BMJ Open

Examining the accuracy of a physical assessment technique For Chronic Fatigue Syndrome/Myalgic Encephalomyelitis.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-017521
Article Type:	Research
Date Submitted by the Author:	27-Apr-2017
Complete List of Authors:	Hives, Lucy; University of Central Lancashire, School of Dentistry Bradley, Alice; Lancashire Teaching Hospitals NHS Foundation Trust, Department of Neurophysiology Richards, Jim; University of Central Lancashire, Allied Health Research Unit Sutton, Chris; University of Central Lancashire, Lancashire Clinical Trials Unit Selfe, James; Manchester Metropolitan University, Department of Health Professions Basu, Bhaskar; University Hospital of South Manchester NHS Foundation Trust, Rehabilitation Medicine; Central Manchester University Hospitals NHS Foundation Trust, Rehabilitation Medicine McGuire, Kerry; The Good Health Centre, Department of Osteopathy Sumner, Gail; Back to Health Physiotherapy Gaber, Tarek; Leigh Infirmary, Taylor Rehabilitation Unit Mukherjee, Annice; Salford Royal NHS Foundation Trust, Department of Endocrinology Perrin, Raymond; University of Central Lancashire, Allied Health Research Unit
Primary Subject Heading:	Diagnostics
Secondary Subject Heading:	Evidence based practice
Keywords:	Chronic fatigue syndrome, screening, assessment, physical signs

Examining the accuracy of a physical assessment technique For Chronic Fatigue Syndrome/Myalgic Encephalomyelitis

Lucy Hives, Alice Bradley, Jim Richards, Chris Sutton, James Selfe, Bhaskar Basu, Kerry Maguire, Gail Sumner, Tarek Gaber, Annice Mukherjee, Raymond N. Perrin
LHives1@uclan.ac.uk; alice.bradley@lthtr.nhs.uk; JRichards@uclan.ac.uk; CJSutton@uclan.ac.uk; J.Selfe@mmu.ac.uk; dr_bhaskarbasu@yahoo.co.uk; gail_sum@yahoo.com; kmaguire92@hotmail.com; annicemukherjee@hotmail.com; Tarek.Gaber@wwl.nhs.uk; drperrin@theperrinclinic.com

Abstract

Objective: To assess 5 physical signs to see whether they can assist in the screening of patients with CFS/ME, and potentially lead to quicker treatment.
Methods: This was a diagnostic accuracy study with inter-rater agreement assessment. Participants recruited from 2 NHS hospitals, local CFS/ME support groups and the community were examined by three practitioners on the same day in a randomized order. Two Allied Health Professionals (AHPs) performed independent examinations of physical signs including; postural/mechanical disturbances of the thoracic spine, breast varicosities, tender Perrin's Point, tender coeliac plexus and dampened cranial flow. A physician conducted a standard clinical neurological and rheumatological assessment, whilst looking for patterns of illness behaviour. Each examination lasted approximately 20 minutes.
Results: Ninety-four participants were assessed, 52 CFS/ME patients and 42 non-CFS/ME controls, aged 18-60. Cohen's kappa revealed agreement between the AHPs was substantial for presence of the tender coeliac plexus (κ=0.65, p<0.001) and moderate for postural/mechanical disturbance of the thoracic spine (κ=0.57, p<0.001) and Perrin's point (κ=0.56, p<0.001). A McNemar's test found no statistically significant bias in the diagnosis by the experienced AHP relative to actual diagnosis, (p=1.0) and a marginally non-significant bias by the newly trained AHP, p=0.052. There was however, a significant bias in the diagnosis made by the physician relative to actual diagnosis, (p<0.001), indicating poor diagnostic utility of the clinical neurological and rheumatological assessment.
Conclusions: Using the physical signs appears to improve the accuracy of identifying people with CFS/ME and shows agreement with current diagnostic techniques, however the present study concludes that only 2 of these may be needed. Examining for physical signs is both quick and simple for the AHP and may be used as an efficient screening tool for CFS/ME. Future research should investigate the physical signs in people with more severe CFS/ME.

Strengths and Limitations:

- This is the first study that explores agreement on the presence of physical signs in the screening of CFS/ME patients.
- The screening method did not involve patient/family history, patient symptoms or any discussion between practitioner and patient, which would increase accuracy further in clinical practice.
- The screening test was simple and quick for Allied Health Professionals to use.
- In total, there were more female participants than males, at a rate similar to other studies, which have found a higher prevalence of CFS/ME in females.
- This study does not include participants with severe CFS/ME, and so that results cannot be generalised to these individuals.

Introduction

Chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) is characterised by severe, debilitating fatigue that is exacerbated by exercise, but does not improve with rest. This condition can lead to a substantial impairment, making every day activities difficult. There is currently no universally accepted method of diagnosing CFS/ME, so other conditions with a similar presentation of symptoms must first be ruled out. Therefore the diagnosis of CFS/ME can often be a long process.

Up until recently the most widely accepted criteria for CFS/ME was the revised United States Centres for Disease Control and Prevention (CDC) definition (Fukuda et al., 1994), which required at least a six month period of fatigue that significantly interferes with a person's everyday activities. In addition to this, four or more of the following symptoms must have persisted or reoccurred within the last six months; impaired memory or concentration, post-exertion malaise, sore throat, tender lymph nodes, aching or stiff muscles, joint pain, headache and unrefreshing sleep. The latest internationally recognized diagnostic criteria for CFS/ME is the International Consensus Criteria (Carruthers et al., 2011) based on the widely adopted Canadian Criteria (Carruthers, 2007).

The Canadian criteria included many of the cardiopulmonary and neurological abnormalities, which were not included in the CDC criteria. In addition, the Canadian criteria selected cases with less psychiatric co-morbidity, more physical functional impairment, more fatigue/weakness, plus neurological symptoms, which were significantly different from psychiatric controls with CFS/ME (Jason et al., 2010). However, in the UK, The National Institute for Health and Care Excellence (NICE) recognises the heterogeneity of the condition advising that diagnosis of CFS/ME should be made after other possible diagnoses have been excluded and the symptoms have persisted for at least four months. The NICE guidance also states that diagnosis should be reassessed if none of the following key features of the disorder are present; post exertional fatigue, cognitive difficulties, sleep problems or chronic pain (Health and Care Excellence, 2007).

Due to the heterogeneity of the disorder, the aetiology of CFS/ME remains unknown with many theories surrounding the pathophysiology of the disorder (Bansal et al., 2012). The literature suggests that a range of possible causes including hormonal disturbances, immune system dysfunction, infectious and viral agents and nervous system abnormalities may all play a role in the pathophysiology of the disease (Shephard, 2001). Early research suggested that infectious agents such as the Epstein-Barr virus (EBV) are associated with CFS/ME, with a persisting EBV infection being seen in those with the condition (Straus et al., 1985). A number of other infectious agents have been linked to the onset of CFS/ME including enteroviruses, which could explain the gastrointestinal symptoms often seen in patients (Chia and Chia, 2008), and also the acute B19 virus infection (Kerr et al., 2002). It has been suggested that viral infections can alter immune response which in turn, can chronically activate the immune system (Lorusso et al., 2009) and lead to many of the symptoms associated with CFS/ME. However, the research within this field is inconsistent with no evidence of a single infection causing CFS/ME, as well as many patients showing no sign of previous infection (Afari and Buchwald, 2014), suggesting infectious agents may only be relevant to a subset of the patient population.

Lymphatic system alterations are suggested to have involvement in CFS/ME with dysfunction within the immune system, causing toxic build up within the central nervous

system (CNS) and leading to engorgement of varicose lymph vessels that can be felt on examination (Perrin et al., 2011). Tender lymph nodes are included in the ICC, Canadian, States Centres for Disease Control and Prevention (CDC) definitions (Fukuda et al., 1994) and the NICE guidance (Health and Care Excellence, 2007) confirming it is a common symptom of the disorder related to immune system abnormalities. Techniques to target these engorgements and stimulate the drainage of toxins in the lymph nodes has been shown to lead to symptom improvement in patients with CFS/ME (Perrin et al., 2011).

Currently there is no definitive way of diagnosing patients, although recent research has suggested that there is a link between CFS/ME, the lymphatic drainage system and the CNS and that, in fact, CFS/ME patients have certain physical signs present that may explain a number of the characteristics of the condition (Perrin et al., 2011). The aim of this study is to assess these physical signs to see whether they can assist in the screening of patients with CFS/ME, which could then subsequently lead to quicker treatment.

