HD CDE Behavior Psychology Subgroup
Psychiatry-Summary Statement

I. Defining Core Instruments

This summary is the background for the NINDS CDE Huntington's Working Group – Psychiatry Table. The "Strongly Recommended" instruments on the Working group Table are included as suggestions for studies with a primary focus on a particular psychiatric disorder. Because of the prominence of psychiatric symptoms in HD, it is recommended that the Problem Behaviors Assessment-Short form (PBA-s) be used in all HD studies with any need for behavioral assessment as a comprehensive screen for the most common psychiatric symptoms in HD. The PBA-s also questions about suicidal behavior, a particular concern in HD. The PBA-s is based on the same set of core behavioral symptoms as the Unified Huntington's Disease Rating Scale (UHDRS) Behavioral questions, which were used previously as the global psychiatric measure in most HD studies; therefore the PBA-s will now take the place of the UHDRS behavioral section. The PBA-s has more detailed questions, more specific guidance on administration and scoring, and has had some validation work done to date. More extensive validation is underway at the time of this writing. If a study has no need for a behavioral measure, e.g. an observational study focusing only measurement of a motor symptom, then the PBA-s does not need to be considered a core measurement. The Working Group encourages investigators to consider including behavioral assessment most HD trials, since behavioral symptoms influence so many other factors in the illness.

Some of these instruments were validated against a DSM-IV-TR diagnosis, and may need to be re-validated when the DSM-V is published. Several of the recommended instruments have not been validated specifically in an HD population, but are the best available assessments in the subcommittee's opinion. Special considerations were made for issues particular to HD when selecting behavioral measures. For example, genetic testing at any point in an individual's lifetime can identify those who have the HD genetic expansion. Psychiatric evaluations for those who are "premanifest" (without motor symptoms, but have the HD genetic expansion) and have at the most mild cognitive deficits can differ greatly from those appropriate for end stages of the illness where dementia is nearly universal. Due to the variability of symptoms across stages, it is not possible at this time to state which instruments are most appropriate for which stage of illness. We also have no information on scales appropriate for juvenile HD patients. Our current recommendation for younger patients is to use pediatric versions of adult assessments if available, but validation studies are needed.

As with all assessments done in studies, it is helpful to record the date and time at which the assessment is performed on the Case Report Form. This is useful if there is a question later about the order in which assessments were performed, and whether this had any effect on psychiatric reporting. We have included with each assessment a recommendation as to who should perform the assessment. In general, a trained rater is recommended for most assessments that are administered to a patient; advanced clinical training, such as an MD or PhD, is not necessary for this purpose. Many are self report, although for more advanced patients who have significant cognitive and motor impairment one to one assistance will be needed. Also, since lack of insight is an issue for many HD patients in all stages of the illness, responses on many self assessments by patients may not reflect actual behavior.

II. Assessing Psychiatric Symptoms in Huntington's Disease

A. Overview – Huntington's disease patients are frequently afflicted with psychiatric symptoms, and most will experience one if not several psychiatric symptoms or disorders during the course of their illness. The most common include depression, anxiety, apathy, irritability, and perseverative/obsessive compulsive symptoms. Disrupted sleep is also reported by many patients. Psychiatric symptoms can be in reaction to illness, loss of function, or the chaotic family environment a genetic condition such as HD can create. They can also be related to neuropathological changes in HD, since the connections between frontal and striatal regions are involved not only in control of movement but also in modulation of emotions. Psychiatric symptoms can have a profound impact on ability to adapt to the illness, quality of life, and the burden placed on families and other carepartners.

B. Core Assessment- the only core behavioral assessment, which should be used in all studies with any relevance to behavioral symptoms, is the PBA-s. The PBA-s provides a thorough overview of the major psychiatric symptoms seen in HD. No behavioral assessment needs to be included in studies where the primary focus is purely non-behavioral. It is recommended that the PBA-s be administered in any study where an intervention is performed that could cause changes in behavior, such as a clinical trial of a novel agent, a clinical trial of a known medication that has not had extensive use in HD, a physical therapy intervention, or an exercise intervention. Other measures for individual psychiatric symptoms would be considered "Core" only for a study in which that symptom is a primary concern, e.g. an apathy scale for a study of apathy progression over disease course.

C. Heterogeneity – Psychiatric symptoms in HD generally do not occur in a predictable manner; any symptom can appear at any time in the illness, including before motor manifestations of the disease. The one exception is apathy, which often worsens with disease progression.

D. Co-morbidity - Psychiatric disorders in HD are highly co-morbid, as is seen in the general population (e.g., depressive and anxiety disorders). If a particular psychiatric disorder is assessed, one might also consider assessment of common co-morbid conditions. The influence of co-morbid psychiatric disorders on scale performance has not been well studied. Co-morbidity of psychiatric symptoms is another reason why the PBA-s, a global assessment of psychopathology, is useful in HD. Some scales, such as the Hospital Anxiety and Depression Scale, do rate more than one symptom on a single scale.

E. Symptom overlap - Psychiatric symptoms occurring in HD may be due either to an underlying co-morbid psychiatric disorder or to the effects of the disease itself. For instance, certain symptoms of depression rated on a scale such as the Hamilton Depression Rating Scale occur often in HD patients without depression (e.g., sleep disturbances, apparent psychomotor agitation due to chorea, termination of employment). This may confound scoring on an instrument.

F. Effects of HD medications - The psychiatric effects of medications used to treat HD symptoms are widespread; e.g. Selective Serotonin Reuptake Inhibitors may cause apathy, neuroleptics may cause akathisia, tetrabenazine may precipitate depression.

G. At risk HD- unlike the vast majority of medical illnesses, individuals at risk for HD can choose to find out if they carry the genetic expansion for HD and will develop the condition at some point in their lives. This ability to know with virtual certainty that one will develop a progressive, incurable

illness can have a negative impact on emotions on its own. Deciding to undergo genetic testing for HD has its own set of psychiatric implications for many who choose to pursue testing.

H. Cognitive impairment – Cognitive decline occurs in most HD patients, and deficits may be detected even in premanifest individuals. The vast majority of HD patients will develop dementia as the condition progresses. Because of this and also due to lack of insight in many patients, input from an informant for many of the assessments is invaluable. The influence of cognitive status on psychiatric scale performance has not been well studied.

III. Individual Psychiatric Disorders

Introduction- Guidance is provided for instruments that should be used for specific psychiatric symptoms. In some cases, more than one instrument is recommended for symptom assessment in a domain. Where published HD data on an instrument are available this is noted; in many cases recommendations are provided based purely on clinical or unpublished research experience with these instruments. There are not sufficient data to recommend one instrument over another in cases where multiple assessments are recommended; investigators should decide based on the population of HD patients they intend to study which instrument is most appropriate.

A. Depression/Suicide

Depression is common in HD, and has a large impact on the ability of patients and families to cope with the condition. Use of standard depression scales and the DSM criteria for depression in HD can be confounded by symptoms of HD itself, apathy, and cognitive decline. The DSM-IV criteria remain the gold standard for assessment of depression at this time. We recommend the Hospital Anxiety and Depression Scale for depression assessment in addition to the question on the PBA-s specific to depression.

