002a). Multiple sclerosis, interferon beta 1-b and depression: a prospective investigation. *Journal of Neurology*, 249, 815 – 820.
- Feinstein, A., O'Connor, P., Gray, T., & Feinstein, K. (1999). The effects of anxiety on psychiatric morbidity in patients with multiple sclerosis. *Multiple Sclerosis*, 5, 323 – 326.
- Feinstein, A., Roy, P., Lobaugh, N., Feinstein, M.A., O'Connor, P., & Black, S. (2002b). Structural brain abnormalities in multiple sclerosis patients with major depression. *Neurology*, 62, 586 – 90.
- Field, A. (2005). *Discovering statistics using SPSS: Second edition*. London: SAGE publications.
- First, M.B., Spitzer, R.L., Gibbon, M., & Williams, J.B. (1995). *Structured clinical interview for DSM-IV Axis I disorders – patient edition (SCID-IP, Version 2.0)*. New York State Psychiatric Institute.
- Fiske, A., Wetherell, J.L., & Gatz, M. (2009). Depression in older adults. *Annual Review in Clinical Psychology*, 5, 363 – 389.

- Fletcher, R.H., & Fletcher, S.W. (2005). *Clinical epidemiology: the essentials. Fourth edition*. Baltimore: Lippincott, Williams & Wilkins.
- Forbes, R.B., Wilson, S.V., & Swingler, R.J. (1999). The prevalence of multiple sclerosis in Tayside, Scotland: do latitudinal gradients really exist? *Journal of Neurology*, 246, 1033 – 1040.
- Forman, A.C., & Lincoln, N.B. (2010). Evaluation of an adjustment group for people with multiple sclerosis: A randomised controlled trial. *Clinical Rehabilitation*, 24(3), 211 – 221.
- Fraley, R.C., Waller, N.G., & Brennan, K.A. (2000). An item response theory analysis of self report measures of adult attachment. *Journal of Personality and Social Psychology*, 78(2), 350 – 365.
- Fydrich, T., Dowdall, D., & Chambless, D.L. (1992). Reliability and validity of the Beck anxiety inventory. *Journal of Anxiety Disorders*, 6, 55 – 61.
- G*Power3. (2010). Retrieved from:
<http://www.psych.uni-duesseldorf.de/abteilungen/aap/gpower3>
- Gallant, M.P. (2003). The influence of social support on chronic illness self management: A review and directions for research. *Health Education and Behaviour*, 30(2), 170 – 195.
- Galeazzi, G.M., Ferrari, S., Giaroli, G., Mackinnon, A., Merellie, E., Motti, L., & Rigatelli, M. (2005). Psychiatric disorders and depression in multiple sclerosis outpatients: Impact of disability

and interferon beta therapy. *Neurological Sciences*, 26(4), 255 – 262.

Garner, M., Wagner, C., & Kawulich, B. (2009). *Research Methods in the Social Sciences*. London: Ashgate

Gaynes, B.N., Rush, J.A., Trivedi, M.H., Wisniewski, S.R., Balsaubramani, G.K., Spencer, D.C., Petersen, T., Klinkman, M., Warden, D., Nicholas, L., & Fava, M. (2007). Major depression in primary care and psychiatric care settings: A cross sectional analysis. *Annals of Family Medicine*, 5, 126 – 134.

Gilbody, S., House, A., & Sheldon, T. (2005). Screening and case finding instruments for depression. *Cochrane Database of Systematic Reviews*. 4, Article Number: CD002792.

Gilbody, S., Sheldon, T., & Wessely, S. (2009). Health policy: Should we screen for depression? *British Medical Journal*, 332, 1027 – 1030.

Gliner, J.A., & Morgan, G.A. (2000). *Research methods in applied settings: An integrated approach to design and analysis*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Gold-Spink, E., Sher, T.A., & Theodos, V. (2000). Uncertainty in illness and optimism in couples with multiple sclerosis. *International Journal of Rehabilitation and Health*, 5(3), 157 – 164.

Gottberg, K., Einarsson, U., Fredrikson, S., von Koch, L., & Holmqvist, L.W. (2007). A population based study of depressive symptoms

in multiple sclerosis in Stockholm county: association with functioning and sense of coherence. *Journal of Neurology, Neurosurgery and Psychiatry*, 78(1), 60 – 65.

Greaves-Lord, K., Ferdinand, R.F., Sondejker, F.E.P.L., Dietrich, A., Oldehinkel, A.J., Rosmalen, J.G.M., Ormel, J., & Verhulst, F.C. (2007). Testing the tripartite model in young adolescents: Is hyperarousal specific for anxiety and not depression? *Journal of Affective Disorders*, 102(1), 55 – 63.

Haffell, G.J., & Howard, G.S. (2010). Self report: Psychology's four letter word. *American Journal of Psychology*, 123(2), 181 – 188.

Hagen, B. (2007). Measuring melancholy: A critique of the Beck Depression Inventory and its use in mental health nursing. *International Journal of Mental Health Nursing*, 16(2), 108 – 115.

Hambelton, R.K., & Jones, R.W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, 12(3), 535 – 566.

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology*, 32, 50 – 55.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry*, 23, 56 – 62.

- Harris, C.A., & D'Eon, J.L. (2008). Psychometric properties of the Beck depression inventory-second edition (BDI-II) in individuals with chronic pain. *Pain*, 127, 609 – 622.
- Haslam, N. (2003). Categorical versus dimensional models of mental disorder: the taxometric evidence. *Australian and New Zealand Journal of Psychiatry*, 37, 696 – 704.
- Hawton, K., Mayou, R., & Feldman, E. (1990). Significance of psychiatric symptoms in the general medical patient with mood disorder. *General Hospital Psychiatry*, 12, 296 – 302.
- Herrmann, C. (1996). International experiences with the Hospital Anxiety and Depression Scale: A review of validation data and clinical results. *Journal of Psychosomatic Research*, 42(1), 17 – 41.
- Hesse-Biber, S.N., & Leavy, P. (2010). *The practice of qualitative research: Second edition*. London, UK: SAGE Publications.
- Higginson, C.I., Fields, J.A., Koller, W.C., & Troster, A.I. (2001). Questionnaire assessment potentially overestimates anxiety in Parkinson's disease. *Journal of Clinical Psychology in Medical Settings*, 8(2), 95 – 99.
- Hinton, P.R., Brownlow, C., McMurray, I., & Cozens, B. (2004). *SPSS explained*. Hove, East Sussex: Routledge.

- Hofmann, S.G. & Smits, J.A.J. (2008). Cognitive behavioural therapy for adult anxiety disorders: A meta-analysis of randomised placebo controlled trials. *Journal of Clinical Psychiatry*, 69(4), 621 – 632.
- Homaifar, B.Y., Brenner, L.A., Gutierrez, P.M., Harwood, J.F., Thompson, C., Filley, C.M., Kelly, J.P., & Adler, L.E. (2006). Sensitivity and specificity of the Beck Depression Inventory-II in persons with traumatic brain injury. *Archives in physical medicine and rehabilitation*, 90(4), 652 – 656.
- Honarmand, K., & Feinstein, A. (2009). Validation of the Hospital Anxiety and Depression Scale for use with multiple sclerosis patients. *Multiple Sclerosis*, 15(12), 1518 – 1524.
- Hoogervorst, E.L.J., Kalkers, N.F., Uitdehaag, B.M.J., & Polman, C.H. (2002). A study validating changes in the Multiple Sclerosis Functional Composite. *Archives of Neurology*, 59, 113 – 116.
- Hopko, D.R., Bourland, S.L., Stanley, M.A., Beck, G., Novy, D.M., Averill, P.M., & Swann, A.C. (2000). Generalised anxiety disorder in older adults: Examining the relation between clinician severity ratings and patient self-report measures. *Depression and Anxiety*, 12(4), 217 – 225.
- International Expert Committee. (2009). International expert committee report on the role of the A1c assay in the diagnosis of diabetes. *Diabetes Care*, 32, 1327 – 1334.
- Jacobs, I.D., Beck, R.W., Simon, J.H., Kinkel, R.P., Brownschidle, C.M., Murray, T.J., Simonain, N.A., Slasor, P.J., & Sandrock,

