OPEN ACCESS

REVIEW

Outcome measurement in functional neurological disorder: a systematic review and recommendations

Susannah Pick ,[1] David G Anderson,[2] Ali A Asadi-Pooya ,[3,4] Selma Aybek ,[5] Gaston Baslet,[6] Bastiaan R Bloem ,[7] Abigail Bradley-Westguard, Richard J Brown ,[8] Alan J Carson ,[9] Trudie Chalder ,[1] Maria Damianova,[2] Anthony S David ,[10] Mark J Edwards ,[11] Steven A Epstein,[12] Alberto J Espay ,[13] Béatrice Garcin ,[14] Laura H Goldstein ,[1] Mark Hallett ,[15] Joseph Jankovic ,[16] Eileen M Joyce ,[17] Richard A Kanaan ,[18] Roxanne C Keynejad ,[1] Kasia Kozlowska ,[19] Kathrin LaFaver ,[20] W Curt LaFrance, Jr. ,[21] Anthony E Lang ,[22] Alex Lehn ,[23] Sarah Lidstone ,[22] Carine W Maurer ,[24] Bridget Mildon ,[25] Francesca Morgante ,[11] Lorna Myers ,[26] Clare Nicholson ,[27] Glenn Nielsen ,[11] David L Perez ,[28] Stoyan Popkirov ,[29] Markus Reuber ,[30] Karen S Rommelfanger ,[31] Petra Schwingenshuh ,[32] Tereza Serranova ,[33] Paul Shotbolt ,[1] Glenn T Stebbins ,[34] Jon Stone ,[9] Marina AJ Tijssen ,[35] Michele Tinazzi ,[36] Timothy R Nicholson [1]

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/jnnp-2019-322180).

For numbered affiliations see end of article.

Correspondence to
Dr Timothy R Nicholson, King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, UK; timothy.nicholson@kcl.ac.uk

Received 7 October 2019
Revised 10 December 2019
Accepted 20 December 2019

To cite: Pick S, Anderson DG, Asadi-Pooya AA, *et al*. *J Neurol Neurosurg Psychiatry* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/jnnp-2019-322180

ABSTRACT

Objectives We aimed to identify existing outcome measures for functional neurological disorder (FND), to inform the development of recommendations and to guide future research on FND outcomes.

Methods A systematic review was conducted to identify existing FND-specific outcome measures and the most common measurement domains and measures in previous treatment studies. Searches of Embase, MEDLINE and PsycINFO were conducted between January 1965 and June 2019. The findings were discussed during two international meetings of the FND-Core Outcome Measures group.

Results Five FND-specific measures were identified—three clinician-rated and two patient-rated—but their measurement properties have not been rigorously evaluated. No single measure was identified for use across the range of FND symptoms in adults. Across randomised controlled trials (k=40) and observational treatment studies (k=40), outcome measures most often assessed core FND symptom change. Other domains measured commonly were additional physical and psychological symptoms, life impact (ie, quality of life, disability and general functioning) and health economics/cost–utility (eg, healthcare resource use and quality-adjusted life years).

Conclusions There are few well-validated FND-specific outcome measures. Thus, at present, we recommend that existing outcome measures, known to be reliable, valid and responsive in FND or closely related populations, are used to capture key outcome domains. Increased consistency in outcome measurement will facilitate comparison of treatment effects across FND symptom types and treatment modalities. Future work needs to more rigorously validate outcome measures used in this population.

INTRODUCTION

Functional neurological disorder (FND or conversion disorder) is defined by motor and sensory symptoms (eg, tremor, dystonia, limb weakness, numbness and seizures) that demonstrate clinical features incompatible with other neurological/medical diagnoses and that are associated with significant distress and functional impairment.[1–3] Until recently, there were few controlled treatment trials and little research effort into outcome measurement in FND. However, there has been a marked increase in the number and quality of intervention studies, so it is now critical to optimise outcome measures for this population, to facilitate comparison of the effectiveness and efficacy of different treatment modalities for specific FND symptoms.

FND has features that make decisions regarding outcome measurement particularly complex.[4 5] These include heterogeneity and variability of symptoms and the marked influence of attention, beliefs and expectations.[6–8] Discrepancy between objective measures and patients' subjective experiences of symptoms can also be a prominent feature.[9] These aspects of FND potentially make objective 'snapshot' measures (eg, clinician-rated scales and objective performance tests) less reliable and valid. They also suggest that patient-rated outcomes may be particularly meaningful in this population.

A wide range of additional physical symptoms (eg, fatigue, pain, sleep disturbance, gastrointestinal and urological problems) are common in people with FND, typically associated with reduced quality of life and greater disability.[10 11] More generally, elevated physical symptom burden is associated

with significant disability, role impairments and high healthcare use.[12 13] Many individuals with FND report significant psychological symptoms/comorbidity (eg, anxiety, depression and dissociation), potentially unhelpful coping behaviours and illness beliefs, and altered emotional processing, which are associated with symptom severity, poorer outcomes and diminished quality of life.[10 11 14–23] These factors, therefore, may be particularly relevant to psychosocial outcomes in FND.[24 25]

There is clearly a wide range of outcome domains relevant to FND, and a key challenge is to determine which are most important to capture. In evidence-based medicine, there has been an increasing emphasis on consistency in outcome measurement and the Core Outcome Measures in Effectiveness Trials (COMET, http://www.comet-initiative.org/) initiative facilitates the development of core outcome measure sets across a range of physical and mental health disorders. The COMET collaboration recommends five key outcome domains, including core physiological/clinical symptoms, life impact (ie, quality of life, functioning or participation and subjective health perception), resource use (health and social), adverse events and mortality. Another initiative, supported by the National Institutes of Health, involves development of 'common data elements', which currently includes core measures for neurological disorders (https://www.nlm.nih.gov/cde/summary_table_1.html).

An important step towards developing a core outcome measure set is to review measures currently used in outcome studies,[26] but there has been no previous review of outcome measures in FND. We aimed to systematically review the currently available FND-specific outcome measures and to identify those measures used most frequently in FND intervention studies to date. We sought to review the outcome domains and measures used in FND treatment research and their quality, not to evaluate the treatments, outcome data or methodologies used. Finally, we aimed to integrate these findings with conclusions derived during two international expert consensus meetings of the Functional Neurological Disorder–Core Outcome Measures (FND-COM) group (http://www.comet-initiative.org/studies/details/951), with the aim of developing recommendations for outcome measurement in future studies.

METHODS

The systematic review was conducted in two parts to identify

1. Publications describing existing FND-specific outcome measures.
2. Randomised controlled trials (RCTs) and prospective intervention studies in FND.

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were followed.[27]

Search strategy

PubMed (MEDLINE), Embase and PsycINFO were searched for both parts of the review. Trial registry websites (eg, clinicaltrials.gov and ISRTCN.com) were also searched to identify intervention studies. The search terms are detailed in box 1. The terms were searched in the abstract, keyword and title fields, mapped to Medical Subject Headings. Reference lists were searched manually for additional sources. Members of the FND-COM group were consulted for additional studies.

Eligibility criteria

Articles written in English and published in peer-reviewed journals between January 1965 and June 2019 were included. We included registered in progress trials, for part 2. Articles referring to core sensorimotor FND symptoms (including seizures) were included, whilst those referring to other less common symptoms (ie, speech and swallowing and cognitive) were excluded.[5]

Box 1 Search terms

Primary terms: 'conversion disorder' OR 'psychogenic' OR 'non*epileptic' OR 'hysteri*' OR 'functional neurological' OR 'functional movement' OR 'functional motor'

Combined with AND:

Part 1: 'outcome*' OR 'measurement instrument' OR 'assessment' OR 'scale*' OR 'outcome measure*' OR 'questionnaire'

Part 2: 'trial' OR 'treatment*' OR 'intervention' OR 'treatment outcome*' OR 'randomi*ed' OR 'therap*'.

For part 1, we included articles describing the psychometric evaluation of an outcome measure explicitly for FND. Instruments designed for diagnosis of FND were excluded as these are not designed to measure change in FND symptoms over time. Measures assessing symptoms across bodily systems (ie, not exclusively neurological) and/or those validated in patients with mixed physical symptoms were excluded.

For part 2, studies were eligible if they prospectively measured outcomes of a specified intervention in patients with FND. Studies were excluded if the target samples included individuals with psychiatric diagnoses without FND or individuals with FND as a secondary or uncertain diagnosis. Case reports, case series and retrospective designs were excluded.

Data extraction and management

Initial screening of the titles and abstracts was conducted by one author (SPi), to exclude those not meeting basic eligibility criteria (ie, duplicates, not FND). The remaining titles, abstracts and/or full texts were screened independently by the lead author (SPi) and a second independent rater (TRN, SPo, RCK and ABW). Disparities were resolved by discussion or by a third independent rater.

