BMJ Open Sport & Exercise Medicine

Performance of current tools used for on-the-day assessment and diagnosis of mild traumatic brain injury in sport: a systematic review

Phoebe Haste ,[1] Leonardo de Almeida e Bueno ,[2] Antoine Jérusalem ,[1,2] Jeroen Bergmann [2,3]

To cite: Haste P, de Almeida e Bueno L, Jérusalem A, *et al*. Performance of current tools used for on-the-day assessment and diagnosis of mild traumatic brain injury in sport: a systematic review. *BMJ Open Sport & Exercise Medicine* 2025;**11**:e001904. doi:10.1136/bmjsem-2024-001904

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjsem-2024-001904).

Accepted 8 January 2025

[1]The Podium Institute for Sports Medicine and Technology, University of Oxford, Oxford, UK
[2]Department of Engineering Science, University of Oxford, Oxford, UK
[3]Department of Technology and Innovation, University of Southern Denmark, Odense, Syddanmark, Denmark

Correspondence to
Professor Antoine Jérusalem; antoine.jerusalem@eng.ox.ac.uk and

Professor Jeroen Bergmann; jberg@iti.sdu.dk

ABSTRACT

Objective The monitoring and diagnosis of sports-related mild traumatic brain injury (SR-mTBI) remains a challenge. This systematic review summarises the current monitoring tools used for on-the-day assessment and diagnosis of SR-mTBI and their performance.

Design Systematic review, using Quality Assessment of Diagnostic Accuracy Studies assessment.

Data sources Embase via Ovid, IEEEXplore, Medline via Ovid, Scopus and Web of Science were searched up to June 2024.

Eligibility criteria Peer-reviewed English-language journal articles which measured athletes using the index test within a day of injury and provided a performance measure for the method used. Studies of all designs were accepted, and no reference methods were required.

Results 2534 unique records were retrieved, with 52 reports included in the review. Participants were 76% male, when reported, and the mean injury-to-measurement time was reported in 10% of reports. 46 different methods were investigated. 38 different reference methods were used, highlighting the lack of gold standard within the field. Area under the curve (AUC), sensitivity and specificity were the most frequent outcome metrics provided. The most frequent index test was the King-Devick (KD) test. However, there were large variations in accuracy metrics between reports for the KD test, for instance, the range of AUC: 0.51–0.92.

Conclusion Combinations of existing methods and the KD test were most accurate in assessing SR-mTBI, despite the inconsistent accuracy values related to the KD test. The absence of a gold-standard measurement hampers our ability to diagnose or monitor SR-mTBI. Further exploration of the mechanisms and time-dependent pathophysiology of SR-mTBI could result in more targeted diagnostic and monitoring techniques. The Podium Institute for Sports Medicine and Technology funded this work.

PROSPERO registration number CRD42022376560.

WHAT IS ALREADY KNOWN

⇒ Sports-related mild traumatic brain injury (SR-mTBI) is a complex and multifaceted condition, which lacks a true gold-standard diagnostic method.

WHAT THIS STUDY ADDS

⇒ The lack of a gold-standard diagnostic method within the field is highlighted by the 38 different reference methods used.

⇒ The highest performing methods within this review are combinations of index tests.

⇒ The exact time between measurement and injury is often poorly reported; as SR-mTBI and its symptoms evolve over time, this is seen here as a significant oversight which should be considered and documented in future studies.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The time between measurement and injury should be recorded and reported more precisely in future studies.

INTRODUCTION

Traumatic brain injury (TBI) occurs in a variety of circumstances and has a myriad of possible presentations. TBI severity is classified using the Glasgow Coma Scale, with a mild TBI (mTBI) scoring of 13–15.[1] The potential for structural and functional deficits to present either simultaneously or separately has led to a heated debate surrounding the best way to diagnose mTBI, or whether it is appropriate to diagnose it at all.

Sports-related mTBI (SR-mTBI) has attracted particular interest as athletes have begun to mount legal challenges against governing bodies in multiple contact sports for misrepresenting the long-term consequences of repeated head impacts or injuries.[2 3] These initiatives have caused governing bodies to act, for example, by implementing new head injury monitoring measures.[4]

An established class of head injury monitoring techniques is combined neuropsychological testing. However, some of the most prominent neuropsychological tests, such as the Sports Concussion Assessment Tool 5 (SCAT5) test, are not officially defined as diagnostic tests.[5] The de facto 'gold standard' in SR-mTBI is a clinical diagnosis[6]; yet this is rarely immediately available for amateur athletes. Another prominent class of monitoring tools is impact monitoring, which aims to detect all impacts that a player experiences over an extended period of time,[7] or classify injurious impacts based on thresholding.[8]

Many existing reviews focus on SR-mTBI assessments (eg, refs.[9–12]). However, no review has explored the performance of SR-mTBI assessment and monitoring tools within a day of injury, which is their intended use. Additionally, these reviews tend to focus on either monitoring or diagnostic measures and can be focused on an individual category of monitoring technique rather than the entire field.

This systematic review aims to identify the monitoring and diagnostic methods that exist for SR-mTBI within a day of injury, how they perform and how this performance is quantified. The primary aim of the review is to investigate the performance of sports monitoring tools that assess and diagnose SR-mTBI during and shortly after injury. The secondary aim is to determine the three most frequently reported performance metrics within the reports.