- A.W. (2000). Intramuscular interferon beta-1a therapy Initiated during a first demyelinating event in multiple sclerosis. *The New England Journal of Medicine*, 343 (13), 898 – 904.
- Janssens, A.C.J.W., van Doorn, P.A., de Boer, J.B., van de Meche, F.G.A., Passchier, J., Hintzen, R.Q. (2003). Impact of recently diagnosed multiple sclerosis on quality of life, anxiety, depression and distress of patients and partners. *Acta Neurologica Scandinavica*, 108, 389 – 395.
- Jared, B., Hancock, L., Arnett, P., & Lynch, S. (2010). Treatment adherence in multiple sclerosis: Association with emotional status, personality and cognition. *Journal of Behavioural Medicine*, 33(3), 219 – 227.
- Joiner, T.E., Walker, R.L., Pettit, J.W., Perez, M., & Cukrowicz, K.C. (2005). Evidence-based assessment of depression in adults. *Psychological Assessment*, 17(3), 267 – 277.
- Jose Sa, M. (2008). Psychological aspects of multiple sclerosis. *Clinical Neurology and Neurosurgery*, 110, 868-877.
- Kaplan, M.L., Asnis, G.M., Sanderson, W.C., Keswani, L., de Lecuona, J.M., & Joseph, S. (1994). Suicide assessment: clinical interview versus self-report. *Journal of Clinical Psychology*, 50(2), 294 – 298.

- Kaplan, R.M., & Saccuzzo, D.P. (2009). *Psychological testing: Principles, applications and issues. Seventh edition.* Belmont, CA, USA: Wadsworth College Learning.
- Kasper, S. (2001). Depression and anxiety: Separate or continuum? *World Journal of Biological Psychiatry*, 2, 162 – 163.
- Katon, W.J. (2003). Clinical and health services relationships between major depression, depressive symptoms and general medical illness. *Biological Psychiatry*, 54, 216 – 226.
- Katz, M., & Klerman, G.L. (1979). Overview of the clinical studies programme. *American Journal of Psychiatry*, 136, 49 – 51.
- Kauhanen, M.L., Korpelainen, J.T., Hiltunen, P., Brusin, E., Mononen, H., Maatta, R., Nieminen, P., Sotaniemi, K.A., Myllyla, V.V. (1999). Poststroke depression correlates with cognitive impairment and neurological deficits. *Stroke*, 20, 1875 – 1880.
- Kendler, K., Gardner, C., & Prescott, C. (2002). Toward a comprehensive developmental model for major depression in women. *American Journal of Psychiatry*, 159, 1133 – 1145.
- Kessler, R.C., Amminger, G.P., Aguilar-Gaxiola, S., Alonson, J., Lee, S., & Ustun, T.B. (2007). Age of onset of mental disorders: a review of recent literature. *Current Opinion In Psychiatry*, 20(4), 359 – 364.
- Kessler, R.C., Berglund, P., Demler, O., Jin, R., Merikangas, K.R., & Walters, E.E. (2005). Lifetime prevalence and age of onset

distributions of DSM-IV disorders in the National Co-morbidity Survey replication. *Archives in General Psychiatry*, 62, 593 – 602.

Kessler, R.C., Merikangas, K.R., Berglund, P., Eaton, W.W., Korte, D.S., & Walters, E.E. (2003). Mild disorders should not be eliminated from the DSM-V. *Archives in General Psychiatry*, 60, 1117 – 1122.

Kline, P. (2000). *Handbook of psychological testing: Second edition*. London: Routledge.

Korostil, M. & Feinstein, A. (2007). Anxiety disorders and their clinical correlates in multiple sclerosis patients. *Multiple Sclerosis*, 13, 67 – 72.

Kreoneke, K., Spitzer, R.K., & Williams, J.B. (2003). The patient health questionnaire II: validity of a two item depression screener. *Medical Care*, 41, 1284 – 1292.

Krupp, L.B., LaRocca, N.G., Muir-Nash, J., & Steinberg, A.D. (1989). The fatigue severity scale: Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology*, 46(10), 1121 – 1123.

Kutzke, J.F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*, 33, 1444 – 1452.

Lasa, L., Ayuso-Mateos, J.L., Vazquez-Barquero, J.L., Diez0Manrique, F.J., & Dowrick, C.F. (2000). The use of the Beck Depression Inventory to screen for depression in the general population: a preliminary analysis. *Journal of Affective Disorders*, 57(1-2), 261 – 265.

Layard, R. (2006). Health policy: the case for psychological treatment centres. *British Medical Journal*, 332, 1030 – 1032.

Leech, N.L., Barrett, K.C., Morgan, G.A. (2005). *SPSS for intermediate statistics: Use and interpretation. Second edition*. New Jersey: Lawrence Erlbaum Associates.

Leeman, E. (1998). Misuse of psychiatric epidemiology. *Lancet*, 351(9116), 1601 – 1602.

Leentjens, A.F.G., Verhey, F.R.J., Lousberg, R., Spitsbergen, H., & Wilmink, F.W. (2000). The validity of the Hamilton and Montgomery-Asberg depression rating scales as screening and diagnostic tools for depression in Parkinson's disease. *International Journal of Geriatric Psychiatry*, 15(7), 644 – 649.

Leichsnring, F., Rabung, S., & Leibing, E. (2004). The efficacy of short term psychodynamic psychotherapy in specific psychiatric disorders: A meta-analysis. *Archives of General Psychiatry*, 61(12), 1208 – 1216.

Leyfer, O.T., Ruberg, J.H., & Woodruff-Borden, J. (2006). An examination of the utility of the Beck Anxiety Inventory and its

factors as screener for anxiety disorders. *Anxiety Disorders*, 20, 444 – 458.

Lezak, M.D., Howieson, D.B., Loring, D.W. (2004). *Neuropsychological assessment (4th edition)*. New York: Oxford University Press.

Lincoln, N.B., Nicholl, C.R., Flannaghan, T., Leonard, M., & Van der Gucht, E. (2003). The validity of questionnaire measures for assessing depression after stroke. *Clinical Rehabilitation*, 17, 842 – 848.

Linden, A. (2006). Measuring diagnostic and predictive accuracy in disease management: an introduction to receiver operating characteristic (ROC) analysis. *Journal of Evaluation in Clinical Practice*, 12(2), 132 – 139.

Livneh, H., & Antonak, R. (2005) Psychological adaptations to chronic illness and disability: A primer for counsellors. *Journal of counselling and development*, 83, 12 – 20.

Lloyd-Williams, M., Friedman, T., & Rudd, N. (2001). An analysis of the validity of the Hospital Anxiety and Depression Scale as a screening tool in patients with Advanced Metastatic Cancer. *Journal of Pain and Symptom Management*, 22(6), 990 – 996.

Lobentanz, I.S., Asenbaum S., Vass, K., Sauter, C., Klosch, G., Kollegger, H., Kristoferitsch, W., & Zeitlhofer, J. (2004). Factors influencing quality of life in multiple sclerosis patients: disability,

depressive mood, fatigue and sleep quality. *Acta Neurologica Scandinavica*, 110, 6 – 13.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635 – 694.