Suicide risk is elevated in HD. More emphasis is being placed in all conditions on standardized assessment of suicidal ideation in treatment trials. Due to impulsivity, HD patients tend to choose high lethality methods of suicide and timing of attempts can be unpredictable. All suicidal ideation should be taken very seriously in HD. The Working Group recommends using the suicide question on the PBA-s as a guide for a clinical assessment of suicidal behaviors and in behavioral trials, where a thorough behavioral interview should be conducted. If more specific documentation is required, the Columbia Suicide Severity Rating Scale (CSSRS) can be added to supplement the interview questions. The CSSRS is recommended by the FDA for clinical trials, in which safety is an issue and suicidal behavior needs to be assessed in a standardized manner for regulatory purposes. An alternative scale, the Concise Health Risk Tracking Scale, has also been approved by the FDA for use in assessment of suicidal behaviors in clinical trials, and is an acceptable alternative for this purpose. Neither of these scales has been validated in HD populations.

B. Anxiety/ Perseverative/Obsessive Compulsive Behavior-

Anxiety disorders and symptoms have received relatively little study in HD, despite clinical opinion that they are common. Anxiety may be worsened by cognitive decline, when loss of memory leads to anxiety about gaps in knowledge. We recommend the Hospital Anxiety and Depression Scale for anxiety assessment in addition to the question on the PBA-s specific to anxiety.

Perseverative/Obsessive Compulsive behaviors are particularly challenging. They are thought to be quite common in HD, but most of these symptoms do not fit any particular DSM category. For example, unlike the OCD profile specified in the DSM, HD patients with obsessive and compulsive

symptoms often do not feel shame or anxiety in connection with their behaviors. This can significantly limit ability to score impairment on assessments due to obsessive/compulsive symptoms. There are no good scales available for perseveration. The PBA-s interview question on perseveration provides guidance on how to assess perseveration clinically. This is an area with definite research needs in the future.

C. Apathy

Apathy is a common syndrome in HD. It is the one psychiatric symptom that generally progresses with advancement of neurological disease. Distinguishing apathy from depression can pose a serious challenge in HD, both in clinical evaluation and when using assessment tools. Presence of apathy can impair ability to report other psychiatric symptoms.

D. Irritability- Irritability is another common symptom in HD. It can be seen very early in the condition. Assessment of irritability often requires input from an informant, due to lack of insight by patients. The PBA-s has questions on both irritability and angry/aggressive behavior.

E. Psychosis/paranoia- these symptoms are less common in HD. If they occur abruptly, evaluation for underlying medical illness is warranted. There are no scales with a good representation of the type of paranoia seen in HD. The PBA-s does have an item on assessment of paranoid and delusional thinking, and another on hallucinations.

F. Sleep Disorders- Sleep problems are common in HD and a variety of sleep problems are seen clinically including latency of sleep onset, frequent awakenings, decreased total sleep time and reversal of circadian rhythms. Few studies to date have systematically examined sleep in HD and even fewer have used formal sleep measures. We recommend two sleep measures as supplemental, the SCOPA-SLEEP and the Pittsburgh Sleep Quality Index (PSQI). The Cambridge Brain Repair Centre HD Sleep Questionnaire is a new screening instrument specifically developed to assess sleep problems in Huntington's disease, but its psychometric properties have not yet been published; thus, it is recommended as an exploratory measure.

NINDS CDE Huntington's Disease Behavioral Psychology Subgroup

Recommended CDE's

Type of Instrument	Instrument	Core?	Comments
Global	Problem Behaviors Assessment-Short	Yes	Should be considered Core in any study that addresses behavioral sx, but not needed for study with no behavioral elements. Note that in contrast to the majority of scales recommended, the Problem Behaviors Assessment uses input from both patient and an informant (if available) to produce one final score for each item rated.
Depression	Hospital Anxiety and Depression Scale	Only for studies where depression is main focus- otherwise use Problem Behaviors Assessment-Short Depression Item to assess depression.	None
Suicidal Ideation	1) Columbia Suicide Severity Rating Scale 2) Concise Health Risk Tracking Scale	1)No 2)No	The FDA requires assessment of suicidal ideation in clinical trials and recommends the Columbia Suicide Severity Rating Scale. The Concise Health Risk Tracking Scale is an alternative measure which has also been accepted by the FDA for use in clinical trials.
Anxiety	Hospital Anxiety and Depression Scale	Only for studies where Anxiety is main focus- otherwise use Problem Behaviors Assessment-Short Anxiety Item to assess anxiety.	None
Apathy	1) Apathy Evaluation Scale 2) Apathy Scale	1) Core for studies where Apathy is main focus. 2) Core for studies where Apathy is main focus.	Based on data available, unable to recommend one apathy scale versus the other for a particular study.

Type of Instrument	Instrument	Core?	Comments
Obsessive/Compulsive Perseverative	1) Padua-Inventory-OCD-Wash-U-Revised 2) Florida Obsessive Compulsive Inventory	1) No 2) No	There are no good scales for the type of O/C and Perseverative symptoms seen in HD- this is an area of research need.
Irritability	1) Irritability Scale	1) Core only if Irritability is main focus.	This is the only irritability scale with significant data in HD populations to support its use.
Sleep	1)Scale for Outcomes of Parkinson's disease- Sleep 2) Pittsburgh Sleep Quality Index 3) Cambridge Brain Repair Centre (BRC) HD Sleep Questionnaire	1) No 2) No 3) No	There are no good scales for evaluation of sleep disorders in HD. This is an area of research need.
Psychosis	None		There are no good scales for psychotic symptoms seen in HD, this is an area of research need.

Description of Apathy Evaluation Scale for HD Common Data Elements

Instrument Name:	Apathy Evaluation Scale
Classification:	Core
Short Description of Instrument:	Summary/ Overview of Instrument: Semi-structures interview with 18 questions assessing apathy in the past four weeks. This scale was originally designed for patients with Parkinson's disease or Alzheimer's disease. Construct measured: Apathy Generic vs. disease specific: Generic Intended use of instrument/ purpose of tool: Assessment of severity of apathy Means of administration: Paper and pencil Location of administration: Clinic or home Intended respondent: Patient/self (AES-S), Informant (AES-I) and Clinician (AES-C) version # of items: 18 # of subscales and names of sub-scales: None
Scoring	Scoring: Rating of each item is based on a semi-structured interview. The interview should begin with a description of the subject's interest, activities and daily routine. The items should be anwered based on the subject's thoughts, emotions, and actions; based on both verbal and non-verbal information of the past 4 weeks. For each item ratings should be judged: 4 possible responses for each question: 'not at all', 'slightly', 'somewhat', 'a lot'. Standardization of scores to a reference population (z scores, T scores, etc): Not available If scores have been standardized to a reference population, indicate frame of reference for scoring (general population, HD subjects, other disease groups, etc). Not available
Measurements	Type of scale used to describe individual items and total/subscale scores: Continuous
Psychometric Properties	Reliability: Test-retest or intra-interview (within rater) reliability: AES-S = 0.76; AES-I = 0.94; AES-C = 0.88 Inter-interview (between-rater) reliability (as applicable): Inter-rater reliability was only tested for the AES-C and w as found to be go od (intraclass correlation coefficient = 0.94 (Marin, 1991). Internal consistency: Coefficient alpha: AES-S = 0.86; AES-I = 0.94; AES-C = 0.90