METHODOLOGY

Eligibility criteria

Participants in the study were required to have sustained the injury in a sporting scenario and to have been assessed using the index test within a day of the initial injury. The selection criteria are presented in table 1. Studies of all designs were accepted. Published journal articles that assess injured athletes within 24 hours of SR-mTBI were considered. Reference tests were not required to have been performed within any specific time frame.

The primary outcome of this study is to define the performance of SR-mTBI diagnostic and monitoring tools used for on-the-day assessment and diagnosis of injury. As the term 'concussive' is often poorly defined, reports discussing 'subconcussive' impacts were not discarded as long as they met the remainder of the inclusion criteria. The performance measures related to the primary outcome were defined by Šimundić[13]. The three most frequently reported performance metrics were identified. Permitted outcome metrics can be found in online supplemental material 1.

Information sources

Searches were conducted in five electronic databases: Embase via Ovid, IEEEXplore, Medline via Ovid, Scopus and Web of Science. The final search included literature up to 14 June 2024.

Search strategy

This review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[14] The completed PRISMA checklists (including those for abstracts and searching) can be found in online supplemental material 2. The review protocol is registered on PROSPERO (protocol ID: CRD42022376560[15]). Amendments to the protocol are provided in online supplemental material 3.

The search terms comprised the keywords related to the following concepts: (a) TBI, (b) sport, (c) monitor and (d) diagnose. If available, search terms were mapped to relevant MeSH terms or subject headings. The concepts were connected by AND terms, and within-concept keywords were connected by OR terms. Further information about the search strategies can be found in online supplemental material 1.

Study selection process

Both the title and abstract, and the full-text screening were carried out in a blinded, standardised manner by the first and second reviewers (PH and LdAeB) independently. The first and second reviewers reviewed all the records and reports against the inclusion and exclusion screening criteria. Any disagreements were resolved by consensus. Unresolvable disputes were resolved by the third and fourth reviewers (AJ and JB). No automated or semiautomated methods were used in the screening process. Records where the abstract lacked sufficient information to be assessed against the screening criteria

Table 1 Inclusion and exclusion criteria of systematic review

	Inclusion criteria	Exclusion criteria
Population	Athletes with sports-related mild traumatic brain injury or subconcussive injury	
Intervention	Measurement within 24 hours of injury	Does not measure within 24 hours of injury
Comparator	No requirement	
Outcome	Data used for diagnosis/prediction/assessment Quantitative outcome metric required	
Setting	Sport	
Document type	Peer-reviewed journal article	Review, conference papers, extended abstracts
Additional requirements	English language	Reports from the same study using duplicate data sets

passed to the full text stage. Authors of primary studies were contacted to clarify eligibility when necessary.

Reports from the same study were considered in the full-text screening stage. When reports were suspected to be from the same study, the reports were manually compared. If reports provided distinct information about the study, they were included. Otherwise, the most recent report from the study was included. Further information about the screening process can be found in online supplemental material 4.

Data extraction and synthesis

Data items collected from the reports were year of study, country where the study was conducted, study design, study population (including sex and age of participants and the size of case/control groups), methodology, reference/comparator measurements, outcomes, source of data, monitoring/diagnostic technique used, processing techniques used, intra-study risk-of-bias assessment, performance summary metrics and their values, study setting (including the sport), the presence of baseline measurements, the presence of recovery monitoring, the underlying injury of interest for the study, whether previous head injury was considered and the time between injury and measurement. The first reviewer collected the data, and the second reviewer cross-checked the results. Data were extracted manually from the text, tables, figures and supplementary materials of reports. Where the data were not available or unclear, the study authors were contacted for additional information.

Reports were grouped for synthesis according to populations, interventions, outcomes and study design for each respective subheading. The included reports were assessed using descriptive statistics. A meta-analysis was not performed as the included reports were not sufficiently heterogeneous in intervention or outcome, and of varying quality. Therefore, no effect measures were used. Forest plots were used to visually represent results, along with a table of evidence.

Study risk-of-bias assessment

To identify the risk of bias in the studies of interest, Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) was chosen for its relevance to diagnostic accuracy.[16] QUADAS-2 assesses the applicability of the study in addition to the risk of bias. It assesses qualitatively, with scoring of 'low', 'unclear' or 'high' in four domains of interest. The first reviewer completed the risk-of-bias assessment independently due to time constraints on the other reviewers.

Equity, diversity and inclusion statement

The author group consists of junior, mid-career and senior researchers. Two authors are members of marginalised communities. Most studies within this review took place in high-income countries, so results may not be generalisable to low-income settings. An analysis of sex of study participants has been undertaken, but none of the reports studied addressed the difference between sex and gender or mentioned non-cisgendered participants.

Patient and public involvement

No patients or members of the public were involved in this work.

RESULTS

Study selection

The PRISMA flow diagram for the study can be seen in figure 1. 5834 records were identified in the initial search, 2534 of which were unique. The full texts of 177 reports were screened. 52 reports, from 50 studies, were included in the final review. Online supplemental material 6 contains examples of reports rejected at full-text stage, and the entire list is available in the online repository provided in the data-sharing section of this review. Quantities which were not consistently reported in paper abstracts were the time between measurement and injury or performance metrics.

