**What are the critical elements of side-line screening that can be used to establish the diagnosis of concussion? A systematic review.**

**Web Appendix**

**CONTENTS**

|  |  |
| --- | --- |
| **Section** | **Page** |
| **Further Methodological Details** |
| •Search strategy for identification of studies | **3** |
| •Development of search strategies | **5** |
| •Study identification and data extraction | **5** |
| •Summary of QUADAS-2 judgement criteria | **6** |
| •Assessment of overall quality of evidence | **7** |
| •Protocol changes | **7** |
| **Results** |
| •Near miss studies | **8** |
| •Diagnostic thresholds used in included sideline screening test studies  | **8** |
| •Detailed results for included sideline screening tests* *Symptoms*
* *Cognition*
* *Balance*
* *Oculomotor*
* *Multimodal*
 | **9****11****12****13****14** |
| •Video analysis and integrated head injury assessment protocols* *Characteristics of Fuller 2016*
 | **15** |
| •Detailed risk of bias assessments* *Symptoms*
* *Cognition*
* *Balance*
* *Oculomotor*
* *Multimodal*
* *Technology*
* *Video and integrated head injury assessment protocol*
 | **16****17****18****19****20****21****22** |
| •Detailed quality of evidence assessments* *Symptoms*
* *Cognition*
* *Balance*
* *Oculomotor*
* *Multimodal*
* *Technology*
* *Integrated head injury assessment protocol*
 | **23****23****24****24****25****26****26** |
| Summary of the sideline head injury assessment protocols used in professional contact and collision sports | **27** |
| Summary of criteria for immediate removal from play or for further assessment used in professional sport. | **28** |
| **Glossary** |
| Glossary of methodical terms | **29** |

**FURTHER METHODOLOGICAL DETAILS**

**Search strategy for identification of studies**

*Electronic information sources*

1. Cochrane Database of Systematic Reviews (via Cochrane library)

2. Cochrane Injuries Group Specialised Register (via Cochrane library)

3. Database of Abstracts of Reviews of Effectiveness (via Cochrane library)

4. Cochrane Central Register of Controlled Trials (via Cochrane library)

5. metaRegister of Controlled Trials (mRCT)

6. ClinicalTrials.gov

7. MEDLINE (via OVID and PubMed platforms)

8. EMBASE (via OVID platform)

9. CINAHL (via OVID platform)

10. SPORTSDiscus (via EBSCO)

11. Science Citation Index (SCI, via Web of Science)

12. SCOPUS

13. ZETOC

14. Conference Proceedings Citation Index – Science (via Web of Science)

15. OpenGrey

16. New York Academy of Medicine Grey Literature Report

17. EThOS: UK E-Theses Online Service

18. ProQuest Dissertation & Theses Database

19. National Clinical Guidelines Clearing House website

20. World wide web

*Non-electronic information sources*

1. Checking reference lists of retrieved articles

2. Checking reference lists of existing literature and systematic reviews

3. Correspondence with experts in the field, and relevant study authors

*Search terms*

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>

Search Strategy:

1 Athletic Injuries/

2 Sports Medicine/

3 exp Sports/

4 (athlete\* or athletic\* or sport\* or player\* or tennis or baseball or football\* or basketball or boxing or boxer or gymnast\* or hockey or soccer or volleyball or netball or wrestler or wrestling).mp.

5 1 or 2 or 3 or 4

6 Craniocerebral Trauma/

7 Brain Concussion/

8 Head Injuries, Closed/

9 Brain Injuries/

10 (blow adj3 head).mp.

11 ((head or brain) adj2 (trauma\* or impact or injur\*)).mp.

12 ((brain or cortical) adj2 contusion\*).mp.

13 ((nonpenetrating or non-penetrating or blunt) adj3 (brain or head)).mp

14 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13

15 Brain Concussion/

16 (commotio cerebri or concuss\*).mp.

17 Ataxia/ (6958)

18 (coordination adj3 (impair\* or lack\*)).mp.

19 (ataxia\* or confusion or confused or dizziness or dizzy).mp.

20 Unconsciousness/

21 (loss ajd2 consciousness or unconscious\*).mp.

22 headache.mp.

23 neurological dysfunction.mp.

24 (change\* adj3 (behav\* or attention or memory)).mp.

25 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24

26 (sideline\* or side-line or side line or touch line or touch-line or touchline or pitch or pitch side or pitchside or pitch-side or court or courtside or court-side or court side or dug out or dugout or dug-out or bench or track or technical area or technical-area or ring or ringside or ring-side or ring side).mp.

27 (field or onfield or on-field or on field or in game or ingame or in-game or in match or inmatch or in-match or in play or inplay or in-play).mp.

28 26 or 27

29 (screen or screening or diagnos\* or assess\* or test\*).mp.

30 Triage/

31 Early diagnosis/

32 Return to Sport/

33 Neuropsychological tests/

34 Vision tests/

35 Vestibular function tests/

36 ((return\* or resume\* or resumption) adj3 play).mp.

37 ((observable or visual) adj3 (sign or signs)).mp.

38 ((saccad\* or psychometric or king-devick or KD or K-D or sensory organi#ation or immediate post-concussion or cognitive) adj2 test\*).mp.

39 post-concussion symptom scale.mp.

40 (balance error scoring system or BESS).mp.

41 (standardi#ed assessment of concussion or SAC).mp.

42 (((sideline or side-line) adj2 concussion assessment tool) or SCAT2 or SCAT3 or SCAT-2 or SCAT-3).mp.

43 sport\* concussion assessment tool or SAC.mp.

44 maddocks.mp.

45 \*\*Add terms for any other sideline screening tests here\*\*

46 29-45/or

47 5 and 14 and 25 and 28 and 46

48 Accelerometry/

49 (accelerometer\* or video analysis or video-analysis or video review or video-review or impact sensor\* or eye-trac advance or mobile app\*).mp.

50 48 or 49

51 5 and 14 and 25 and 28 and 50

**Development of search strategies**

The search strategies were developed by the research team together with an information services expert from University College London based on expert subject knowledge and existing published search strategies. The search strategy was then further peer reviewed by librarians at the University of Sheffield. Searches were run research team members in conjunction with librarians from the University of Pretoria and University College London.