In part 1, the following were extracted: (1) lead author, (2) publication date, (3) name of instrument, (4) type of instrument (eg, self-report questionnaire and clinical evaluation), (5) content of measure, (6) validation sample and (7) measurement properties. For part 2, the data extracted were (1) the lead author, (2) publication date, (3) geographical location, (4) nature of intervention(s), (5) design (eg, parallel and cross-over), (6) details of blinding (ie, participants and intervention deliverer), (7) FND symptoms, (8) sample size, (9) timescale, (10) outcome measures and (11) measurement properties. The specific measurement properties (validity, reliability and responsiveness) that were systematically extracted (where available) are defined in online supplementary file 1.[28–31]

FND-COM group consensus meetings

The findings of the reviews were discussed at two international consensus meetings (September 2017, Edinburgh, UK; September 2018, Atlanta, USA) involving 43 members from 12 countries, all with expertise in FND treatment or outcome research. The members represented a range of relevant health professions (neurology, (neuro)psychiatry, (neuro)psychology, physiotherapy and occupational therapy) and patient representation.

RESULTS

Existing FND-specific outcome measures

Five articles were eligible (online supplementary appendix 1 for the PRISMA diagram), which are summarised in table 1. Three

articles described clinician-rated scales, two assessing functional movement disorder (FMD) symptoms: Psychogenic Movement Disorder Rating Scale (PMDRS),[32] Simplified Functional Movement Disorder Rating Scale (S-FMDRS)[33] and one clinician-rated measure for functional (ie, psychogenic non-epileptic/dissociative) seizures.[34] Two patient-rated measures assessed a range of FND symptoms in children (Conversion Disorder Scale (CDS) and Conversion Disorder Scale—Revised (CDS-R)).[35 36] There was no single outcome measure for use across FND symptom subtypes in adults and none designed specifically for functional weakness/paralysis or sensory symptoms. The measurement properties of the scales are outlined in online supplementary file 2 and summarised as follows.

Validity

The content validity of the measures has not been confirmed according to established standards.[28 30] Only two studies referred to consultation with independent experts during scale development,[35 36] and none reported consultation with patients. Regarding construct validity, exploratory factor analysis was conducted for both patient-reported measures.[35 36] The CDS had three underlying factors (disability, pain and seizures) and the CDS-R had five factors (swallowing and speech, motor symptoms, sensory symptoms, weakness and fatigue, and mixed symptoms).

Criterion validity for new measures of FND symptoms is difficult to assess due to a lack of validated gold standards. However, moderate–strong correlations (r>0.3) with measures of similar or related constructs were noted for some scales, suggesting acceptable convergent validity.[32–34 36] The CDS and CDS-R also differentiated patients with FND from healthy controls.[35 36]

Reliability

Inter-rater reliability (IRR) statistics for the clinician-rated scales were generally in an acceptable range (coefficients of >0.8 for total scores), except for the seizure rating scale, which had slightly lower IRR (several coefficients of <0.7).[34] The internal consistency of the patient-rated measures was satisfactory to good (Cronbach's alpha values of >0.8 for total scores).[35 36]

Responsiveness

The PMDRS detected a significant change in patients' scores between preintervention and postintervention.[32] The S-FMDRS also displayed good sensitivity to change, with a large effect size (Cohen's d=0.79) between post-treatment scores in an intervention group compared with controls.[33] Data on responsiveness were not reported for the other measures.

Limitations

While the studies indicated some sound measurement properties of these scales, there were limitations to the psychometric evaluations conducted, according to the recommended guidelines.[28 30 31] None of the measures were evaluated in more than one study by independent research teams, although additional data on the measurement properties of the PMDRS have been provided in subsequent intervention studies. Validity data were variable and often incomplete, most importantly for content/face validity. No study included confirmatory factor analysis or item response theory and Rasch analyses to assess construct validity.[37] Similarly, none of the measures were validated cross-culturally, and there was limited evidence relating to ecological

Table 1 Existing outcome measures developed for use in functional neurological disorder

Study details	Measure	Description of measure
Cianci *et al* (Italy)[34]	Rating scale for functional 'psychogenic non-epileptic' seizures (clinician-rated)	Motor phenomena (tremor/oscillation, tonic, clonic/jerking, hypermotor/agitation, atonic/akinetic and automatisms) Associated features (incontinence, tongue-biting, drooling, eye closure, hyperventilation and lamenting/crying) Responsiveness 13 body regions Ratings: present/absent, severity, duration, global severity by phenomenon Total phenomenology, total associated features and total seizure scores
Hinson *et al* (USA)[32]	PMDRS (clinician-rated)	Motor phenomena (rest tremor, action tremor, dystonia, chorea, bradykinesia, myoclonus, tics, athetosis, ballism and cerebellar incoordination) Functions (gait and speech) 14 body regions Ratings: present/absent, severity, duration, global severity and incapacitation by phenomenon/function (0–4 scale) Total phenomenology, total function and total PMDRS score
Ijaz *et al* (Pakistan)[36]	CDS-R (patient-rated)	40 items (phenomena): five factors identified (swallowing/speech, motor, sensory, weakness/fatigue and mixed) Ratings: 0–2 (not at all, sometimes and most of the time) for each item CDS-R total scores (clinical cut-off score of 19)
Nielsen *et al* (UK)[33]	S-FMDRS (clinician-rated)	Seven body regions Two functions (speech and gait) Ratings: presence/absence of abnormal movement, severity and duration (0–3 scale) Total S-FMDRS score
Sarfraz and Ijaz (Pakistan)[35]	CDS (patient-rated)	19 items (phenomena): three factors (disability, pain and seizures)

CDS, Conversion Disorder Scale; CDS-R, Conversion Disorder Scale—Revised; PMDRS, Psychogenic Movement Disorder Rating Scale; S-FMDRS, Simplified Functional Movement Disorders Rating Scale.

Table 2 Primary outcome measures used across randomised controlled trials for FND subgroups

FND symptoms	Primary outcome measures
Seizures (k=15)	Core symptoms ► Seizure frequency/occurrence (k=11)[38–45] (S1 and S2) (Fobian NCT02801136). Life impact ► SF-36/QoLIE-31 (k=1) (Hingray NCT02311829). ► WSAS (k=1).[45] Other ► Treatment adherence/HCU (k=3) (S3 and S4) (Schommer NCT03329703).
Movement disorder (k=7)	Core symptoms ► PMDRS (k=4) or other FMD clinician rating scale (k=1) (S5–S9). ► CGI-I (clinician) (k=1) (S10). Other ► Unspecified (k=1) (S11).
Paralysis/weakness (k=4)	Core symptoms ► Subjective symptom severity ratings (k=1) (S12) or CGI-I (patient) (k=1) (Nicholson ISRCTN51225587). ► Clinician-rated symptom severity (Chastan NCT01352910). ► Dynamometry (k=1) (S13). Life impact ► Disability (SF-12, MRS, k=1) (S12).
Mixed motor symptoms (k=4)	Core symptoms ► CGI-I (k=1) (Koning-Tijssen NCT02589886). Life impact ► SF-36 Physical Functioning Scale (k=1) (S14). Other ► Unspecified (k=1) (S15).
Other mixed symptoms (k=10)	Core symptoms ► Clinician-rated symptom severity (k=3) (S16–S18) ► CGI-I (k=2, 1=clinician, 1=patient) (S19 and S20). ► SDQ-20 (k=1) (S19). ► Recovery (unspecified, k=1) (S21). Life impact ► SF-36 (k=1) (S22). ► MRS (k=1) (S19). Other ► Treatment adherence/HCU (k=1) (Bullock NCT02764476). ► Interoception/attention tasks (k=1) (S23).

References S1-S23 are available in Supplementary File 7
CGI-I, Clinical Global Impression–Improvement; FMD, functional movement disorder; FND, functional neurological disorder; HCU, healthcare use; k, number of studies; MRS, Modified Rankin Scale; PMDRS, Psychogenic Movement Disorder Rating Scale; QoLIE-31, Quality of Life In Epilepsy-31; SDQ-20, Somatoform Dissociation Questionnaire-20 item; SF-12, 12-Item Short Form Health Survey; SF-36, 36-Item Short Form Health Survey; WSAS, Work and Social Adjustment Scale.

validity. Sample sizes were relatively small for validations (ie, 120 or less) and a priori sample size calculation or justification was omitted, except for one study which reported a large effect size (Cohen's d=0.79) for comparison of control and intervention groups.[33] Only three studies included data from control groups.[33 35 36] Bias was potentially present in most samples due to recruitment within specialist healthcare clinics and/or stringent inclusion criteria. While all studies presented at least one appropriate measure of reliability, there were several omissions across studies and data were generally incomplete. For example, none of the articles provided statistics on test–retest reliability.

Outcome domains and measures in existing FND treatment studies

Randomised controlled trials

Forty RCTs were identified (31=published and 9=ongoing or unpublished) (online supplementary appendix 2 for the PRISMA diagram). Online supplementary files 3 and 4 detail the study designs and outcome measures adopted. Table 2 summarises the primary outcome measures in these RCTs across FND symptom subtypes.[38–45] Most RCTs focused on FND symptom change as the primary outcome measure (k=26). The most consistent primary outcome measure was functional seizure frequency; however, studies differed in how this was operationalised (eg, weekly or monthly seizure counts), and there was no single standardised way of recording the seizures (eg, patient seizure logs or clinical records). For FMD, the most common approach was a structured clinician-rated scale (eg, PMDRS), which can include standardised video protocols assessed by blind raters. Patient-rated symptom severity was often the primary outcome for functional limb weakness or mixed symptoms. The Clinical Global Impression–Improvement (CGI-I) scale[46] was used as the primary outcome measure in five studies (clinician-rated=2, patient-rated=2 and unspecified=1). At least one of the Clinical Global Impression (CGI) scales (improvement and/or severity) was included as a primary or secondary outcome in 13 studies.