Risk of bias in studies

Online supplemental material 5 shows the results of the QUADAS-2 risk-of-bias assessment, along with an example of an ideal study for the question of this review. Both the highest risk of bias and the greatest concerns regarding applicability were in the reference standard domain, where 15% and 19% of studies were judged as high risk, respectively. The index test domain contains the signalling question ('Were the index test results interpreted without knowledge of the results of the reference standard?') which was negatively answered for 35% of reports.

Study characteristics

43/52 reports (83%) collected their own data. Nine reports (9/52, 17%) drew on existing databases. Three (3/52, 6%) reports used a data set produced by the Concussion Assessment, Research and Education (CARE) consortium study, with variations on the exact data items considered.[17–19] The year of publication of the reports in this review ranges from 2001 to 2024.

Most reports were from studies performed in the USA (33/52. 63%), with the next most frequent location being Australia (5/52, 10%). 46/52 (88%) of the reports were published in or after 2015. The majority of the study designs within this review were prospective cohort studies (37/50, 74%). Retrospective cohort studies were the next most common (6/50, 12%). The remaining study designs present were retrospective case-control (3/50, 6%), retrospective case series (2/50, 4%), prospective case report (1/50, 2%) and randomised controlled trial (1/50, 1%). Further details of study characteristics are present in online supplemental materials 7 and 8.

Setting

All the studies focused on were related to sports. 40 (42/50, 84%) studies considered a single sport. Ten (10/50, 20%) studies involved multiple sports. Overall,

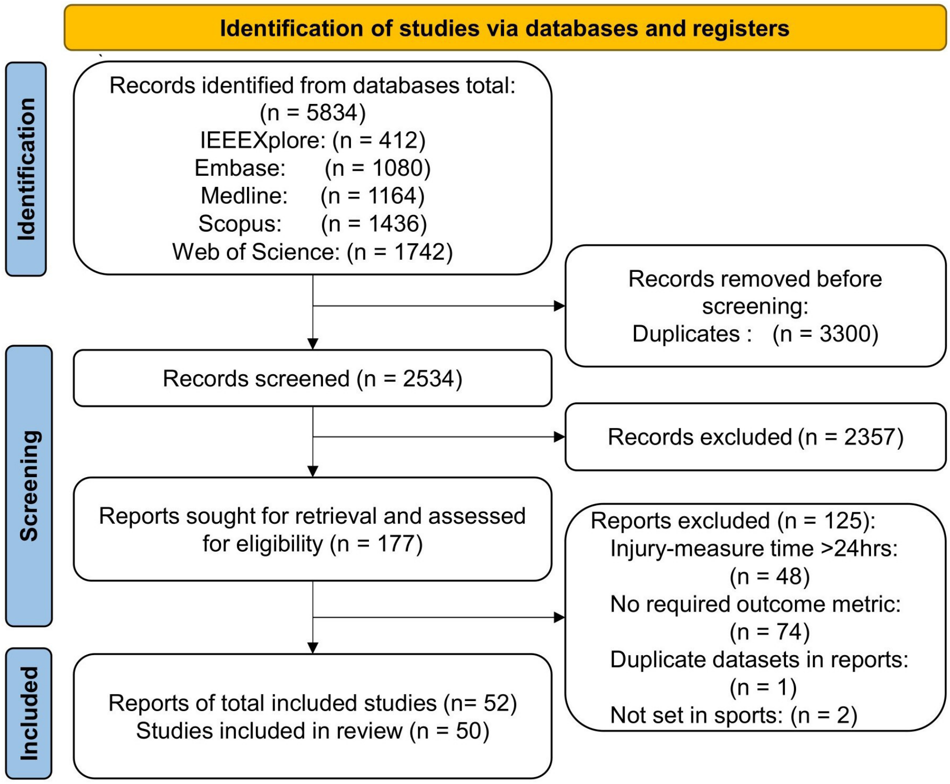

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram demonstrating the screening process.

American Football (17/50, 34%) and Rugby Union (11/50, 22%) were the two most frequent sports in the studies in this review.

47 (47/50, 94%) studies used data from a live sporting environment, with three exceptions (3/50, 6%). Kahouadji *et al*'s study[20] was carried out in a hospital's emergency department, and Wirsching *et al*'s study[21] was conducted in a controlled sporting environment. Rowson and Duma[22] used a Head Impact Telemetry System (HITS) data set from the National Football League (NFL) matches in addition to laboratory-constructed impacts.

The setting for three (3/50, 68%) studies was entirely amateur sports settings, and 14 (14/50, 28%) of the studies were professional sports settings. Reyes *et al*[23] incorporated both amateur and professional athletes (2%), while two studies (2/50, 4%) had only semiprofessional athletes as participants. The setting of studies with amateur athletes fell into two categories: community settings (12/34, 32%) or settings related to education (21/34, 62%). One study included both community settings and settings related to education. Online supplemental materials 7 and 8 contain further information about study settings.

Time of measurement

Though the selection criteria included the time of measurement, 44 (44/50, 90%) studies did not precisely measure the average time period between injury and assessment. Two studies (2/50, 4%) noted the mean and SD of the injury to assessment period. One study (1/50, 2%) provided the median time for a subset of participants. In the final three (3/50, 6%) studies, measurement and the injury were simultaneous. A further breakdown of time of measurement is available in online supplemental material 9.