**Study identification and data extraction**

Although not eligible for inclusion, identified review articles were examined to provide a strategic overview and cross-check references. Where necessary study authors were contacted to provide additional information. Where appropriate, data were extracted to allow analysis consistent with the review questions and a standard diagnostic accuracy study design, rather than the investigators primary results. A single unblinded reviewer extracted information on study characteristics, methodology and results using a standardised data extraction form; and a second reviewer independently checked data for consistency and accuracy.

**Summary of QUADAS-2 Risk of Bias Judgement criteria**



**Assessment of overall quality of evidence**

The overall quality of evidence for each outcome was assessed using the consensus Grades of Recommendation, Assessment, Development and Evaluation Working Group (GRADE) approach. This specifies four outcome-specific levels of quality (high, moderate, low, and very low). For comparative effectiveness studies RCTs initially are initially rated as high quality, and observational studies as low quality evidence; for diagnostic accuracy studies cohort studies begin as high quality. The body of evidence is downgraded in the presence of within-study risk of bias, indirectness of evidence, heterogeneity, imprecision of effect/diagnostic accuracy estimates, and risk of publication bias; or up-graded due to large effect sizes, dose-response gradients, or plausible biases all working to undermine effect/accuracy estimates.

**Protocol changes**

There was a single protocol modification. The Newcastle-Ottawa risk of bias tool was used instead of a hierarchical level of evidence for non-diagnostic cohort studies in response to peer review.

**results**

**Near miss articles**

Seven potentially eligible sideline studies were identified which recorded data on sideline tests and concussion, but did not report useable data on diagnostic accuracy (McCrory 2000 – Digital Subtraction Test and symptoms; Daniel 2002 – SAC; Nassiri 2002 –SAC; McCrea 1997 – SAC; McCrea 1998 – SAC; McCrea 2010 – Concussion Severity Inventory, BESS; Barr 2012 – Concussion Severity Inventory, BESS; McCrea 2013 – GSC, SAC). Six potentially eligible technology studies were also identified, which recorded data on technology use in concussed and non-concussed athletes, but did not report useable data on diagnostic accuracy or effectiveness, including: iPad software applications for concussion screening (Alberts 2014, McKenzie 2014); Head Impact Telemetry Systems (Duma 2005, Brolinson 2006, Eckner 2011); and a portable computerised neuropsychological assessment tool (Espinoza 2014).

**Diagnostic thresholds used in included sideline screening test studies**

|  |  |  |
| --- | --- | --- |
| **Study** | **Index tests** | **Test Threshold** |
| Maddocks 1995 | •Symptoms•Orientation, recent memory  | •Present / not present•Correct / incorrect |
| McCrory 2000 | Symptoms | Present / not present |
| Barr 2001 | SAC | Any worsening from baseline |
| Erlanger 2003  | Symptoms | Present / not present |
| McCrea 2001 | SAC | Any worsening from baseline |
| McCrea 2002 | SAC | <10th percentile of normal performance |
| McCrea 2005 | GSC, SAC, BESS | Standardized regression based indices for detection of significant change in test scores |
| Echlin 2010 | SAC, BESS | Any worsening from baseline |
| Galetta K 2011 | KD | Any worsening from baseline |
| Galetta K 2011b | KD | Any worsening from baseline |
| Barr 2012 | CSI, SAC, BESS | Any worsening from baseline |
| King 2012 | KD | >3 seconds prolongation from baseline  |
| Galetta M 2013 | SCAT2, KD | Any worsening from baseline |
| Dhawan 2014 | KD | Any worsening from baseline |
| Fuller 2014 | •Symptom Checklist•Mental status evaluation •PSCA•Tandem Stance Test | •Any present•Any abnormality•Any abnormality•>4 errors in 20 seconds |
| Leong 2014 | KD | Any worsening from baseline |
| Galetta K 2015 | •SAC•Timed Tandem Gait, KD | •≥2 point drop in SAC compared to baseline •Any worsening from baseline |
| Leong 2015 | KD | Any worsening from baseline |
| Marinides 2015 | •SAC•KD•BESS | •≥2 point drop in SAC from baseline•Any worsening from baseline•≥3 point worsening form baseline |
| Putukian 2015 | SCAT2 symptom checklist, SAC, SCAT 2, Modified BESS | <5th centile of normative performative. |
| Seidman 2015 | KD | Any worsening from baseline |

**Detailed results for included sideline screening tests***Symptoms*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Index test** |  | **TP** | **FN** | **FP** | **TN** | **Sensitivity (%)** | **LCL** | **UCL** | **Specificity****(%)** | **LCL** | **UCL** |
| Maddocks 1995\*\* | Dizziness |  | 18 | 8 | 1 | 27 | **69.2** | 48.2 | 85.7 | **96.4** | 81.7 | 99.9 |
|  | Nausea |  | 17 | 9 | 2 | 26 | **65.4** | 44.3 | 82.8 | **92.9** | 76.5 | 99.1 |
|  | Headache |  | 26 | 2 | 5 | 23 | **92.9** | 76.5 | 99.1 | **82.1** | 63.1 | 93.9 |
| McCrory 2000 | Dizziness |  | 15 | 8 | NM | NM | **65.2** | 42.7 | 83.6 | **-** | - | - |
|  | Nausea |  | 5 | 18 | NM | NM | **21.7** | 7.5 | 43.7 | **-** | - | - |
|  | Headache |  | 23 | 0 | NM | NM | **100.0** | 85.2 | 100..0 | **-** | - | - |
| McCrea 2005\* | GSC |  | 84 | 10 | 0 | 56 | **89.4** | 81.3 | 94.8 | **100.0** | 93.6 | 100.0 |
| Erlanger 2003 | Dizziness |  | 40 | 7 | - | - | **85.1** | 71.7 | 93.8 | **-** | - | - |
|  | Nausea |  | 25 | 22 | - | - | **53.2** | 38.1 | 67.9 | **-** | - | - |
|  | Headache |  | 44 | 3 | - | - | **93.6** | 82.5 | 98.7 | **-** | - | - |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Fuller 2014 | Symptom Checklist |  | 50 | 15 | 23 | 77 | **76.9** | 64.8 | 86.5 | **77.0** | 67.5 | 84.8 |
|  | Mental status evaluation |  | 30 | 25 | 5 | 95 | **54.5** | 40.6 | 68.0 | **95.0** | 88.7 | 98.4 |
| Putukian 2015† | SCAT2 symptom checklist – number |  | 27 | 5 | 0 | 23 | **84.4** | 67.2 | 94.7 | **100.0** | 85.2 | 100.0 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | SCAT2 symptom checklist – severity |  | 24 | 8 | 0 | 23 | **80.0** | 61.4 | 92.3 | **100.0** | 85.2 | 100.0 |

\* McCrea 2005 (i)Numbers for TP, FN, FP, TN, and 95% confidence intervals calculated from reported sensitivity and specificity estimates derived from standardized regression based indices for detection of significant change in test scores. † Numbers for TP, FN, FP, TN, and 95% confidence intervals calculated from reported diagnostic accuracy data for impairment in symptom number and severity <5th centile of normative performative. \*\* A range of symptoms studied, representative results for common symptoms presented.