Ten RCTs included a measure of additional physical symptoms, with the Patient Health Questionnaire-15 (PHQ-15)[47] (k=3) and Symptom Checklist-90 (SCL-90)[48] somatic subscale (k=3) used in several studies. Twenty-eight RCTs included a measure of psychological symptoms (primary or secondary), most commonly depression and/or anxiety. Patient-rated measures included the Hospital Anxiety and Depression Scale (HADS, k=9),[49] the Beck Depression Inventory (BDI or BDI-II, k=7)[50]

and the Beck Anxiety Inventory (BAI, k=5).[51] The Hamilton Rating Scale for Depression (HAM-D, k=8)[52] and the Hamilton Rating Scale for Anxiety (HAM-A, k=4)[53] were commonly used clinician-rated measures. Psychological dissociation (eg, depersonalisation, derealisation and amnesia) was measured with the Dissociative Experiences Scale[54] in six trials.

Life impact was assessed in most studies. The Short Form Health Survey (SF) scales (12-Item Short Form Health Survey (SF-12)=3 and 36-Item Short Form Health Survey (SF-36)=1)[55 56] were often included to measure health-related quality of life and/or disability. The Quality of Life in Epilepsy (QoLIE, k=6) scales (originally devised for epilepsy studies) were often used for patients with seizures.[57 58] Disability was assessed specifically in 16 studies; however, while most studies used clinician-rated scales, there was little consistency in the scales adopted. Measures of general (ie, social and occupational) functioning were included in 12 studies, most often with the Work and Social Adjustment Scale (WSAS, k=6)[59] or the Global Assessment of Functioning (GAF, k=5).[60]

Data for health economic/cost–utility analyses were also obtained in many trials. For example, healthcare resource use was monitored in a notable proportion (k=15). Four studies used a validated measure (Client Services Receipt Inventory, CSRI).[S24] A few studies provided data on quality-adjusted life years (QALYs) using the EQ-5D-5L.[S25]

Ten studies assessed illness perceptions or symptom attributions often using study-specific questions or variants of the Illness Perception Questionnaire (IPQ, k=5).[S26–S28] Adverse events were reported in 11 trials.

Prospective cohort observational studies

Forty prospective observational treatment studies were identified (34=published and 6=ongoing/unpublished) (online supplementary appendix 2 and online supplementary file 5). The pattern of outcome measures was very similar to the RCTs. The most common outcome domain was core symptom change (k=33). Patient-rated functional seizure frequency was measured often (k=18), and clinician-rated scales were used to assess FMD symptoms (eg, PMDRS k=5 and S-FMDRS k=2) (k=11). Self-reported symptom severity was assessed in nine studies, often using the CGI-I scale (k=5). Clinician-rated CGI scales were used in a small proportion of studies (k=3).

Additional physical (k=6) and psychological (k=21) symptoms were commonly assessed. The PHQ-15 was the most frequently used measure of additional physical symptoms (k=4). Common measures of psychological symptoms included the HADS (k=5), the BAI, BDI (k=8), the HAM-D (k=5) and/or the HAM-A (k=3).

In terms of life impact, quality of life and disability were measured in 19 studies, with the SF-36 being the most consistently adopted measure (k=11). Thirteen studies included measures of general functioning, including the WSAS (k=4) and GAF (k=3). A proportion of studies (k=7) monitored healthcare resource use or included other measures yielding data for health economic analyses. Illness beliefs were assessed in eight studies, again with the IPQ (or revised versions) in the majority (k=5). Three studies reported adverse events.

Measurement properties of outcome measures used in FND intervention studies

The review identified several outcome domains and measures often included in FND treatment research (table 3). Available psychometric data detailing the measurement properties of these measures were extracted (online supplementary file 6).

Table 3 Common outcome domains and measures in functional neurological disorder intervention studies

Outcome domain	Commonly used measures
Core symptoms	CGI-I: patient-rated=9, clinician-rated=7, both=4 Seizure frequency/remission=30 PMDRS=10
Other physical symptoms	PHQ-15/extended PHQ-15*=7 SCL-90 Somatic Symptom Scale=6
Psychological symptoms	Depression: BDI/BDI-II=12, HAM-D=13 Anxiety: BAI=8, HAM-A=7 Depression and anxiety: HADS=14
Life impact: quality of life/disability/general functioning	SF-36=22, SF-12=3 QoLIE-31=7, QoLIE-10=2 WSAS=10 GAF=8
Health economics/cost–utility	Healthcare resource use=20 (CSRI=4) QALYs (SF-36/EQ-5D-5L)=3
Illness beliefs/attributions	IPQ/modified IPQ=10

*Extended PHQ-15 (S29).
BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; CGI-I, Clinical Global Impression–Improvement; CSRI, Client Service Receipt Inventory; DES, Dissociative Experiences Scale; GAF, Global Assesment of Functioning; HADS, Hospital Anxiety & Depression Scale; HAM-A, Hamilton Anxiety Rating Scale; HAM-D, Hamilton Depression Rating Scale; IPQ, Illness Perception Questionnaire; PHQ-15, Patient Health Questionnaire-15; PMDRS, Psychogenic Movement Disorder Rating Scale; QALY, quality-adjusted life years; QoLIE-10, Quality of Life in Epilepsy-10; QoLIE-31, Quality of Life in Epilepsy-31; SCL-90, Symptom Checklist-90; SF-12, 12-Item Short Form Health Survey; SF-36, 36-Item Short Form Health Survey; WSAS, Work and Social Adjustment Scale.

Core symptoms

Clinical Global Impression–Improvement Scale

Few studies reported data pertaining to the validity or reliability of the CGI-I scale in FND samples. In one study, patient-rated CGI-I scores were significantly predicted by post-treatment change in Health of the Nation Outcomes Scale scores (p=0.033).[S30] However, another noted no significant associations between clinician-rated CGI-I and SF-36 or HADS scores,[S31] suggesting weak convergent validity with these outcome domains. Only one study examined reliability, noting moderate IRR between clinician-rated CGI-I and CGI-severity scores (kappa=0.65).[S10]

Some data suggested that the CGI-I scale is sensitive to change in this population. Post-treatment improvements were reported in most studies. The proportion of patients reporting improvement after an intervention was 60%–96% in 6/9 studies.[S9, S15, S30, S32–S34] In two studies, 30%–48% of the active intervention groups reported improvements,[S10, S20] whereas in one study, only average scores were reported.[S7] Average reported patient-rated and clinician-rated CGI-I scores were in the minimally to much improved range.[S7, S31, S33, S35] Significant treatment effects in patient-rated and clinician-rated CGI-I scores were observed,[43, S2, S19, S20] although not consistently.[42, S6, S10] Effect sizes varied from small to large.[43, S2, S10, S20] These findings suggest that the CGI-I scale has adequate responsiveness in patients with FND, but more data are needed for different symptom types and treatment modalities.

Seizure frequency/freedom

There were no data available on reliability for seizure outcomes and few studies reported data relevant to validity in these FND samples. De Barros *et al* noted significant correlations between post-treatment seizure frequency and HAM-D, HAM-A and alexithymia scores (p<0.05).[S36] Furthermore, Kuyk *et al* observed

a significant negative correlation between post-treatment seizure frequency and SF-36 energy–vitality scores ($p<0.05$).[S37] Seizure-free patients showed significantly greater improvement on several SF-36 subscales (mental health, energy–vitality and pain), BDI and STAI (anxiety) scores, relative to patients who were not seizure-free ($p<0.05$).

Post-treatment decreases in seizure frequency or increases in seizure freedom were observed in a proportion of patients across all studies. A number of studies reported statistically significant treatment effects on seizure frequency [39, 40, 42–44, S2, S4, S36–S44] or seizure freedom.[38, S45] Effect sizes were generally small–medium.[40, 42, S2, S4] Future studies should standardise the methods of logging/comparing seizure outcomes to provide more comparable data.

Psychogenic Movement Disorder Rating Scale

Dreissen *et al* reported adequate IRR for the PMDRS (average intraclass correlation coefficient=0.76),[S10] which is lower than the original values presented by Hinson *et al* (0.87–0.89),[32] but both are in an acceptable range. Regarding validity, Taib *et al*[S7] observed significant negative correlations between PMDRS total and SF-36 physical role scores ($r=-0.77$, $p=0.0002$), and between PMDRS total and SF-36 general health scores ($r=-0.67$, $p=0.02$). PMDRS total scores also correlated positively with CGI severity ($r=0.88$, $p=0.001$). Similarly, the original validation study noted a positive correlation between CGI and total PMDRS scores.[32] These relationships suggest good convergent validity. Seven out of 10 studies reported significant treatment effects for the PMDRS,[32, S5, S7, S32, S46–S48] and one noted a significant improvement in a placebo arm but not for the active intervention.[S8] Together, these findings suggest that the PMDRS shows adequate sensitivity to change in people with FMD.