MacHale, S. (2002). Managing depression in physical illness. *Advances in Psychiatric Treatment*, 8, 297 – 305.

Macintosh-Michaelis, S.A., Roberts, M.H., Wilkinson, S.M., Diamond, I.D., McLellan, D.L., Martin, J.P., & Spackman, A.J. (1991). The prevalence of Cognitive Impairment in a community survey of Multiple Sclerosis. *British Journal of Clinical Psychology*, 30, 333 – 348.

Magno, C. (2009). Demonstrating the difference between classical test theory and item response theory using derived test data. *The International Journal of Educational and Psychological Assessment*, 1, 1 – 11.

Mapstone, J., Elbourne, D. & Roberts, I. (2007). Strategies to improve recruitment to research studies. *Cochrane Database Systematic Review*, 18(2), MR000013.

Marinus, J., Leentjens, A.F.G., Visser, M., Stiggelbout, A.M., & van Hilten, J.J. (2002). Evaluation of the Hospital Anxiety and Depression Scale in patients with Parkinson's disease. *Clinical Neuropharmacology*, 25(6), 318 – 324.

- Martin, C.R. (2005). What does the Hospital Anxiety and Depression Scale measure in liaison psychiatry settings? *Current Psychiatry Reviews*, 1, 69 – 73.
- Maser, J., & Cloninger, C.R. (1990). *Co-morbidity in anxiety and mood disorders*. Washington, D.C: American Psychiatric Press.
- McGuigan, C., & Hutchinson, M. (2006). Unrecognised symptoms of depression in a community based population with multiple sclerosis. *Journal of Neurology*, 253, 219 – 223.
- McManus, S., Meltzer, H., Brugha, T., Bebbington, P. Jenkins, R. (2009). *Adult psychiatric morbidity in England, 2007: Results of a household survey*. London: National Centre for Social Research.
- Mental Capacity Act. (2005). *Mental Capacity Act*. London: The Stationary Office.
- Moffic, H.S., & Paykel, E.S. (1975). Depression in medical in-patients. *The British Journal of Psychiatry*, 126, 346 – 353.
- Mohr, D.C., Cox, D., Epstein, L., & Boudewyn, A. (2002). Teaching patients to self-inject: pilot study of a treatment for injection anxiety and phobia in multiple sclerosis patients prescribed injectable medications. *Journal of Behaviour Therapy and Experimental Psychiatry*, 33(1), 39 – 47.
- Mohr, D.C., & Goodkin, D.E. (1999). Treatment of depression in Multiple Sclerosis: Review and Meta-analysis. *Clinical Psychology: Science and Practice*, 6(1), 1-9.

- Mohr, D.C., Goodkin, D.E., Likosky, W., Gatto, N., Baumann, K.A., & Rudick, R.A. (1997a). Treatment of depression improves adherence to interferon beta-1b therapy for multiple sclerosis. *Archives of Neurology*, 54(5), 531 – 533.
- Mohr, D.C., Goodkin, D.E. Likosky, W., Beutler., Gatto, N., & Langan, M. K. (1997b). Identification of Beck Depression Inventory Items related to Multiple Sclerosis. *Journal of Behavioural Medicine* 20(4), 407 – 414.
- Mohr, D.C., Hart, S.L., & Goldberg, A. (2003). Effects of treatment for depression on fatigue in multiple sclerosis. *Psychosomatic medicine*, 65, 542 – 547.
- Mohr, D.C., Hart, S.L., Julian, L., & Tasch, E.S. (2007). Screening for depression among patients with multiple sclerosis: two questions may be enough. *Multiple Sclerosis*, 13, 215 – 219.
- Moran, P.J., & Mohr, D.C. (2005). The validity of Beck depression Inventory and Hamilton rating scale for depression items in the assessment of depression among patients with multiple sclerosis. *Journal of Behavioural Medicine*, 28, 35 – 41.
- Motl, R.W., Suh, Y., & Weikert, M. (2010). Symptom cluster and quality of life in multiple sclerosis. *Journal of Pain Symptom Management*, 39(6), 1025 – 1032.

- Murray, C.J., & Lopez, A.D. (1997). Global mortality, disability and the contribution of risk factors: Global burden of disease study. *Lancet*, 349, 1436 – 1442.
- Mykletun, A., Stordal, E., & Dahl, A.A. (2001). Hospital Anxiety and Depression Scale: factor structure, item analyses and internal consistency in a large population. *The British Journal of Psychiatry*, 179, 540 – 544.
- Nathan, P.E., & Langenbucher, J. (2003). Diagnosis and classification. In I.B.Weiner, D.K.Freedhein, G.Stricker. J.A.Schinka & T.A.Widger (Eds.) *Handbook of psychology: Clinical Psychology*. (pp. 3 – 26). New Jersey: John Wiley & Sons.
- National Institute for Health and Clinical Excellence (NICE). (2004). *Anxiety*. London: NICE
- National Institute for Health and Clinical Excellence (NICE) (2009). *Depression in adults with a chronic physical health problem*. London: NICE.
- National Screening Committee. (2003). *The UK National Screening Committee's criteria for appraising the validity, effectiveness and appropriateness of a screening programme*. London: HMSO.
- Nicholl, C.R., Lincoln, N.B., Francis, V.M., & Stephan, T.F. (2001). Assessment of emotional problems in people with multiple sclerosis. *Clinical Rehabilitation*, 15, 657 – 688.

- Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231 – 239.
- Norton, P.J., Hayes, S.A., & Hope, D.A. (2004). Effects of a transdiagnostic group treatment for anxiety on secondary depression. *Depression and Anxiety*, 20, 198 – 202.
- Noy, S., Achiron, A., Gabbay, U., Barak, Y., Rotstein, Z., Laor, N., & Sarova-Pinhas, I. (1995). A new approach to affective symptoms in relapsing remitting multiple sclerosis. *Comprehensive Psychiatry*, 36(5), 390 – 395.
- O'Boyle, M., & Slef, D. (1990). A comparison of two interviews for DSM-III-TR personality disorders. *Psychiatry Research*, 32, 85 – 92.
- O'Connor, E.A., Whitlock, E.P, Beil, T.L., & Gaynes, B.N. (2009). Screening for depression in adult patients in primary care settings: A systematic evidence review. *Annals of Internal Medicine*, 151, 793 – 803.
- Olek, M.J. (2005). *Multiple Sclerosis: Etiology, diagnosis and new treatment strategies*. New Jersey: Humana Press Inc.
- Osman, A., Hoffman, J., Barrios, F.X., Kopper, B.A., Breitenstein, J.L., & Hahn, S.K. (2002). Factor structure, reliability, and validity of the Beck anxiety inventory in adolescent psychiatric inpatients. *Journal of Clinical Psychology*, 58, 443 – 456.

- Pallant, J., (2006). *SPSS survival manual: A step by step guide to data analysis using SPSS version 12. Second edition*. Maidenhead, Berkshire: Open University Press.
- Pallant, J.F. & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale. *British Journal of Clinical Psychology*, 46, 1 – 18.
- Paradiso, S., Vaidya, J., Tranel, D., Kosier, T., & Robinson, R.G. (2008). Non- dysphoric depression following stroke. *Neuropsychiatry and Clinical Neuroscience*, 20, 50 – 61.
- Parker, I. (1999). Against relativism in psychology, on balance. *History of the Human Sciences*, 12(4), 61 – 78.
- Parvizi, J., Archiniegas, D.B., Bernardini, G.L., Hoffmann, M.W., Mohr, J.P., Rapoport, M.J., Schmahmann, J.D., Silver, J.M., & Tuhim, S. (2006). Diagnosis and management of pathological laughter and crying. *Mayo Clinic Proceedings*, 81(11), 1482 – 1486.
- Patel, A. (2009). The cost of mood disorders: Service considerations. *Psychiatry*, 8(2), 76 – 80.
- Patel, M.X., Doku, V., & Tennakoon, L. (2003). Challenges in recruitment of research participants. *Advances in Psychiatric Treatment*, 9, 229 – 238.