TP: True positives; FN: False negatives; FP: False positives; TN: True negatives; LCI: Lower confidence interval; UCI: Upper confidence interval.

The presence of individual symptoms in concussed and non-concussed athletes was investigated by Maddocks 1995, McCrory 2000 and Erlanger 2003. Headache was a sensitive indicator of concussion with point estimates reported between 92.9% and 100.0%. Nausea and dizziness were less sensitive, but more specific (92.9% to 96.4% respectively). Diagnostic accuracy results for symptoms checklists were imprecise and heterogeneous. McCrea 2005 (GCS) and Putukian 2015 (SCAT2 symptom checklist) reported moderate sensitivity of 89.4% and 84.4% respectively for the presence of any symptoms, with excellent specificities of 100%. However, these results were not replicated in Fuller 2014 (PSCA symptom checklist) where sensitivity and specificity of 76.9% and 77.0% were reported. Clinical signs of abnormal mentation were found to be specific (95.0%), but not sensitive (54.5%) for concussion.

*Cognition*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Index test** | **TP** | **FN** | **FP** | **TN** | **Sensitivity (%)** | **LCL** | **UCL** | **Specificity****(%)** | **LCL** | **UCL** |
| **Orientation** |
| Maddocks 1995\* | Orientation | 6 | 22 | 2 | 26 | **21.4** | 8.3 | 41.0 | **92.9** | 76.5 | 99.1 |
| **Maddock’s Questions** |
| Maddocks 1995\* | Recent memory  | 21 | 7 | 4 | 24 | **75.0** | 55.1 | 89.3 | **85.7** | 67.3 | 96.0 |
| Fuller 2014 | Maddock’s Questions | 22 | 43 | 7 | 93 | **33.8** | 22.6 | 46.6 | **93.0** | 86.1 | 97.1 |
| **Standardised Assessment of Concussion** |
| Barr 2001\*\* | SAC | 47 | 3 | 16 | 52  | **94.0** | 83.5 | 98.7 | **76.5** | 64.6 | 85.9 |
| McCrea 2001\*\* | SAC | 60 | 3 | 13 | 42 | **95.2** | 86.7 | 99.0 | **76.4** | 63.0 | 86.8 |
| McCrea 2002† | SAC | 68 | 23 | NM | NM | **79.1** | 69.3 | 86.9 | **-** | - | - |
| McCrea 2005ⱡ | SAC | 75 | 19 | 5 | 51  | **79.8** | 70.2 | 87.4 | **91.1** | 80.4 | 97.0 |
| Echlin 2010\*\* | SAC | 7 | 6 | NM | NM | **53.8** | 25.1 | 80.8 |  |  |  |
| Marindes 2015§ | SAC | 15 | 14 | NM | NM | **55.6** | 35.3 | 74.5 | **-** | - | - |
| Galetta K 2015§ | SAC | 2 | 8 | 3 | 14 | **20.0** | 2.5 | 55.6 | **82.4** | 56.6 | 96.2 |
| Putukian 2015\*\*\* | SAC | 13 | 19 | 2  | 20 | **40.6**  | 23.7 | 59.4 | **90.9** | 70.8 | 98.9 |
|  |  |  |  |  |  |  |  |  |  |  |  |

\* Diagnostic accuracy reported separately for a range of orientation and recent memory questions. Representative data for ‘What month is it?’ and ‘How far in the quarter?’ presented.

\*\* Sensitivity and specificity presented for ≥1 point drop in SAC compared to baseline

† Sensitivity calculated for SAC score below 10th percentile of normal performance

ⱡ Numbers for TP, FN, FP, TN, and 95% confidence intervals calculated from reported sensitivity and specificity estimates derived from standardized regression based indices for detection of significant change in test scores.

§ Sensitivity and specificity presented for ≥2 point drop in SAC compared to baseline

\*\*\* Numbers for TP, FN, FP, TN, and 95% confidence intervals calculated from reported diagnostic accuracy data for impairment in symptom number and severity <5th centile of normative performative.

TP: True positives; FN: False negatives; FP: False positives; TN: True negatives; LCI: Lower confidence interval; UCI: Upper confidence interval.

Diagnostic accuracy for orientation questions was available from Maddocks 1995, reporting a range of low and imprecise estimates for sensitivity between 3.6% and 57.1%, and 73.1% and 100% for specificity. Maddocks also provided estimates for individual sports-related recent memory questions (‘Maddock’s Questions) with sensitivity varying from 34.1% to 75.0%, and specificity of 85.7% to 100.0%. Fuller reported a contrasting sensitivity of 33.8% (95% CI 22.6 – 46.6) and specificity of 93.0% (95% CI 86.1 to 97.1) for all Maddock’s Questions taken together. Studies examining the SAC used a wide variety of cut-points for positivity including a ≥1 or ≥2 drop in baseline score, regression based indices for detection of significant change in test scores, or scores <5th or 10th percentile of normal performance. Unsurprisingly, accuracy results varied widely , with lowest estimates for sensitivity and specificity of 20.0% and 76.4%, and highest estimates of 95.1% and 91.1% respectively (I2 90.1%).