Other physical symptoms

The two most common measures of additional physical symptoms were the PHQ-15 and the SCL-90 somatic symptoms scale. There were no data available regarding the reliability or validity of these measures in these FND intervention studies.

Patient Health Questionnaire-15

Five studies presented post-treatment data for the PHQ-15. Four studies reported significant reductions in post-treatment scores.[S5, S20, S30, S49] Williams *et al*[22] also noted a decline in symptoms, but the statistical significance did not withstand correction for multiple testing ($p=0.03$). One study reported a medium effect size ($d=0.53$).[S30] Together, these findings suggest that the PHQ-15 shows good responsiveness across FND subgroups.

Symptom Checklist-90

Five studies reported outcomes on the SCL-90 somatic scale, with two observing significant improvements[S2, S40] and one reporting a large effect size.[S2] The responsiveness of this scale in FND samples requires further examination.

Psychological symptoms

The HADS, BDI (I or II) and BAI were common patient-rated outcome measures for anxiety and depression, and the HAM-D and HAM-A were often used as clinician-rated measures. There were no data regarding reliability for these measures in these FND samples.

Hospital Anxiety and Depression Scale

Conwill *et al*[S31] reported no significant association between HADS and CGI-I (clinician) scores, suggesting poor convergent validity with core symptom change. Outcome data for the HADS were provided in 12 studies; however, significant treatment effects were only reported in four (33%) and reported effect sizes were variable.[43, S30, S39, S44]

Beck Depression Inventory & Beck Anxiety Inventory (BDI and BAI)

There were no data on the validity of the BDI/BAI in these FND samples. Nine studies reported outcomes for the BDI (I or II), with six (67%) reporting significant treatment effects.[39, S2, S19, S40, S43, S50] Large effect sizes were reported in two studies.[39, S2] Six studies provided outcome data for the BAI, with significant treatment effects observed in four[32, S2, S5, S46] and a large effect size in one.[S2] One study reported a significant improvement in two treatment arms.[S6]

Hamilton Anxiety Rating Scale and Hamilton Depression Rating Scale

HAM-A and HAM-D scores positively correlated with post-treatment seizure frequency in one study.[S36] Demartini *et al*[S32] found that HAM-A and HAM-D scores were significantly higher in patients with FND relative to healthy controls at baseline, suggesting adequate known-groups validity.

Three of five studies reporting outcome data for HAM-A reported significant treatment effects.[38, S36, S44] Two studies showed non-significant trends for improvement following treatment.[S8, S32] Thirteen studies reported post-treatment outcomes for HAM-D scores, with six reporting significant effects.[32, S2, S5, S36, S44, S46] Vizcarra *et al*[S8] noted a significant improvement in a placebo arm but not an intervention arm, whereas Kompoliti *et al*[S6] noted significant improvements in two treatment arms. Effect sizes were moderate–large ($d=0.71$–1.8).[S2, S36]

Life impact

Short Form Health Survey-36 & Short Form Health Survey-12 (SF-36 and SF-12)

The SF scales were the most consistently used of all measures in these studies (most often SF-36). There were no data available on their reliability in the FND samples, but three studies examined correlations with other relevant measures. Conwill *et al*[S31] found no significant associations between SF-36 scores and CGI-I (clinician) ratings. However, Kuyk *et al*[S37] noted significantly greater improvements on SF-36 domains (mental health, energy–vitality and pain) for seizure-free patients at follow-up. Williams *et al*[22] observed a significant positive correlation between post-treatment SF-36 mental health component scores and emotional processing improvements ($p<0.008$).

Significant treatment effects were observed on domains of the SF-36/SF-12 in 14/18 studies presenting outcome data.[22, S11, S15, S19, S20, S22, S31, S32, S34, S36, S37, S42, S49, S51] Effect sizes were generally moderate.[22, S15, S34, S36] Significant effects were most often observed for mental health (k=9), emotional role (k=4), general health (k=4), social function (k=4) and physical function (k=4). Several studies showed differences in the physical component (k=3), physical role (k=3) and energy/vitality (k=3) domains.

Quality of Life in Epilepsy-31 & Quality of Life in Epilepsy-10 (QoLIE-31 and QoLIE-10)

No data on validity or reliability were available for the QoLIE scales in these functional seizure samples. However, post-treatment outcome data were available in seven studies. Four reported significant treatment effects [39, S2, S4, S40] and one reported a non-significant trend towards improvement.[S52] Effects sizes were moderate–large.[39, S2, S4]

Work and Social Adjustment Scale

The WSAS was included in 10 studies. There were no data on validity and reliability of the WSAS in the FND samples included. Of eight studies with outcome data, three reported significant treatment effects.[45, S34, S39] Reported effect sizes were small. [S15, S45] The measurement properties of the WSAS in FND require further examination.

Global Assessment of Functioning

In the five FND intervention studies providing data on the GAF, there were no details on its validity or reliability. Four studies, however, showed significant treatment effects.[32, S2, S40, S53] LaFrance *et al*[S2] reported large effect sizes (d=1.2–1.8).

Health economics and cost–utility

Healthcare resource use

A proportion of studies reported data on healthcare resource use. The most common way to measure this was to obtain a frequency count of healthcare contacts (eg, total healthcare contacts, emergency department, inpatient hospital days, outpatient contacts and mental health contacts), either by self-report or obtained from clinical records. No data were available for validity and reliability and few significant differences were noted following treatment in these FND samples.

In patients with functional seizures, significant post-treatment reductions in emergency department visits and overall healthcare use [S2, S41] and greater probability of contact with a mental health professional [S3] were reported in a few studies. In patients with mixed FND symptoms, Aybek *et al*[S51] observed reduced postintervention medical follow-up, and Hubschmid *et al*[S19] reported fewer postintervention inpatient hospital days. Greater consistency is needed in healthcare use measurement in future studies.

Quality-adjusted life years

A minority of studies reported data on QALYs (k=3). Nielsen *et al*[S15, S34] calculated QALY utility scores from the EQ-5D-5L, reporting post-treatment gains of 0.125 and 0.08. One study reported the average cost per QALY as £12 087.[S15] Reuber and colleagues reported QALY data derived from the SF-36, reporting post-treatment gains of 0.04 and an average cost per QALY of £5328.[S49] More data are needed on QALYs and their associated costs for different FND treatments.

Illness perceptions/attributions

Illness perceptions and attributions were measured often with the IPQ or revised versions. There was no evidence of the reliability or validity of the IPQ in these FND samples. Demartini *et al* noted that IPQ-Revised (IPQ-R) scores at baseline were not predictive of patient-rated CGI-I outcomes (p=0.77).[S30]

All seven studies providing post-treatment data on the IPQ (or Brief-IPQ and IPQ-R) reported statistically significant post-treatment improvements. Some studies reported post-treatment reductions in composite scores.[22, S15, S45] Other studies reported significant improvements on individual items, including beliefs about illness permanence (timeline, k=3), perceived negative consequences (k=2) and level of concern (k=2). Individual studies also reported significant differences in illness understanding, emotional representations, physical attributions, psychological attributions and belief in possible cure/management.

Several studies explored illness beliefs with study-specific methods. In patients with functional seizures, belief that the seizures could be helped (OR=3.9, p=0.003), subjective control (OR=3.3, p=0.021) and an internal locus of control (OR=7.5, p<0.001) significantly predicted seizure freedom at follow-up. [S54] Significant post-treatment changes in beliefs relating to functional seizures have been noted, including perceived control, being bothered by the seizures, perceived impact, ability to avoid triggers and perceived understanding.[39, 45, S45] Additional data are needed to explore the importance of these changes for patients' outcomes.

FND-COM CONSENSUS OPINION

Key points that emerged from the FND-COM group meetings and subsequent interchanges were as follows:

General issues

- Challenges for FND outcome measurement include the temporal variability of symptoms and the impact of attention, beliefs and expectations.
- Patient-rated outcome measures are important when assessing outcomes in this population.
- Discrepancies between objective and patient-rated outcomes may provide insights into the mechanisms underlying symptoms and treatment responses.