- Patten, S.B., Beck, C.A., Williams, J.V.A., Barbui, C., & Metz, L.M. (2003). Major depression in multiple sclerosis: A population based perspective. *Neurology*, 61(11), 1524 – 1527.
- Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression. *British Journal of Psychiatry*, 177, 486 – 492.
- Pignone, M.P., Gaynes, B.N., Rushton, J.L., Burchell, C.M., Orleans, C.T., Murlow, C.D., & Lohr, K.N. (2002). Screening for depression in adults: A summary of the evidence for the US Preventive Services Task Force. *Annals of Internal Medicine*, 136(10), 765 – 776.
- Pinquart, M., & Shen, Y. (2011). Depressive symptoms in children and adolescents with chronic physical illness: An updated meta analysis. *Journal of Paediatric Psychology*, 36(4), 375 – 384.
- Poder, K., Ghatavi, K., Fisk, J.D., Campbell, T.L., Kisely, S., Sarty, I., Stadnyk, K., & Bhan, V. (2009). Social anxiety in a multiple sclerosis clinic population. *Multiple Sclerosis*, 15, 393 – 398.
- Poppe, A.Y., Wolfson, C., & Zhu, B. (2008). Prevalence of multiple sclerosis in Canada: a systematic review. *Canadian Journal of Neurological Science*, 35(5), 593 – 601.
- Power, M. (2004). *Mood disorders: A handbook of science and practice*. Chichester: Wiley
- Rachman, S. (1998). *Anxiety*. Hove: Psychology Press Ltd.

- Radloff, L.S. (1977). The CES-D Scale: A self report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385 – 401.
- Ramjeet, J., Koutantji, M., Barrett, E.M., & Scott, D.G.I. (2005). Coping and psychological adjustment in recent-onset inflammatory polyarthritis: the role of gender and age. *Rheumatology*, 44(9), 1166 – 1168.
- Razavi, D., Delvaux, N., Farvacques, C., & Robaye, E. (1990). Screening for adjustment disorders and major depressive disorders in cancer inpatients. *British Journal of Psychiatry*, 156, 79 – 83.
- Reiss, H.T., & Judd, C.M. (2000). *Handbook of research methods in social and personality psychology*. Cambridge: Cambridge University Press.
- Revenson, T.A., Schiaffion, K.M., Majerovitz, D., & Gibofsky, A. (1991). Social support as a double edged sword: The relation of positive and problematic support to depression among rheumatoid arthritis patients. *Social Science and Medicine*, 33(7), 807 – 813.
- Rickards, H. (2005). Depression in neurological disorders: Parkinson's disease, multiple sclerosis and stroke. *Journal of Neurology, Neurosurgery and Psychiatry*, 76(s1), i48 – i52.
- Rijnders, C.A., van den Berg, J.F.M., Hodinamont, P.P.G., Neinhuis, F.J., Furer, J.W., Mulder, J., & Giel, R. (2000). Psychometric

properties of the schedules for clinical assessment in neuropsychiatry (SCAN-2.1). *Social psychiatry and psychiatric epidemiology*, 25, 348 – 352.

Robertson, N., Dean, J., Fraser, M. & Compston, D.A. (1995). Multiple sclerosis in the north Cambridgeshire districts of East Anglia. *Journal of Neurology, Neurosurgery and Psychiatry*, 59(1), 71 – 76.

Robins, L.N., Wing, J., Wittchen, H.U., Helzer, J.E., Babor, T.F., Burke, J., Farmer, A., Jablensky, A., Pickens, R., Regler, D.A., Sartorius, N., & Towle, L.H. (1989). The Composite international Diagnostic interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry*, 45, 1069 – 1077.

Robinson, R.G., Parikh, R.M., Lipsey, J.R., Starkstein, S.E., & Price, T.R. (1993). Pathological laughing and crying following stroke: validation of a measurement scale and a double blind treatment study. *American Journal of Psychiatry*, 150, 286 – 293.

Romer, L., Orsillo, S.M., & Salters-Pedneault, K. (2008). Efficacy of an acceptance based behavioural therapy for generalised anxiety disorder: Evaluation in a randomised controlled trial. *Journal of Consulting and Clinical Psychology*, 76(6), 1083 – 1089.

- Rosenman, S.J., Korten, A.E., & Levings, C.T. (1997). Computerised diagnosis in acute psychiatry: Validity of CIDI-Auto against routine clinical diagnosis. *Journal of Psychiatric Research*, 31(5), 581 – 592.
- Rossier, P., & Wade, D.T. (2002). The Guy's neurological disability scale in patients with multiple sclerosis: A clinical evaluation of its reliability and validity. *Clinical Rehabilitation*, 16, 75 – 95.
- Rust, J., & Golombok, S. (2009). *Modern Psychometrics: The science of psychological assessment. Third Edition*. London: Taylor and Francis.
- Sadovnick, A.D. (2009). European Charcot Foundation Lecture: The natural history of multiple sclerosis and gender. *Journal of the Neurological Sciences*, 186, 1 – 5.
- Sareen, J., Cox, B.J., Clara, I., & Asmundson, G.J.G. (2005). The relationship between anxiety disorders and physical disorders in the US national co-morbidity survey. *Depression and anxiety*, 21, 193 – 202.
- Sayer, A. (2000). *Realism and social science*. London: SAGE Publications
- Schrag, A., Barone, P., Brown, R., Leentjens, A.F.G., McDonald, W., Starksten, S., Weintraub, D., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G.T. & Goetz, C.G. (2007). Depression rating

scales in Parkinson's disease: Critique and recommendations.
Movement disorders, 22(8), 1077 – 1092.

Schumacher, G.A., Beebe, G., Kibler, R.F., Kurland, L.T., Kurtzke, J.F., McDowell, F., Nagler, B., Sibley, W.A., Tourtellotte, W.W., & Willmon, T.L. (1965). Problems of experimental trials of therapy in multiple sclerosis: A report by the panel on the evaluation of experimental trials of therapy in multiple sclerosis. *Annals of the New York Academy of Sciences*, 122, 552 – 568.

Segal, D.L., Coolidge, F.L., Cahill, B.S., & O'Riley, A.A. (2008). Psychometric properties of the Beck Depression Inventory-II among community dwelling older adults. *Behaviour Modification*, 32(3), 3 – 20.

Siebert, R.J., & Abernethy, D.A. (2005). Depression in multiple sclerosis: A review. *Journal of Neurology, Neurosurgery and Psychiatry*, 76, 469 – 475.

Sharrack, B., & Hughes, R.A. (1999). The Guy's neurological disability scale (GNDS): a new disability measure for multiple sclerosis. *Multiple Sclerosis*, 5, 223 – 233.

Smith, G.T. (2005). On construct validity: Issues of method and measurement. *Psychological Assessment*, 17(4), 396 – 408.

Smith, S.J., & Young, C.A. (2000). The role of affect on the perception of disability in multiple sclerosis. *Clinical Rehabilitation*, 14, 50 – 54.