	Statistical methods used to assess reliability: Validity: Content validity: not available in reviewed references Construct validity: not available in reviewed references Convergent validity: Intercorrelations among the three scales (AES-S, AES-I, AES-C): AES-C and AES-I: r = 0.62; AES-C and AES-S: r = 0.72; AES-S and AES-I: r = 0.43. Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): Not available in reviewed references Known Relationships to Other Variables: Depression and use of medication (especially neuroleptics, antidepressants, and benzodiazepines) are related to apathy. Diagnostic Sensitivity and Specificity, if applicable: Not available in reviewed references
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: This instrument assesses multiple aspects of apathy and has been used in a variety of neuropsychiatric disorders, and allows for comparison between patient/self, informant, and clinician reports. Weaknesses: The AES may not discriminate apathy from depression. Availability (copyright): http://www.dementia-assessment.com.au/symptoms/ Special Requirements for administration: None Administration Time: Likely 15-30 minutes Translations available: Available in English, German, Dutch, French, Spanish
References:	Key Reference: Marin RS, Biedrzycki RC, Firinciogullari S: Reliability and validity of the Apathy Evaluation Scale. Psychiatry Research 1991;38:143-162

Description of Apathy Scale for HD Common Data Elements

Instrument Name:	Apathy Scale
Classification:	Core
Short Description of Instrument:	Summary/ Overview of Instrument: The Apathy Scale is an abridged version of the Apathy Evaluation Scale (Marin, 1990). The AS consists of 14 items regarding different dimensions of apathetic behavior. The score for each item ranges from 0 to 3. Rating of each item is based on a semi-structured interview: each question should be read by the examiner, and the patient is provided with the four possible answers. Construct measured: Apathy Generic vs. disease specific: Generic Intended use of instrument/ purpose of tool: Assessment of severity Means of administration: Paper and Pencil Location of administration: Clinic or at home Intended respondent: Patient/self and informant # of items: 14 # of subscales and names of sub-scales: None
Scoring	Scoring: Ratings should be based on both verbal and non-verbal information of the past 4 weeks (sometime 2 weeks!). For each item ratings should be judged: 4 possible responses for each question: 'not at all', 'slightly', 'somewhat', 'a lot'. With a cutoff score of 14 points a sensitivity 66% and specificity of 100% has been reported in patients with Alzheimer's disease. Standardization of scores to a reference population (z scores, T scores, etc): Not available If scores have been standardized to a reference population, indicate frame of reference for scoring (general population, HD subjects, other disease groups, etc). Not available
Measurements	Type of scale used to describe individual items and total/subscale scores: Continuous
Psychometric Properties	Reliability: Test-retest or intra-interview (within rater) reliability: The AS showed test-retest reliability ($r = 0.90$, $df = 10$, $p < 0.01$) (Starkstein, 1992) Inter-interview (between-rater) reliability: The AS showed good interrater reliability ($r = 0.81$, $df = 10$, $p < 0.01$) (Starkstein, 1992). Inter-interview (between-rater) reliability (as applicable): Interrater agreement for the presence of apathy above a median score in a HD population ranged from poor for the most cognitively impaired subjects to good for the less cognitively impaired subjects

	(Chatterjee, 2005). Internal consistency: Not available in reviewed references Statistical methods used to assess reliability: Not available in reviewed references Validity: Content validity: Not available in reviewed references Construct validity: Not available in reviewed references Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): Not available in reviewed references Known Relationships to Other Variables: Subjects with depression score higher on the AS. Diagnostic Sensitivity and Specificity, if applicable (in general population, HD population- premanifest/ manifest, other disease groups): Not available
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: This instrument assesses multiple aspects of apathy and has been used in a variety of neuropsychiatric disorders, and allows for comparison between patient/self and informant reports. Weaknesses: The AS may not discriminate apathy from depression. Availability (copyright): Available freely Special Requirements for administration: None Administration Time: Likely 15-30 minutes Translations available: Available in English, Dutch, German, French, Spanish and multiple other languages.
References:	Key Reference: Starkstein SE, Mayberg HS, Preziosi TJ, Andrezejewski P, Leiguarda R, Robinson RG: Reliability, validity, and clinical correlates of apathy in Parkinson's disease. J Neuropsychiatry Clin Neurosci 1992; 4: 134-139 Other References: Chatterjee A, Anderson KE, Moskowitz CB, Hauser WA, Marder KS: A comparison of self-report and caregiver assessment of depression, apathy, and irritability in Huntington's disease. J Neuropsychiatry Clin Neurosci 2005; 17: 378-383 [scoring of the IS in this article is incorrect!] Starkstein, SE, Migliorelli R, Manes F, Tesón A, Petracca G, Chemerinski E, Sabe L, Leiguarda R: The prevalence and clinical correlates of apathy and irritability in Alzheimer's disease. Eur J Neurology 1995; 2: 540-546

Description of Cambridge Brain Repair Centre (BRC) HD Sleep Questionnaire for HD Common Data Elements

Instrument Name:	Cambridge Brain Repair Centre (BRC) HD Sleep Questionnaire
Classification:	Exploratory
Short Description of Instrument:	Summary/ Overview of Instrument: A newly created instrument specifically designed for use with HD patients, this questionnaire was, in part, based on recent questionnaires used in Parkinson's disease. It contains 45 questions that focus on many aspects of sleep: duration, quality, quality of life, and abnormal sleep behaviors. The authors group the questions into four themed subcategories: quality of sleep, motor activity, abnormal nocturnal behavior and other aspects of disturbed sleep. The authors of this instrument have written comprehensive reviews of the literature in HD and have conducted prior studies of sleep in HD. Construct measured: The total score is conceptualized as a measure of "sleep disturbance", but individual questions address a broad range of sleep problems Generic vs. disease specific: Developed specifically for Huntington's disease populations, though the questions do not specifically mention HD such that it could potentially be studied for possible use in other populations. Intended use of instrument/ purpose of tool: To date the instrument has only been used in one cross-sectional study (primary reference) but the authors suggest that it might be useful in longitudinal studies. It is primarily proposed as a screen for sleep disturbance. Means of administration: Paper and pencil Location of administration: Clinic or home Intended respondent: Patient # of items: 45 # of subscales and names of sub-scales: None
Scoring	Scoring: A total sleep disturbance score is calculated using a number (but not all) of the questions in the questionnaire. The questions that count toward the total score were, in part, selected as they distinguished HD patients from controls in the primary reference study. A scoring sheet is appended to the original paper which identifies the items that count toward the total score as well as the points assigned to various options. Scores range from 0-19 and the authors recommend the following subgroup classifications: normal (0-3), mild (4-6) and significant sleep disturbance (7 and greater) with significant sleep disturbance being the classification that warrants further investigation and/or treatment when used as a screening measure in clinical settings. Standardization of scores to a reference population (z scores, T scores, etc): Insufficient research

	If scores have been standardized to a reference population, indicate frame of reference for scoring (general population, HD subjects, other disease groups, etc). Scoring was developed specifically for HD subjects
Measurements	Type of scale used to describe individual items and total/subscale scores: Some questions are ordinal, others are dichotomous (yes/no)
Psychometric Properties	Reliability: The primary reference study statistically tested for differences in individual questions and total scores between subject group (HD, carers and controls) but no analysis of the psychometric properties of this new instrument was reported to determine its reliability. Test-retest or intra-interview (within rater) reliability (as applicable): Not available Inter-interview (between-rater) reliability (as applicable): Not available Internal consistency: Not available Statistical methods used to assess reliability: Not available Validity: While a number of questions effectively distinguished between an HD group and controls, no analyses of validity were reported in the primary reference study. Content validity: Not available Construct validity: Not available Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): Unknown Known Relationships to Other Variables (e.g. gender, education, age, etc): Unknown Diagnostic Sensitivity and Specificity, if applicable (in general population, HD population- premanifest/ manifest, other disease groups): Not intended for diagnosis of specific sleep disorders, but rather for clinical screening for sleep disturbance in Huntington's disease.
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: Specifically designed to screen for sleep disturbances in HD, simple to administer and brief Weaknesses: Psychometric properties (internal consistency, test-retest reliability, inter-interview reliability, contruct validity) are unknown at present. Other than the primary reference study, no studies to date have used this instrument so that comparisons cannot be m ade with the literature of other populations. The sleep disturbance score, while based on items that distinguished HD patients and controls, may not be an effective measure for outcome studies as many items are dichotomous (yes/no) and do not allow for gradations of severity. Availability (copyright): Free—appended to the original paper. D etails are not specified, so contacting the authors for permission is recommended. Special Requirements for administration: None