Methods

This was a diagnostic accuracy cohort study with inter-rater agreement assessment. Before the study began, a protocol was developed, which is available upon request from the corresponding author. CFS/ME participants were recruited from hospitals and clinics of the Wrightington, Wigan and Leigh NHS Foundation and Salford Royal NHS Foundation Trusts and support groups within the North West, and circulated to the wider community by the research team. Healthy participants were recruited from non-blood relatives and friends of people with CFS/ME, as well as staff and students from the University of Central Lancashire. All participants voluntarily contacted the research team to register their interest.

On receiving their completed consent form, participants were contacted by telephone in order to assess their eligibility to take part in the study. Potential participants, aged 18-60, were assessed using two forms: a recruitment screening form based on the NICE guidance 2007 & 2011, and a form based on the International Consensus Criteria (Carruthers et al., 2011). Included participants were assigned to either the CFS/ME or healthy (control) group. Each participant was allocated a participant ID number with the clinical team being blinded to the groupings. The study received ethical approval from NRES Committee North West - Lancaster, REC reference 12/NW/0877, R&D approval was also obtained from each participating NHS trust.

CFS/ME group

Briefly, participants needed to have received a formal diagnosis of CFS/ME and have had persistent or recurrent fatigue for at least the past 4 months. They were also required to be able to attend their 1-hour session without the use of a wheelchair. In addition, there needed to be 1) a clear starting point to the fatigue, 2) the fatigue should be unexplained by any other conditions, 3) the fatigue should reduce the amount of physical activity each person could do, and 4) the fatigue should feel worse after physical activity.
Additionally, participants had at least one of the following symptoms; difficulty sleeping or insomnia, joint pain without swelling, headaches, painful lymph nodes that are not enlarged, recurrent sore throats, muscle pain without swelling, poor mental function, such as difficulty thinking, symptoms getting worse after physical or mental exertion, feeling unwell or having flu-like symptoms, dizziness or nausea, heart palpitations.

Exclusionary illnesses included, anaemias, autoimmune diseases, cardiac disease, endocrine disorders, infectious diseases, intestinal diseases, malignancies, neurological disorders, primary psychiatric disorders, significant pulmonary disease and primary sleep disorders.

Control group

Participants were either non-blood relatives or friends of people with CFS/ME or students or staff from The University of Central Lancashire. Each participant answered "no" to the question, "have you suffered persistent or recurrent fatigue for at least the past 4 months?". The same exclusionary illnesses applied.

Data collection

For each participant, all data collection was performed on a single assessment day. They were allowed to bring a friend or relative along for support, or if requested, a chaperone was provided by the research team. Participants were re-briefed on what would be involved and consent was confirmed.

Each of the participants were examined by three practitioners; one Allied Health Professional (AHP) with experience of using the Perrin technique and working with CFS/ME patients (experienced AHP), another AHP newly trained in the Perrin technique with no prior experience of CFS/ME (newly trained AHP), and a physician with experience of working in NHS clinics for CFS/ME but with no experience of the Perrin technique (physician). The order of examinations was randomised. No conversation took place between the participants and practitioners except to determine if there was any pain or tenderness in certain regions. The two AHPs performed an examination of the following physical signs:

A. Participant standing: observation and palpation of thoracic spine for any postural defects; regions of redness, temperature change or skin rashes or eruptions e.g. acne/boils.
B. Participant lying supine: observation and palpation of breast tissue for varicosities in the surface lymphatics and abnormal breast tenderness at 'Perrin's Point' which is a superficial tender area found at around 2-3 cm lateral and superior to the left nipple (Puri et al, 2011).
C. With the participant remaining supine, palpation of the region of the coeliac plexus just below the xiphoid in the upper central area of the abdomen for any abnormal tenderness with possible temperature change in the region.
D. With the participant still remaining in a supine position, cradle the head and examine the quality of the cranial rhythmic impulse.

These assessments resulted in identification of the following signs as present or not present; 1) postural/mechanical disturbances of the thoracic spine (assessment A); 2) breast varicosities (assessment B); 3) tender Perrin's Point (assessment B); 4) tender coeliac plexus (assessment C); and 5) dampened cranial flow (assessment D). If all five signs were present then the participant was classified as having CFS/ME. If one or more of the signs was absent, then the participant was classified as not having CFS/ME.

The physician conducted a standard clinical neurological and rheumatological assessment, whilst observing the participant for any signs of illness behaviour but no clinical history was taken. The neurological examination included muscle strength testing, examination of muscle

tone in arms and legs, coordination including the finger nose test, heel-shin test, heel-toe walking, reflexes and sensation with eyes closed. The rheumatological examination examined joint swelling, wasting of regional muscles, deformity of joint, redness in joints or tendons, and the palpation of the margin of joints in hands and feet. If all the tests were normal with no observed illness behaviour, the patient was classified as not having CFS/ME, whereas if abnormal observations were made, the physician used their clinical experience to decide if the participant had CFS/ME.

Data analysis

The sensitivity and specificity of the diagnosis of CFS/ME (relative to the reference standard) were each estimated as simple proportions, accompanied by exact (binomial) 95% confidence intervals.McNemar's test was used to investigate whether any of the practitioners systematically under- or over-diagnosed CFS/ME. Agreement in the diagnosis of CFS/ME between AHPs using the Perrin Technique was estimated using Cohen's Kappa (κ) coefficient; an approximate 95% confidence interval for κ was obtained using bias-corrected non-parametric bootstrapping. Agreement between their identification of the individual physical signs was also estimated using the same methods.

Sample size

The target sample size was 50 with CFS/ME and 50 controls to enable estimation of sensitivity and specificity of each Perrin technique AHP's diagnosis of CFS/ME (relative to the imperfect reference standard) to within ±9.9% with 95% confidence if the sensitivity and specificity were each at least 85%. It would also enable the estimation of κ (for inter-rater agreement between each pair), with 95% confidence, to within ±0.140 providing κ were at least 0.7.

Results

Ninety-four participants were recruited: 52 CFS/ME patients and 42 non-CFS/ME controls. There was no missing data. The gender ratio in the CFS/ME group (shown in table 1) is in keeping with epidemiological studies which have shown a larger number of CFS/ME patients to be female with a ratio of 2:1 or more (Jason et al,1999). Results show that, on average, the experienced AHP was most accurate (86%) at correctly diagnosing participants. This was followed by the newly trained who correctly diagnosed 77%, and the physician who correctly diagnosed 69% of participants.

	CFS/ME	Control
Male	9 (17%)	25 (60%)
Female	43 (83%)	17 (40%)

Table 1: Gender balance of the two groups

Sensitivity and Specificity:

Table 2 shows the prevalence, sensitivity (the proportion of positive results in people with CFS/ME), specificity (the proportion of negative CFS/ME results in healthy controls), positive predictor value (the proportion with CFS/ME in participants with a positive result)

and negative predictive value (the proportion of healthy controls with negative CFS/ME results).

Practitioner	True Positives	False Negatives	False Positives	True Negatives	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Accuracy (%)
Experienced AHP	46	6	7	35	0.88	0.83	0.87	0.85	86.2
Newly trained AHP	36	16	6	36	0.69	0.86	0.86	0.69	76.6
Physician	23	29	0	42	0.44	1.00	1.00	0.57	69.1

Table 2: Diagnostic performance of the three practitioners

When using the five physical signs of the Perrin Technique, the sensitivity for the experienced AHP was 0.88 (95% CI 0.77 to 0.96) and the specificity was 0.83 (95% CI 0.69 to 0.93). Similarly, when the newly trained AHP used the same technique, the specificity was 0.86 (95% CI 0.71 to 0.95), however sensitivity was lowered to 0.69 (95% CI 0.55 to 0.81). This shows that although using the same technique, the newly trained AHP struggled more frequently to identify correctly all five physical signs in people with a positive diagnosis of CFS/ME. There was no statistically significant bias in the diagnosis by the experienced AHP relative to actual diagnosis ($p = 1.0$). There was also a marginally non-significant evidence of biased diagnosis by the newly trained AHP relative to actual diagnosis ($p = 0.052$) potentially favouring a non-CFS/ME diagnosis.

When using the standard clinical neurological and rheumatological examination, the sensitivity of the physician was 0.44 (95% CI 0.30 to 0.59) and the specificity was 1.0 (95% CI 0.92 to 1.0). These results show that whilst able to identify correctly all healthy controls, the physician struggled the most out of all three practitioners to identify correctly people with a positive diagnosis of CFS/ME. There was a significant bias in the diagnosis by the physician relative to actual diagnosis ($p < 0.001$), also favouring a non-CFS diagnosis.