*Balance*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Index test** | **TP** | **FN** | **FP** | **TN** | **Sensitivity (%)** | **LCL** | **UCL** | **Specificity****(%)** | **LCL** | **UCL** |
| McCrea 2005 (i)\*McCrea 2005 (ii)\*\* | BESSBESS | 34 | 60 | 3 | 53 | **36.0****34.0** | 26.5NR | 46.7NR | **94.6****91.0** | 85.1NR | 98.9NR |
|  |  |  |  |  |  |  |  |  |  |  |  |
| Echlin 2010§ | BESS | 4 | 1 | - | - | **80.0** | 28.4 | 99.5 | **-** | - | - |
|  |  |  |  |  |  |  |  |  |  |  |  |
| Fuller 2014\*\*\* | Tandem Stance | 18 | 47 | 5 | 95 | **27.7** | 17.3 | 40.2 | **95.0** | 88.7 | 98.4 |
| Putukian 2015† | Modified BESS | 8 | 24 | 0 | 23 | **25.0** | 11.5 | 43.5 | **100** | 85.2 | 100.0 |
| Marindes 2015ⱡ | BESS | 16 | 4 | NM | NM | **80.0** | 56.3 | 94.3 | **-** | - | - |
| Galetta K 2015§ | Timed Tandem Gait | 10 | 2 | 5 | 9 | **83.3** | 51.6 | 97.9 | **64.3** | 35.1 | 87.2 |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

\* McCrea 2005 (i)Numbers for TP, FN, FP, TN, and 95% confidence intervals calculated from reported raw data for any impairment of BESS from baseline. \*\* McCrea 2005 (ii) Point estimates for sensitivity and specificity from standardized regression based indices for detection of significant change in test scores. \*\*\*>4 errors in 20 seconds. † Numbers for TP, FN, FP, TN, and 95% confidence intervals calculated from reported diagnostic accuracy data for impairment of modified BESS <5th centile of normative performative. ⱡ≥3 point worsening in BESS. §Any worsening from baseline.

TP: True positives; FN: False negatives; FP: False positives; TN: True negatives; LCI: Lower confidence interval; UCI: Upper confidence interval.

Individual sensitivity estimates for the BESS were heterogenous and imprecise, with point estimates ranging from 34.0 to 80.0%, I2 87.4%. BESS specificity, reported in a single study, was high 94.6% (95% CI 85.1 – 98.9). A range of accuracy results were calculated for the modified BESS by Putukian 2015 based on reliable change indices and comparison to normative performance. A representative sensitivity of 25.0% (95% CI 11.5 – 43.4) and specificity of 100.0% (95% CI 85.2 to 100.0) was reported for performance compared to normative values below the 5th percentile. The Tandem Stance Test demonstrated poor sensitivity (27.7%, 95% CI 17.3 – 40.2) and good specificity (95.0%, 95% CI 88.7 – 98.4) in the single study available. The Timed Tandem Gait demonstrated moderate sensitivity and specificity of 83.3% (95% CI 51.6 -97.9) and 64.3% (95% CI 35.1-87.2) respectively.

*Oculomotor*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **TP** | **FN** | **FP** | **TN** | **Sensitivity (%)** | **LCL** | **UCL** | **Specificity****(%)** | **LCL** | **UCL** |
| **Galetta K 2011** | 5 | 0 | 2 | 0 | **100.0** | 47.8 | 100.0 | **0.0** | 0.0 | 84.2 |
| **Galetta K 2011b** | 9 | 1 | - | - | **90.0** | 55.5 | 99.7 | **-** | - | - |
| **King 2012\*** | 3 | 0 | 0 | 0 | **100.0** | 29.2 | 100.0 | **-** | - | - |
| **Galetta M 2013** | 2 | 0 | - | - | **100.0** | 15.8 | 100.0 | **-** | - | - |
| **Dhawan 2014** | 20 | 0 | 11 | 110 | **100.0** | 83.2 | 100 | **90.9** | 84.3 | 95.4 |
| **Leong 2014†** | 1 | 0 | 0 | 5 | **100.0** | 2.5 | 100.0 | **100.0** | 47.8 | 100.0 |
| **Galetta K 2015** | 9 | 3 | 1 | 13 | **75.0** | 42.8 | 94.8 | **92.9** | 66.1 | 100.0 |
| **Leong 2015†** | 8 | 1 | 2 | 0 | **88.9** | 51.8 | 99.7 | **0.0** | 0.0 | 84.2 |
| **Marinides 2015** | 23 | 6 | NM | NM | **79.3** | 60.3 | 92.0 | **-** | - | - |
| **Seidman 2015** | 9 | 0 | 0 | 328 | **100.0** | 66.4 | 100.0 | **100.0** | 98.9 | 100.0 |

-: No data available to allow calculation

\* Data for witnessed head impact events undergoing side-line testing used only.

† Results reconstructed from side-line SCAT2 reference standard, not original case control study as per protocol.

TP: True positives; FN: False negatives; FP: False positives; TN: True negatives; LCI: Lower confidence interval; UCI: Upper confidence interval.

Data allowing calculation of sensitivity of the post-head impact event KD time for side-line identification of concussion was measured in all included studies and varied widely from 71.4% (Galetta K 2011) to 100.0% (King 2012, Galetta M 2013, Dhawan 2014, Leong 2014, King 2015, Seidman 2015). Individual estimates were very imprecise secondary to small sample sizes, with lower 95% confidence limits as low as 2.5% calculated (Leong 2014). This diversity was reflected in a high I2 statistic (52.1%). Data for specificity estimates was measured in six studies with similarly imprecise and heterogeneous results calculated, ranging from 0.0% (Leong 2015) to 100.0% (Leong 2014, Seidman 2015), I2 statistic 89.3%. KD test errors were reported in five studies (Galetta K 2011, Galetta K 2011b, Leong 2014, Leong 2015, Seidman 2015) and were found to be infrequent as shown in Table 5. Errors in isolation appeared to be specific, but non-sensitive, for the identification of concussion. However, results were very heterogeneous and imprecise with sensitivity point estimates ranging from 9.1 to 100.0%. 95% confidence limits for specificity varied from 47.8 to 100.0%. Insufficient data was reported to allow assessment of the diagnostic accuracy of both prolonged KD test times and errors in combination