Measuring FND symptom change

- There are few validated outcome measures for FND symptoms across symptom types.
- There was some dissatisfaction with the PMDRS among the FND-COM members. Key limitations of this scale have been outlined elsewhere.[33]
- Of the clinician-rated scales for FMDs, the S-FMDRS was preferred by the FND-COM group due its ease of administration and lack of requirement for expertise in movement disorder phenomenology. However, this measure currently is not used widely, possibly due to time constraints, cost and the limitations of snapshot clinician-rated measures.
- Assessing the criterion validity of new FND-specific symptom measures is challenging due to a lack of existing 'gold standard(s)'.
- An individualised approach could be considered, whereby patients select the most troublesome symptoms and then use an individualised scale to assess change on these difficulties, or in which patient-specific goals are identified and progress towards them is subsequently monitored (eg, Psychological Outcome Profiles (PSYCHLOPS), Goal Attainment Scaling). [S55, S56]

Additional outcome domains

- Measures of the impact of FND (ie, quality of life, disability, functioning and psychological distress) are important domains as they are relevant across all symptom types.
- Existing, generic outcome measures, known to be reliable and valid in related populations, can be used for adjunctive key outcome domains in FND (eg, quality of life, additional physical symptoms and global functioning).
- Illness beliefs and attributions may be important mediators of treatment response and/or might represent a common domain of change across different treatments.

DISCUSSION

We aimed to identify FND-specific outcome measures, describe the outcome domains and measures most commonly included in previous FND intervention studies, integrate these findings with FND-COM group expert opinion and present preliminary recommendations for outcome measurement in future studies.

Existing FND-specific outcome measures

Only five FND-specific outcome measures were eligible: three clinician-rated scales[32–34] and two self-report measures.[35 36] There were no measures specifically for functional limb weakness or sensory symptoms. Importantly, there is no single outcome measure suitable for use across all adult FND symptom types.

Of the identified scales, content validity was not established rigorously and the validation studies were limited by homogeneous or unrepresentative samples, cultural specificity and a lack of evidence of ecological validity. Data on responsiveness were not available for all scales.[34–36] Due to the limited evidence for satisfactory measurement properties of the scales identified,[28 30] it is not possible to recommend their use as core outcome measures at present.

There is a clear absence of FND-specific outcome measures to capture the full spectrum of FND symptoms. It is therefore unclear whether any single measure could adequately capture all relevant domains for the FND population, given its semiologic and symptomatic heterogeneity. Furthermore, it is possible that FND-specific outcome measures may be unnecessary and that existing measures used in other related disorders are valid and reliable for measuring outcomes in FND.

Outcome domains and measures in previous FND intervention studies

Outcome domains measured commonly in previous intervention studies included core FND symptoms, other physical symptoms, psychological symptoms, life impact (quality of life, disability and general functioning), and health economics or cost–utility (healthcare resource use and QALYs). These domains and the associated measurement tools are discussed further.

Core FND symptom change

The CGI scales and patient-reported seizure frequency or freedom (functional seizures) were the most consistently used measures of symptom change. The main advantage of the CGI-I scale is that it can be used across the full range of FND symptoms, allowing for direct comparisons of treatments across diverse presentations. It is brief and simple to administer and can assess specific symptoms or global clinical status. However, there are scarce data on the validity and reliability of this scale in FND samples, and mixed findings regarding its responsiveness. Another limitation is the lack of evidence for inter-rater reliability of the CGI. Furthermore, the scale is ordinal in nature (ie, scores are relative to baseline state only) and does not provide an absolute quantification of the utility of an intervention. There are data from other relevant patient populations suggesting acceptable measurement properties,[S57-S61] but future studies should examine further the properties of this scale in FND samples, including correlations between patient and clinician ratings. A modified version of the CGI for use in FND could also be considered, similarly to its use in other disorders.[S62, S63]

The assessment of seizure frequency and/or freedom was the most common outcome for studies involving functional seizures. However, there was variability in how these were defined and a lack of data on their reliability and validity. Some preliminary evidence of convergent validity with other outcome domains was available. Importantly, improvements in seizure outcomes were observed consistently, with many studies reporting significant treatment effects. A standardised approach to operationalising seizure outcomes would be beneficial for future studies.[S64] It may be relevant to monitor duration/severity and specific seizure symptoms, as well as event frequency.

Clinician-rated scales were adopted most often for FMD. The PMDRS was used frequently, and findings suggested some sound measurement properties. However, the face validity of the measure has not been examined sufficiently and doubts have been raised about its practical utility, along with other limitations.[33]

Other physical symptoms

Additional physical symptoms were often measured in previous studies. The FND-COM group agreed that these symptoms are important outcomes for patients with FND, particularly pain and fatigue, which are prevalent and impactful. The two most common measures of additional physical symptoms (PHQ-15 and SCL-90 somatic scale) lacked evidence of reliability and validity in studies of FND samples. Responsiveness was superior for the PHQ-15; this measure is brief and it has been validated across a wide range of physical and mental health diagnoses in various clinical and cultural settings.[47, S65-S68] The PHQ-15, therefore, is a potentially useful measure, but more data on its measurement properties are needed in FND. Extended versions of the PHQ-15 [S29] can also be considered in future research.

Psychological symptoms

Many studies included one or more measures of psychological symptoms, most often depression and/or anxiety, measured by the HADS or BDI and BAI (patient-rated) or HAM-A and HAM-D (clinician-rated).

There are several advantages of using the HADS, including its brevity, combined anxiety and depression subscales in a single measure, and the omission of physical manifestations, reducing the likelihood of confounding physical with psychological symptoms. The measurement properties of the HADS in FND samples are not well described and more data are needed. However, the HADS has shown good measurement properties across other physical and mental health diagnoses and cross-culturally.[S69-S74]

Evidence for the reliability and validity of the BDI (BDI-II) and BAI in FND was lacking in the FND intervention studies reviewed. Responsiveness was stronger for these measures, but interpretation is potentially hindered by the inclusion of physical symptoms in these scales (eg, fatigue, sleep disturbance and tremor). Nevertheless, both the BAI and BDI/BDI-II have also been validated in a range of clinical and non-clinical populations, and across cultures.[S75-S79] The HAM-A and HAM-D seem to be the most useful clinician-rated measures of psychological symptoms. Although reliability and validity have not been rigorously examined in FND samples, these scales show preliminary evidence of adequate responsiveness.

Life impact

The most common measure of life impact was the SF-36. This measure has the advantage of assessing patients' subjective perceptions of overall health status, aspects of disability, and physical and mental health in a single measure, and can also be used for the derivation of QALYs. A key benefit of the SF-36 for use in FND is its multidimensional nature; it is possible to examine outcomes on eight specific subscales, depending on the goals and modality of a given treatment. Again, there was a paucity of reliability and validity data in FND samples, but there is considerable evidence for responsiveness, with most studies reporting significant treatment effects on at least one SF-36 domain. The SF-36 has also been well-validated across clinical populations and cultures.[S80-S82] The QoLIE scales serve a

similar purpose in functional seizures, but more data are needed on the measurement properties in FND samples.

Additional measures of life impact include measures of general (social and occupational) functioning. The WSAS is often used to capture patient-reported general functioning and the GAF for clinician ratings. No evidence is available for the reliability or validity of either scale in FND samples. The responsiveness of the GAF appeared slightly stronger in the studies presented here.

Health economics and cost–utility

Health economic data were primarily derived from healthcare resource use frequency, with the CSRI or study-specific questions about healthcare contacts being the most common methods. The CSRI was developed in the UK, so it requires adaptation to local regions and is a relatively lengthy measure to administer; therefore, it is unlikely to be appropriate for widespread international use. Instead, specific questions about healthcare contacts (eg, emergency department or outpatient visits, inpatient hospital admissions and inpatient days) appear to be more generalisable methods. The EQ-5D-5L and SF-36 allow the derivation of QALYs and associated costs.[S83, S84] Future research is needed to compare the relative cost–utility of treatment modalities for different subgroups of patients with FND.

Illness perceptions/symptom attributions

Illness perceptions/beliefs were noted by the FND-COM group as also important to measure in FND treatment studies, as they might mediate change across different treatment modalities and symptom types. Of note, 100% of studies including a version of the IPQ noted significant treatment effects on composite IPQ scores or specific items. The original IPQ is relatively lengthy, but the B-IPQ or IPQ-R could be adopted (or modified) to better assess FND-relevant cognitions (eg, acceptance/understanding of the diagnosis, beliefs about causation and subjective control).

FND-COM recommendations for outcome measurement in FND

In the absence of rigorously validated, widely endorsed FND-specific outcome measures, we recommend an interim approach involving the assessment of core outcome domains with existing patient-rated measures that are valid, reliable and sensitive to change in FND or other relevant populations. The core domains and recommended measures are detailed in table 4, alongside additional supplementary patient-rated or clinician-rated measures. We recommend that these measures (or a subset) be assessed in future intervention research in adult FND populations.

There are similarities between the domains identified here and those described in recommendations for other related disorders (eg, chronic pain, fibromyalgia, complex regional pain syndrome and somatic symptom disorders).[S85-S89] The outcome domains and measures recommended here could be supplemented with additional tools, depending on the specific intervention and patient characteristics within individual studies. We also acknowledge that adaptations to these recommendations might be necessary due to cultural/context-specific factors. Furthermore, cost considerations may affect the use of the recommended measures, as some are licensed.

Limitations

This article had some limitations. Articles not written in English were excluded. Future research on outcome measurement in FND should aim to examine cultural variation in practices, and outcome measures should be validated across cultural contexts. The focus of this article was on identifying and evaluating outcome measures specifically, so we did not conduct a quality appraisal of the methods used to assess the effectiveness and efficacy of the interventions, nor did we evaluate the quality and significance of the outcome data in relation to the treatment modalities tested.

