Evidence Review D - Identifying & Diagnosing ME/CFS

Review Question 1: Diagnostic Criteria

1.1 Review question - *In people with suspected ME/CFS what are the criteria used to establish the diagnosis?* [page 8, lines 2-3]

> *"This review examines the criteria currently in use in clinical practice and research to assess which of those criteria are most appropriate for suspecting and then establishing an ME.CFS diagnosis for clinical practice." [lines 4-6]*

Nine sets of criteria are evaluation on six parameters and an 'overall rating' obtained.

The rating system used was the 'AGREE' instrument, *which "was developed to address the issue of variability in guideline quality".* The immediate problem is that drawing up clinical diagnostic criteria to delineate a disorder and providing guidance on its management and/or treatment are not the same thing. The task in hand here is very different from assessing the validity of a clinical guideline regarding a disorder where diagnostic criteria and processes are well established and the disorder is both well and widely understood. This particular task concerns appraising the value of diagnostic criteria specifically, and where there is a clear need for enhanced clarity.

We do not seek to question the value or otherwise of the AGREE instrument *per se*, but the relevance of using this rating scheme to evaluate clinical diagnostic criteria. Appraisal in 'AGREE' terms in the present circumstances is questionable at best, and has led to some very odd outcomes.

Every document is rated as having 'very serious limitations' with the exception of the IoM (2015) and NICE (2007), which both attract a 'serious limitations' rating.

In terms of the six components which feed into an overall rating on the 'AGREE' instrument[1], the Carruthers 2003 (Canadian Consensus) Criteria paper rates identically to Sharpe 1991 (Oxford) in all respects with the exception that the 'Oxford' paper is rated 'partial' in terms of clarity of presentation whereas 'Canadian' are considered to have met this criterion. Yet Carruthers *et al.* has impressive credentials: based on the consensus panel's collective extensive clinical experience diagnosing and/or treating more than twenty thousand (20,000) ME/CFS patients, the emerging clinical case definition was reviewed and refined with regard to analysis of symptoms in over 2,500 patients: De Becker P, McGregor N, De Meirleir K. *A definition-based analysis of symptoms in a large cohort of patients with chronic fatigue syndrome.* J Intern Med 2001; 250: 234-240; also the criteria do require post exertional deterioration.

Note too that this study be de Becker et al., reference 29, is excluded from consideration in terms of diagnostic criteria on the grounds that it is "not original publication". We have a copy of the paper and it does indeed appear to be an original publication.

Likewise this paper: Davenport TE, Stevens SR, Baroni K, Van Ness M, Snell CR. *Diagnostic accuracy of 35 symptoms characterising chronic fatigue syndrome. Disability and Rehabilitation.* 36 2011; 33(19-20): 1768-1775 (reference 28), is excluded from consideration on these same grounds *i.e.* that it is "not an original publication". This is very odd, this does indeed appear to be an original publication: https://www.tandfonline.com/doi/abs/10.3109/09638288.2010.546936

Note that the purpose of this paper was: *To determine the diagnostic accuracy for single symptoms and clusters of symptoms to distinguish between individuals with and without chronic*

[1] Scope & purpose; stakeholder involvement; rigour of development; clarity of presentation; applicability in practice and editorial independence.

fatigue syndrome (CFS). and that Davenport *et al.* conclude: *A cluster of associated symptoms distinguishes between individuals with and without CFS. Fewer associated symptoms may be necessary to establish a diagnosis of CFS than currently described.*

A further problem is the prioritising of sensitivity (*i.e.* making sure that all patients with the disorder in question are identified) over specificity (*i.e.* ensuring that people who do not have the disorder in question are excluded):

> *For this guideline, sensitivity was considered more important than specificity on the basis that at an early point in the diagnostic process, it is of greater importance to avoid false negative results and excluding people from a diagnosis. [Supporting Documentation 4: Methods - P17, lines 45-47]*

This may be all very well at an early point in the diagnostic process (though that is arguable) - but what are the implications further down the road? Is there any point where the guideline will provide guidance on firming up with a more specific protocol?

It takes only a small lack of specificity for a patient population that is relatively small (genuine ME/strictly defined CFS) to be swamped by the inclusion of even a small proportion of patients with another disorder (generalised 'fatigue') if the number of patients suffering the other disorder is much greater.

Review Question 2: Diagnostic Testing

2.1 Review question - *What is the diagnostic accuracy of specific tests to identify ME/CFS in people with suspected ME/CFS?* [page 60, lines 2-4]

This differs from the draft question as set out in the final guideline scope, which was:

What tests are clinically and cost effective in diagnosing of ME/CFS?

The alteration has facilitated an unfortunate end result: <u>all</u> of the studies identified as possibly relevant in this regard were filtered out, no published information on potential biomarkers whatsoever was appraised.

If permitted to stand, this peculiar approach to identification of relevant evidence eliminates any opportunity to make progress on the issue of diagnostic testing - or indeed the identification of a biomarker pertinent to any stage of the illness.