Agreement between the experienced and newly trained AHPs

There was moderate agreement between the experienced and newly trained AHPs on overall diagnosis using the 5 physical signs of the Perrin Technique ($\kappa = 0.56$, 95% CI 0.40 to 0.72, $p < 0.001$). Regarding the identification of the individual physical signs, there was substantial agreement between the AHPs on the presence of the tender coeliac plexus ($\kappa = 0.65$; 95% CI 0.48 to 0.80, $p < 0.001$) and agreement was moderate both on the presence of postural/mechanical disturbance of the thoracic spine ($\kappa = 0.57$; 95% CI 0.39 to 0.73, $p < 0.001$) and on the presence of Perrin's point ($\kappa = 0.56$; 95% CI 0.37 to 0.73, $p < 0.001$). However, there was only fair agreement between the AHP's identification of the dampened cranial flow ($\kappa = 0.35$; 95% CI 0.15 to 0.54, $p = 0.001$) and there was non-significant 'slight' agreement on the presence of breast varicosities ($\kappa = 0.03$; 95% CI -0.12 to 0.22, $p = 0.75$).

Diagnostic properties of alternative criteria

Given the relatively low sensitivity for the newly trained AHP, alternative criteria were applied using lower number of physical signs. Using the Kappa statistics for agreement between the experienced and newly trained AHPs on the presence of the Perrin technique physical signs, each physical sign, in order starting with the sign with the least agreement, was excluded and sensitivity and specificity were recalculated. Overall, the AHPs were more accurate at diagnosing participants when using only 2 of the 5 signs (tender coeliac plexus and postural/ mechanical disturbance of the thoracic spine). Tables 3 and 4 show the accuracy

of the experienced and newly trained AHPs, respectively, when using each number of physical signs. The accuracy of the experienced AHP is the same using 3-5 of the physical signs (86.2%) with the highest accuracy using only 2 of the physical signs (88.3%). The accuracy of the newly trained AHP is highest using only 1 or 2 of the physical signs (80.9%). Therefore, accuracy for both AHPs, overall, is highest when using only tests of tender coeliac plexus and postural/ mechanical disturbance of the thoracic spine.

Table 3: Experienced AHP

Number of Items	True Positives	False Positives	True Negatives	False Negatives	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Accuracy (%)
5*	46	7	35	6	0.88	0.83	0.87	0.85	86.2
4*	46	7	35	6	0.88	0.83	0.87	0.85	86.2
3*	46	7	35	6	0.88	0.83	0.87	0.85	86.2
2*	48	7	35	4	0.92	0.83	0.87	0.90	88.3
1*	49	12	30	3	0.94	0.71	0.80	0.91	84.0

Table 4: Newly trained AHP

Number of Items	True Positives	False Positives	True Negatives	False Negatives	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Accuracy (%)
5*	36	6	36	16	0.69	0.86	0.86	0.69	76.6
4*	36	7	35	16	0.69	0.83	0.84	0.69	75.5
3*	40	7	35	12	0.77	0.83	0.85	0.74	79.8
2*	42	8	34	10	0.81	0.81	0.84	0.77	80.9
1*	48	14	28	4	0.92	0.67	0.77	0.88	80.9

5* includes all 5 Perrin technique physical signs; 4* includes postural/mechanical disturbances of the thoracic spine, tender Perrin's Point, tender coeliac plexus, dampened cranial flow; 3* includes postural/mechanical disturbances of the thoracic spine, tender Perrin's Point, tender coeliac plexus; 2* includes tender coeliac plexus and postural/mechanical disturbance of the thoracic spine; 1* includes tender coeliac plexus.

Cohen's κ was re-computed for the level of agreement between the experienced and newly trained AHPs on whether they believed the 94 individuals had CFS/ME or were healthy controls, using the reduced 2 physical signs of the Perrin Technique. There was substantial agreement between the two AHPs on overall diagnosis using the 2 physical signs, $\kappa = 0.61$ (95% CI, 0.45 to 0.74), $p < 0.001$. There was no statistically significant bias in the diagnosis by the experienced AHP and actual diagnosis ($p = 0.55$), or by the newly trained AHP and actual diagnosis ($p = 0.63$) when using the reduced 2 item Perrin Technique, showing that the revised criteria no longer favoured a non-CFS/ME diagnosis.

Discussion

Between the AHPs, the AHP with prior experience of using the Perrin technique was the most accurate at correctly diagnosing individuals with CFS/ME, whereas the AHP with no prior experience of CFS/ME or the Perrin technique was better at correctly recognising healthy individuals.

The AHP experienced in the Perrin technique, was able to identify 88% of patients with CFS/ME using all five physical signs, and 83% of healthy controls who did not display all five signs. However, for the AHP with no prior Perrin technique experience, they were able to detect 86% of healthy controls, but identified only 69% of patients with CFS/ME using all five signs; there was borderline non-significant evidence of this AHP underdiagnosing rather than over-diagnosing CFS/ME. This highlights that the newly trained AHP was not able to

identify all 5 signs in some people with a diagnosis of CFS/ME. However, this does not necessarily mean that the signs were not present; it could mean that the newly trained AHP found these signs more difficult to detect. Despite this, there was moderate agreement between both AHPs on overall diagnosis.

The physician was able to correctly identify 100% of the healthy controls using the standard clinical neurological and rheumatological examination. However, they were only able to correctly identify 44% of patients with CFS/ME and the tendency to under-diagnose CFS/ME was highly statistically significant ($p<0.001$). This affirms the current approach used in CFS diagnostics based on NICE guidance in that clinical examination is most useful in identifying alternative diagnoses and to exclude the diagnosis of CFS/ME, but that clinical examination is not a useful modality for confirming diagnosis of CFS/ME.

The agreement of the AHPs on the presence of each of the five physical signs varied from substantial agreement on the presence of the tender coeliac plexus to non-significant 'slight' agreement on the presence of breast varicosities. From the results presented, it would seem that the physical signs can improve the accuracy of diagnosing CFS/ME, although not all of the 5 physical signs are necessary. Even with the experienced AHP, who, on the whole, was able to identify the 5 signs, breast varicosities and dampened cranial flow did not improve accuracy of diagnosis. Further exploration of the sensitivity, specificity and accuracy found that using only 2 of the 5 physical signs (tender coeliac plexus and postural/mechanical disturbance of the thoracic spine) was the most accurate and efficient method of correctly diagnosing the participants for AHPs with differing levels of prior experience of CFS/ME and the Perrin technique.

Previous work by Puri et al (2011) found Perrin's point to have a diagnostic accuracy of 80% in patients with CFS/ME. This was very similar to the accuracy when including Perrin's point in the current study (Accuracy; experienced AHP = 86.2% and newly trained AHP = 79.8%). However it was found that the omission of Perrin's point marginally increased the accuracy of the AHP with prior experience of the Perrin Technique by 2.1% and the AHP with no prior experience of the Perrin Technique by 1.1%.

Limitations and suggestions for future research

Although this study shows clearly that diagnostic accuracy for CFS/ME increases using the physical signs of the Perrin technique, there are some limitations which should be highlighted. Firstly, this study recruited two groups of participants; people with a prior diagnosis of CFS/ME and healthy controls with no symptoms of the condition. Therefore, the study did not explore how accurate the Perrin physical signs would be, when presented with patients with different conditions with similar presenting symptoms.

The gender balance in the CFS/ME patients was similar to that of previously published work, although the healthy volunteers were recruited as a convenience sample which was not gender matched, therefore any prior knowledge of the expected male-female ratio amongst CFS/ME patients could have influenced the results. It should be noted, however, that none of the practitioners knew if the control group was matched or not for gender and that the AHP with no prior experience of CFS/ME was unaware of the gender balance.

Future research should investigate whether the physical signs are more apparent in people with more severe CFS/ME. The present study did not collect data how long each participant

had had CFS/ME for or on the severity of their symptoms, which could add further understanding.

Conclusion

Current methods for diagnosing CFS/ME are challenging. The use of standard clinical neurological and rheumatological examination to examine illness behaviour is more likely to have a false negative result than a true positive one. Using certain physical signs appears to improve the accuracy of identifying people with CFS/ME and shows agreement with current diagnostic techniques, although not all of the physical signs were useful, and it is suggested that only 2 of these are needed. Examining for physical signs is both quick and simple for the AHP and may be used as an efficient screening tool for CFS/ME. This study did not include patient/family history or the patient talking about their symptoms, which should increase accuracy in clinical practice.

Contributership Statement:

LH recruited participants, collected data and analysed the data. AB recruited participants and collected data. JR initiated the project, designed the research methods and analysed the data. CS wrote the statistical analysis plan and analysed the data. JS initiated the project and designed the research methods. BB, KM and GS collected data. TG and AM recruited participants. RP initiated the project and designed the research methods. All authors drafted and revised the paper.