Population

The total cohort size of this review was 117170 across 52 reports. However, the largest study[22] was an impact monitoring study which reported the total number of impacts within the study rather than the number of participants. Excluding this, 54101 subjects were included in this review. Population statistics are stated without this large study; however, a complete breakdown including it can be found in online supplemental material 7.

The sex of athletes within the reports was predominantly male. 12 (12/51, 24%) reports did not mention the sex of the participants; therefore, it was unspecified in 4614 subjects (9% of the total subjects). 16 (16/51, 31%) reports only had male participants, in contrast to the two (2/51, 4%) reports with only female participants. Overall, 13407 females (25% of subjects) participated in the included reports compared with 36 080 males (67% of subjects). A preliminary analysis of a subset of this review was presented at the International Research Council on Biomechanics of Injury Conference.[24] 6904 injuries were monitored or diagnosed, with a mean of 135 cases per experiment, and a median of 28 cases. 11 studies (10/51, 20%) did not have any control subjects. The total number of control participants was 7171, with a mean of 141 controls per paper, or a median of 31 controls.

Ten reports (10/52, 19%) did not provide information pertaining to the age of the participants, while 37 (37/52, 71%) reports provided the mean ages of participants. The combined mean age across all reports was 21.9 years old.

Summary tables for the condition investigated, characteristics of injury and demographics can be found in online supplemental materials 8 and 9, respectively.

Intervention

46 methods of monitoring and diagnosing TBI were implemented as the index test within this review. Over two thirds (32/46, 70%) of these methods only appeared once, while an eighth appeared twice (6/46, 13%). The Standardized Assessment of Concussion (SAC), Neurofilament Light biomarkers and S100b biomarkers were implemented four times (3/46, 7%). Versions of the head injury assessment (HIA) were used five times.

The King-Devick (KD) test was used 18 times within the reports. Versions of the SCAT test were used six times. Two accelerometry devices were used: the X-Patch once,[23] and HITS was used four times.

38 (38/52, 73%) reports used a single index test. Six (6/52, 12%) reports used two. Four (4/52, 8%) reports used three. Three reports (3/52, 6%) used four. One report (1/52, 2%) used ten methods.[17] Methods that occurred twice or less and additional information can be found in online supplemental material 7. Online supplemental material 10 contains a report-by-report breakdown of index and reference tests.

Comparators

36 different techniques were used as reference methods. 13/36 (36%) of the reference methods were also used as index tests.

32 (32/52, 62%) reports compared against assessments by team doctors or certified athletic trainers. Two of these 32 reports (2/32, 6%) used the Military Acute Concussion Evaluation (MACE) test. Two reports compared against diagnoses by the subject's own physicians (2/52, 4%). One study[20] used CT scans (1/52, 4%), and 21 (21/52, 40%) reports used versions of the SCAT test. The Rivermead Post Concussion Symptom Questionnaire (RPQ) was used twice (2/32, 6%) in assessments by team doctors or certified athletic trainers, with impacts also being counted in one of the uses.[25]

21/52 (38%) reports compared against event monitoring methods. Five (5/21, 24%) were from the HIA test, where the event/player is visually assessed, and a video review takes place. Seven (7/21, 33%) used visual event monitoring, one study as the sole comparator,[26] another along with a self-report logbook,[27] two using the RPQ,[28] two more studies used observation of the injury along with physicians assessments and the final study used MACE and the assessment of a physician.[29] Five (5/21, 24%) studies used video review; three leveraging video review to confirm impacts recorded by accelerometers[23 30 31] and two studies used video to confirm loss of consciousness or fight verdict.[32]

Eight (8/21, 41%) studies used accelerometry methods as comparators. Notably, Rowson *et al*[22] recategorised 207 subconcussive impacts as concussive impacts based on prevalence statistics to combat under-reporting.[22]

Three studies used the same method as both index and comparator; two using the KD test[33 34] and the third using clinical judgement within the HIA01 test.[35] Wirsching *et al*[21] conducted a study in a controlled sporting environment, and no comparator was used (1/49, 2%). CARE Consortium Investigators *et al* noted that there was 'no gold standard to compare measurements to'.[17] Additional comparator information can be found in online supplemental materials 7 and 10.

Results of individual studies

Outcome

All reports considered head injury, 16 (16/52, 31%) reports monitored injury, 32 (32/52, 61%) reports diagnosed injury and four (4/52, 8%) reports predicted head injury. Additional information about the outcome metrics can be found in online supplemental material 5. The three most frequently provided metrics, which will be discussed here, are the area under the receiver operating curve (AUC), sensitivity and specificity. The AUC, sensitivity and specificity values provided in this report are shown in table 2.

Area under the curve

19 reports within this sample provided the AUC statistics for 23 methods. Additionally, Reyes *et al*[36] provided sufficient information to infer the AUC figure (Mann-Whitney U and population sizes). Figure 2a shows the forest plot for the KD tests values provided. In Hecimovich *et al*,[37] five statistics were provided for the subtests of KD/ET, but no overall summary statistics were given. The highest AUC statistic provided was 0.982, produced using HITS.