Table 4 Recommendations for outcome measurement in FND

Outcome domain	Recommended measures	Supplementary/alternative measures
Core FND symptoms	► CGI-I patient (+/-clinician, +/-carer).	► Seizure frequency, prospective daily seizure log (for seizures). ► S-FMDRS or PMDRS (for FMD).
Other physical symptoms	► PHQ-15.	► Extended PHQ-15. ► SCL-90.
Psychological symptoms	► HADS or ► BDI/BDI-II+BAI.	► HAM-D+HAM-A.
Life impact	Quality of Life ► SF-36. And/or General (social/occupational) functioning ► WSAS.	► QoLIE-31/10 (for seizures). ► GAF.
Health economics/ cost–utility	Healthcare resource use ► Healthcare contacts (total, inpatient admissions, inpatient days, ED/ outpatient visits).	► CSRI. ► Scale validated for QALY conversion (EQ-5D-5L or SF-36).
AEs	AEs, SAEs and mortality.	

AE, adverse event; BAI, Beck Anxiety Inventory; BDI, Beck Depression inventory; CGI-I, Clinical Global Impression–Improvement; CSRI, Client Service Receipt Inventory; ED, emergency department; FMD, functional movement disorder; FND, functional neurological disorder; GAF, Global Assessment of Functioning; HADS, Hospital Anxiety Depression Scale; HAM-A, Hamilton Rating Scale for Anxiety; HAM-D, Hamilton Rating Scale for Depression; PHQ-15, Patient Health Questionnaire-15; PMDRS, Psychogenic Movement Disorder Rating Scale; QALY, quality-adjusted life years; QoLIE-10, Quality of Life in Epilepsy-10; QoLIE-31, Quality of Life in Epilepsy-31; SAE, serious adverse event; SCL-90, Symptom Checklist-90; SF-36, Short Form Health Survey-36; S-FMDRS, Simplified Functional Movement Disorder Rating Scale; WSAS, Work and Social Adjustment Scale.

Future directions

This review highlighted several directions for future research and development. First, FND experts should convene to discuss further the need for new FND-specific outcome measures for use across all symptom types and additional tools for individual symptom types. Additional measures could include patient-rated, clinician-rated, caregiver-rated and/or objective measures. If deemed appropriate, the development and validation of such scales should be prioritised. Consideration of content and face validity should be given during the design and development of any new scales. Second, future FND intervention studies should aim to present data on the measurement properties of the outcome measures included (where possible), to strengthen the evidence base for the use of specific measures in FND. Outcome measures that are used frequently (eg, seizure frequency, SF-36 and PMDRS) should undergo more extensive psychometric evaluation, where feasible, in larger samples of patients with FND, ideally across different cultures.

Future research may explore additional avenues for outcome measurement in FND, such as the use of objective/performance-based measures (eg, actigraphy and 5 m walk test), or personalised symptom rating or goal attainment scales. Discrepancies between patient-rated and objective/clinician-rated measures may provide important insights into mechanisms underlying

symptoms or therapeutic change. Additional psychological domains that could be relevant for future investigation include illness beliefs, psychological dissociation (eg, depersonalisation and derealisation) and aspects of emotional processing (eg, stress reactivity and alexithymia). It will also be important to explore which outcome domains and measures are most relevant for additional functional neurological symptoms not covered in this review, such as cognitive (eg, consciousness and memory), speech and swallowing symptoms.

An important direction for further research is to examine the perspectives of patients, carers and other relevant stakeholders on outcome measurement in FND, particularly exploring views on the relative merits of subjective and objective measures. Additional work is needed to develop recommendations for outcome measurement in children with FND, given that many of the recommended measures may not be validated or relevant for younger patients.

CONCLUSIONS

There is a need for greater consistency in outcome measurement in FND research. At present, we recommend an interim set of outcome measures to routinely assess core outcome domains, including core FND symptom change, other physical and psychological symptoms, life impact, health economics/cost–utility and adverse events. These recommendations could be supplemented with additional measures as appropriate to individual studies. Further research is needed to provide more extensive psychometric evaluation of outcome measures for FND and to explore the views of patients, carers and other stakeholders on optimising outcome measurement in this disorder.

Author affiliations
[1]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[2]Donald Gordon Medical Centre, University of the Witwatersrand, Johannesburg, South Africa
[3]Epilepsy Research Center, Shiraz Medical School, Shiraz University of Medical Sciences, Shiraz, Iran, Islamic Republic of
[4]Jefferson Comprehensive Epilepsy Center, Department of Neurology, Thomas Jefferson University, Philadelphia, Pennsylvania, USA
[5]Department of Neurology, University Hospital Bern & University of Bern, Bern, Switzerland
[6]Department of Psychiatry, Brigham & Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA
[7]Department of Neurology, Radboud University Nijmegen Medical Centre, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands
[8]School of Health Sciences, The University of Manchester, Manchester, UK
[9]Department of Clinical Neurosciences, School of Molecular and Clinical Medicine, Western General Hospital, University of Edinburgh, Edinburgh, UK
[10]Institute of Mental Health, Division of Psychiatry, Faculty of Brain Sciences, University College London, London, United Kingdom
[11]Neuroscience Research Centre, Institute of Molecular and Clinical Sciences, St George's University, London, UK
[12]Department of Psychiatry, Georgetown University, Washington, District of Columbia, USA
[13]Gardner Family Center for Parkinson's Disease and Movement Disorders, Department of Neurology, University of Cincinnati, Cincinnati, Ohio, USA
[14]Department of Neurology, Hopital Avicenne, Assistance Publique, Hôpitaux de Paris, Paris, Île-de-France, France
[15]Human Motor Control Section, National Institute of Neurological Disorders and Stroke, Bethesda, Maryland, USA
[16]Parkinson's Disease Center and Movement Disorders Clinic, Department of Neurology, Baylor College of Medicine, Houston, Texas, USA
[17]University College London Institute of Neurology, National Hospital for Neurology and Neurosurgery, London, UK
[18]Department of Psychiatry, Austin Health, University of Melbourne, Melbourne, Victoria, Australia
[19]Discipline of Psychiatry and Child and Adolescent Health, The Children's Hospital at Westmead, Sydney Medical School, Sydney, New South Wales, Australia
[20]Department of Neurology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA
[21]Departments of Psychiatry and Neurology, Rhode Island Hospital, Brown Medical School, Providence, RI, USA
[22]Movement Disorders Clinic and the Edmond J. Safra Program in Parkinson's Disease, University Health Network, Toronto, Ontario, Canada
[23]Mater Neurosciences Centre, Brisbane, Queensland, Australia
[24]Department of Neurology, Stony Brook University Renaissance School of Medicine, Stony Brook, New York, USA
[25]FND Hope International, Salmon, Idaho, USA
[26]Northeast Regional Epilepsy Group, New York, New York, USA
[27]Therapy Services, The National Hospital for Neurology and Neurosurgery, London, UK
[28]Departments of Neurology and Psychiatry, Therapy Services, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States
[29]Department of Neurology, University Hospital Knappschaftskrankenhaus Bochum, Ruhr University Bochum, Bochum, Germany
[30]Academic Neurology Unit, University of Sheffield, Royal Hallamshire Hospital, Sheffield, UK
[31]Departments of Neurology and Psychiatry and Behavioral Sciences, Emory Centre for Ethics, Emory University School of Medicine, Atlanta, Georgia, USA
[32]Department of Neurology, Medical University of Graz, Graz, Austria
[33]Department of Neurology and Center of Clinical Neuroscience, Charles University in Prague, Prague, Czech Republic
[34]Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA
[35]Department of Neurology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands
[36]Department of Neuroscience, Biomedicine, and Movement, University of Verona, Verona, Italy

Twitter Susannah Pick @susannah_pick, Selma Aybek @SelmaAybek, Bastiaan R Bloem @basbloem, Alan J Carson @AlanCarson15, Anthony S David @ProfTonyDavid, Mark J Edwards @drmarkedwards, Alberto J Espay @AlbertoEspay, Béatrice Garcin @beagarcin1, Mark Hallett @MarkHallett007, Joseph Jankovic @JankovicJoseph, Roxanne C Keynejad @RoxanneKeynejad, Kathrin LaFaver @LaFaverMD, Sarah Lidstone @SarahLidstone, Bridget Mildon @FNDHopeInternational, Francesca Morgante @franci_morgante, Lorna Myers @dramyers1, Clare Nicholson @CNicholsonOT, Glenn Nielsen @GNielsen_Physio, David L Perez @PerezMGHLab, Stoyan Popkirov @popkirov, Markus Reuber @MarkusReuber, Paul Shotbolt @Paul_Shotbolt, Jon Stone @jonstoneneuro and Timothy R Nicholson @Tim_R_Nicholson

Contributors TRN and SPi formulated the idea for the review; TRN, SPi and AJC planned the overall structure, with input from the Functional Neurological Disorder–Core Outcome Measures (FND-COM) group. SPi conducted the literature reviews, extracted the data, prepared the tables/figures and wrote the manuscript (including revised versions). TRN contributed to the tables in supplementary files 3 and 4. ABW, TRN, RCK and SPo independently screened a subset of titles/abstracts during the literature review. All other authors contributed to the FND-COM meetings or discussions in person, by email or teleconference. All authors reviewed the manuscript for intellectual content and/or suggested revisions.