The change was not merely semantic *i.e.* 'suspected' does not simply imply that ME/CFS is 'suspected' in the absence of a definitive diagnostic test, identification of which then permits confirmation. No, the review population selected when seeking evidence was:

> *Adults, children and young people who are suspected of having ME/CFS by their primary clinician, <u>but who are yet to be formally diagnosed</u>*. [our emphasis] [Table 5: PICO characteristics of review question, page 60]

The thinking behind the change in the framing of the review question and the relevant study population is neither discussed nor explained.

However the impact is clear - and predictable: precisely no relevant studies were identified.

Why would an evidence review on diagnostic testing be conducted in respect of patients where diagnosis - presently made on clinical grounds alone - was merely 'suspected' and not in terms of studies involving patients who had had a sound and thorough appraisal of their clinical picture, on the basis of which they were found to fulfil clinical criteria for 'ME/CFS'? It is almost perverse to look at testing in respect of patients in whom the illness is merely suspected and decide on

that basis if the findings are fit to use to diagnose, rather than using the findings of biomarker studies on patients with confirmed diagnosis to be extrapolated for use at the diagnostic ('suspected') stage in future.

The exclusion of all studies identified in the evidence search on this most peculiar basis has as an end result that the evidence review has failed to meet its stated aim in this regard, *vis:*

> *"This review aims to identify up to date evidence in relation to tests which may help to identify ME/CFS, and to assess which of these may be useful to incorporate into clinical practice"* [P60, lines 6-7].

It follows that in *'The committee's discussion and interpretation of the evidence'* [Section 2.2, P64-65] the Guideline Committee (GC) are thrown back on their own knowledge as "no evidence was identified in the review" [P64, line 32].

It is notable that, when discussing diagnostic tests in the draft, the committee consider the relevant patient population to be those *who have been diagnosed*:

> *The committee identified 2-day cardiopulmonary exercise testing, grip strength, immunosignature, cytokine profile, erythrocyte sedimentation rate, mitochondrial function tests, postural hypotension test and C-reactive protein as potential diagnostic tests. These tests were considered to be emerging areas of research that have been identified as potentially showing differences in people diagnosed with ME/CFS compared to people without ME/CFS.* [P65, lines 5-14]

The approach to exclusion of studies discussed above has allowed at least a dozen to be disregarded, solely on the grounds that they looked at patients with confirmed CFS or ME when attempting to ascertain if a biomarker could be found.

There is a further issue - the list of tests to be considered was delineated in advance, rather than being allowed to emerge from the research evidence.

Had it not been that all the studies were filtered out in any case, this would constitute a considerable flaw.

Table 5: PICO characteristics of review question has a list labelled 'index tests' [P60 line 10]

It would appear that any evidence pertaining to a test not on this list was to be set aside. This is prejudging the outcome of consideration of the evidence.

What's more, the twin issues here come together in an exclusion *i.e.* Excluded: studies where the index test informs the eventual diagnosis (Appendix A, P85)

This exclusion is bizarre, given that study population is people where the diagnosis is merely suspected, and not people who have a confirmed clinical diagnosis.

Review Question 3: Clinical Signs and Symptoms

3.1 Review question - *What are the predictive accuracies of specific clinical symptoms/signs to identify people who will subsequently be given a clinical diagnosis of ME/CFS?* [page 66, lines 204)

> *"This review aims to identify up to date evidence in relation to symptoms and signs which may help to identify ME/CFS early, and to assess which of these may assist in making a clinical diagnosis."* [page 66, lines 6-8]

This is a new review question, not contained in the final scope. Like question 2, it has been framed in such a way as to facilitate filtering out many studies - one study and one study only made it through the sifting process: (82) Jason LA *et al. CFS prevalence and risk factors over time* Journal of Health Psychology 2011; 16 (3); 445-456. However it is considered of little value:

> *"Evidence for the accuracy of muscle weakness, insomnia, hypersomnia, irritable bowel syndrome, unrefreshing sleep, impairment of memory/concentration and post-exertional malaise for predicting later diagnosis of ME/CFS was based on a single study and was of very low quality. This was due to risk of bias, imprecision and methodological limitations."*[page 72, lines 9-12]

The paper De Becker P, McGregor N, De Meirleir K. *A definition-based analysis of symptoms in a large cohort of patients with chronic fatigue syndrome.* J Intern Med 2001; 250: 234-240 - ref 29, as discussed above - is not referenced at all in connection with this review question, not even to then be excluded. Likewise the paper by Davenport et al. (ref 28). This is very odd.

As with diagnosis, the evidence on clinical signs and symptoms in 'suspected' patients is viewed in terms of a pre-determined list of 'index tests' - in this case 6 pre-determined signs / symptoms [Table 6, page 66]

As with diagnostic testing, this is putting cart before horse.

Summary regarding procedure regarding identification of evidence regarding diagnosis

Competent results could be obtained by the simple process of:

(i) identifying patients by the best clinical criteria

(ii) ascertaining what biomarkers emerge; and

(iii) looking at clinical signs and symptoms in large cohort of competently diagnosed patients.

In contrast, the procedures followed have downgraded and filtered out evidence, resulting in very little evidence being put before the Guideline Committee and a key opportunity to make progress in terms of the review questions posed.

The 25% ME Group

26th February 2021