Competing interests:

This research explored the findings by one of the co-authors, Dr Perrin. To avoid any conflict of interest he was not involved in any of the recruitment of participants, clinical examinations, data collection or analysis. Dr Perrin's role in the study was to assist with developing the design, writing the protocol, setting up the project coordinating committee with the different clinical recruitment centres, applying for ethical approval, and assisting with the writing of the introduction and methods of the paper. There were no other conflicts of interest.

Funding:

This work was funded by: The Fund for Osteopathic Research into Myalgic Encephalomyelitis (F.O.R. M.E.) Registered charity number: 1045005

Data sharing statement:

The relevant anonymised patient level data are available on request from the corresponding author.

References

Afari N, Buchwald D. Chronic Fatigue Syndrome: A Review. American Journal Of Psychiatry 2003;160(2):221-36.

Bansal A, Bradley A, Bishop K, Kiani-Alikhan S, Ford B. Chronic Fatigue Syndrome, The Immune System And Viral Infection. Brain Behavior and Immunity 2012;26,24-31.

Carruthers BM. Definitions and aetiology of myalgic encephalomyelitis: how the Canadian consensus clinical definition of myalgic encephalomyelitis works. J Clin Pathol. 2007 Feb; 60(2): 117-9

Carruthers BM, van de Sande MI, De Meirleir KL, Klimas NG, Broderick G, Mitchell T, et al. Myalgic encephalomyelitis: International consensus criteria. J Intern Med. 2011; 270;4: 327-338

Chia JK, Chia AY. Chronic fatigue syndrome is associated with chronic enterovirus infection of the stomach. Journal of clinical pathology 2008;61:43-48.

Cleare AJ. The HPA axis and the genesis of chronic fatigue syndrome. Trends in Endocrinology & Metabolism 2004;15:55-59.

Cleare AJ, Blair D, Chambers S, Wessely S. Urinary free cortisol in chronic fatigue syndrome. American Journal of Psychiatry 2001;158(4):641-643

Cleare AJ, Heap E, Malhi GS, Wessely S, O'keane V, Miell J.. Low-dose hydrocortisone in chronic fatigue syndrome: a randomised crossover trial. The Lancet 1999;353:455-458.

De Lange FP, Kalkman JS, Bleijenberg G, Hagoort P, Van Der Meer JW, Toni I. Gray matter volume reduction in the chronic fatigue syndrome. Neuroimage 2005:26:777-781.

Fries E, Hesse J, Hellhammer J, Hellhammer DH. A new view on hypocortisolism. Psychoneuroendocrinology 2005;30:1010-1016.

Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff, A.. The chronic fatigue syndrome: a comprehensive approach to its definition and study. Annals of internal medicine 1994:;121:953-959.

Gaab J, Rohleder N, Heitz V, Engert V, Schad T, Schürmeyer TH, Ehlert U. Stress-induced changes in LPS-induced pro-inflammatory cytokine production in chronic fatigue syndrome. Psychoneuroendocrinology 2005;30:188-198.

Hardcastle SL, Brenu EW, Johnston S, Nguyen T, Huth T, Wong N, Ramos S, Staines D, Marshall-Gradisnik, S.. Characterisation of cell functions and receptors in Chronic Fatigue Syndrome/Myalgic Encephalomyelitis (CFS/ME). BMC immunology 2015;16:35.

Jason LA, Richman JA, Rademaker AW, Jordan KM, Plioplys AV, Taylor RR, McCready W, Huang C, Plioplys S. A Community-Based Study of Chronic Fatigue Syndrome. Arch Intern Med 1999;159(18):2129-2137.

Kerr JR, Bracewell J, Laing I, Mattey DL, Bernstein RM, Bruce IN, Tyrrell DA. Chronic fatigue syndrome and arthralgia following parvovirus B19 infection. The Journal of rheumatology 2002;29: 595-602.

Lange G, Deluca J, Maldjian JA. Lee HJ, Tiersky LA, Natelson BH. Brain MRI abnormalities exist in a subset of patients with chronic fatigue syndrome. Journal of the neurological sciences, 1999;171:3-7.

Lorusso L, Mikhaylova SV, Capelli E, Ferrari D, Ngonga GK, Ricevuti G.. Immunological aspects of chronic fatigue syndrome. Autoimmunity reviews 2009;8:287-291.

Mckenzie R, O'Fallon A, Dale J, Demitrack M, Sharma G, Deloria M, Garcia-Borreguero D, Blackwelder W, Straus SE.. Low-dose hydrocortisone for treatment of chronic fatigue syndrome: a randomized controlled trial. JAMA 1998;280:1061-1066.

NICE Clinical Guidelines. Chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy): diagnosis and management of chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy) in adults and children. NICE Clinical Guidelines 2007;53.

Parker A, Wessely S, Cleare, A.. The neuroendocrinology of chronic fatigue syndrome and fibromyalgia. Psychological medicine 2001;31:1331-1345.

Patarca-Montero R, Antoni M, Fletcher MA, Klimas NG. Cytokine and other immunologic markers in chronic fatigue syndrome and their relation to neuropsychological factors. Applied Neuropsychology 2001;8:51-64.

Perrin RN, Richards JD, Pentreath V, Percy DF. Muscle fatigue in chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) and its response to a manual therapeutic approach: A pilot study. International Journal of Osteopathic Medicine 2011;14:96-105.

Peterson D, Brenu E, Gottschalk G, Ramos S, Nguyen T, Staines D, Marshall-Gradisnik S. Cytokines in the Cerebrospinal Fluids of Patients with Chronic Fatigue Syndrome/Myalgic Encephalomyelitis. Mediators of inflammation 2015.

Puri BK, Gunatilake KD, Fernando KA, Gurusinghe AI, Agour M, Treasaden IH.. Increased tenderness in the left third intercostal space in adult patients with myalgic encephalomyelitis: a controlled study. J Int Med Res 2011;39(1):212-4

Scott L, Medbak S, Dinan T. Blunted adrenocorticotropin and Cortisol responses to corticotropin-releasing hormone stimulation in chronic fatigue syndrome. Acta Psychiatrica Scandinavica 1998;97:450-457.

Shephard RJ. Chronic fatigue syndrome. Sports Medicine 2001;31:167-194.

Straus SE, Tosato G, Armstrong G, Lawley T, Preble OT, Henle W, Davey R, Pearson G, Epstein J, Brus I. Persisting illness and fatigue in adults with evidence of Epstein-Barr virus infection. Annals of internal medicine 1985;102:7-16.

Strickland P, Morriss R, Wearden A, Deakin B. A comparison of salivary cortisol in chronic fatigue syndrome, community depression and healthy controls. Journal of Affective Disorders 1998;47:191-94.

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	1
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	2-3
	4	Study objectives and hypotheses	3
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	3
Participants	**6**	Eligibility criteria	3-4
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	3-4
	8	Where and when potentially eligible participants were identified (setting, location and dates)	3
	9	Whether participants formed a consecutive, random or convenience series	4
Test methods	**10a**	Index test, in sufficient detail to allow replication	4-5
	10b	Reference standard, in sufficient detail to allow replication	3
	11	Rationale for choosing the reference standard (if alternatives exist)	3
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	3
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	3
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	4
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	N/A
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	5
	15	How indeterminate index test or reference standard results were handled	5
	16	How missing data on the index test and reference standard were handled	5
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	5
	18	Intended sample size and how it was determined	5
RESULTS			
Participants	**19**	Flow of participants, using a diagram	N/A
	20	Baseline demographic and clinical characteristics of participants	5
	21a	Distribution of severity of disease in those with the target condition	N/A
	21b	Distribution of alternative diagnoses in those without the target condition	N/A
	22	Time interval and any clinical interventions between index test and reference standard	N/A
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	6-7
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	6-7
	25	Any adverse events from performing the index test or the reference standard	NONE
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	8
	27	Implications for practice, including the intended use and clinical role of the index test	9
OTHER INFORMATION			
	28	Registration number and name of registry	N/A
	29	Where the full study protocol can be accessed	3
	30	Sources of funding and other support; role of funders	9

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

BMJ Open

Can physical assessment techniques aid diagnosis in people with chronic fatigue syndrome/myalgic encephalomyelitis? A diagnostic accuracy study.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-017521.R1
Article Type:	Research
Date Submitted by the Author:	23-Aug-2017
Complete List of Authors:	Hives, Lucy; University of Central Lancashire, School of Dentistry Bradley, Alice; Lancashire Teaching Hospitals NHS Foundation Trust, Department of Neurophysiology Richards, Jim; University of Central Lancashire, Allied Health Research Unit Sutton, Chris; University of Central Lancashire, Lancashire Clinical Trials Unit Selfe, James; Manchester Metropolitan University, Department of Health Professions Basu, Bhaskar; University Hospital of South Manchester NHS Foundation Trust, Rehabilitation Medicine; Central Manchester University Hospitals NHS Foundation Trust, Rehabilitation Medicine McGuire, Kerry; The Good Health Centre, Department of Osteopathy Sumner, Gail; Back to Health Physiotherapy Gaber, Tarek; Leigh Infirmary, Taylor Rehabilitation Unit Mukherjee, Annice; Salford Royal NHS Foundation Trust, Department of Endocrinology Perrin, Raymond; University of Central Lancashire, Allied Health Research Unit
Primary Subject Heading:	Diagnostics
Secondary Subject Heading:	Evidence based practice
Keywords:	Chronic fatigue syndrome, screening, assessment, physical signs

Can physical assessment techniques aid diagnosis in people with chronic fatigue syndrome/myalgic encephalomyelitis? A diagnostic accuracy study.