- Snaith, R.P. (2003). The Hospital Anxiety and Depression Scale. *Health and quality of life outcomes*, 1, 29 – 33.
- Solari, A., Acquarone, N., Pucci, E., Marinelli, V., Marrosu, M.G., Trojano, M., et al. (2007). Communicating the diagnosis of multiple sclerosis: A qualitative study. *Multiple Sclerosis*, 13(6), 763 – 769.
- Solari, A., Mattarozzi, K., Vignatelli, L., Giordano, A., Russo, P.M., Messmer Uccelli, M., & D'Alessandro, R. (2010). Development and validation of a patient self-assessed questionnaire on satisfaction with communication of the multiple sclerosis diagnosis. *Multiple Sclerosis*, 16(10), 1237 – 1247.
- Somers, J.M., Goldner, E.M., Waraich, P., & Hsu, L. (2006). Prevalence and incidence studies of anxiety disorders: A systematic review. *Canadian Journal of Psychiatry*, 51, 100 – 113.
- Spalletta, G., Ripa, A., & Caltagirone, C. (2005). Symptom profile of DSM-IV major and minor depressive disorders in first ever stroke patients. *American Journal of Geriatric Psychiatry*, 13(2), 108 – 115.
- Spielberger, C.D. (1989). *State-trait Anxiety Inventory: a comprehensive bibliography*. Palo Alto, CAL Consulting Psychologists Press.
- Sprinkle, S.D., Lurie, D.L., Insko, S.L., Atkinson, G., Jones, G.L., Logan, A.R., & Bissada, N.N. (2002). Criterion validity, severity cut scores and test retest reliability of the Beck Depression

Inventory-II in a university counselling centre sample. *Journal of Counselling Psychology*, 49(3), 381 – 385.

Stanton, A.L., Revenson, T.A. & Tennen, H. (2007). Health psychology: Psychological adjustment to chronic disease. *Annual Review of Psychology*, 58, 565 – 592.

Stanton, B.R., Barnes, F., & Sibling, E. (2006). Sleep and fatigue in multiple sclerosis. *Multiple Sclerosis*, 12, 481 – 486.

Steer, R.A., Geetha, K., Ranieri, W.F., & Beck, A.T. (1998). Use of the Beck Depression Inventory-II with adolescent psychiatric outpatients. *Journal of Psychopathology and Behavioural Assessment*, 20(2), 127 – 137.

Steer, R.A., Ranieri, W.F., Beck, A.T., & Clark, D.A. (1993). Further evidence for the validity of the Beck Anxiety Inventory with psychiatric outpatients. *Journal of Anxiety Disorders*, 7, 192 – 205.

Steer, R.A., Rissmiller, D.J., & Beck, A.T. (2000). Use of the Beck Depression Inventory-II with depressed geriatric inpatients. *Behavioural Research & Therapy*, 38, 311 – 318.

Strober, L.B., & Arnett, P.A. (2009). Assessment of depression in three medically ill, elderly populations: Alzheimer's disease, Parkinson's disease and stroke. *The Clinical Neuropsychologist*, 23(2), 205 – 230.

- Sullivan, L.W. (2009). *The SAGE glossary of the social and behavioural sciences*. London: Sage publications ltd.
- Sullivan, M.J., Weinshenker, B., Mikail, S., & Bishop, S.R. (1995). Screening for major depression in the early stages of multiple sclerosis. *Canadian Journal of Neurological Sciences*, 22, 228 – 231.
- Sullivan, M.J., Weinshenker, B., Mikaril, S., & Edgley, K. (2008). Depression before and after the diagnosis of multiple sclerosis. *Multiple Sclerosis*, 1, 104 – 108.
- Sutcliffe, J.P. (1965). A probability model for errors of classification I: General considerations. *Psychometrika*, 30, 73 – 96.
- Svanborg, P., & Asberg, M. (2001). A comparison between the Beck Depression Inventory and the self-rating version of the Montgomery Asberg Depression Rating Scale. *Journal of Affective Disorders*, 64, 203 – 216..
- Swets, J.A., Dawes, R.M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological science in the public interest*, 1, 1-26.
- Stanton, B.R., Barnes, F., & Sibling, E. (2006). Sleep and fatigue in multiple sclerosis. *Multiple Sclerosis*, 12, 481 – 486.
- Trojano, M. & Paolicelli, D. (2001). The differential diagnosis of multiple sclerosis: Classification and clinical features of relapsing and

progressive neurological syndromes. *Neurological Sciences*, 22(8), S98 – S102.

Tsivgoulis, G., Triantafyllou, N., Papageorgiou, C., Evangelopoulis, M.E., Kararizou, E., Sfagos, C., Vassilopoulos, D. (2007). Associations of the Expanded Disability Status Scale with anxiety and depression in multiple sclerosis outpatients. *Acta Neurolgica Scandinavia*, 115, 67 – 72.

Vanheule, S., Desmet, M., Groenvynk, H., Rosseel, Y., & Fontaine, J. (2008). The factor structure of the Beck Depression Inventory II: An evaluation. *Assessment*, 15, 177 – 187.

Walker, J., Postma, K., McHugh, G.S., Rush, R., Coyle, B, Strong, C. & Sharpe, M. (2007). Performance of the Hospital Anxiety and Depression Scale as a screening tool for major depressive disorder in cancer patients. *Journal of Psychosomatic Research*, 63(1), 83 – 91.

Warren, S., & Warren, K.G. (2001). *Multiple Sclerosis*. Geneva: World Health Organization.

Watson, D. (2009). Differentiating the mood and anxiety disorders: a quadripartite model. *Annual Review of Clinical Psychology*, 5, 221 – 247.

Watson, D., Kendall, P.C. (1989). Understanding anxiety and depression: Their relation to negative and positive affective states. In P.C. Kendall & D. Watson (Eds.), *Anxiety and*

depression: Distinctive and overlapping features (pp. 3 – 26).
San Diego, CA: Academic Press.

Watson, D., Gamez, W., & Simms, L.J. (2005). Basic dimensions of temperament and their relation to anxiety and depression: A symptom based perspective. *Journal of Research in Personality*, 39, 46 – 66.

Watson, D., Weber, K., Assenheimer, J.S., Clark, L.A., Struass, M.E., & McCormick, R.A. (1995). Testing a tripartite model: 1. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology*, 104(1), 3 – 14.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, 98, 219 – 235.

Wetherell, J.L., & Arean, P.A. (1997). Psychometric evaluation of the Beck Anxiety Inventory with older medical patients. *Psychological Assessment*, 9(2), 136 – 144.

Wilhelm, K., Parker, G., & Hazdi-Pavlic, D. (1997). Fifteen years on: Evolving ideas in researching gender differences in depression. *Psychological Medicine*, 27, 875 - 883.

Wilson, K.A., Chambless, D.L. & de Beurs, E. (2006). Beck anxiety inventory. In M.E. Maurish (Ed.) *The use of psychological testing for treatment planning and outcomes assessment: Third edition*.

Volume three: Instruments for adults.(pp. 399 – 420). New Jersey: Lawrence Erlbaum Associates Inc.

Wittkamp, K.A., van Zwieten, M., Smits, F.T., Schene, A.H., Huyser, J., & van Weert, H.C. (2008). Patients' view of screening for depression in general practice. *Family Practice*, 57, 438 – 444.

Win. J.K., Babor, T., Brugha, T., Burke, J (1990). SCAN: Schedules for clinical assessment in neuropsychiatry. *Archives of General Psychiatry*, 47(6), 589 – 593.

World Health Organisation (WHO). (1992). *International Classification of Diseases (ICD- 10)*. Geneva: WHO

Yin, P. & Fan, X. (2000). Assessing the reliability of Beck depression inventory scores: Reliability generalization across studies. *Educational and Psychological measurement*, 60, 201 – 223.

Yonkers, K.A., Bruce, S.E., Dyck, I.R., & Keller, M.B. (2003). Chronicity, relapse and illness – course of panic disorder, social phobia and generalised anxiety disorder: Findings in men and women from eight years follow up. *Depression and Anxiety*, 17, 173 – 179.