*Multimodal*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **TP** | **FN** | **FP** | **TN** | **Sensitivity (%)** | **LCL** | **UCL** | **Specificity****(%)** | **LCL** | **UCL** |
| **Sports Concussion Assessment Tool 2** |
| Galetta M 2013\* | 2 | 0 | 0 | 0 | **100.0** | 15.8 | 100.0 | **-** | - | - |
| Putukian 2015† | 25 | 7 | 1 | 22 | **78.1** | 60.0 | 90.7 | **95.7** | 78.1 | 99.9 |
| **Pitchside Concussion Assessment Tool** |
| Fuller 2014 | 55 | 10 | 26 | 74 | **84.6%** | 73.5 | 92.4 | **74.0** | 64.3 | 82.3 |
| **Sports Concussion Assessment Tool 2, King-Devick Test\*** |
| Galetta M 2013 | 2 | 0 | 0 | 0 | **100.0** | 15.8 | 100.0 | **-** | - | - |
| **Timed Tandem Gait, Standardised Assessment of Concussion, King-Devick Test\*** |
| Galetta K 2015 | 24 | 0 | NR | NR | **100.0** | 85.8 | 100.0 | **-** | - | - |
| **Balance Error Scoring System, Standardised Assessment of Concussion, King-Devick Test\*\*** |
| Marinides 2015 | 20 | 0 | NM | NM | **100.0** | 83.2 | 100 | **-** | - | - |
| **Graded Symptom Checklist, Balance Error Scoring System, Standardised Assessment of Concussion** |
| McCrea 2005\*  | 89 | 5 | 6 | 49 | **94.7** | 88.0 | 98.3 | **89.1** | 77.8 | 95.9 |

\* Any worsening from baseline in any sub-test.

† Numbers for TP, FN, FP, TN, and 95% confidence intervals calculated from reported diagnostic accuracy data for impairment in symptom number and severity <15th centile of normative performative.

\*\* From baseline: any increase in KD test, ≥2 points worsening on SAC, ≥3 points worsening on BESS

TP: True positives; FN: False negatives; FP: False positives; TN: True negatives; LCI: Lower confidence interval; UCI: Upper confidence interval.

Point estimates for the sensitivity of combined use of individual sideline screening tools were high, but imprecise, reaching 100% for combinations of SCAT2/KD, TTG/SAC/KD, and BESS/SAC/KD; and 94.7% for joint use of GCS/BESS/SAC. The specificity of joint use of individual screening tests was available for a single study (McCrea 2005, GCS/BESS/SAC), at 89.1% (95% CI 77.8-95.9). The diagnostic accuracy of multifaceted sideline screening tests appeared lower, with sensitivity and specificity of 78.1% and 95.7%, and 84.6% and 74.0% reported for the SCAT2 and PSCA instruments respectively.

**Video analysis and integrated head injury assessment protocol**

**Characteristics of Fuller 2016**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Setting** | **Design** | **Sample Size****(n=)** | **Sport(s)** | **Level** | **Mean age (years±SE)** | **Technology** | **Risk of Bias / evidence level** | **Applicability concerns** | **Primary finding(s)** |
| Fuller 2016 | UK | PCS | 49 | Rugby Union | Professional | 26.5 (SD 3.5) | Sideline video review | Level 2b | Low | •Contributed to identification of 61.% of significant head impact events•21% of all diagnosed concussions presented post game |

**Detailed risk of bias assessments**

*Symptoms*

|  |  |  |
| --- | --- | --- |
| **Study** | **Risk of bias** | **Applicability concerns** |
|  | **Patient selection** | **Index test** | **Reference standard** | **Flow and timing** | **Overall** | **Patient selection** | **Index test** | **Reference standard** | **Overall** |
| Maddocks 1995 | HighCase-control design | Low | Low | Low | **High** | Low | Low | Low | **Low** |
| McCrory 2000 | HighCase-control design | UnclearTest review bias? | Low | Low | **High** | Low | Low | Low | **Low** |
| McCrea 2005  | HighCase-control design | HighTest review bias  | HighNon-physician assessment  | Low | **High** | Low | Low | Low | **Low** |
| Erlanger 2003 | HighCase-control design | UnclearTest review bias? | UnclearDiagnostic review bias? Non-physician assessment? | Low | **High** | Low | Low | Low | **Low** |
| Fuller 2014 | Low | Low | HighDiagnostic review bias | Low | **High** | Low | Low | Low | **Low** |
| Putukian 2015  | HighCase-control design | HighTest review bias | Low | HighDelayed index test | **High** | Low | Low | Low | **Low** |

*Cognition*

|  |  |  |
| --- | --- | --- |
| **Study** | **Risk of bias** | **Applicability concerns** |
|  | **Patient selection** | **Index test** | **Reference standard** | **Flow and timing** | **Overall** | **Patient selection** | **Index test** | **Reference standard** | **Overall** |
| Maddocks 1995 | HighCase-control design | Low | Low | Low | **High** | Low | Low | Low | **Low** |
| Barr 2001 | HighCase-control design | UnclearTest review bias? | HighNon-physician assessment | Low | **High** | Low | Low | Low | **Low** |
| McCrea 2001 | HighCase-control design | UnclearTest review bias? | HighNon-physician assessment  | Low | **High** | Low | Low | Low | **Low** |
| McCrea 2002 | HighCase-control design | UnclearTest review bias? | HighNon-physician assessment  | Low | **High** | Low | Low | Low | **Low** |
| McCrea 2005  | HighCase-control design | HighTest review bias  | HighNon-physician assessment  | Low | **High** | Low | Low | Low | **Low** |
| Echlin 2010 | HighCase-control design | HighTest review bias | HighIncorporation bias | HighVery high missing data levels | **High** | Low | Low | Low | **Low** |
| Galetta M 2013 | HighCase-control design | UnclearTest review bias? | UnclearDiagnostic review bias?Non-physician assessment? | Low | **High** | Low | Low | Low | **Low** |
|  |  |  |  |  |  |  |  |  |  |
| Fuller 2014 | Low | Low | HighDiagnostic review bias | Low | **High** | Low | Low | Low | **Low** |
| Marinides 2015 | HighCase-control design | UnclearTest review bias? | UnclearDiagnostic review bias?Non-physician assessment? | HighDelayed index test | **High** | Low | Low | Low | **Low** |
| Putukian 2015  | HighCase-control design | HighTest review bias | Low | HighDelayed index test | **High** | Low | Low | Low | **Low** |
| Galetta K 2015 | HighCase-control design | UnclearTest review bias? | UnclearTiming of reference standard? | Low | **High** | Low | Low | Low | **Low** |