Description of Cambridge Brain Repair Centre (BRC) HD Sleep Questionnaire for HD Common Data Elements

	Administration Time: Brief (approximately 5 minutes) Translations available: None known
References:	Key Reference: Goodman AO, Morton AJ, Barker RA. Identifying sleep disturbances in Huntington's disease using a simple disease-focused questionnaire. PLoS Currents 2010; October 15. (Online access: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2957697/)

Description of Florida Obsessive Compulsive Inventory for HD Common Data Elements

Instrument Name:	Florida Obsessive Compulsive Inventory (FOCI)
Classification:	Supplemental
Short Description of Instrument:	Summary/ Overview of Instrument: Self report questionnaire that has Checklist for symptom enumeration and Severity Scale. Construct measured: Obsessions and compulsions Generic vs. disease specific: Generic Intended use of instrument/ purpose of tool: Evaluation of severity of obsessive and compulsive symptoms. Means of administration: Paper and pencil, electronic online version available. Location of administration: Clinic or home Intended respondent: Patient # of items: Checklist (20 items), Severity Scale (5 items). # of subscales and names of sub-scales: 2 – Checklist and Severity Scale
Scoring	Scoring: For Checklist: yes or no responses, no total score, just rating presence of a symptom; For Severity Scale, add total for items 1-5 to get severity measure. Standardization of scores to a reference population (z scores, T scores, etc): Not available. If scores have been standardized to a reference population, indicate frame of reference for scoring (general population, HD subjects, other disease groups, etc). N/A
Measurements	Type of scale used to describe individual items and total/subscale scores: Checklist is nominal (yes/no), Severity Scale is ordinal. If ordinal or continuous, explain if ceiling or floor effects are to be expected if the measure is used in specific HD Subgroups. Long number of questions on checklist may limit utility in patients with more cognitive impairment.
Psychometric Properties	Reliability: Test-retest or intra-interview (within rater) reliability (as applicable): N/A Inter-interview (between-rater) reliability (as applicable): N/A Internal consistency: very good, alpha = 0.89. Statistical methods used to assess reliability: Not available. Validity: Content validity: Not available. Construct validity: Very good; high correlation with Yale Brown Obsessive Compulsive Scale. Not available.

	Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): Aldea et al, 2009- Severity Scale showed lower scores over time in response to an intervention. Known Relationships to Other Variables (e.g. gender, education, age, etc): N/A Diagnostic Sensitivity and Specificity, if applicable (in general population, HD population- premanifest/ manifest, other disease groups):
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: Quick evaluation of both types of OCD symptoms and severity. Weaknesses: Does not measure severity of individual obsessive or compulsive symptoms, only global severity. Availability (copyright): Available for free online but note: The FOCI cannot be reprinted, reproduced or modified without written permission of Dr. Goodman wayne.goodman@mssm.edu. Likewise, those individuals interested in clinical or research use of the FOCI need to obtain permission from Dr. Goodman. http://www.mssm.edu/research/centers/center-of-excellence-for-ocd/rating-scales Special Requirements for administration: None Administration Time: 15-20 minutes Translations available (e.g. Spanish, French, Other languages):
References:	Key Reference: Storch EA, Kaufman DA, Bagner D, Merlo LJ, Shapira NA, Geffken GR, Murphy TK, Goodman WK.J Clin Psychol. 2007 Sep;63(9):851-9. Florida Obsessive-Compulsive Inventory: development, reliability, and validity. Other References: Aldea MA, Geffken GR, Jacob ML, Goodman WK, Storch EA. Further psychometric analysis of the Florida Obsessive-Compulsive Inventory. J Anxiety Disord. 2009 Jan;23(1):124-9.

Description of Frontal Systems Behavior Scale for HD Common Data Elements

Instrument Name:	Frontal Systems Behavior Scale™ (FrSBe™)
Classification:	Supplemental
Short Description of Instrument:	Summary/ Overview of Instrument: Formerly the Frontal Lobe Personality Scale (FLOPS), the FrSBe was designed to identify and quantify behavioral problems associated with frontal lobe dysfunction. This scale assesses behavior related to frontal systems damage. It also quantifies behavioral changes over time by including both baseline (retrospective) and current assessments of behavior. Forms are available for both patient and family member to complete, with separate norms for each informant. There is potential for discrepancy between the information collected from the informant and the participant. Construct measured: Assesses behavior Generic vs. disease specific: Generic Intended use of instrument/ purpose of tool: Cross-sectional or longitudinal assessment of symptoms commonly seen in patients with 'frontal' disorders Means of administration (paper and pencil, computerized): Paper and pencil Location of administration: Clinic or home (self-report) Intended respondent: Patient and/or caregiver # of items: Apathy (14 items), Disinhibition (15 items), Executive Dysfunction (17 items) # of subscales and names of sub-scales: 3 – Apathy, Disinhibition, Executive Dysfunction
Scoring	Scoring: Each item is rated on a 5-point Likert scale. Totals are generated for each subscale and normative data is referenced (based on patient gender, age and education) and standardized T scores are determined (mean: 50, SD:10). Interpretation of results require training and coursework in psychological assessment. Standardization of scores to a reference population (z scores, T scores, etc): Previously validated in patients with a variety of neuropsychiatric disorders. If scores have been standardized to a reference population, indicate frame of reference for scoring (general population, HD subjects, other disease groups, etc). Not available
Measurements	Type of scale used to describe individual items and total/subscale scores: Ordinal
Psychometric Properties	Reliability: Acceptable based on normative sample data (Grace).

	Validity: Construct validity: Reviewed in manual and acceptable. Convergent validity with other behavioral measures was high (NPI, r=.64). Discriminant validity also good (Grace). Feasibility: Informants completing the Family Rating Form should have at least weekly contact with the patient to ensure accurate behavioral observation. Patients must have cognitive capacity to read and complete the form. Factor structure: An exploratory principal component factor analysis using the family version with 324 neurological outpatients (mainly HD, PD and Alzheimer's disease patients) confirmed a factor structure consistent with the subscales originally proposed on theoretical grounds (Stout). Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): This measure was designed in part to assess change over time. Known Relationships to Other Variables: There have been no published reports of patients with manifest HD using the FrSBe, other than the factor analysis referred to above (Stout). In the PREDICT-HD study, 745 mutation-positive subjects, 163 mutation-negative control subjects and their companions completed subject and family versions respectively of the FrSBE (Duff). Mutation-positive subjects reported more frontal behaviours than mutation-negative controls, even though most subjects were more than 10 years from predicted motor onset. However, discrepancies between self-report and companion scores suggested impaired insight in those closest to predicted disease onset. In non-HD studies, Apathy and Executive Dysfunction subscale scores are correlated with IADL's (Grace), and the Disinhibition scale score is strongly related to caregiver burden (Grace). Diagnostic Sensitivity and Specificity, if applicable: N/A
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: Assesses multiple domains of frontal lobe functioning and allows for comparison of premorbid behavior with current status. Also allows for comparison between patient and caregiver reports. Weaknesses: Large number of items may be a problem for more cognitively impaired subjects. Scoring requires normative database and understanding of T scores. Availability (copyright): Through Psychological Assessment Resources, Inc. This measure is copyrighted and cannot be reproduced without permission. http://www4.parinc.com/Products/Product.aspx?ProductID=FRSBE Special Requirements for administration: None Administration Time: The scale takes 10 minutes to administer and 10-15 minutes to score.