Young, M.A., Scheftner, W.A., Fawcett, J., & Klerman, G.L. (1990). Gender differences in the clinical features of unipolar major depressive disorder. *Journal of Nervous and Mental Disease*, 178, 200 – 203.

Zigmond, A.S., & Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67(6), 361 – 370.

Zorzon, M., de Massi, R., Nasuelli, D., Ukma, M., Mucelli, R.P., Cazzato, G., Bratina, A., & Zivadinov, R. (2001). Depression and anxiety in multiple sclerosis: A clinical and MRI study in 95 subjects. *Journal of Neurology*, 248, 416 – 421.

Zou, K.H., O'Malley, A.J., & Mauri, L. (2007). Receiver operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 115, 654 – 657.

APPENDIX A: Confirmation of ethical approval

Leicestershire, Northamptonshire & Rutland Research Ethics Committee 1

1 Standard Court

Park Row

Nottingham

NG1 6GN

Telephone: 0115 8839428

Facsimile: 0115 9123300

25 November 2009

Professor Nadina Lincoln

University of Nottingham

I-WHO, Jubilee Campus,

Wollaton Road

Nottingham

NG2 1BB

Dear Professor Lincoln

Study Title: Validation of mood measures for patients with Multiple Sclerosis

REC reference number: 09/H0406/112

Protocol number: 1

Thank you for your letter of 24 November 2009, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see “Conditions of the favourable opinion” below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

For NHS research sites only, management permission for research (“R&D approval”) should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Where the only involvement of the NHS organisation is as a Participant Identification Centre, management permission for research is not required but the R&D office should be notified of the study. Guidance should be sought from the R&D office where necessary.

Sponsors are not required to notify the Committee of approvals from host organisations.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

<i>Document</i>	<i>Version</i>	<i>Date</i>	
Covering Letter		09 October 2009	
REC application	22577/67171/1/845	12 October 2009	
Protocol	1	23 September	

		2009	
Evidence of insurance or indemnity		28 July 2009	
Letter from Sponsor		06 October 2009	
Investigator CV	Chief Investigator	04 May 2009	
Participant Consent Form: opt-in slip	2	01 October 2009	
Referees or other scientific critique report			
Questionnaire: Demographic Information	1	01 April 2009	
Covering Letter	2	01 August 2009	
Investigator CV	Student	08 October 2009	
Questionnaire: Hospital Anxiety and Depression Scale			
Questionnaire: BAI			
Investigator CV	Supervisor	11 September 2009	
Participant Information Sheet	5	01 November 2009	
Participant Consent Form	5	01 November 2009	
Consultant Letter	1	01 November 2009	
Guy's Neurological Disability Scale			
Response to Request for Further Information		24 November 2009	

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

09/H0406/112	Please quote this number on all correspondence
---------------------	---

Yours sincerely

Dr Carl Edwards / Miss Jeannie D McKie

Chair / Committee Coordinator

Email: jeannie.mckie@nottspct.nhs.uk

Enclosures: *"After ethical review – guidance for researchers" SL- AR2 for other studies*

Copy to: *Paul Cartledge, Reserach Innovation Services*
R&D office for NHS care organisation at lead site

APPENDIX B: Confirmation of ethics amendment



National Research Ethics Service

Leicestershire, Northamptonshire & Rutland Research Ethics Committee 1

1 Standard Court
Park Row
Nottingham
NG1 6GN

Tel: 0115 8839435
Fax: 0115 9123300

12 March 2010

Professor Nadina Lincoln
I-WHO, Jubilee Campus,
Wollaton Road
Nottingham
NG2 1BB

Dear Professor Lincoln

Study title: Validation of mood measures for patients with Multiple Sclerosis
REC reference: 09/H0406/112
Protocol number: 1
Amendment number: 1 Dr Nikos Evangelou New Principal Investigator
Amendment date: 11 March 2010

Thank you for your letter of 11 March 2010, notifying the Committee of the above amendment.

The Committee does not consider this to be a "substantial amendment" as defined in the Standard Operating Procedures for Research Ethics Committees. The amendment does not therefore require an ethical opinion from the Committee and may be implemented immediately, provided that it does not affect the approval for the research given by the R&D office for the relevant NHS care organisation.

Documents received

The documents received were as follows:

Document	Version	Date
Notification of a Minor Amendment	1 Dr Nikos Evangelou New Principal Investigator	11 March 2010

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

This Research Ethics Committee is an advisory committee to East Midlands Strategic Health Authority.
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England.

APPENDIX C: Confirmation of ethical approval from NHS trust research and development department



Please reply to:

Telephone:

Fax:

E-mail:

30 March 2010

Prof Nadina B Lincoln
Institute of Work, Health
Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB

Dear Prof Lincoln

ID: 09NS014

The validation of mood measures for use with patients with Multiple Sclerosis

The R&D Department has considered the following documents:

- . IRAS Application form, version 2.5
- . Protocol, version 1 dated 23/09/09
- . Participant consent Form: Option Slip version 2 dated 01/10/09
- . Participant Information Sheet version 5 dated 01/11/09
- . Participant consent form version 5 dated 01/11/09
- . Consultant letter version 1 dated 01/11/09

Your study now has R&D approval, on the understanding and provision that you will follow the conditions set out below.

Conditions of Approval

That you:

1. Comply with all relevant laws, regulations and codes of practice applicable to the trial including but not limited to, the UK Clinical Trials Regulations, Medicines for Human Use (Clinical Trial) Regulations 2004, principles of Good Clinical Practice, the World Medical Association Declaration of Helsinki entitled 'Ethical Principles for Medical Research Involving Human Subjects' (1996 version), the Human Rights Act 1998, the Data Protection Act 1998 the Medicines Act 1968, the NHS Research Governance Framework for Health and Social Care (version 2 April 2005). Should any of these be revised and reissued the latest version of the relevant laws and regulations will apply. Copies of the regulations are available from the R&D Office or via the R&D website <http://nuhrise.org>
2. For NUH sponsored studies accept the responsibilities as outlined in the "Clinical Trial Delegation of Sponsor responsibilities to Chief Investigator" agreement.
3. Request written approval from the R&D department, Ethics Committee and MHRA (as appropriate) for any Protocol Amendments, changes to study documentation or changes to study team.
4. Ensure all study personnel, not employed by the Nottingham University Hospitals NHS Trust hold either honorary contracts/letters of access with this [redacted] any patients or staff, their data, tissue or organs or any NUH facilities.
5. According to R&D SOP 11 - "Adverse Event Monitoring, Recording and Reporting for investigators" report any Serious Adverse Events to the R&D department.

6. According to R&D SOP 12 - "Protocol Violations and Serious Breach Reporting" report any Serious Breach of the UK Clinical Trial regulations in connection with the trial or Serious Breach of the protocol, immediately after becoming aware of the breach to R&D.

7. complete Annual Safety, Progress reports and End of Study reports as required by R&D, Ethics Committee and the MHRA.

8. Notify R&D within 7 calendar days of the first patient or healthy volunteer recruited onto the study, as well as the detail of the specific recruitment date. Please email the recruitment notification to rdmon@nuh.nhs.uk.

This approval letter constitutes a favourable Site Specific Assessment (SSA) for this site.

Please note that the R&D department has a database containing study related information, and personal information about individual investigators e.g. name, address, contact details etc. This information will be managed according to the principles established in the Data Protection Act.