*Balance*

|  |  |  |
| --- | --- | --- |
| **Study** | **Risk of bias** | **Applicability concerns** |
|  | **Patient selection** | **Index test** | **Reference standard** | **Flow and timing** | **Overall** | **Patient selection** | **Index test** | **Reference standard** | **Overall** |
| McCrea 2005  | HighCase-control design | HighTest review bias  | HighNon-physician assessment  | Low | **High** | Low | Low | Low | **Low** |
| Echlin 2010 | HighCase-control design | HighTest review bias | HighIncorporation bias | HighVery high missing data levels | **High** | Low | Low | Low | **Low** |
| Fuller 2014 | Low | Low | HighDiagnostic review bias | Low | **High** | Low | Low | Low | **Low** |
| Galetta K 2015 | HighCase-control design | Low | UnclearDiagnostic review bias?Timing of reference standard? | Low | **High** | Low | Low | Low | **Low** |
| Marinides 2015 | HighCase-control design | UnclearTest review bias? | UnclearDiagnostic review bias?Non-physician assessment? | HighDelayed index test | **High** | Low | Low | Low | **Low** |
| Putukian 2015  | HighCase-control design | HighTest review bias | Low | HighDelayed index test | **High** | Low | Low | Low | **Low** |

*Oculomotor*

|  |  |  |
| --- | --- | --- |
| **Study** | **Risk of bias** | **Applicability concerns** |
|  | **Patient selection** | **Index test** | **Reference standard** | **Flow and timing** | **Overall** | **Patient selection** | **Index test** | **Reference standard** | **Overall** |
| Galetta K 2011\* | Low | UnclearDiagnostic review bias? | UnclearTest review bias? | HighDelayed index test | **High** | Low | Low | Low | **Low** |
| Galetta K 2011b | HighCase-control design | Low | HighNon-physician assessmentTest review bias? | Low | **High** | Low | Low | Low | **Low** |
| King 2012 | Low | UnclearDiagnostic review bias? | UnclearTest review bias? | UnclearTiming of index test? | **Unclear** | Low | Low | Low | **Low** |
| Galetta M 2013 | HighCase-control design | UnclearDiagnostic review bias? | UnclearTest review bias?Non-physician assessment? | Low | **High** | Low | Low | Low | **Low** |
| Dhawan 2014 | HighCase-control design | UnclearDiagnostic review bias? | UnclearTest review bias?Non-physician assessment?Accurate reference standard? | Low | **High** | UnclearSample not described | Low | UnclearReference standard not described | **Unclear** |
| Leong 2014 | Low | Low | Low | HighDelayed index test | **High** | Low | Low | Low | **Low** |
| Galetta K 2015 | HighCase-control design | Low | UnclearTest review bias?Timing of reference standard? | Low | **High** | Low | Low | Low | **Low** |
| Leong 2015 | Low | Low | HighTest review biasNon-physician assessment? | Low | **High** | Low | Low | Low | **Low** |
| Marinides 2015 | HighCase-control design | UnclearDiagnostic review bias? | UnclearTest review bias?Non-physician assessment? | Low | **High** | Low | Low | Low | **Low** |
| Seidman 2015 | HighCase control design | Low | Low | HighDelayed index test | **High** | Low | Low | Low | **Low** |

*Multimodal*

|  |  |  |
| --- | --- | --- |
| **Study** | **Risk of bias** | **Applicability concerns** |
|  | **Patient selection** | **Index test** | **Reference standard** | **Flow and timing** | **Overall** | **Patient selection** | **Index test** | **Reference standard** | **Overall** |
| McCrea 2005  | HighCase-control design | HighTest review bias  | HighNon-physician assessment  | Low | **High** | Low | Low | Low | **Low** |
| Galetta M 2013 | HighCase-control design | UnclearTest review bias? | UnclearDiagnostic review bias?Non-physician assessment? | Low | **High** | Low | Low | Low | **Low** |
| Fuller 2014 | Low | Low | HighDiagnostic review bias | Low | **High** | Low | Low | Low | **Low** |
| Putukian 2015  | HighCase-control design | HighTest review bias | Low | HighDelayed index test | **High** | Low | Low | Low | **Low** |
| Marinides 2015 | HighCase-control design | UnclearTest review bias? | UnclearDiagnostic review bias?Non-physician assessment? | HighDelayed index test | **High** | Low | Low | Low | **Low** |
| Galetta K 2015 | HighCase-control design | UnclearTest review bias? | UnclearTiming of reference standard? | Low | **High** | Low | Low | Low | **Low** |

*Technology*

|  |  |  |
| --- | --- | --- |
| **Study** | **Risk of bias** | **Applicability concerns** |
|  | **Patient selection** | **Index test** | **Reference standard** | **Flow and timing** | **Overall** | **Patient selection** | **Index test** | **Reference standard** | **Overall** |
| Guskiewicz 2007 | Low | Low | Low  | Low | **Low** | Low | Low | Low | **Low** |
| Mihalak 2007 | Low | Low | UnclearDiagnostic review bias?Non-physician assessment? | Low | **Unclear** | Low | Low | Low | **Low** |
| Greenwald 2008 | Low | Low | Low  | Low | **Low** | Low | Low | Low | **Low** |
| Broglio 2010 | Low | Low | Low  | Low | **Low** | Low | Low | Low | **Low** |

*Video and integrated head injury assessment protocols*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Study** | **Patient selection** | **Comparability** | **Outcome** | **Overall** |
| Fuller 2016 | Low•Census sample•Comprehensive identification of head impact events •Healthy athletes at start of study•No attrition | Not applicable•Not comparative effectiveness/diagnostic accuracy/aetiological study | Low•Comprehensive outcome assessment•Follow up beyond acute period  | **Low** |

**Detailed quality of evidence assessments**

These table summarise the strength of evidence for sensitivity and specificity estimates in each sub-topic domain according to GRADE criteria.