Sensitivity

Sensitivity was provided in 26 reports for 25 distinct methods. These can be found in table 2 and figure 2b. Hubbard *et al*[26] only provided the range for the KD test sensitivity (0.63–0.87) and was not included in the forest plot. The highest sensitivity statistic provided was 1, related to the KD test.[38]

Specificity

Specificity was provided in 24 reports for 29 separate methods. Figure 2c contains the KD test statistics report for specificity. The highest specificity statistic reported was 0.972.[17]

DISCUSSION

Summary of evidence

This systematic review explores the current state of the field of sport monitoring tools used for on-the-day assessment and diagnosis of SR-mTBI. The novel aspect of this

Table 2 AUC, sensitivity and specificity statistics provided within this review

Reference	Index test	Reference test	AUC	Sensitivity	Specificity	Study size; case; control
22	Accelerometry (HITS)	207 subconcussive impacts were transformed to concussive data points to account for underreporting. 37 diagnoses made by certified athletic trainers/doctors.	**0.982 (95% CI: 0.818 to 0.947**)			63 069; 244; 62 800
	Accelerometry (NFL)	Diagnosis made by team doctors/athletic trainers, with accelerations obtained by impact reconstruction.	0.892 (95% CI: 0.801 to 0.983)			58; 25; 23
47	Accelerometry (HITS) (PLA)	Diagnosis made by clinician 'in addition to objective assessment measures'.	0.904 (95% CI: 0.842 to 0.951)			128; 15; 113
	Accelerometry (HITS) (PRA)		0.824 (95% CI: 0.662 to 0.918)			
	Accelerometry (HITS) (GAM-CY)		0.894 (95% CI: 0.818 to 0.947)			
8	Accelerometry (HITS)	Certified athletic trainer/physician assessment against concussion definition.		0.213 (range: 0.004–0.556)		78; 13; 65
48	DETECT	Accelerometry (HITS) and athletic trainer/team doctors' assessment.	0.778 (95% CI: 0.544 to 1)	0.867 (95% CI: 0.595 to 0.983)	0.667 (95% CI: 0.223 to 0.957)	91; 15; 6
49	HIA01 (normative thresholds)	The HIA02 and HIA03 assessments, including diagnosis from a medical doctor.	0.8 (95% CI: 0.7 to 0.8)	0.804 (95% CI: 0.764 to 0.839)	0.69 (95% CI: 0.653 to 0.724)	1118; 448; 670
	HIA01 (baseline thresholds)		0.6 (95% CI: 0.6 to 0.6)	0.896 (95% CI: 0.864 to 0.923)	0.339 (95% CI: 0.301 to 0.378)	
	Clinical judgement/real-life removal from play decisions (HIA01)		0.8 (95% CI: 0.8 to 0.8)	0.766 (95% CI: 0.724 to 0.804)	0.866 (95% CI: 0.837 to 0.891)	
50	Diplopia subtest (HIA01)	The HIA02 and HIA03 assessments, including diagnosis from a medical doctor.		0.051 (95% CI: 0.011 to 0.141)	0.971 (95% CI: 0.847 to 0.999)	119; 64; 43
	Passive neck movement subtest (HIA01)			0.085 (95% CI: 0.028 to 0.187)	0.912 (95% CI: 0.762 to 0.981)	
51	Abridged Stroop Test	The HIA02 and HIA03 assessments, including diagnosis from a medical doctor.	0.66 (95% CI: 0.54 to 0.76)			644; 33; 67
	HIA01			0.788 (95% CI: 0.611 to 0.91)	0.642 (95% CI: 0.515 to 0.755)	
	Abridged Stroop test and HIA01		0.72	0.848	0.582	
	All HIA01 subtests (normative thresholds)			Range provided (0.121–0.242)	Range provided (0.657–0.869)	
	All HIA01 subtests (baseline thresholds)			Range provided (0.292–0.75)	Range provided (0.5 to 0.738)	
35	Clinical judgement (HIA01)	The HIA02 and HIA03 assessments, including diagnosis from a medical doctor.		0.889 (90% CI: 0.865 to 0.914)	0.77 (90% CI: 0.736 to 0.803)	1149; 63; 326
36	Visible signs (community HIA01)	SCAT3, Cogstate and DASS-21 item and the BPI-18.	0.74			83; 58; 35