	Translations available: Not available.
References:	Key Reference: Grace J, Malloy PF. Frontal Systems Behavior Scale Professional Manual. Lutz, FL: Psychological Assessment Resources, Inc. 2001. Other References: Stout JC, Ready RE, Grace J, Malloy PF, Paulsen JS. Factor Analysis of the Frontal Systems Behavior Scale (FrSBe). *Assessment* 2003; 10: 79-85. Duff K, Paulsen JS, Beglinger LJ *et al*. 'Frontal Behaviors' before the diagnosis of Huntington's disease and their relationship to markers of disease progression: evidence of early lack of awareness. *J Neuropsychiatry Clin Neurosci*. 2010; 22(2): 196-207.

Description of Hospital and Anxiety Depression Scale for HD Common Data Elements

Instrument Name:	Hospital Anxiety and Depression Scale (HADS)
Classification:	Supplemental
Short Description of Instrument:	Summary/ Overview of Instrument: The scale was designed to screen for mood disorders in general (non-psychiatric) medical outpatients. It focuses on subjective disturbances of mood rather than physical signs, and aims at distinguishing depression from anxiety. Compared to other instruments scales, it focuses on emotional aspects of anxiety disturbances, as opposed to somatic and cognitive symptoms. Construct measured: Anxiety and depression Generic vs. disease specific: Generic Intended use of instrument/ purpose of tool: Clinical Trials, Observational Studies Means of administration: Self- administered Location of administration: Clinic, home, telephone Intended respondent: Patient # of items: 14 – Anxiety (7 items), Depression (7 items) # of subscales and names of sub-scales: 2 – Anxiety, Depression
Scoring	Scoring: Items are rated on a 4 -point Likert-type scale ranging from 0 t o 3, generating a scale range of 0 to 42 points, with higher scores representing greater symptom severity. The anxiety subscale has 3 items that refer to panic and 4 to generalized anxiety. Add the A questions to get a score for anxiety and the D questions for depression. Scores of 0-7 indicate normal levels of anxiety and depression; 8-10 indicate borderline abnormal anxiety and depression levels and 11-21 suggest abnormal levels of anxiety and depression. Standardization of scores to a reference population (z scores, T scores, etc): Not available. If scores have been standardized to a reference population, indicate frame of reference for scoring (general population, HD subjects, other disease groups, etc). N/A
Measurements	Type of scale used to describe individual items and total/subscale scores: Ordinal
Psychometric Properties	Reliability: Internal consistency described for patients with cancer (Moorey et al 1991): Anxiety subscale Cronbach's alpha = 0.93; Depression subscale alpha= 0.9. In healthy UK sample, internal consistency for Anxiety, Depression and Total scores

	were 0.82, 0.77 and 0.86 respectively (Crawford et al 2001). Test-retest reliability for healthy sample: correlation for Depression scale= 0.92; Anxiety subscale 0.89 (Snaith & Zigmond, test manual) Validity: Concurrent validity established in a number of studies (see Snaith & Zigmond, test manual). Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): Not available. Known Relationships to Other Variables: HADS depression scores differentiate between patients taking/ not taking antidepressants, and male patients and older patents at time of diagnosis had higher HADS depression scores; HADS anxiety scores differentiated between patients with and without a psychiatric history and those taking/ not taking antidepressants (Wicks et al 2007). HADS Depression scores correlated with limb impairment, overall disease severity scores and, also with Anxiety scores with impairment on domains of the Sickness Impact Scale (Goldstein et al 1998). Anxiety and depression subscale scores correlated with subscales of the Sickness Impact Scale; Depression subscale scores correlated with speech and mobility scores on the Barthel Index and Anxiety scores correlated with Barthel speech items (Hogg et al 1994). Diagnostic Sensitivity and Specificity, if applicable (in general population, HD population- premanifest/ manifest, other disease groups): Not available.
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: Serves as a good screening measure. Has been widely used. Relatively simple to complete. Weaknesses: This scale is not designed for HD; however, it is a quick screen. Requires insight to provide accurate reflection. No proxy verification. Availability (copyright): www.gl-assessment.co.uk Special Requirements for administration: None Administration time: About 2-5 minutes Translations available: Over 80 translations available
References:	Key Reference: Zigmond AS and Snaith RP: The Hospital Anxiety And Depression Scale. *Acta Psychiatr Scand* 1983, 67:361-70. Other References: Crawford, J. R., Henry, J. D., Crombie, C. & Taylor, E. P. Normative data for the HADS from a large non-clinical sample. British Journal of Clinical Psychology 2001; 40: 429–434.

Description of Irritability Scale for HD Common Data Elements

Instrument Name:	Irritability Scale
Classification:	Core
Short Description of Instrument:	Summary/ Overview of Instrument: The Irritability Scale consists of 14 items regarding different dimensions of apathetic behavior. The score for each item ranges from 0 to 3. Rating of each item is based on a semi-structured interview. Construct measured: Irritability Generic vs. disease specific: Generic Intended use of instrument/ purpose of tool: Assessment of severity of irritability Means of administration: Paper and Pencil Location of administration: Clinic or Home Intended respondent: Patient and Caregiver/Informant # of items: 14 # of subscales and names of sub-scales: None
Scoring	Scoring: Ratings should be based on both verbal and non-verbal information of the past 4 weeks. For each item ratings should be judged: 4 possible responses for each question: 'not at all', 'slightly', 'somewhat', 'a lot'. Standardization of scores to a reference population (z scores, T scores, etc): Using ROC analysis, a score of ≥ 14 points on the IS-self was identified as a robust indicator for irritability (Reedeker, submitted). The IS cut-off score of≥ 14 points yielded an acceptable sensitivity and high specificity for all three cut-off points. If scores have been standardized to a reference population, indicate frame of reference for scoring: First-degree non-carriers
Measurements	Type of scale used to describe individual items and total/subscale scores: Continuous
Psychometric Properties	Reliability: Test-retest or intra-interview (within rater) reliability (as applicable): Not available Inter-interview (between-rater) reliability (as applicable): Interrater agreement for the presence of irritability above a median score in a HD population ranged from poor for the most cognitively impaired subjects to good for the less cognitively impaired subjects. Agreement between IS-self and IS-inf scores was assessed using one-way random, single measure intraclass correlation coefficients (ICCs). The overall ICC for IS-self and I S-inf scores was 0.61 (95% CI = 0.50 – 0.72, $p < 0.001$) (Reedeker, submitted).