Lucy Hives, Alice Bradley, Jim Richards, Chris Sutton, James Selfe, Bhaskar Basu, Kerry Maguire, Gail Sumner, Tarek Gaber, Annice Mukherjee, Raymond N. Perrin.

Corresponding Author: Jim Richards, Allied Health Research Unit, University of Central, Lancashire, Preston UK; JRichards@uclan.ac.uk; tel: 01772 894575.

Lucy Hives, University of Central Lancashire, School of Dentistry, Preston, UK.
Alice Bradley, Lancashire Teaching Hospitals NHS Foundation Trust, Department of Neurophysiology, Preston, UK.
Chris Sutton, University of Central Lancashire, Lancashire Clinical Trials Unit, Preston, UK.
James Selfe, Manchester Metropolitan University, Department of Health Professions, Manchester, UK.
Bhaskar Basu, University Hospital of South Manchester NHS Foundation Trust/ Central Manchester University Hospitals NHS Foundation Trust, Rehabilitation Medicine, Manchester, UK.
Kerry Maguire, The Good Health Centre, Department of Osteopathy, County Clare, Ireland.
Gail Sumner, Back to Health Physiotherapy, Bolton, UK.
Tarek Gaber, Leigh Infirmary, Taylor Rehabilitation Unit Leigh, Greater Manchester, UK.
Annice Mukherjee, Salford Royal NHS Foundation Trust, Department of Endocrinology, Salford, UK.
Raymond N. Perrin, University of Central Lancashire, Allied Health Research Unit, Preston, UK.

Keywords: Chronic fatigue syndrome, screening, assessment, physical signs.

Word count: 4,148

Abstract

Objective: To assess 5 physical signs to see whether they can assist in the screening of patients with CFS/ME, and potentially lead to quicker treatment.
Methods: This was a diagnostic accuracy study with inter-rater agreement assessment. Participants recruited from 2 NHS hospitals, local CFS/ME support groups and the community were examined by three practitioners on the same day in a randomized order. Two Allied Health Professionals (AHPs) performed independent examinations of physical signs including; postural/mechanical disturbances of the thoracic spine, breast varicosities, tender Perrin's Point, tender coeliac plexus and dampened cranial flow. A physician conducted a standard clinical neurological and rheumatological assessment, whilst looking for patterns of illness behaviour. Each examination lasted approximately 20 minutes.
Results: Ninety-four participants were assessed, 52 CFS/ME patients and 42 non-CFS/ME controls, aged 18-60. Cohen's kappa revealed agreement between the AHPs was substantial for presence of the tender coeliac plexus (κ=0.65, p<0.001) and moderate for postural/mechanical disturbance of the thoracic spine (κ=0.57, p<0.001) and Perrin's point (κ=0.56, p<0.001). A McNemar's test found no statistically significant bias in the diagnosis by the experienced AHP relative to actual diagnosis, (p=1.0) and a marginally non-significant

bias by the newly trained AHP, p=0.052. There was however, a significant bias in the diagnosis made by the physician relative to actual diagnosis, (p<0.001), indicating poor diagnostic utility of the clinical neurological and rheumatological assessment.
Conclusions: Using the physical signs appears to improve the accuracy of identifying people with CFS/ME and shows agreement with current diagnostic techniques, however the present study concludes that only 2 of these may be needed. Examining for physical signs is both quick and simple for the AHP and may be used as an efficient screening tool for CFS/ME. This is a small single centre study and therefore further validation in other centres and larger populations is needed.

Strengths and Limitations:

- This is the first study that explores agreement on the presence of physical signs in the screening of CFS/ME patients and demonstrates proof-of-concept of these signs.
- This study did not assess the performance of physical signs in diagnosing CFS/ME amongst people reporting with illness in clinical practice.
- The screening method did not involve patient/family history, patient symptoms or any discussion between practitioner and patient; including these would be likely to increase accuracy in clinical practice.
- There were more female participants than males, at a rate similar to other studies, which have found a higher prevalence of CFS/ME in females; however, there were similar percentages of males and females amongst the controls.
- This was a small study which included only two AHPs using the Perrin technique and did not include participants with severe CFS/ME; this limits the generalisation of the findings.

Introduction

Chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) is characterised by severe, debilitating fatigue that is exacerbated by exercise, but does not improve with rest. This condition can lead to a substantial impairment, making every day activities difficult. There is currently no universally accepted method of diagnosing CFS/ME, so other conditions with a similar presentation of symptoms must first be ruled out. Therefore the diagnosis of CFS/ME can often be a long process.

Up until recently the most widely accepted criteria for CFS/ME was the revised United States Centres for Disease Control and Prevention (CDC) definition [1], which required at least a six month period of fatigue that significantly interferes with a person's everyday activities. In addition to this, four or more of the following symptoms must have persisted or reoccurred within the last six months; impaired memory or concentration, post-exertion malaise, sore throat, tender lymph nodes, aching or stiff muscles, joint pain, headache and unrefreshing sleep. The latest internationally recognized diagnostic criteria for CFS/ME is the International Consensus Criteria [2] based on the widely adopted Canadian Criteria [3].

The Canadian criteria included many of the cardiopulmonary and neurological abnormalities, which were not included in the CDC criteria. In addition, the Canadian criteria selected cases with less psychiatric co-morbidity, more physical functional impairment, more fatigue/weakness, plus neurological symptoms, which were significantly different from psychiatric controls with CFS/ME [4]. However, in the UK, The National Institute for Health and Care Excellence (NICE) recognises the heterogeneity of the condition advising that diagnosis of CFS/ME should be made after other possible diagnoses have been excluded and the symptoms have persisted for at least four months. The NICE guidance also states that diagnosis should be reassessed if none of the following key features of the disorder are present; post exertional fatigue, cognitive difficulties, sleep problems or chronic pain [5].

Due to the heterogeneity of the disorder, the aetiology of CFS/ME remains unknown with many theories surrounding the pathophysiology of the disorder [6]. The literature suggests that a range of possible causes including hormonal disturbances, immune system dysfunction, infectious and viral agents and nervous system abnormalities may all play a role in the pathophysiology of the disease [7]. Early research suggested that infectious agents such as the Epstein-Barr virus (EBV) are associated with CFS/ME, with a persisting EBV infection being seen in those with the condition [8]. A number of other infectious agents have been linked to the onset of CFS/ME including enteroviruses, which could explain the gastrointestinal symptoms often seen in patients [9], and also the acute B19 virus infection [10]. It has been suggested that viral infections can alter immune response which in turn, can chronically activate the immune system [11] and lead to many of the symptoms associated with CFS/ME. However, the research within this field is inconsistent with no evidence of a single infection causing CFS/ME, as well as many patients showing no sign of previous infection [12], suggesting infectious agents may only be relevant to a subset of the patient population.

Lymphatic system alterations are suggested to have involvement in CFS/ME with dysfunction within the immune system, causing toxic build up within the central nervous system (CNS) and leading to engorgement of varicose lymph vessels that can be felt on examination [13]. Tender lymph nodes are included in the ICC, Canadian, States Centres for Disease Control and Prevention (CDC) definitions [1] and the NICE guidance [5] confirming it is a common symptom of the disorder related to immune system abnormalities. Techniques

to target these engorgements and stimulate the drainage of toxins in the lymph nodes has been shown to lead to symptom improvement in patients with CFS/ME [13].