*Symptoms*

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **Study designs** | **Factors decreasing quality of evidence** | **Overall GRADE rating** |
|  |  | **Risk of bias** | **Indirectness** | **Inconsistency** | **Imprecision** | **Publication bias** |  |
| **Graded Symptom Scale** |
| **Sensitivity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Very Low** |
| **Specificity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | No Concerns | Not detected | **Low** |
| **Individual Symptoms** |
| **Sensitivity** | 3 PCS | Serious concerns  | No concerns | Serious concerns | Serious concerns | Not detected | **Very Low** |
| **Mental Status Evaluation** |
| **Sensitivity** | 1 PCS | Some concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Low** |
| **Specificity** | 1 PCS | Some concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Low** |
| **PSCA symptom checklist** |
| **Sensitivity** | 1 PCS | Some concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Low** |
| **Specificity** | 1 PCS | Some concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Low** |
| **SCAT2 Symptom Checklist** |
| **Sensitivity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Very Low** |
| **Specificity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Low** |

*Cognition*

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **Study designs** | **Factors decreasing quality of evidence** | **Overall GRADE rating** |
|  |  | **Risk of bias** | **Indirectness** | **Inconsistency** | **Imprecision** | **Publication bias** |  |
| **Orientation Questions** |
| **Sensitivity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Very Low** |
| **Specificity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Very Low** |
| **Maddock’s Questions** |
| **Sensitivity** | 2 PCS | Serious concerns  | No concerns | Serious concerns | Serious concerns | Not detected | **Very Low** |
| **Specificity** | 2 PCS | Serious concerns | No concerns | Serious concerns | Serious concerns | Not detected | **Very Low** |
| **Standardised Assessment of Concussion** |
| **Sensitivity** | 6 PCS1 RCS | Serious concerns | No concerns | Serious concerns | Serious concerns | Not detected | **Very Low** |
| **Specificity** | 5 PCS | Serious concerns | No concerns | Serious concerns | Serious concerns | Not detected | **Very Low** |

*Oculomotor*

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **Study designs** | **Factors decreasing quality of evidence** | **Overall GRADE rating** |
|  |  | **Risk of bias** | **Indirectness** | **Inconsistency** | **Imprecision** | **Publication bias** |  |
| **King-Devick Test** |
| **Sensitivity** | 10 PCS1 RCS | Serious concerns | No concerns | Serious concerns | Serious concerns | Not detected | **Very Low** |
| **Specificity** | 6 PCS | Serious concerns | No concerns | Serious concerns | Serious concerns | Not detected | **Very Low** |

*Balance*

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **Study designs** | **Factors decreasing quality of evidence** | **Overall GRADE rating** |
|  |  | **Risk of bias** | **Indirectness** | **Inconsistency** | **Imprecision** | **Publication bias** |  |
| **Balance Error Scoring System** |
| **Sensitivity** | 2 PCS1RCS | Serious concerns | No concerns | Serious concerns | Serious concerns | Not detected | **Very Low** |
| **Specificity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Low** |
| **Tandem Stance Test** |
| **Sensitivity** | 1 PCS | Some concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Low** |
| **Specificity** | 1 PCS | Some concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Low** |
| **Modified BESS** |
| **Sensitivity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Very Low** |
| **Specificity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Low** |
| **Timed Tandem Gait** |
| **Sensitivity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Very Low** |
| **Specificity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Very Low** |

*Multimodal tests*

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **Study designs** | **Factors decreasing quality of evidence** | **Overall GRADE rating** |
|  |  | **Risk of bias** | **Indirectness** | **Inconsistency** | **Imprecision** | **Publication bias** |  |
| **Sports Concussion Assessment Tool 2** |
| **Sensitivity** | 2 PCS | Serious concerns | No concerns | Serious concerns | Serious concerns | Not detected | **Very Low** |
| **Specificity** | 2 PCS | Serious concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Very Low** |
| **Pitchside Concussion Assessment Tool** |
| **Sensitivity** | 1 PCS | Serious concerns  | No concerns | Unknown (single study) | Some concerns | Not detected | **Low** |
| **Specificity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Low** |
| **Sports Concussion Assessment Tool 2, King-Devick Test\*** |
| **Sensitivity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Very Low** |
| **Timed Tandem Gait, Standardised Assessment of Concussion, King-Devick Test\*** |
| **Sensitivity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Low** |
| **Balance Error Scoring System, Standardised Assessment of Concussion, King-Devick Test\*\*** |
| **Sensitivity** | 1 RCS | Serious concerns | No concerns | Unknown (single study) | Serious concerns | Not detected | **Very Low** |
| **Graded Symptom Checklist, Balance Error Scoring System, Standardised Assessment of Concussion** |
| **Sensitivity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | Some concerns | Not detected | **Very Low** |
| **Specificity** | 1 PCS | Serious concerns | No concerns | Unknown (single study) | No concerns | Not detected | **Low** |

*Technology*

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **Study designs** | **Factors decreasing quality of evidence** | **Overall GRADE rating** |
|  |  | **Risk of bias** | **Indirectness** | **Inconsistency** | **Imprecision** | **Publication bias** |  |
| **Head Impact Telemetry System** |
| **Positive predictive value** | 4 PCS | No concerns | No concerns | No concerns | Unknown(not reported) | Not detected | **Moderate** |
| **Side-line video review** |
| **Identification of significant head impact events** | 1 PCS | No concerns | No concerns | Unknown (single study) | Some concerns(small sample size) | Not detected | **Low** |

*Integrated head injury assessment protocol*

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **Study designs** | **Factors decreasing quality of evidence** | **Overall GRADE rating** |
|  |  | **Risk of bias** | **Indirectness** | **Inconsistency** | **Imprecision** | **Publication bias** |  |
| **Identification of significant head impact events and concussion** | 1 PCS | No concerns | No concerns | Unknown (single study) | Some concerns(small sample size) | Not detected | **Low** |

**Summary of the sideline head injury assessment protocols used in professional contact and collision sports**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sporting body** | **Tool /****protocol** | **Person/s who can request sideline screening** | **Person/s conducting the assessment** | **Use of video review** | **Location /duration of testing** | **Other key components** |
| AFL/NRL | Sport-specific HIA Form | Team doctor | Team doctor | Mandatory | Off-fieldMinimum of 15 mins | Other club support staff must report observations to the team doctor.SCAT3 used for further assessment.HIA forms are collected for audit and injury surveillance purposes. |
| FIFA | Immediate removal criteria |  |  |  | On-field/pitchside | 3-minute injury time following head impact.Pitch-Side assessment performed (based on a number of immediate removal criteria) |
| IIHF | Concussion protocol |  | Team doctor and/or AT(Team doctor solely responsible for determining concussion diagnosis) |  | Off-pitch | Observations made by team medical staff (or by any other team personnel and passed on to team medical staff). |
| NFL | Side-line concussion assessment tool | Coach, player, teammate, official, team doctor, AT, AT in the media booth or UNC | Team doctor, ATC or UNC | Mandatory | Off-pitch | Booth ATC, UNC, officials and the team doctor are connected by radio communication.The team doctor will review the video of the incident and (at a minimum) assess the player with a focussed neurological assessment (asking what happened, reviewing the “Go/No Go” signs and symptoms; and asking the Maddock’s questions.If the diagnosis is unclear, the player will undergo a full NFL sideline Concussion Assessment in the team locker room. |
| World Rugby | HIA process | Match official, team doctor or independent match day doctor  | Certified medical professional | Mandatory | Off-pitch10 minutes | Mandatory online education program for relevant personnel.Where the diagnosis is not immediately apparent, players removed & assessed. HIA forms are collected for audit & research |