	Internal consistency: The Cronbach's alphas were 0.90 for the IS-self and 0.93 for the IS-inf (Reedeker, submitted). Statistical methods used to assess reliability: Kruskal-Wallis tests were conducted to compare IS-self and IS-inf scores (Reedeker, submitted). Validity: Content validity: Not available Construct validity: Not available Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): The IS cut-off score of ≥ 14 points yielded an acceptable sensitivity and high specificity for all three cut-off points. Known Relationships to Other Variables: Use of medication. Diagnostic Sensitivity and Specificity, if applicable (in general population, HD population- premanifest/ manifest, other disease groups): Not available.
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: Self report and Informant report Weaknesses: Presence and severity are roughly scored on a 5-points likert-scale. The cut-off score is not validated (no external validity). Availability (copyright): Available freely Special Requirements for administration: None Administration Time: Likely 15-30 minutes Translations available: Available in English, Dutch
References:	Key Reference: Chatterjee A, Anderson KE, Moskowitz CB, Hauser WA, Marder KS: A comparison of self-report and caregiver assessment of depression, apathy, and irritability in Huntington's disease. J Neuropsychiatry Clin Neurosci 2005; 17: 378-383 [scoring of the IS in this article is incorrect] Other References: Klöppel S, Stonnington CM, Petrovic P, Mobbs D, Tüscher O, Craufurd D, Tabrizi SJ, Frackowiak RSJ: Irritability in pre-clinical Huntington's disease. Neuropsychologia 2010; 48: 549-557 Reedeker W, Bouwens JA, Giltay EJ, Le Mair SE, Roos RAC, van der Mast RC, van Duijn E: Irritability in Huntington's disease. submitted

Description of Padua Inventory- Washington State University Revision for HD Common Data Elements

Instrument Name:	Padua Inventory – Washington State University Revision
Classification:	Supplemental
Short Description of Instrument:	Summary/ Overview of Instrument: 39 item inventory of 5 factors in OCD. Construct measured: Obsessive Compulsive Disorder (OCD). Note this revision was done to ensure the scale was specific to OCD, and did not measure worry, which was an issue with the original Padua Inventory. Generic vs. disease specific: Generic OCD scale, not specific to HD. Intended use of instrument/ purpose of tool: Not specified Means of administration: Paper and Pencil Location of administration: Clinic or home. Intended respondent: Patient # of items: 39 items # of subscales and names of sub-scales: 5 – Contamination Obsessions and Washing Compulsions; Dressing/Grooming Compulsions; Checking Compulsions; Obsessional Thoughts of Harm to Self/Others; Obsessional Impulses to Harm Self/Others
Scoring	Scoring (include reference to detailed scoring instructions, including calculation of a total score and subscale scores, and any limitations of scale or scoring posed by item nonresponse): Each item is rated on a 5 -point scale according to the degree of disturbance caused by the thought or behavior (0 = "not at all" to 4 = "very much"). S ubscales are simply scored by summing scores for all items included in the subscale. Standardization of scores to a reference population (z scores, T scores, etc): Burns et al, 1996 reference provides a t able with normative data on the PI--WSUR for the sample of 5010 individuals. If scores have been standardized to a reference population, indicate frame of reference for scoring (general population, HD subjects, other disease groups, etc). General population
Measurements	Type of scale used to describe individual items and total/subscale scores: Ordinal
Psychometric Properties	Reliability: Test-retest or intra-interview (within rater) reliability (as applicable): test-retest correlation for the total PI-WSUR was 0.76 with the values for the 5 subscales varying from 0.61 for the OTAHSO subscale and 0.84 for the OITHSO subscale. Inter-interview (between-rater) reliability (as applicable): Not applicable- only patient reports symptoms on the scale. Internal consistency: Not available.

	Statistical methods used to assess reliability: test – retest correlation with Bonferroni correction. Validity: Several articles available on validity, (e.g. Burns, Formea, Keortge & Sternberger, 1995; Sanavio, 1988; Sternberger & Burns, 1990, 1991; Van Oppen, 1992). This instrument has not been validated in HD. Content validity: Principal component analysis Construct validity: correlations showed stronger relationship to items within scale than to items on a Worry scale (Burns et al, 1996) Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): Not studied Known Relationships to Other Variables (e.g. gender, education, age, etc): There is some gender difference on subscales with Women scoring significantly higher than men on the OTAHSO (3%), COWC (< 1%), the DRGRC (< 1%) subscales, and with men scoring higher than women on the OITHSO (3%) subscales. Diagnostic Sensitivity and Specificity, if applicable (in general population, HD population- premanifest/ manifest, other disease groups): Would be best for premanifest and early symptomatic groups; length of scale may limit use in more cognitively impaired populations.
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: Self report Weaknesses: May be issues about whether constructs hold. See Sascha Gönner, Willi Ecker, and Rainer Leonhart. The Padua Inventory: Do Revisions Need Revision? Assessment March 2010 17: 89-106. Availability (copyright): Available freely Special Requirements for administration: None Administration Time: Likely 15-30 minutes Translations available: Available in German, French, Persian, Spanish and multiple other languages.
References:	Key Reference: Burns GL, Keortge SG, Formea GM, Sternberger LG.Behav Res Ther. 1996 Feb;34(2):163-73.Revision of the Padua Inventory of obsessive compulsive disorder symptoms: distinctions between worry, obsessions, and compulsions. Other References: Burns, Formea, Keortge & Sternberger, 1995; Sanavio, 1988; Sternberger & Burns, 1990, 1991; Van Oppen, 1992

Description of Pittsburgh Sleep Quality Index for HD Common Data Elements

Instrument Name:	Pittsburgh Sleep Quality Index (PSQI)
Classification:	Supplemental
Short Description of Instrument:	Summary/ Overview of Instrument: A self-rated questionnaire that primarily assesses nighttime sleep problems. It focuses on sleep experiences over the past month. It has 19 self-rated questions and 5 additional questions for a bed partner or roommate. Construct measured: Sleep quality, sleep habits and sleep disturbances. Seven component scores: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, daytime dysfunction Generic vs. disease specific: Has been used in many different populations; it is not disease specific. Intended use of instrument/ purpose of tool: Can be used as a screening instrument for nighttime sleep disturbance or for clinical studies. It cannot be used to diagnose specific sleep disorders, but instead may help distinguish "good" versus "poor" sleepers. Means of administration: Paper and Pencil Location of administration: Clinic, Home Intended respondent: Patient (with 5 supplemental questions for a bed partner or roommate) # of items: 24 (19 self-rated items, and 5 supplemental items to be rated by a bed partner or roommate) # of subscales and names of sub-scales: 7 – Subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, daytime dysfunction
Scoring	Scoring: Seven component scores are calculated, each scored from 0 to 3, the total score ranges from 0 to 21, with higher scores indicating more severe sleep problems in many areas. Scoring requires following closely a complex algorithm and is not a simple summation of answers. A cutoff of 5/6 for the total score is used in general populations to distinguish between "good" and "poor" sleepers. Scoring can be time consuming. Standardization of scores to a reference population (z scores, T scores, etc): The PSQI scores are not standardized to a particular population but this instrument has been used in many different populations If scores have been standardized to a reference population, indicate frame of reference for scoring (general population, HD subjects, other disease groups, etc). (See above.) While the scores are not standardized to a particular reference population, the cutoff of 5/6 for "good" versus "poor" sleepers was developed from

	general population samples and thus it may not carry over as the best screening cutoff for specific populations such as HD subjects.
Measurements	Type of scale used to describe individual items and total/subscale scores (nominal, ordinal, or [essentially] continuous): Ordinal If ordinal or continuous, explain if ceiling or floor effects are to be expected if the measure is used in specific HD Subgroups. Unknown due to insufficient use in HD populations
Psychometric Properties	Reliability: Test-retest or intra-interview (within rater) reliability (as applicable): The Pearson correlation coefficient for test-retest reliability in a non-HD population was 0.87 and is stable over time (Högl et al., 2010). Inter-interview (between-rater) reliability (as applicable): not available in reviewed references Internal consistency: A Cronbach's alpha of 0.72 was found in a one HD study (Aziz et al. 2010); Cronbach's alphas of between 0.80 and 0.83 have been reported for the PSQI in different studies of non-HD populations Statistical methods used to assess reliability: (as above) Validity: Content validity: Not available in reviewed references Construct validity: In the original study, the instrument successfully discriminated between clinical populations of good sleepers (normal healthy controls) and patients from a sleep evaluation clinic. In a HD sample, the measure correlated highly with another sleep measure, the SCOPA-SLEEP Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): Unknown Known Relationships to Other Variables: Not available in reviewed references Diagnostic Sensitivity and Specificity, if applicable (in general population, HD population- premanifest/ manifest, other disease groups): Not useful for diagnosis of sleep disorders
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: Extensive literature of its use in other populations. Includes a number of questions for bed partners (though these are not comprehensive and are not used in the scoring.) Weaknesses: Primarily assesses nighttime sleep problems; wording might be confusing; does not directly address changes in circadian rhythms (sleep time shifting to the day and awake all night) that clinically is often observed in HD patients; the wording of certain questions is likely problematic for patients with HD and measures other constructs such as mood or motivation, e.g., "during the past month, how much of a problem has it been for you to keep up enough enthusiasm to get things done." One study (Aziz et al., 2010) in an HD population found the