Currently there is no definitive way of diagnosing patients, although recent research has suggested that there is a link between CFS/ME, the lymphatic drainage system and the CNS and that, in fact, CFS/ME patients have certain physical signs present that may explain a number of the characteristics of the condition [13]. The Perrin technique is a system of manual diagnosis and treatment, that is based on the hypothesis that CFS/ME is a disorder of the lymphatic drainage of the central nervous system which leads to five physical signs [14].

The first aim of this study was to see whether the five physical signs of the Perrin Technique can assist in the screening of patients with CFS/ME, which could then subsequently lead to quicker treatment. Secondly, the study aimed to see whether the diagnostic accuracy was similar for a newly trained allied health professional with no prior experience of CFS/ME compared to an experienced allied health professional.

Methods

This was a diagnostic accuracy study with inter-rater agreement assessment. The study received ethical approval from NRES Committee North West - Lancaster, REC reference 12/NW/0877, R&D approval was also obtained from each participating NHS trust. The full study protocol has been made available at the same publisher.

Recruitment

CFS/ME participants were recruited from 2 hospital clinics and local support groups within the North West. Social media and posters displayed around the University of Central Lancashire were used to advertise the study. Healthy participants were recruited from non-blood relatives and friends of people with CFS/ME, staff and students from the university, and from those who had heard about the study over social media. All participants voluntarily contacted the research team, by email, telephone or post, to register their interest.

Those who contacted the researcher were sent a participant information sheet via email or post. Potential participants were given time to consider participation, during which they could contact the researcher to ask any questions about the research. The researcher then sent out a consent form to each person, which were then returned if they were happy to take part.

Participant eligibility

On receiving their completed consent form, the researcher contacted each person by telephone in order to assess their eligibility to take part in the study. Potential participants who were aged between 18 and 60, were assessed using two forms: a recruitment screening form based on the NICE guidance [5], and a form based on the International Consensus Criteria [2] (reference standard). These eligibility criteria were used to ensure that each CFS/ME patient had received a correct diagnosis of CFS/ME and to ensure that control participants did not have undiagnosed CFS/ME.

Inclusion Criteria for CFS/ME group

To be included CFS/ME patients needed to have a prior formal diagnosis of CFS/ME [5] at an NHS hospital specialised clinic, persistent or recurrent fatigue for at least the past 6 months, a clear starting point to the fatigue, the fatigue should be unexplained by any other

conditions, the fatigue should reduce the amount of physical activity each person could do, the fatigue should feel worse after physical activity.

Additionally CFS/ME patients needed to have at least one of the following symptoms: difficulty sleeping or insomnia, joint pain without swelling, headaches, painful lymph nodes that are not enlarged, recurrent sore throats, muscle pain without swelling, poor mental function, such as difficulty thinking, symptoms getting worse after physical or mental exertion, feeling unwell or having flu-like symptoms, dizziness or nausea or heart palpitations.

Exclusion Criteria for both CFS/ME and Control groups

The following were excluded from taking part: people needing to use a wheelchair and pregnant and lactating women. In addition: comorbidities including: Anaemias, autoimmune diseases, cardiac disease, endocrine disorders, infectious diseases, intestinal diseases, malignancies, neurological disorders, primary psychiatric disorders, significant pulmonary disease, primary sleep disorders. Additional exclusion criteria for control group included: a diagnosis of CFS/ME or a family history of CFS/ME.

Each participant was allocated a participant ID number with the clinical team being blinded to the groupings.

Assessment methods:

Perrin Technique:

The examination comprised the following four assessments:

A. Participant standing: observation and palpation of thoracic spine for any postural defects; regions of redness, temperature change or skin rashes or eruptions e.g. acne/boils.
B. Participant lying supine: observation and palpation of breast tissue for varicosities in the surface lymphatics and abnormal breast tenderness at 'Perrin's Point' which is a superficial tender area found at around 2-3 cm lateral and superior to the left nipple [15].
C. With the participant remaining supine, palpation of the region of the coeliac plexus just below the xiphoid in the upper central area of the abdomen for any abnormal tenderness with possible temperature change in the region.
D. With the participant remaining in a supine position, cradle the head and examine the quality of the cranial rhythmic impulse [16].

These assessments resulted in identification of the following signs as present or not present; 1) postural/mechanical disturbances of the thoracic spine (assessment A); 2) breast varicosities (assessment B); 3) tender Perrin's Point (assessment B); 4) tender coeliac plexus (assessment C); and 5) dampened cranial flow (assessment D).

If all five signs were present then the participant is classified as having CFS/ME. If one or more of the signs was absent, then the participant is classified as not having CFS/ME [16].

The examination was performed by two Allied Health Professionals. One had 10 years of experience of using the Perrin technique and working with CFS/ME patients (experienced AHP); the other was newly trained in the Perrin technique with no prior experience of CFS/ME (newly trained AHP). The newly trained AHP received training, especially for this study, which involved being taught how to examine patients for the 5 physical signs, and having hand-on experience of practicing the technique.

Rheumatological Assessment:

A standard clinical neurological and rheumatological assessment was performed by a physician whilst observing the participant for any signs of illness behaviour, but no clinical history was taken. The neurological examination included muscle strength testing, examination of muscle tone in arms and legs, coordination including the finger nose test, heel-shin test, heel-toe walking, reflexes and sensation with eyes closed. The rheumatological examination examined joint swelling, wasting of regional muscles, deformity of joint, redness in joints or tendons, and the palpation of the margin of joints in hands and feet.

If all the tests were normal with no observed illness behaviour, the patient was classified as not having CFS/ME, whereas if abnormal observations were made, the physician used their clinical experience to decide if the participant had CFS/ME. The physician performing these assessments had experience of working in NHS clinics for CFS/ME but had no experience of the Perrin technique (physician).

Data collection

For each participant, all data collection was performed on a single assessment day. They were allowed to bring a friend or relative along for support, or if requested, a chaperone was provided by the research team. Participants were re-briefed on what would be involved and consent was confirmed.

Each of the participants were separately examined by the three practitioners in different rooms; The order of examinations was randomised. No conversation took place between the participants and practitioners except to determine if there was any pain or tenderness in certain regions.

Data analysis

A priori, the Perrin technique required all five symptoms to be present for a patient to be diagnosed as CFS/ME. Using this criterion, the sensitivity and specificity of the diagnosis of CFS/ME relative to the reference standard were estimated as simple proportions, accompanied by exact (binomial) 95% confidence intervals. McNemar's test was used to investigate whether any of the practitioners systematically under- or over-diagnosed CFS/ME. Agreement in the diagnosis of CFS/ME between AHPs using the Perrin Technique was estimated using Cohen's Kappa (κ) coefficient; an approximate 95% confidence interval for κ was obtained using bias-corrected non-parametric bootstrapping. Agreement between their identification of the individual physical signs was also estimated using the same methods.

Exploratory analysis of the performance of using different numbers of symptoms was then performed. Symptoms were removed based on the observed agreement between AHPs; so the four-symptom test excluded the symptom with lowest agreement, the three-symptom test excluded the two symptoms with the two lowest agreements, and so on.

Sample size

The target sample size was 50 with CFS/ME and 50 controls to enable estimation of sensitivity and specificity of each Perrin technique AHP's diagnosis of CFS/ME (relative to the imperfect reference standard) to within ±9.9% with 95% confidence if the sensitivity and specificity were each at least 85%. It would also enable the estimation of κ (for inter-rater agreement between each pair), with 95% confidence, to within ±0.140 providing κ were at least 0.7.

Results

Ninety-four participants were recruited: 52 CFS/ME patients and 42 non-CFS/ME controls. The gender ratio in the CFS/ME group (shown in table 1) is in keeping with epidemiological studies which have shown a larger number of CFS/ME patients to be female with a ratio of 2:1 or more [17]. Results show that, on average, the experienced AHP was most accurate (86%) at correctly diagnosing participants. This was followed by the newly trained who correctly diagnosed 77%, and the physician who correctly diagnosed 69% of participants.

	CFS/ME	Control
Male	9 (17%)	25 (60%)
Female	43 (83%)	17 (40%)

Table 1: Gender balance of the two groups

Sensitivity and Specificity:

Table 2 shows the prevalence, sensitivity (the proportion of positive results in people with CFS/ME), specificity (the proportion of negative CFS/ME results in healthy controls), positive predictor value (the proportion with CFS/ME in participants with a positive result) and negative predictive value (the proportion of healthy controls with negative CFS/ME results).

