

APPENDIX 1: EVIDENCE AND GUIDELINE RECOMMENDATION GRADING SYSTEM USED FOR THIS GUIDELINE

QUALITY

The quality of each study was assessed using the GATE (Generic appraisal tool for epidemiology) appraisal tools (www.health.auckland.ac.nz/comhealth/epiq/epiq.htm).

Each study was graded according to the following criteria:

+	Good: Very low risk of bias; met all criteria
∅	Fair: Low risk of bias
-	Poor: Risk of bias; most criteria not met

LEVELS OF EVIDENCE FOR THERAPY

The guideline development team ranked the evidence according to the revised system of the Scottish Intercollegiate Guidelines Network (SIGN).²²⁹ Evidence statements relating to interventions have been assigned a grading according to the 'strength' of the supporting evidence where 1 is the best quality evidence and 4 is expert opinion.

Qualitative material was systematically appraised for quality, but was not ascribed a level of evidence.

ADAPTED SIGN GRADING SYSTEM USED TO IDENTIFY THE LEVEL OF EVIDENCE	
1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well conducted meta-analysis, systematic reviews of RCTs or RCTs with a low risk of bias
1-	Well conducted meta-analysis, systematic reviews of RCTs or RCTs with a high risk of bias
2++	High quality systematic review or case-control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal.
2+	Well conducted case-control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal.
2-	Case-control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg, case reports, case series
4	Expert Opinion eg, narrative reviews.

LEVELS OF EVIDENCE FOR DIAGNOSTIC TEST

The key criteria considered for inclusion were:

- number of participants >35
- blind independent assessment of the new test and the reference standard
- comparison of a reference test with the new test in > 90% of people
- an appropriate spectrum of people or a defined clinical group.

LEVELS OF EVIDENCE

SINGLE DIAGNOSTIC STUDIES	
D++	Good: All diagnostic tests criteria met
D+	Fair: One or two of the criteria not met
D-	Poor: None of the criteria met.

DIAGNOSTIC SYSTEMATIC REVIEWS	
DSR++	High quality meta-analysis or systematic review of diagnostic studies
DSR+	Fair quality meta-analysis or systematic review of diagnostic studies
DSR-	Poor quality meta-analysis or systematic review of diagnostic studies.

GRADING

The final step in grading is to consider the **WHOLE BODY OF EVIDENCE** ie, all the studies relevant to the issue and decide on a recommendation and grade based on all of the individual studies.

GRADES OF RECOMMENDATIONS	
A	The recommendation is supported by good evidence
B	The recommendation is supported by fair evidence
C	The recommendation is supported by expert opinion only, based on level 4 evidence in the text and the expertise within the multidisciplinary team.
I	No recommendation can be made because the evidence is insufficient (ie, evidence is lacking, of poor quality or conflicting and the balance of benefits and harms cannot be determined).

The grades A to C, I are a measure of the strength of evidence underlying the recommendations and should not be construed as an indication of the relative importance of the recommendations.

CRITERIA FOR LEVEL ONE EVIDENCE FOR CLINICAL DECISIONS RULES²⁴

Derivation of the rules by statistically evaluating the predictive value of variables from the history and physical examination

Validation of the rules in an appropriate setting, with blind assessment of the reference standard

Assessment of the impact of the rules on a person's management to ensure its acceptability to people and physicians, and to determine that it improves a person's outcomes and reduces costs

Validation of the rule in local and international settings to establish its applicability.