	SCOPA-S more internally consistent, and much easier to score and use than the PSQI. The scoring algorithm is unusually complex. Availability (copyright): Copyrighted. T he author, Dr. Buysse, should be contacted for permission to use the instrument. If used for commercial research purposes, the Office of Technology Management at the University of Pittsburgh (412-648-2206) must be contacted for licensing information. http://www.sleep.pitt.edu/content.asp?id=1484&subid=2316 Special Requirements for administration: None Administration Time: Likely 5-10 minutes Translations available: It has been translated in over 56 languages, according to the University of Pittsburgh Sleep Medicine Institute website where these versions can be requested.
References:	Key Reference: Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. Psychiatry Research 1989;28:193-213. Other References: Videnovic A, Leurgans S, Fan W, Jaglin J, Shannon K. Daytime somnolence and nocturnal sleep disturbances in Huntington disease. Parkinsonism and Related Disorders 2009;15:471-4. Aziz NA, Anguelova GV, Marinus J, Lammers GJ, Roos RAC. Sleep and Circadian rhythm alterations correlate with depression and cognitive impairment in Huntington's disease. Parkinsonism and Related Disorders 2010;16:345-50.

Description of Problem Behaviours Assessment for HD- short version for HD Common Data Elements

Instrument Name:	Problem Behaviours Assessment for HD – short version (PBA-s)
Classification:	Core
Short Description of Instrument:	Summary/ Overview of Instrument: Brief semi-structured interview covering the most common behavioural and psychiatric manifestations of HD Construct measured: The interview is not restricted to a single construct, but rather covers several broad symptom domains (affect, irritability, loss of motivation, perseverative phenomena and psychotic symptoms) relevant to HD. Generic vs. disease specific: Specific to HD Intended use of instrument/ purpose of tool: Cross-sectional or longitudinal studies. The interview is focused on symptoms rather than diagnoses and does not make assumptions about the diagnostic significance of these symptoms in the presence of organic brain disease (i.e.HD). Means of administration: Face to face semi-structured interview. The suggested approach is to complete the PBA-s with the companion and participant together. After completeing the scale the interviewer should speak with the companion without the participant present. Location of administration: Clinic or home Intended respondent: Patient and knowledgeable informant (e.g. spouse or caregiver) together; the informant should be briefly re-interviewed afterwards to elicit any additional information which could not be discussed openly in presence of the patient. # of items: 11 items (low mood, suicidal ideation, anxiety, irritability, anger/aggressive behaviour, loss of motivation, perseverative thinking or behaviour, obsessive-compulsive behaviours, paranoid thinking, hallucinations, behaviour suggestive of disorientation) # of subscales and names of sub-scales: The PBA-s has not been formally divided up into subscales; however, principal components analysis of data obtained with the original 40-item Problem Behaviours Assessment for HD (PBA-HD) in both English (Craufurd *et al*., 2001)and Dutch (Kingma *et al*., 2007) translations identified three main factors corresponding to affective symptoms, irritability and apathy respectively. It would therefore be reasonable to add the scores for low mood, suicidal ideation and anxiety to create a single 'affect' score, and to add the scores for irritability and anger to create a composite 'irritability and aggression' score, as was done in the TRACK-HD study (Tabrizi et al., 2009).

Description of Problem Behaviours Assessment for HD- short version for HD Common Data Elements

Scoring	Scoring: Each symptom is rated for severity on a 5-point scale according to detailed scoring criteria which roughly correspond to the following: 0 = "not at all"; 1 = trivial; 2 = mild; 3 = moderate (disrupting everyday activities) and 4 = severe or intolerable. Each symptom is also scored for frequency on a 5-point scale as follows: 0 = symptom absent; 1 = less than once weekly; 2 = at least once a week; 3 = most days (up to and including some part of every day); and 4 = all day, every day. Severity and frequency scores are multiplied to produce an overall 'PBA score' for each symptom. Although it would be possible to sum the individual PBA symptom scores to derive an overall total score, it is unlikely that this is very meaningful. Standardization of scores to a reference population (z scores, T scores, etc): Not applicable
Measurements	Type of scale used to describe individual items and total/subscale scores (nominal, ordinal, or [essentially] continuous): Ordinal If ordinal or continuous, explain if ceiling or floor effects are to be expected if the measure is used in specific HD Subgroups: Significant ceiling effects were observed in the TRACK-HD study (Tabrizi *et al*.); for most PBA symptoms, between 25% - 50% of pre-symptomatic HD mutation carriers scored zero for severity on the more common PBA symptoms (e.g. low mood, anxiety or irritability) with a much higher proportion scoring zero for less common items such as aggression, suicidal ideation or hallucinations.
Psychometric Properties	Reliability: Test-retest or intra-interview (within rater) reliability (as applicable): N/A Inter-interview (between-rater) reliability (as applicable): inter-rater reliability was measured in the TRACK-HD study by video-recording interviews and re-scoring of a random sample by an expert rater. Overall unweighted kappa scores for the severity ratings were 0.70 and for the frequency ratings, 0.77. Kappa scores in the range 0.61 – 0.80 are usually considered to represent 'substantial agreement'. Internal consistency: N/A Statistical methods used to assess reliability: Cohen's kappa. Validity: Difficult to assess because of the lack of an alternative 'gold standard' measure. The selection of symptoms for inclusion in the measure, and the development of the detailed scoring criteria used, were carried out by a group of experts from the EHDN Behavioural Phenotype Working Group. The range of symptoms covered by the measure is very similar to the behavioural section of the UHDRS, and the PBA-s is probably best thought of as the latest elaboration of the UHDRS behavioural interview. Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): The longitudinal rate of change in PBA-s apathy scores over 24 months' follow-up was significantly greater in manifest HD subjects than in controls (Tabrizi et al., manuscript under review). Analysis of the TRACK-HD 36-