Continued

Table 2 Continued

Reference	Index test	Reference test	AUC	Sensitivity	Specificity	Study size; case; control
45	HIA01 (baseline thresholds)	The HIA02 and HIA03 assessments, including diagnosis from a medical doctor.	0.83 (95% CI: 0.78 to 0.88)	0.748 (95% CI: 0.656 to 0.825)	0.913 (95% CI: 0.820 to 0.967)	274; 148; 51
	HIA01 and KD combined			0.926 (95% CI: 0.859 to 0.967)	0.333 (95% CI: 0.208 to 0.479)	
	HIA01 and KD (time only) combined			0.889 (95% CI: 0.814 to 0.941)	0.373 (95% CI: 0.241 to 0.519)	
	HIA01 and KD (errors only) combined			0.806 (95% CI: 0.718 to 0.875)	0.765 (95% CI: 0.625 to 0.872)	
	KD		0.51 (95% CI: 0.41 to 0.61)	0.6 (95% CI: 0.52 to 0.68)	0.39 (95% CI: 0.26 to 0.54)	
19	KD: ≥2.6 s increase	Concussions diagnosed by sports medicine teams using 'current consensus guidelines'.	0.72 (95% CI: 0.68 to 0.77)	0.8 (95% CI: 0.73 to 0.85)	0.46 (95% CI: 0.43 to 0.49)	1559; 1239; 159
27	KD/ET: time	Logbook filled in by participant. Visual event monitoring. No official concussion tool.		0.17 (95% CI: 0.01 to 0.43)	0.88 (95% CI: 0.82 to 0.98)	19; 8; 6
	KD/ET: total saccades			0.33 (95% CI: 0.06 to 0.73)	0.44 (95% CI: 0.34 to 0.59)	
	KD/ET: total blinks			0.67 (95% CI: 0.27 to 0.94)	0.69 (95% CI: 0.54 to 0.79)	
33	KD	If a concussion was suspected, subjects' own health practitioner made formal diagnosis.		**1.00 (95% CI: 0.73 to 1.00**)	0.85 (95% CI: 0.42 to 0.97)	19; 7; 13
37	KD/ET: time	Certified athletic trainer assessment and SCAT5 test.		0.4 (95% CI: 0.05 to 0.85)	0.86 (95% CI: 0.68 to 0.96)	49; 8; 28
	KD/ET: total saccades		0.55 (95% CI: 0.23 to 0.77)	0.50 (95% CI: 0.12 to 0.88)	0.43 (95% CI: 0.25 to 0.63)	
38	KD	Physicians' diagnosis using the definition of concussion or KD (>3 s time increase) and SCAT3.		**1.00 (95% CI: 0.93 to 1.00**)	0.94 (95% CI: 0.84 to 0.99)	104; 52; 52
52	KD	Clinical examination by player's primary or emergency room physician, recommended to be within 24 hours of concussion.	0.91	0.9	0.91	141; 20; 121
53	KD: ≥1 s increase	Team physicians' diagnosis 'using established guidelines' and either SCAT3 or SCAT5.	0.69 (95% CI: 0.6 to 0.79)	0.62	0.72	295; 128; 166
54	KD	Physiotherapist assessment using SCAT3 and history.		0.98 (95% CI: 0.87 to 1.00)	0.96 (95% CI: 0.80 to 1.00)	110; 7; 103
55	KD: any time increase	SCAT5 and physical examination by a sports medical physician.	0.73 (95% CI: 0.56 to 0.90)	0.72 (95% CI: 0.47 to 0.90)	0.79 (95% CI: 0.64 to 0.89)	65; 18; 47
	KD: ≥6 s increase			0.44 (95% CI: 0.22 to 0.69)	0.94 (95% CI: 0.83 to 0.99)	
56	KD	Team physicians' assessment using PSCA2 and CogSport.		0.53	0.69	176; 19; 33
57	KD	Team athletic therapist/physician's clinical examination, SCAT3 and ImPACT.		0.62	0.84	1233; 84; 63

Continued

Table 2 Continued

Reference	Index test	Reference test	AUC	Sensitivity	Specificity	Study size; case; control
58	SAC	Athletic trainers or volunteer parents' judgement, confirmed by an expert physician.	0.68			332; 14; 12
	TTG		0.87			
	SAC, KD and TTG		0.97			
	KD		0.92			
59	SAC	Certified athletic trainer assessment.		0.95	0.76	1325; 66; 55
17	SCAT, SAC and BESS (same year baseline thresholds)	Assessment by local medical staff.		0.612 (range: 0.5–0.705)	**0.972 (range: 0.902 to 0.982**)	24 586; 1335; 4166
	SCAT, SAC, BESS, computer testing and BSI (same year baseline thresholds)			0.51	0.971	
18	SCAT, SAC and BESS (CART analysis)	Team physicians' or athletic trainers' assessment using the definition of concussion.		Range (0.8942–0.9852)	Range (0.2321 to 0.716)	1085; 1085; 0
60	SCAT2 (baseline thresholds)	SCAT2 and team physician's evaluation.	0.91	0.96	0.81	263; 32; 23
61	Concussion Symptom Inventory	Clinical assessment.	0.867 (95% CI: 0.85 to 0.88)			16 350; 641; 0
20	S100B (<1 day)	CT scan and Glasgow Coma Scale score.	0.7 (95% CI : 0.56 to 0.84)	0.95 (95% CI: 0.751 to 0.999)	0.1 (95% CI: 0.043 to 0.203)	130; 87; 0
25	Plasma CK-BB	RPQ and number of blows (visual event monitoring).	0.902	0.778	0.824	35; 18; 17
28	Plasma von Willebrand factor	RPQ-3 and number of blows received.	0.83	0.941	tme 0.765	38; 17; 21
62	NfL (1 hour)	RPQ and team physicians' diagnosis.	0.82			288; 87; 31
	NfL (12 hours)		0.72			
	Ttau (1 hour)		0.67			
	Ttau (12 hours)		0.6			
	NSE (1 hour)		0.55			
	NSE (12 hours)		0.66			
	S100B (1 hour)		0.57			
	S100B (12 hours)		0.54			

Grey cells indicate that the statistic was not provided in report. Bold indicates the highest statistic provided in review.
Further acronyms can be found in online supplemental material 11.
AUC, area under the curve; BESS, Balance Error Scoring System; BPI, Brief Pain Inventory; BSI, Brief Symptom Inventory; CART, Classification and Regression Tree; CK-BB, creatine kinase B type; CSI, Concussion Symptom Inventory; DASS, Depression, Anxiety and Stress Scale; DETECT, Display Enhanced Testing for Cognitive Impairment and mild Traumatic Brain Injury; GAM-CY, Generalized Acceleration Model for Concussion in Youth; HIA, head injury assessment; HITS, Head Impact Telemetry System; ImPACT, Immediate post-concussion assessment and cognitive testing; KD, King-Devick test; NfL, neurofilament light; NFL, National Football League; PLA, peak linear acceleration; PRA, peak rotational acceleration; PSCA2, Pitch-Side Concussion Assessment Version 2; RPQ, Rivermead Post-concussion Questionnaire; SAC, Standardized Assessment of Concussion; S100b, S100 calcium-binding protein B; SCAT, Sports Concussion Assessment Tool; Ttau, total tau; TTG, timed tandem gait.