Yours sincerely



Dr Brian Thomson / Dr Maria Koufali

Director of R&D / Assistant Director Research and Innovation

cc [REDACTED]

APPENDIX D: Conformation of sponsorship



Our reference:

RIS 09079

Your reference:

09/H0406/112

0115 9515679

paul.cartledge@nottingham.ac.uk

Research Innovation Services

University of Nottingham

King's Meadow Campus

Lenton Lane Nottingham

NG7 2NR

Leicestershire, Northampton and Rutland

REC 1

1 Standard Court

Park Row

Nottingham

NG1 6GN

6th October 2009

Professor Lincoln, Chief Investigator

University of Nottingham

IWHO, International House

Jubilee Campus

Nottingham

NG8 1BB

Dear sir or madam,

Sponsorship Statement

Re: Validation of mood measures for patients with Multiple Sclerosis

I can confirm that this research proposal has been discussed with the Chief Investigator and agreement to sponsor the research is in place.

An appropriate process of scientific critique has demonstrated that this research proposal is worthwhile and of high scientific quality.*

Any necessary indemnity or insurance arrangements will be in place before this research starts. Arrangements will be in place before the study starts for the research team to access resources and support to deliver the research as proposed.

Arrangements to allocate responsibilities for the management, monitoring and reporting of the research will be in place before the research starts.

The duties of sponsors set out in the NHS Research Governance Framework for Health and Social Care will be undertaken in relation to this research.**

* Not applicable to student research (except doctoral research).

** Not applicable to research outside the scope of the Research Governance Framework.

Yours faithfully

A handwritten signature in dark ink, appearing to read "P. Cartledge".

Paul Cartledge

Head of Research Grants and Contracts

University of Nottingham



APPENDIX E: Cover letter of information pack sent to participants



The University of
Nottingham

Tessa Hopkins
Institute for Work, Health and Organisations

Jubilee Campus

Wollaton Road

University of Nottingham

NG8 1BB

Telephone: 0115 8467523

To Whom It May Concern,

**RE: Study to assess the appropriateness of measures of mood
within a Multiple Sclerosis patients.**

Thank you for enquiring about the above study. Please find enclosed:

- An information sheet for you to read to find out more about the study and allow you to decide if you want to take part
- A consent form for you to look at if you are willing to take part in the interview
- An opt-in slip for the interview and to receive a summary of the results once the study is complete
- Five Questionnaires labelled: BDI, BAI, Hospital Anxiety and Depression Scale, Demographic Information and Guy's Neurological and Disability Scale.
- A stamped addressed envelope

Please read through the information sheet. If having read the information sheet you decide you do not wish to take part in the study then please return the contents of this pack to the above address. If you decide you do wish to take part then please complete the questionnaires before returning them in the stamped addressed envelope. By completing the questionnaires you are implying that you consent to the information you provide in the questionnaires being used in the study.

If you are willing to take part in an interview as well as complete the questionnaires, please indicate this on 'opt-in' slip and return it with the completed questionnaires. You will be contacted to arrange a time and place to complete the interview or to inform you that no interview is necessary. Consent for the interview will be discussed with you prior to the interview taking place.

If you have questions or concerns please contact using the details above. Thank you for your time,

Tessa Hopkins (study co-ordinator)

APPENDIX F: Consent Form

CONSENT FORM

(Draft Version 3 / Final version 1.0: September)

Title of Study: The validation of mood measures for use with patients with Multiple Sclerosis

REC ref:

Name of Researcher: Tessa Hopkins

Name of Participant:

Please initial box

1. I confirm that I have read and understand the information sheet version number 3 dated September 2009 for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without my medical care or legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that this information may still be used in the project analysis. ☐
3. I understand that relevant sections of my medical notes and data collected in the study may be looked at by authorised individuals from the University of Nottingham where it is relevant to my taking part in this study. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential. ☐
4. I understand that the interview will be recorded and that anonymous direct quotes from the interview may be used in the study reports. ☐
5. I agree to take part in the above study.

Name of Participant

Date

Signature

Name of Person taking consent Date

Signature

(Study Co-ordinator)

APPENDIX G: Opt in slip



The University of
Nottingham

The validation of mood measures for use with patients with Multiple Sclerosis

OPT-IN SLIP

If you are willing to take part in an interview or you would like to receive a summary of the results when the research is complete then please provide your contact details below. Your contact details will be kept confidential and secure at the University of Nottingham and will be destroyed once they are no longer needed.

Name: _____

Address: _____

Contact Number: _____

Please tick to indicate whether you are willing to take part in an interview or if you would like a summary of the results.

☐

I am willing to take part in an interview

☐

I would like to be sent a summary of the results once the research is complete

Signed _____

Date _____

Please return with questionnaires in stamped envelope provided.

APPENDIX H: Information sheet sent to participants



Participant Information Sheet

Study to assess the appropriateness of measures of mood within a Multiple Sclerosis patients.

I would like to invite you to take part in a research study. Before you decide you need to understand why the research is being done and what it would involve for you. Please take the time to read the following information carefully. Talk to others about the study if you wish.

Please contact me if there is anything that is not clear or if you would like more information. Take the time to decide whether or not you wish to take part.

What is the purpose of the study?

The study is looking at how anxiety and depression are assessed using questionnaires for people who have a diagnosis of multiple sclerosis.

The study's purpose is to ensure the questionnaire assessments used to assess anxiety and depression are appropriate to be used with people with multiple sclerosis. The completed research will go on to form part of a qualification for a Doctorate in Clinical Psychology.

Why have I been invited?

You have been invited because I am interested in people with a diagnosis of multiple sclerosis, some with anxiety and/or depression and some without. In total 18 people will be recruited to take part in the study.

Do I have to take part?

It is up to you to decide. The study is described in this information sheet. We will then ask you to sign a consent form to show you have agreed to take part. You are free to withdraw at any time, without giving a reason. This would not affect the standard of any care you receive.

What will happen to me if I take part?

You will be involved in the research for approximately 50 minutes if you are completing the questionnaires alone. You will be involved for a further 2 hours and 20 minutes if you complete the interview and repeat questionnaires. The research itself will last approximately 18 months. You will be asked to complete some questionnaires which ask about your mood and how you have been feeling over the last two weeks. You will also be asked to give some information such as your age, gender and when you were diagnosed with multiple sclerosis. You have been provided with a pre-paid envelope for you to return the completed information to a Research Associate.

If you decide you are willing to take part in an interview then please complete the opt-in slip included, you may be contacted once your questionnaires have been returned. Not everyone that has stated they are willing to take part in an interview will be doing so. If you have agreed to take part in an interview but it is not necessary you will be informed by letter. If you have agreed to take part in an interview and it is felt it would be useful to the study the Research Associate will give your contact details to the Study Co-ordinator. The Study Co-ordinator will contact you to complete the interview. Once the interview is complete you will be asked to fill out a further set of questionnaires and return them by pre-paid envelope which will be provided.

All the information you give will be identified by a unique study identity code to you in the study. Information that identifies you will be kept securely and separately from the information that is used in the study. The interview with the Study Co-ordinator will involve an audio recording, this will be copied onto a compact disc labelled with your unique number and kept securely, the original on the audio recorder will be deleted. The recording may be listened to by a Research Associate, this will be anonymous and the person listening to the recording will not have any other information about you.

Expenses and payments

You will not receive any payments for taking part in the study. The Study Co-ordinator will be travelling to meet you for the interview and so this will not be an expense. Any information that needs to be sent by post will be paid for by the study.

What are the possible disadvantages and risks of taking part?

There is a possible risk that you may become distressed when taking part in the study. The study will require you to think about how you are feeling at the moment and there is a potential that this may be distressing for you. A list of organisations that may help you if you do become distressed is at the end of this information sheet. The study co-ordinator will also bring the list of organisations to the interview. If you feel distressed following the study you may also wish to talk to staff within the Multiple Sclerosis service.