Funding AJE received grant support from the National Institutes of Health (NIH) and the Michael J Fox Foundation. LHG reported salary support from the UK National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre at the South London and Maudsley NHS Foundation Trust and King's College London. TRN and SPi were also funded by an NIHR clinician scientist fellowship. GN also received funding from the NIHR. RCK received an NIHR Academic Clinical Fellowship in General Adult Psychiatry and a Royal College of Psychiatrists Gosling Fellowship. She also received PhD funding from King's College London and a King's Institute of Psychiatry, Psychology and Neuroscience (IoPPN) Clinician Investigator Scholarship. MH was supported by the National Institute of Neurological Disorders and Stroke Intramural Programme (NIH, USA). TS received support from Ministry of Health of the Czech Republic (grant number AZV ČR 16-29651).

Disclaimer The views expressed in this publication are those of the authors and not necessarily those of the NHS (UK), the National Institute for Health Research (UK), the Department of Health and Social Care (UK) or the National Institutes of Health/National Institute of Neurological Disorders and Stroke (USA).

Competing interests AAA-P reports honoraria from Cobel Daruo, Sanofi and RaymandRad, and royalty from Oxford University Press (book publication). AJC reports independent expert testimony work for personal injury and medical negligence claims, is a paid associate editor of JNNP and runs a free non-profit self-help website (www.headinjurysymptoms.org). AJE has received personal compensation as a consultant/scientific advisory board member for Abbvie, Adamas, Acadia, Acorda, Neuroderm, Impax/Amneal, Sunovion, Lundbeck, Osmotica Pharmaceutical and US WorldMeds; publishing royalties from Lippincott Williams & Wilkins, Cambridge University Press and Springer; and honoraria from US WorldMeds, Lundbeck, Acadia, Sunovion, the American Academy of Neurology and

the Movement Disorders Society. BM has received honoraria from The Cleveland Clinic and runs a free non-profit self-help website (www.fndhope.org). WCLF has served on the editorial boards of *Epilepsia, Epilepsy & Behavior*; *Journal of Neurology*, *Neurosurgery and Psychiatry* and *Journal of Neuropsychiatry and Clinical Neurosciences*; receives editor's royalties from the publication of Gates and Rowan's Nonepileptic Seizures, 3rd ed. (Cambridge University Press, 2010) and 4th ed. (2018); author's royalties for *Taking Control of Your Seizures: Workbook and Therapist Guide* (Oxford University Press, 2015); has received research support from the Department of Defense (DoD W81XWH-17-0169), National Institutes of Health (NIH) (NINDS 5K23NS45902 [PI]), Providence VAMC, Center for Neurorestoration and Neurorehabilitation, Rhode Island Hospital, the American Epilepsy Society (AES), the Epilepsy Foundation (EF), Brown University and the Siravo Foundation; serves on the Epilepsy Foundation New England Professional Advisory Board; received honoraria for the American Academy of Neurology Meeting Annual Course; served as a clinic development consultant at University of Colorado Denver, Cleveland Clinic, Spectrum Health, Emory University and Oregon Health Sciences University; and provided medicolegal expert testimony. DLP has received honoraria from the American Academy of Neurology, Movement Disorder Society and Harvard Medical School. JS reports independent expert testimony work for personal injury and medical negligence claims, receives royalties from UpToDate for articles on functional neurological disorder and runs a free non-profit self-help website (www.neurosymptoms.org). KLF has received honoraria from the American Academy of Neurology and the Movement Disorder Society. MH may accrue revenue on a US patent for an Immunotoxin (MAB-Ricin) for the treatment of focal movement disorders and for a coil for magnetic stimulation and methods for using the same (H-coil); in relation to the latter, he has received licence fee payments from the NIH (from Brainsway). He is on the medical advisory boards of CALA Health, Brainsway and Cadent. He is on the editorial board of approximately 15 journals and receives royalties and/or honoraria from publishing from Cambridge University Press, Oxford University Press, Springer and Elsevier. Grant research funds have come from Merz for treatment studies of focal hand dystonia; Allergan for studies of methods to inject botulinum toxins; Medtronic, Inc. for a study of DBS for dystonia; and CALA Health for studies of a device to suppress tremor. MJE reports independent expert testimony work for personal injury and medical negligence claims and receives royalties from the *Oxford Specialist Handbook of Parkinson's Disease and Other Movement Disorders*. He has received honoraria from Merz Pharma and Boeringher Ingleheim. AEL reports consultancy support from Abbvie, AFFiRis, Biogen, Janssen, Lilly, Lundbeck, Merck, Paladin, Roche, Sun Pharma, Theravance, and Corticobasal Degeneration Solutions; advisory board support form Jazz Pharma, PhotoPharmics, Sunovion; other honoraria from Sun Pharma, AbbVie, Sunovion, American Academy of Neurology and the International Parkinson and Movement Disorder Society; grants from Brain Canada, Canadian Institutes of Health Research, Corticobasal Degeneration Solutions, Edmond J Safra Philanthropic Foundation, Michael J. Fox Foundation, the Ontario Brain Institute, Parkinson Foundation, Parkinson Canada, and W. Garfield Weston Foundation and royalties from Elsevier, Saunders, Wiley-Blackwell, Johns Hopkins Press, and Cambridge University Press.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Susannah Pick http://orcid.org/0000-0003-2001-6723
Ali A Asadi-Pooya http://orcid.org/0000-0002-2598-7601
Selma Aybek http://orcid.org/0000-0002-7877-6760
Bastiaan R Bloem http://orcid.org/0000-0002-6371-3337
Richard J Brown http://orcid.org/0000-0003-3961-2023
Alan J Carson http://orcid.org/0000-0002-7425-0964
Trudie Chalder http://orcid.org/0000-0003-0775-1045
Anthony S David http://orcid.org/0000-0003-0967-774X
Mark J Edwards http://orcid.org/0000-0002-8283-9015
Alberto J Espay http://orcid.org/0000-0002-3389-136X
Béatrice Garcin http://orcid.org/0000-0001-9109-9904
Laura H Goldstein http://orcid.org/0000-0001-9387-3035
Mark Hallett http://orcid.org/0000-0002-3180-6811
Joseph Jankovic http://orcid.org/0000-0003-4129-3674
Eileen M Joyce http://orcid.org/0000-0003-0469-2844
Richard A Kanaan http://orcid.org/0000-0003-0992-1917
Roxanne C Keynejad http://orcid.org/0000-0003-4434-3526
Kasia Kozlowska http://orcid.org/0000-0002-9578-0753
Kathrin LaFaver http://orcid.org/0000-0001-5416-894X
W Curt LaFrance, Jr. http://orcid.org/0000-0002-4901-3852
Anthony E Lang http://orcid.org/0000-0003-1229-3667
Alex Lehn http://orcid.org/0000-0003-2780-6496
Sarah Lidstone http://orcid.org/0000-0002-0147-0202
Carine W Maurer http://orcid.org/0000-0002-7892-7127
Bridget Mildon http://orcid.org/0000-0001-7191-3163
Francesca Morgante http://orcid.org/0000-0002-9834-3639
Lorna Myers http://orcid.org/0000-0001-7706-9952
Clare Nicholson http://orcid.org/0000-0002-8047-6820
Glenn Nielsen http://orcid.org/0000-0001-6053-5670
David L Perez http://orcid.org/0000-0003-2721-583X
Stoyan Popkirov http://orcid.org/0000-0001-6168-0036
Markus Reuber http://orcid.org/0000-0002-4104-6705
Karen S Rommelfanger http://orcid.org/0000-0002-2667-1190
Petra Schwingenshuh http://orcid.org/0000-0002-3294-6824
Tereza Serranova http://orcid.org/0000-0001-6525-3971
Paul Shotbolt http://orcid.org/0000-0001-8868-837X
Glenn T Stebbins http://orcid.org/0000-0001-7905-9336
Jon Stone http://orcid.org/0000-0001-9829-8092
Marina AJ Tijssen http://orcid.org/0000-0001-5783-571X
Michele Tinazzi http://orcid.org/0000-0002-2370-7063
Timothy R Nicholson http://orcid.org/0000-0002-2350-2332