Practitioner	True Positives	False Negatives	False Positives	True Negatives	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Accuracy (%)
Experienced AHP	46	6	7	35	0.88	0.83	0.87	0.85	86.2
Newly trained AHP	36	16	6	36	0.69	0.86	0.86	0.69	76.6
Physician	23	29	0	42	0.44	1.00	1.00	0.57	69.1

Table 2: Diagnostic performance of the three practitioners

When using the five physical signs of the Perrin Technique, the sensitivity for the experienced AHP was 0.88 (95% CI 0.77 to 0.96) and the specificity was 0.83 (95% CI 0.69 to 0.93). Similarly, when the newly trained AHP used the same technique, the specificity was 0.86 (95% CI 0.71 to 0.95), however sensitivity was lowered to 0.69 (95% CI 0.55 to 0.81). This shows that although using the same technique, the newly trained AHP struggled more frequently to identify correctly all five physical signs in people with a positive diagnosis of

CFS/ME. There was no statistically significant bias in the diagnosis by the experienced AHP relative to actual diagnosis ($p = 1.0$). There was also a marginally non-significant evidence of biased diagnosis by the newly trained AHP relative to actual diagnosis ($p = 0.052$) potentially favouring a non-CFS/ME diagnosis.

When using the standard clinical neurological and rheumatological examination, the sensitivity of the physician was 0.44 (95% CI 0.30 to 0.59) and the specificity was 1.0 (95% CI 0.92 to 1.0). These results show that whilst able to identify correctly all healthy controls, the physician struggled the most out of all three practitioners to identify correctly people with a positive diagnosis of CFS/ME. There was a significant bias in the diagnosis by the physician relative to actual diagnosis ($p < 0.001$), also favouring a non-CFS diagnosis.

Agreement between the experienced and newly trained AHPs

There was moderate agreement between the experienced and newly trained AHPs on overall diagnosis using the 5 physical signs of the Perrin Technique ($\kappa = 0.56$, 95% CI 0.40 to 0.72, $p < 0.001$). Regarding the identification of the individual physical signs, there was substantial agreement between the AHPs on the presence of the tender coeliac plexus ($\kappa = 0.65$; 95% CI 0.48 to 0.80, $p < 0.001$) and agreement was moderate both on the presence of postural/mechanical disturbance of the thoracic spine ($\kappa = 0.57$; 95% CI 0.39 to 0.73, $p < 0.001$) and on the presence of Perrin's point ($\kappa = 0.56$; 95% CI 0.37 to 0.73, $p < 0.001$). However, there was only fair agreement between the AHP's identification of the dampened cranial flow ($\kappa = 0.35$; 95% CI 0.15 to 0.54, $p = 0.001$) and there was non-significant 'slight' agreement on the presence of breast varicosities ($\kappa = 0.03$; 95% CI -0.12 to 0.22, $p = 0.75$).

Diagnostic properties of alternative criteria

Given the relatively low sensitivity for the newly trained AHP, alternative criteria were applied using lower number of physical signs. Using the Kappa statistics for agreement between the experienced and newly trained AHPs on the presence of the Perrin technique physical signs, each physical sign, in order starting with the sign with the least agreement, was excluded and sensitivity and specificity were recalculated. Overall, the AHPs were more accurate at diagnosing participants when using only 2 of the 5 signs (tender coeliac plexus and postural/ mechanical disturbance of the thoracic spine). Tables 3 and 4 show the accuracy of the experienced and newly trained AHPs, respectively, when using each number of physical signs. The accuracy of the experienced AHP is the same using 3-5 of the physical signs (86.2%) with the highest accuracy using only 2 of the physical signs (88.3%). The accuracy of the newly trained AHP is highest using only 1 or 2 of the physical signs (80.9%). Therefore, accuracy for both AHPs, overall, is highest when using only tests of tender coeliac plexus and postural/ mechanical disturbance of the thoracic spine.

Table 3: Experienced AHP

Number of Items	True Positives	False Positives	True Negatives	False Negatives	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Accuracy (%)
5*	46	7	35	6	0.88	0.83	0.87	0.85	86.2
4*	46	7	35	6	0.88	0.83	0.87	0.85	86.2
3*	46	7	35	6	0.88	0.83	0.87	0.85	86.2
2*	48	7	35	4	0.92	0.83	0.87	0.90	88.3
1*	49	12	30	3	0.94	0.71	0.80	0.91	84.0

Table 4: Newly trained AHP

Number of Items	True Positives	False Positives	True Negatives	False Negatives	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Accuracy (%)
5*	36	6	36	16	0.69	0.86	0.86	0.69	76.6
4*	36	7	35	16	0.69	0.83	0.84	0.69	75.5
3*	40	7	35	12	0.77	0.83	0.85	0.74	79.8
2*	42	8	34	10	0.81	0.81	0.84	0.77	80.9
1*	48	14	28	4	0.92	0.67	0.77	0.88	80.9

5* includes all 5 Perrin technique physical signs; 4* includes postural/mechanical disturbances of the thoracic spine, tender Perrin's Point, tender coeliac plexus, dampened cranial flow; 3* includes postural/mechanical disturbances of the thoracic spine, tender Perrin's Point, tender coeliac plexus; 2* includes tender coeliac plexus and postural/mechanical disturbance of the thoracic spine; 1* includes tender coeliac plexus.

Cohen's κ was re-computed for the level of agreement between the experienced and newly trained AHPs on whether they believed the 94 individuals had CFS/ME or were healthy controls, using the reduced 2 physical signs of the Perrin Technique. There was substantial agreement between the two AHPs on overall diagnosis using the 2 physical signs, $\kappa = 0.61$ (95% CI, 0.45 to 0.74), $p < 0.001$. There was no statistically significant bias in the diagnosis by the experienced AHP and actual diagnosis ($p = 0.55$), or by the newly trained AHP and actual diagnosis ($p = 0.63$) when using the reduced 2 item Perrin Technique, showing that the revised criteria no longer favoured a non-CFS/ME diagnosis.

Discussion

Between the AHPs, the AHP with prior experience of using the Perrin technique was the most accurate at correctly diagnosing individuals with CFS/ME, whereas the AHP with no prior experience of CFS/ME or the Perrin technique was better at correctly recognising healthy individuals.

The AHP experienced in the Perrin technique, was able to identify 88% of patients with CFS/ME using all five physical signs, and 83% of healthy controls who did not display all five signs. However, for the AHP with no prior Perrin technique experience, they were able to detect 86% of healthy controls, but identified only 69% of patients with CFS/ME using all five signs; there was borderline non-significant evidence of this AHP underdiagnosing rather than over-diagnosing CFS/ME. This highlights that the newly trained AHP was not able to identify all 5 signs in some people with a diagnosis of CFS/ME. However, this does not necessarily mean that the signs were not present; it could mean that the newly trained AHP found these signs more difficult to detect. Despite this, there was moderate agreement between both AHPs on overall diagnosis.

The physician was able to correctly identify 100% of the healthy controls using the standard clinical neurological and rheumatological examination. However, they were only able to correctly identify 44% of patients with CFS/ME and the tendency to under-diagnose CFS/ME was highly statistically significant ($p<0.001$). This affirms the current approach used in CFS diagnostics based on NICE guidance in that clinical examination is most useful in identifying alternative diagnoses and to exclude the diagnosis of CFS/ME, but that clinical examination is not a useful modality for confirming diagnosis of CFS/ME.

The agreement of the AHPs on the presence of each of the five physical signs varied from substantial agreement on the presence of the tender coeliac plexus to non-significant 'slight' agreement on the presence of breast varicosities. From the results presented, it would seem that the physical signs can improve the accuracy of diagnosing CFS/ME, although not all of the 5 physical signs may be necessary. Even with the experienced AHP, who, on the whole, was able to identify the 5 signs, breast varicosities and dampened cranial flow did not improve accuracy of diagnosis. Further exploration of the sensitivity, specificity and accuracy found that using only 2 of the 5 physical signs (tender coeliac plexus and postural/mechanical disturbance of the thoracic spine) was the most accurate and efficient method of correctly diagnosing the participants for both AHPs despite their differing levels of prior experience of CFS/ME and the Perrin technique.

Previous work [15] found Perrin's point to have a diagnostic accuracy of 80% in patients with CFS/ME. This was very similar to the accuracy when including Perrin's point in the current study (Accuracy; experienced AHP = 86.2% and newly trained AHP = 79.8%). However it was found that the omission of Perrin's point marginally increased the accuracy of the AHP with prior experience of the Perrin Technique by 2.1% and the AHP with no prior experience of the Perrin Technique by 1.1%.