What are the possible benefits of taking part?

You will be given the opportunity to talk about how you are currently feeling and your mood at the moment. The information we get from this study will also help improve the treatment of people with a diagnosis of Multiple Sclerosis.

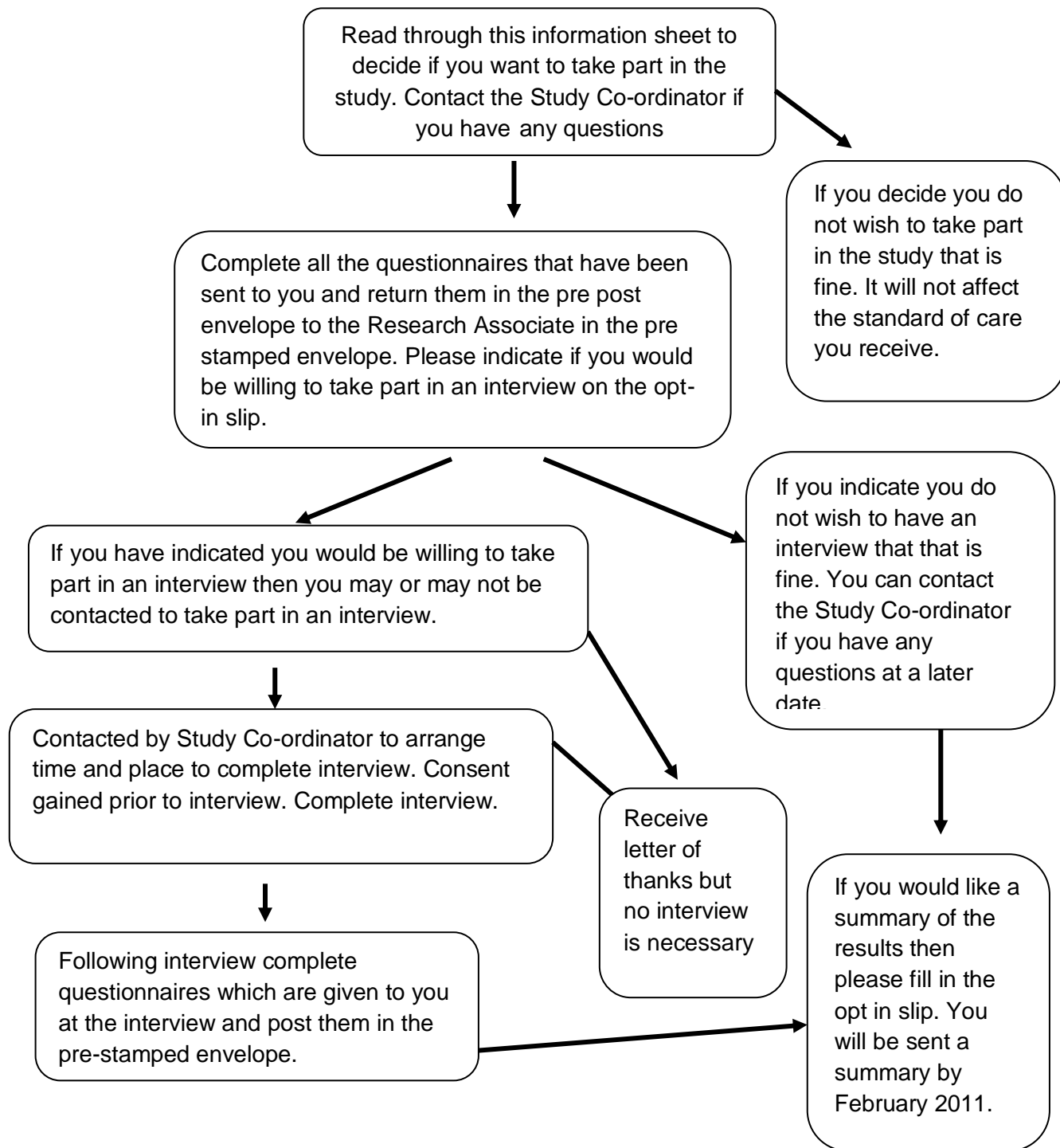
What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. If you have a concern about any aspect of this study, you should ask to speak to the chief investigator who will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure by contacting the Patient Advice and Liaison Service on 0115 9249924 extension 65412, or the University of Nottingham on 0115 8467523.

In the event that something does go wrong and you are harmed during the research due to someone's negligence then you may have grounds for a legal action for compensation against the University of Nottingham but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

What will I have to do?

If you choose to take part in the study you will be asked to do a number of things. The steps involved are shown overleaf:



Will my taking part in the study be kept confidential?

Yes. We follow ethical and legal practice and all information about you will be handled in confidence. All information which is collected about you during the course of the research will be kept strictly confidential and will be marked with a unique study identity code so no identifiable information will be on the questionnaires or interview recording. Your contact details and consent form will be placed in a sealed envelope, labelled with the name of the Study Co-ordinator, the identifier code given to you and the date it can be destroyed. This will be kept in a

locked filing cabinet at the University of Nottingham. The other information you provide will be kept in the same manner although parts of it may be placed onto a computer. If this is the case, then it will be encrypted and only the Research Associate and Study Co-ordinator will have access to it. Any information that has your name or address on it will not be accessed by anyone other than the Research Associate and Study Co-ordinator. All the data you provide will be kept for seven years after the study is complete, it will then be destroyed securely.

What will happen if I don't want to carry on with the study?

You are able to withdraw from the study at any time without giving a reason, you just need to let the study co-ordinator know. Once the data has been collected by questionnaires or interview it cannot be erased but it will remain anonymous.

What will happen to the results of the research study?

The results of the study will be reported as part of the Study Co-ordinator's doctoral thesis. If you wish to receive a summary of the results when the study has been completed then tick the box on the opt-in slip and provide your contact details.. The Study Co-ordinator will use your contact details to send you a summary of the results in February 2011.

Who is organising and funding the research?

The sponsor of the study is the University of Nottingham.

Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This occurs before the study begins and so they will not have access to any information about you.

Further Information and Contact Details

You may still have some questions about this research project. If you wish to have further information the please use the contact details below for the study co-ordinator.

Chief Investigator: Nadina Lincoln

Study Co-ordinator: Tessa Hopkins

Address: Institute of Work, Health and Organisations
Jubilee Campus
Wollaton Road
University of Nottingham
NG8 1BB

Telephone Number: 0115 8467523

E-mail: lwxt4@nottingham.ac.uk

If you have any questions about participating in research in general or wish to make a complain using the NHS Complaints Procedure please contact the Patient Advice and Liaison Service on: 0115 9249924 extension 65412.

If you wish to make a complaint through the University of Nottingham please contact the Institute of Work and Organisations on: 0115 8467523.

If you feel distressed at any point due to your participation in the study the following are organisations that may help:

- Samaritans: 08457 909090
- Nottingham Multiple Sclerosis Society: 0115 9786745
- Nottingham Counselling Service: 0115 9501743
- Focus Line, support for anyone affected by mental health issues: 0800 027 2127
- NHS Direct: 0845 46 47

You may also wish to discuss your participation in the study with family, friends, and clinicians from the Multiple Sclerosis service or your GP.

Thank you for reading this information sheet. If you have decided you want to take part in the study then please sign the consent form and complete all the questionnaires. These will be picked up by the study co-ordinator at the interview.

APPENDIX I: Demographic questions sent to participants

DEMOGRAPHIC INFORMATION

Please complete the following:

Age (years): _____

Gender: Male / Female

Type of Multiple Sclerosis: Primary Progressive

Relapsing Remitting

Secondary Progressive

Time since diagnosis of Multiple Sclerosis:

_____ years