REFERENCES

1 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5)*. 5th edn. Washington: American Psychiatric Association, 2013.
2 Espay AJ, Aybek S, Carson A, *et al*. Current concepts in diagnosis and treatment of functional neurological disorders. *JAMA Neurol* 2018;75:1132–41.
3 Carson A, Lehn A. Epidemiology. *Handb Clin Neurol* 2016;139:47–60.
4 Baizabal-Carvallo JF, Hallett M, Jankovic J. Pathogenesis and pathophysiology of functional (psychogenic) movement disorders. *Neurobiol Dis* 2019;127:32–44.
5 Nicholson TR, Carson A, Edwards MJ, *et al*. Outcome measures for functional neurological disorder: a review of the theoretical complexities. *J Neuropsychiatry Clin Neurosci*:appi.neuropsych.
6 Kenney C, Diamond A, Mejia N, *et al*. Distinguishing psychogenic and essential tremor. *J Neurol Sci* 2007;263:94–9.
7 McWhirter L, Stone J, Sandercock P, *et al*. Hoover's sign for the diagnosis of functional weakness: a prospective unblinded cohort study in patients with suspected stroke. *J Psychosom Res* 2011;71:384–6.
8 Benbadis SR. Provocative techniques should be used for the diagnosis of psychogenic nonepileptic seizures. *Epilepsy Behav* 2009;15:106–9.
9 Parees I, Saifee TA, Kassavetis P, *et al*. Believing is perceiving: mismatch between self-report and actigraphy in psychogenic tremor. *Brain* 2012;135:117–23.
10 Gelauff JM, Kingma EM, Kalkman JS, *et al*. Fatigue, not self-rated motor symptom severity, affects quality of life in functional motor disorders. *J Neurol* 2018;265:1803–9.
11 Věchetová G, Slovák M, Kemlink D, *et al*. The impact of non-motor symptoms on the health-related quality of life in patients with functional movement disorders. *J Psychosom Res* 2018;115:32–7.
12 Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. *Arch Gen Psychiatry* 2005;62:903–10.
13 Harris AM, Orav EJ, Bates DW, *et al*. Somatization increases disability independent of comorbidity. *J Gen Intern Med* 2009;24:155–61.
14 Bodde NMG, Janssen AMAJ, Theuns C, *et al*. Factors involved in the long-term prognosis of psychogenic nonepileptic seizures. *J Psychosom Res* 2007;62:545–51.
15 Gelauff J, Stone J, Edwards M, *et al*. The prognosis of functional (psychogenic) motor symptoms: a systematic review. *J Neurol Neurosurg Psychiatry* 2014;85:220–6.
16 Pick S, Mellers JDC, Goldstein LH. Dissociation in patients with dissociative seizures: relationships with trauma and seizure symptoms. *Psychol Med* 2017;47:1215–29.
17 Reuber M, Pukrop R, Bauer J, *et al*. Outcome in psychogenic nonepileptic seizures: 1 to 10-year follow-up in 164 patients. *Ann Neurol* 2003;53:305–11.
18 LaFrance WC, Syc S. Depression and symptoms affect quality of life in psychogenic nonepileptic seizures. *Neurology* 2009;73:366–71.
19 Lawton G, Mayor RJ, Howlett S, *et al*. Psychogenic nonepileptic seizures and health-related quality of life: the relationship with psychological distress and other physical symptoms. *Epilepsy Behav* 2009;14:167–71.
20 Feinstein A, Stergiopoulos V, Fine J, *et al*. Psychiatric outcome in patients with a psychogenic movement disorder: a prospective study. *Neuropsychiatry Neuropsychol Behav Neurol* 2001;14:169–76.
21 Rawlings GH, Brown I, Reuber M. Predictors of health-related quality of life in patients with epilepsy and psychogenic nonepileptic seizures. *Epilepsy Behav* 2017;68:153–8.
22 Williams IA, Howlett S, Levita L, *et al*. Changes in emotion processing following brief augmented Psychodynamic interpersonal therapy for functional neurological symptoms. *Behav Cogn Psychother* 2018;46:350–66.
23 Pick S, Goldstein LH, Perez DL, *et al*. Emotional processing in functional neurological disorder: a review, biopsychosocial model and research agenda. *J Neurol Neurosurg Psychiatry* 2019;90:704–11.

24 Reuber M, Mitchell AJ, Howlett S, *et al*. Measuring outcome in psychogenic nonepileptic seizures: how relevant is seizure remission? *Epilepsia* 2005;46:1788–95.
25 Jones B, Reuber M, Norman P. Correlates of health-related quality of life in adults with psychogenic nonepileptic seizures: a systematic review. *Epilepsia* 2016;57:171–81.
26 Williamson PR, Altman DG, Blazeby JM, *et al*. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;13:132.
27 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
28 Boateng GO, Neilands TB, Frongillo EA, *et al*. Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front Public Health* 2018;6:149.
29 Prinsen CAC, Mokkink LB, Bouter LM, *et al*. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147–57.
30 Prinsen CAC, Vohra S, Rose MR, *et al*. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" – a practical guideline. *Trials* 2016;17:449.
31 Mokkink LB, de Vet HCW, Prinsen CAC, *et al*. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1171–9.
32 Hinson VK, Cubo E, Comella CL, *et al*. Rating scale for psychogenic movement disorders: scale development and Clinimetric testing. *Mov Disord* 2005;20:1592–7.
33 Nielsen G, Ricciardi L, Meppelink AM, *et al*. A simplified version of the psychogenic movement disorders rating scale: the simplified functional movement disorders rating scale (S-FMDRS). *Mov Disord Clin Pract* 2017;4:710–6.
34 Cianci V, Ferlazzo E, Condino F, *et al*. Rating scale for psychogenic nonepileptic seizures: scale development and Clinimetric testing. *Epilepsy Behav* 2011;21:128–31.
35 Sarfraz N, Ijaz T. Development of conversion disorder scale for children. *J Behav Sci* 2014;24:44–56.
36 Ijaz T, Nasir A, Sarfraz N, *et al*. Psychometric properties of conversion disorder Scale-revised (CdS) for children. *J Pak Med Assoc* 2017;67:725–30.
37 Mokkink LB, Terwee CB, Patrick DL, *et al*. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–49.
38 Ataoglu A, Ozcetin A, Icmeli C, *et al*. Paradoxical therapy in conversion reaction. *J Korean Med Sci* 2003;18:581–4.
39 Drane DL, LaRoche SM, Ganesh GA, *et al*. A standardized diagnostic approach and ongoing feedback improves outcome in psychogenic nonepileptic seizures. *Epilepsy Behav* 2016;54:34–9.
40 Goldstein LH, Chalder T, Chigwedere C, *et al*. Cognitive-Behavioral therapy for psychogenic nonepileptic seizures: a pilot RCT. *Neurology* 2010;74:1986–94.
41 Goldstein LH, Mellers JDC, Landau S, *et al*. Cognitive behavioural therapy vs standardised medical care for adults with dissociative non-epileptic seizures (codes): a multicentre randomised controlled trial protocol. *BMC Neurol* 2015;15:98.
42 LaFrance WC, Keitner GI, Papandonatos GD, *et al*. Pilot pharmacologic randomized controlled trial for psychogenic nonepileptic seizures. *Neurology* 2010;75:1166–73.
43 Khattak T, Farooq S, Jan B. Behavior therapy in dissociative convulsions disorder. *J Coll Physicians Surg Pak* 2006;16:359–63.
44 Oto M, Espie CA, Duncan R. An exploratory randomized controlled trial of immediate versus delayed withdrawal of antiepileptic drugs in patients with psychogenic nonepileptic attacks (PNEAs). *Epilepsia* 2010;51:1994–9.
45 Chen DK, Maheshwari A, Franks R, *et al*. Brief group psychoeducation for psychogenic nonepileptic seizures: a neurologist-initiated program in an epilepsy center. *Epilepsia* 2014;55:156–66.
46 Guy W. *US dept health, education and welfare publication (ADM) 76-338*. Rockville, MD: National Institute of Mental Health, 1976: 218–22.
47 Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* 2002;64:258–66.
48 Derogatis LR. SCL-90-R: symptom Checklist-90-R: administration, scoring, and procedures manual. NCS Pearson 1996.
49 Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361–70.
50 Beck AT, Steer RA, Brown GK. . Beck depression inventory-II. Pearson 1996.
51 Beck AT, Epstein N, Brown G, *et al*. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol* 1988;56:893–7.
52 Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62.
53 Hamilton M. The assessment of anxiety states by rating. *Br J Med Psychol* 1959;32:50–5.
54 Bernstein EVEM, Putnam FW. Development, reliability, and validity of a dissociation scale. *J Nerv Ment Dis* 1986;174:727–35.
55 Ware Jr JE, Kosinski M, Keller SD. A 12-Item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220–33.
56 Ware JE, Sherbourne CD. The mos 36-item short-form health survey (SF-36). I. conceptual framework and item selection. *Med Care* 1992;30:473–83.
57 Cramer JA, Perrine K, Devinsky O, *et al*. Development and cross-cultural translations of a 31-Item quality of life in epilepsy inventory. *Epilepsia* 1998;39:81–8.
58 Cramer JA, Perrine K, Devinsky O, *et al*. A brief questionnaire to screen for quality of life in epilepsy the QOLIE-10. *Epilepsia* 1996;37:577–82.
59 Mundt JC, Marks IM, Shear MK, *et al*. The work and social adjustment scale: a simple measure of impairment in functioning. *Br J Psychiatry* 2002;180:461–4.
60 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders - fourth edition (DSM-IV)*. Washington, DC: American Psychiatric Association, 1994.