Limitations and suggestions for future research

Although this study shows clearly that diagnostic accuracy for CFS/ME increases using the physical signs of the Perrin technique, there are some limitations, which should be highlighted. Firstly, this study recruited two groups of participants, people with a prior diagnosis of CFS/ME and healthy controls with no symptoms of the condition, as the purpose was to establish 'proof-of-concept' of the Perrin technique. Therefore, the study did not explore how accurate the Perrin Technique physical signs would be, when presented with patients with different conditions with similar presenting symptoms such as fibromyalgia. However, in a clinical setting, knowing the history and symptoms together with the physical signs would help to differentially diagnose CFS/ME from other possible illnesses. Now that we have established proof-of-concept, it will be important to identify whether the physical signs of the Perrin technique, combined with history,

The gender balance in the CFS/ME patients was similar to that of previously published work [17], although the healthy volunteers were recruited as a convenience sample which was not gender matched, therefore any prior knowledge of the expected male-female ratio amongst CFS/ME patients could have influenced the results. It should be noted, however, that none of the practitioners knew if the control group was matched or not for gender and that the AHP with no prior experience of CFS/ME was unaware of the gender balance. Although all participants were aged between 18 and 60, individual age data was not collected for each participant. However, it is reported that CFS/ME develops more commonly in those between their mid-twenties and mid-forties [18].

Future research should investigate whether the physical signs are more apparent in people with more severe CFS/ME. The present study did not collect data how long each participant had had CFS/ME for or on the severity of their symptoms, which could add further understanding.

Only 3 practitioners (1 in each category) were used. We therefore have very limited information on agreement between practitioners and whether diagnostic accuracy is

substantially affected by experience; we have no information on within-category variation. Future research should involve a study design whereby there are multiple experienced AHPs, newly trained AHPs and physicians.

A further limitation is that we selected specific symptoms for exclusion from the diagnostic criterion based on the agreement between practitioners. This was a pragmatic decision based on the estimated agreement between practitioners; it does not invalidate our findings, but there may be alternative criteria which have better performance. Again, optimisation of the set of symptoms for diagnosis merits further investigation in a larger study, in which additional information around acceptability and performance of individual physical sign assessments could be performed.

Conclusion

Current methods for diagnosing CFS/ME are challenging. The use of standard clinical neurological and rheumatological examination to examine illness behaviour is more likely to have a false negative result than a true positive one. Using certain physical signs appears to improve the accuracy of identifying people with CFS/ME and shows agreement with current diagnostic techniques, although not all of the physical signs were useful, and it is suggested that only 2 of these are needed. Examining for physical signs is both quick and simple for the AHP and may be used as an efficient screening tool for CFS/ME. This study did not include patient/family history or the patient talking about their symptoms, which should increase accuracy in clinical practice.

This work was supported by: The Fund for Osteopathic Research into Myalgic Encephalomyelitis (F.O.R. M.E.) Registered charity number: 1045005.

Contributorship statement: LH recruited participants, collected data and analysed the data. AB recruited participants and collected data. JR initiated the project, designed the research methods and analysed the data. CS wrote the statistical analysis plan and analysed the data. JS initiated the project and designed the research methods. BB, KM and GS collected data. TG and AM recruited participants. RP initiated the project and designed the research methods. All authors drafted and revised the paper.

Competing interests: This research explored the findings by one of the co-authors, Dr Perrin. To avoid any conflict of interest he was not involved in any of the recruitment of participants, clinical examinations, data collection or analysis. Dr Perrin's role in the study was to assist with developing the design, writing the protocol, setting up the project coordinating committee with the different clinical recruitment centres, applying for ethical approval, and assisting with the writing of the introduction and methods of the paper. There were no other conflicts of interest.

Data sharing statement: The relevant anonymised patient level data are available on request from the corresponding author.

References

[1] Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff, A. The chronic fatigue syndrome: a comprehensive approach to its definition and study. *Annals of internal medicine*. 1994;121(12): 953-959.

[2] Carruthers BM, van de Sande MI, De Meirleir KL, Klimas NG, Broderick G, Mitchell T, et al. Myalgic encephalomyelitis: International consensus criteria. *J Intern Med*. 2011;270: 327-338.

[3] Carruthers BM. Definitions and aetiology of myalgic encephalomyelitis: How the Canadian consensus clinical definition of myalgic encephalomyelitis works. *J Clin Pathol*. 2007;60(2): 117-9.

[4] Jason LA, Richman JA, Rademaker AW, Jordan KM, Plioplys AV, Taylor RR, McCready W, Huang C, Plioplys S. A Community-Based Study of Chronic Fatigue Syndrome. *Arch Intern Med*. 1999;159(18): 2129-2137.

[5] National Institute for Health and Care Excellence. *Chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy): diagnosis and management of chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy) in adults and children*. London: NICE; 2007.

[6] Bansal A, Bradley A, Bishop K, Kiani-Alikhan S, Ford B. Chronic Fatigue Syndrome, the immune system and viral infection. *Brain Behavior and Immunity*. 2012;26(1): 24-31.

[7] Shephard RJ. Chronic fatigue syndrome. *Sports Medicine*. 2001;31(3): 167-194.

[8] Straus SE, Tosato G, Armstrong G, Lawley T, Preble OT, Henle W, Davey R, Pearson G, Epstein J, Brus I. Persisting illness and fatigue in adults with evidence of Epstein-Barr virus infection. *Annals of internal medicine*. 1985;102(1): 7-16.

[9] Chia JK, Chia AY. Chronic fatigue syndrome is associated with chronic enterovirus infection of the stomach. *Journal of clinical pathology*. 2008;61(1): 43-48.

[10] Kerr JR, Bracewell J, Laing I, Mattey DL, Bernstein RM, Bruce IN, Tyrrell DA. Chronic fatigue syndrome and arthralgia following parvovirus B19 infection. *The Journal of rheumatology*. 2002;29(3): 595-602.

[11] Lorusso L, Mikhaylova SV, Capelli E, Ferrari D, Ngonga GK, Ricevuti G. Immunological aspects of chronic fatigue syndrome. *Autoimmunity Reviews*. 2009;8(4): 287-291.

[12] Afari N, Buchwald D. Chronic Fatigue Syndrome: A Review. *American Journal Of Psychiatry*. 2003;160(2):221-36.

[13] Perrin RN, Richards JD, Pentreath V, Percy DF. Muscle fatigue in chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) and its response to a manual

therapeutic approach: A pilot study. *International Journal of Osteopathic Medicine* 2011;14(3): 96-105.

[14] Perrin R. The involvement of cerebrospinal fluid and lymphatic drainage in chronic fatigue syndrome/ME. PhD [thesis]. Salford: University of Salford; 2005.

[15] Puri BK, Gunatilake KD, Fernando KA, Gurusinghe AI, Agour M, Treasaden IH.. Increased tenderness in the left third intercostal space in adult patients with myalgic encephalomyelitis: a controlled study. *J Int Med Res*. 2011;39(1): 212-4

[16] Perrin RN. Lymphatic drainage of the neuraxis in chronic fatigue syndrome: A hypothetical model for the cranial rhythmic impulse. *Journal of the American Osteopathic Association*. 2007;107(6): 218-224.

[17] Jason LA, Richman JA, Rademaker AW, Jordan KM, Plioplys AV, Taylor RR, McCready W, Huang C, Plioplys S. A Community-Based Study of Chronic Fatigue Syndrome. *Arch Intern Med.* 1999;159(18):2129-2137.

[18] NHS Choices. Chronic fatigue syndrome. Available from: http://www.nhs.uk/conditions/Chronic-fatigue-syndrome/Pages/Introduction.aspx. [Accessed 18 Aug 2017].

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	1
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3-4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	4-5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	4
	8	Where and when potentially eligible participants were identified (setting, location and dates)	4
	9	Whether participants formed a consecutive, random or convenience series	4
Test methods	**10a**	Index test, in sufficient detail to allow replication	5-6
	10b	Reference standard, in sufficient detail to allow replication	4
	11	Rationale for choosing the reference standard (if alternatives exist)	4
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	5
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	N/A
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	N/A
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	6
	15	How indeterminate index test or reference standard results were handled	6
	16	How missing data on the index test and reference standard were handled	N/A
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	6
	18	Intended sample size and how it was determined	7
RESULTS			
Participants	**19**	Flow of participants, using a diagram	N/A
	20	Baseline demographic and clinical characteristics of participants	
	21a	Distribution of severity of disease in those with the target condition	N/A
	21b	Distribution of alternative diagnoses in those without the target condition	N/A
	22	Time interval and any clinical interventions between index test and reference standard	N/A
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	7
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	7-9
	25	Any adverse events from performing the index test or the reference standard	NONE
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	10-11
	27	Implications for practice, including the intended use and clinical role of the index test	11
OTHER INFORMATION			
	28	Registration number and name of registry	N/A
	29	Where the full study protocol can be accessed	4
	30	Sources of funding and other support; role of funders	11

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.