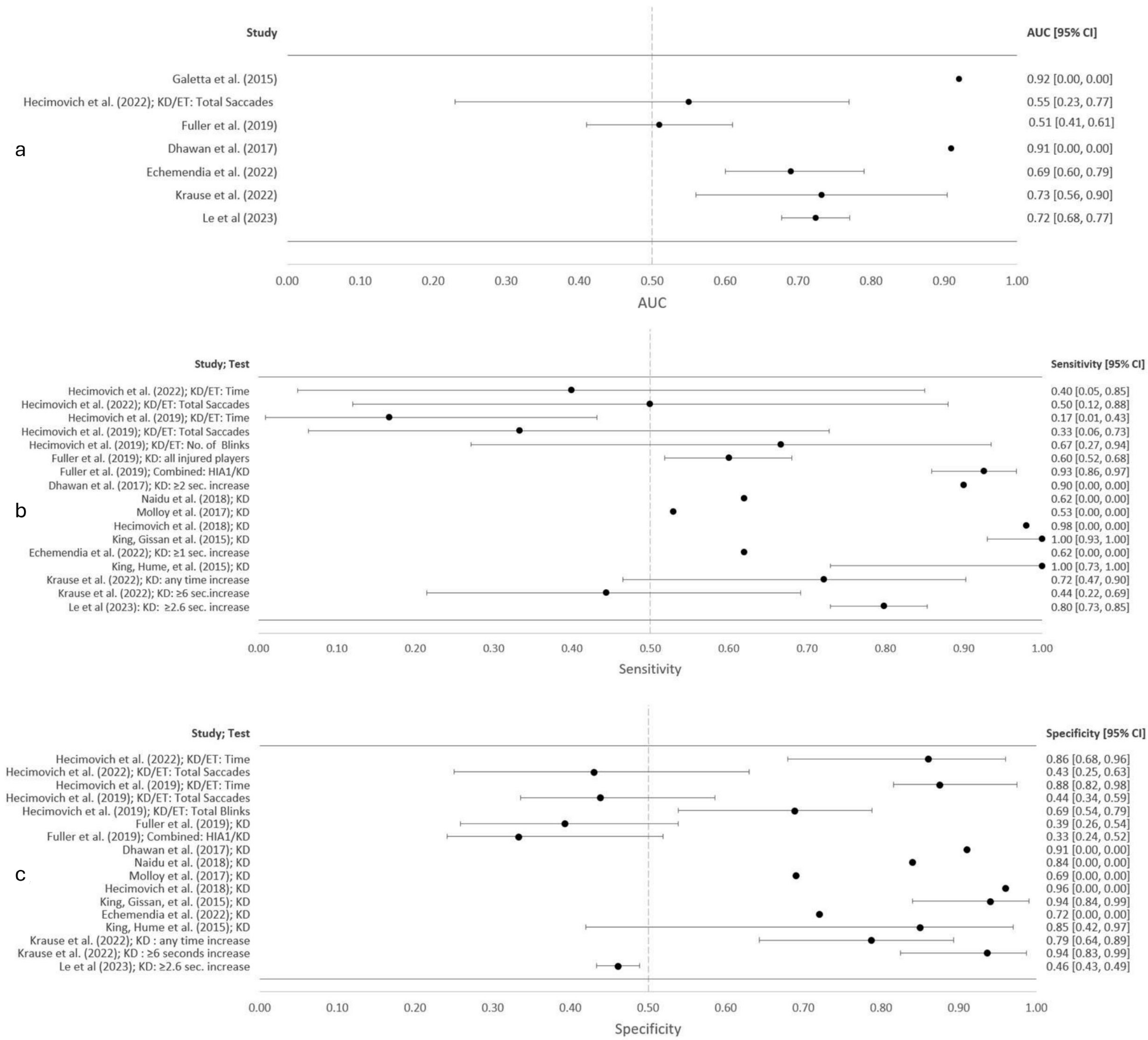

Figure 2 Forest plots for accuracy metrics provided for the KD test. (a) AUC values; (b) sensitivity values; (c) specificity values. AUC, area under the curve; HIA, head injury assessment; KD/ET, King-Devick eye tracking.

study is the consideration of the field in its entirety, along with the time frame of interest and the evaluation of study characteristics and performance.

It has captured a wide range of techniques used to assess concussion and the diverse reporting of these techniques. The performance of the tests within this review varies considerably.

Study characteristics

The studies within this sample were relatively recent, with 88% of the reports published in the last ten years. American Football, Ice Hockey and Rugby Union were the most frequent sports occurring in the sample. However, athletes in these sports wear different levels of protective equipment; for instance, Ice Hockey helmets and American Football helmets are designed differently, and no helmets are worn in rugby. This should be considered in biomechanical investigations conducted between sports as the variation in protective equipment may affect the mechanics of collisions.