	month longitudinal data is in progress. Known Relationships to Other Variables: A paper describing relationships between PBA-s scores and those obtained using other symptom-specific measures in the TRACK-HD study is in preparation Diagnostic Sensitivity and Specificity, if applicable (in general population, HD population- premanifest/ manifest, other disease groups): Developed for use with manifest HD population
Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: Obvious face validity (semi-structured interview covering the same core behavioural symptoms as the UHDRS behavioural section); demonstrated excellent inter-rater reliability in TRACK-HD study; suitable for use by all professionals familiar with HD (physician, psychologist, study nurse, etc.). Weaknesses: Investment of staff time for 25-minute interview; some training required. Availability (copyright): Available freely Special Requirements for administration: None Administration Time: 20-25 minutes Translations available: Available in English, Dutch, French, German, Norwegian and Spanish versions. Translation into other European languages in progress.
References:	Key Reference: A paper describing the reliability of the PBA-s (in English, Dutch and French) based on data from the TRACK-HD study is in preparation and will be submitted shortly. Other References: Craufurd D, Thompson J, Snowden JS. Behavioural changes in Huntington's disease: the Problem Behaviours Assessment. *Neuropsychiatry, Neuropsychology and Behavioural Neurology* 2001; 14 (4): 219-226. Kingma EM, van Duijn E, Timman R, van der Mast RC, Roos RAC. Behavioural problems in Huntington's disease using the Problem Behaviours Assessment. *General Hospital Psychiatry* 2008; 30: 155-161. Tabrizi SJ, Langbehn DR, Leavitt BR, Roos RAC, Durr A, Craufurd D, Kennard C, Hicks SL, Fox NC, Scahill RI, Borowsky B, Tobin A, Rosas HD, Johnson H, Reilmann R, Landwehrmeyer GB, Stout JC and the TRACK-HD Investigators. Biological and clinical manifestations of Huntington's disease before and after diagnosis – the TRACK-HD study. *Lancet Neurology* 2009; 8 (9): 791-801. epub 2009 July 29.

Description of Scale for Outcomes of Parkinson's disease-Sleep for HD Common Data Elements

Instrument Name:	Scale for Outcomes of Parkinson's disease – Sleep [SCOPA-SLEEP (SCOPA-S)]
Classification:	Supplemental
Short Description of Instrument:	Summary/ Overview of Instrument: The SCOPA-S (Scale for Outcomes of Parkinson's disease-Sleep) is a self-rated questionnaire developed originally for research in Parkinson's disease that addresses both nighttime sleep problems and daytime sleepiness. Its questions focus on experiences over the past month. It has two questions on the use of sleeping medications, one question on an overall global perception of sleep at night, 6 questions on sleep patterns at night and 6 questions on sleep in the day. Construct measured: Two constructs are measured by its two subscales: 1.) Nighttime sleep problems (NS), and 2.) Daytime sleepiness (DS) Generic vs. disease specific: Developed for use in Parkinson's disease research, but its questions do not refer at all to Parkinson's disease with the original intent for it to be potentially useful across other conditions. It has been used in one study of Huntington's disease patients and controls. Intended use of instrument/ purpose of tool: For research or clinical purposes—it could be used as a screening instrument or for rating severity of nighttime sleep disturbances or daytime sleepiness. Means of administration: Paper and Pencil Location of administration: Clinic or home Intended respondent: Patient # of items: 14 (2 questions on use of sleeping tablets, 1 global perception of sleep at night question, and the 11 items of the two subscales # of subscales and names of sub-scales: 2 – Nighttime sleep problems (NS:5 items; score range:0-15), Daytime sleepiness (DS:6 items; score range 0-18)
Scoring	Scoring: With the exception of the one question on use of sleep medications in which the names of medications are listed with doses and frequency of use, all 13 other items are Likert-type, with 4 to seven options such as "not at all," "a little," "quite a bit," and "a lot." All 11 questions from the two subscales, NS and DS have 4 options which are scored 0-3. Subscale totals are easily calculated though summing the totals of all items in that subscale. Standardization of scores to a reference population (z scores, T scores, etc): Insufficient research data to convert subscale scores to standardized scores, either with reference to normal population or manifest HD patients, presymptomatic gene carriers or controls. If scores have been standardized to a reference population, indicate frame of reference for scoring (general population, HD subjects, other disease groups, etc). Not available.

Description of Scale for Outcomes of Parkinson's disease-Sleep for HD Common Data Elements

Measurements	Type of scale used to describe individual items and total/subscale scores (nominal, ordinal, or [essentially] continuous): Ordinal If ordinal or continuous, explain if ceiling or floor effects are to be expected if the measure is used in specific HD Subgroups. Unknown. (One study of its use in HD (Aziz, Anguelova, Marinus et al. 2010) used the measure in HD patients, premanifest mutation carriers, bed partners and controls; the authors did not report ceiling or floor effects in their small samples.)
Psychometric Properties	Reliability: Test-retest or intra-interview (within rater) reliability (as applicable): not assessed in HD populations, but good test-retest reliability in Parkinson's disease population (intraclass correlation coefficients for NS 0.94, and DS 0.89 (Marinus et al, 2003) Inter-interview (between-rater) reliability (as applicable): Not assessed in HD populations Internal consistency: Cronbach's alpha for SCOPA-NS 0.89 and for SCOPA-DS 0.85 in HD sample (Aziz et al.2010); in PD sample these values were 0.88 and 0.91. Statistical methods used to assess reliability: (as above) Validity: Content validity: Not reported in reviewed references Construct validity: found to correlate highly with other scales that measure similar constructs: in HD subjects the NS subscale correlated highly with the Pittsburgh Sleep Quality Index (r=0.77) in a small HD sample (Aziz et al., 2010) and the DS subscale correlated highly with a measure of daytime sleepiness, the Epworth Sleepiness Scale (r=0.75). Sensitivity to Change/ Ability to Detect Change (over time or in response to an intervention): Unknown in HD populations Known Relationships to Other Variables (e.g. gender, education, age, etc): No known relationship to demographic variables in HD due to paucity of data using this instrument. Diagnostic Sensitivity and Specificity, if applicable (in general population, HD population- premanifest/ manifest, other disease groups): Not appropriate for diagnosis of sleep disorders; rather, it is useful for screening and for measurement of the constructs of nighttime sleep problems and daytime sleepiness

Description of Scale for Outcomes of Parkinson's disease-Sleep for HD Common Data Elements

Rationale/ Justification (include any information on language and countries/ cultures/ ethnic groups where tested)	Strengths: Easy to administer and to score. The two subscales have been shown in other populations to each measure one factor. The questions are clear and it is quite comprehensive. One small study in HD (Aziz et al 2010) found it more reliable in HD patients than the Pittsburgh Sleep Quality Index or the Epworth Sleepiness Scale. Weaknesses: Limited use to date in the sleep literature of other populations compared to other measures; does not directly address changes in circadian rhythms (sleep time shifting to the day and awake all night) that clinically is often observed in HD patients. Aziz et al., 2010 suggested that one item in the DS subscale "falling asleep while talking" may not be suitable for HD populations as it was not endorsed by any of their subjects. Availability (copyright): Free, with permission of the authors, to all researchers whether nonprofit or commercial. http://www.scopa-propark.eu/index.php?page=1&navRight=3&doc=3&group=Yes&taal=eng&language=eng&show=yes Special Requirements for administration: None Administration Time: Short (less than 5 minutes) Translations available: A Spanish translation was studied in a population of Parkinson's disease patients. Unclear if other translations are available.
References:	Key Reference: Marinus J, Visser M, van Hilten JJ, Lammers GJ, Stieggelbout AM. Assessment of Sleep and Sleepiness in Parkinson Disease. Sleep 2003; 26:1049-54. Other References: Aziz NA, Anguelova GV, Marinus J, Lammers GJ, Roos RAC. Sleep and Circadian rhythm alterations correlate with depression and cognitive impairment in Huntington's disease. Parkinsonism and Related Disorders 2010;16:345-50. Hogl B, Arnulf I, Comella C, et al. Scales to Assess Sleep Impairment in Parkinson's Disease:Critique and Recommendations. Movement Disorders 2010;25:2704-16.