AFL = Australian Football League; FIFA = Federation Internationale de Football Association; HIA + Head Injury Assessment; IIHF = International Ice Hockey Federation; NFL = National Football League; NRL = National Rugby League. AT=Athletic trainer. UNC= unaffiliated neurotrauma consultant. HIA= Head Injury Assessment

**Summary of criteria for immediate removal from play or for further assessment used in professional sport.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Clinical criteria** | **AFL/****NRL** | **FIFA** | **IIHF** | **NFL** | **World****Rugby** |
| Confirmed loss of consciousness |  |  |  |  |  |
| Definite confusion/disorientation |  |  |  |  |  |
| Any balance disturbance (e.g. ataxia) or motor incoordination |  |  |  |  |  |
| Impact seizure/convulsions or tonic posturing |  |  |  |  |  |
| Player reports significant, new or progressive/persistent concussion symptoms |  |  |  |  |  |
| Clearly dazed, “dinged”, blank or vacant stare |  |  |  |  |  |
| Behavioural change atypical of the player |  |  |  |  |  |
| Any clinical impression that the player is not quite right following trauma (i.e. “physician’s decision”) |  |  |  |  |  |
| Loss of responsiveness/suspected loss of consciousness |  |  |  |  |  |
| Memory impairment/amnesia |  |  |  |  |  |
| No protective action when falling to the ground (can be either tonic or hypotonic) – observed on video |  |  |  |  |  |
| Dangerous mechanism of trauma |  |  |  |  |  |
| Cross eyes (strabismus) or spontaneous nystagmus |  |  |  |  |  |
| Possible impact seizure or tonic posturing on video review |  |  |  |  |  |
| Possible balance disturbance |  |  |  |  |  |
| Slow to get up following a hit to the head |  |  |  |  |  |
| Possible behavioural changes |  |  |  |  |  |
| Possible confusion |  |  |  |  |  |
| Head impact event with the potential to result in concussion |  |  |  |  |  |
| Diagnosis not apparent |  |  |  |  |  |

AFL = Australian Football League; IIHF = International Ice Hockey Federation; NRL = National Rugby League; NFL = National Football League, FIFA = Federation Internationale de Football Association

 = Criteria for immediate removal and no return (i.e. diagnosis of concussion)

 = Criteria for further assessment

 = Criteria not specified

**GLOSSARY OF METHODOLOGICAL TERMS**

|  |  |  |
| --- | --- | --- |
| **Term** | **Definition** | **Ref** |
| **Grey Literature** | Grey literature (or gray literature) are materials and research produced by organizations outside of the traditional commercial or academic publishing and distribution channels e.g. websites, conference proceedings, PhD theses, etc. |  |
| **Current awareness search** | Literature searches conducted after the initial manuscript draft and just prior to submission to keep up-to-date with the most recently published information and developments. |  |
| **Forest plots** | A graphical representation of the individual results of each study included in systematic review, presenting point estimates of effect estimates/diagnostic accuracy metrics (represented as squares) together with their precision (95% confidence intervals, represented as lines). The forest plot provides a quick visual representation of overall effect estimates, how certain these results are, and heterogeneity in results across studies. |  |
| **Imprecision** | Imprecision is a measure of statistical variability. It is typically quantified by a confidence interval providing an estimated range of values which is likely to include the unknown population parameter in question, estimated from a given set of sample data. The width of the confidence interval indicates how uncertain we are about the unknown parameter. A very wide interval may indicate that more data should be collected before anything very definite can be said about the parameter. |  |
| **Heterogeneity** | Statistical variability of results among studies included in a systematic review is termed heterogeneity. This may occur due to :* Variability in the participants, interventions and outcomes studied, described as clinical diversity or clinical heterogeneity.
* Variability in study design and risk of bias, described as methodological diversity or methodological heterogeneity.

Statistical heterogeneity manifests itself as the observed intervention results being more different from each other than one would expect due to random error (chance) alone.  |  |
| **I2 statistic** | A useful statistic for quantifying inconsistency across studies included in a systematic review.The importance of the observed value of I2 depends on (i) magnitude and direction of effects and (ii) strength of evidence for heterogeneity e.g. a confidence interval for I2.A rough guide to interpretation is as follows:I2 0% to 40%: might not be important;I2 30% to 60%: may represent moderate heterogeneityI2 50% to 90%: may represent substantial heterogeneityI2 75% to 100%: considerable heterogeneity |  |
| **Meta-analysis** | A statistical analysis that combines the results of multiple scientific studies into a single weighted average. |  |
| **Narrative synthesis** | The results of studies included in a systematic review are summarised, described, explained and interpreted qualitatively using words and text. |  |
| **Test review bias** | Test review bias may be present when the results of the reference standard are known to those interpreting the index test. Results in overestimation of sensitivity.. |  |
| **Diagnostic review bias** | Diagnostic review bias may be present when the results of the index test are known to those interpreting the reference standard. Results in overestimation of sensitivity and specificity. |  |
| **Incorporation bias** | Systematic error in calculated diagnostic accuracy metrics occurring when the result of the index test is used in establishing the final diagnosis (i.e. it forms part of the reference standard). Results in overestimation of sensitivity and specificity. |  |
| **Attrition bias** | Non-random loss to follow up or withdrawal from the study can result in a non-representative sample and biased results if the withdrawal rate depends on the results of the index test or reference standard. |  |
| **Delayed index testing bias** | A systematic error in diagnostic accuracy results arising from conducting the index test later than would be expected in practice (e.g. performing ‘sideline’ screening rests for concussion after completion of sporting participation). Could result in different estimates of diagnostic performance due to disease progression (e.g. transient concussions could have resolved). |  |
| **Inaccurate reference standard assessment** | The error in diagnoses derived from an imperfect reference standard can result in underestimation of the performance of the index test. |  |