33% of all studies took place in professional or semi-professional sports settings, 27% were conducted in American college sports environments, while 23% of studies were conducted in community settings. As many factors, such as athlete conditioning, resources available and concussion education, can vary between settings, it is important that methods are validated in the environment where they are going to be used.

Risk of bias was judged to be 'low' (42%) or 'unclear' (50%) within the 'flow and timing' domain of the risk assessment. The high percentage of 'unclear' risk of bias is due to the lack of reporting as 90% of reports did not precisely give the average period between injury and assessment, despite all reports providing a minimum time period to measurement. Additionally, the time period between the index and reference standard was unclear in 29% of reports.

Concussive injuries and their symptoms are known to evolve over time.[39] At different time points, the injured subject will be at a different point within the neurometabolic cascade of SR-mTBI, a series of acute physiological changes within the brain in response to the injury. These physiological changes are significant within a day of injury and have been linked to symptom burden and cognitive performance.[39] Consequently, this review recommends that the time point of each measurement of the subject relative to the time of injury should be precisely reported in future reports.

Population

A key result from this review is that the sex of participants of studies is not reported in all the studies and, where sex was reported, only 25% of participants were female. Preliminary evidence indicates a sex-related disparity in heart rate variability, symptom burden, neurocognitive performances and recovery time.[40] Koerte *et al*[40] suggest that the response of the brain to concussive injuries differs between males and females due to neck size, strength and other biological and hormonal differences.[40] Due to these differences, it is important that authors of future studies report the sex of the participants and ensure that females are appropriately represented when assessing the performance of SR-mTBI assessment and diagnostic methods.

Where age was provided, a wide age range was present within the included studies. However, the mean ages of participants were not reported in 30% of reports, and the SD not reported in 40% of reports. A third of the studies was conducted in American college settings, and college-age participants may still be physically developing, which could influence test results. It is recommended that the distribution of ages of participants should be reported in future studies.

20% of reports did not have any control subjects. It is important to ensure that studies have control subjects, when considering diagnostic or monitoring methods, to optimise reliability and reduce bias in studies.[41]

Interventions

A wide variety of interventions was implemented in this review, which fell under the index test domain in the quality assessment. Only 9% of reports were judged to have a ‘high’ risk of bias. Similarly, concerns regarding applicability were mostly judged to be ‘low’ or ‘unclear’. The QUADAS-2 signalling question with the highest percentage of ‘no’ answers (35%) fell within this domain: ‘Were the index test results interpreted without knowledge of the results of the reference standard?’. In populations of SR-mTBI studies, athletes have often been identified as sustaining a head impact and consequently been removed from the field, at which point the index test is performed. Proper blinding of this process may allow for studies which are more suited to validate SR-mTBI monitoring and assessment methods.

Comparators

A pervasive issue in this field is the lack of true diagnostic gold standard method for concussions, which was explicitly identified by CARE Consortium Investigators.[17] This is evident from the large number of reference methods used within this review, and the crossover between reference and index tests examined. The medical definition of concussion is a ‘traumatically induced transient disturbance of brain function’.[42] However, this is not an explicit mechanical criterion. It is thus difficult to mechanistically differentiate what is a concussion, or TBI, from what is not. Studies have sought to define a mechanical criterion for concussion; however, accelerometry-based methods are subject to high error.[43] Even with this error, these studies aim to measure the motion of the skull, not the brain within the skull. Brain modelling has also attempted to define biomechanical injury criteria, but the validation of brain models has proven difficult.[44]

The reference standard domain had the largest proportion of reports that were judged to have a ‘high’ risk of bias and ‘high’ concerns regarding applicability within the quality assessment. The nebulous definition of concussion influences this; without a gold standard it is extremely difficult to quantify what reference standard is or is not acceptable. The de facto gold standard within the SR-mTBI field is an assessment by a medical doctor, which was used by most studies investigating concussion in this review.

The lack of clarity surrounding this debate makes it difficult to judge whether studies are using the appropriate index tests. This ambiguity allows some studies to use the same tests as both index and reference. Future investigations should consider how this impacts the validity of their outcomes.

Outcomes

The variation in the outcome of interest for each report was reflected in the outcome reporting, as a wide range of metrics was provided. The highest AUC statistic was related to using accelerometry to collect kinematic data.[22] However, this study reclassified 207/244 concussive impacts based on prevalence without additional evidence, casting doubt on the validity of this result. The next highest AUC statistic, 0.97, was for the SAC, KD and timed tandem gait test combined using a Classification and Regression Tree analysis.[45] The highest sensitivity value was related to the KD test, which was provided by King *et al*.[34 38] The highest specificity statistic reported was for the SCAT, SAC and BESS tests combined.[17] Therefore, the highest performance summary statistics were related to either the use of the KD test or combinations of other index tests. This indicates the wider trend of combining tests in order to create more accurate, multimodal tests.

These multimodal tests, such as the HIA01, have been thoroughly investigated within this review and are deemed to generally have good accuracy. However, the practicalities of performing such test batteries limit their utility in low-resource environments. Two of the reports which analysed data from the CARE Consortium used methods such as decision trees to search for the most significant indicators of concussion.[17 18] However, this search may be impaired by a lack of consideration of the physiological mechanisms of concussion. Leveraging machine learning techniques alongside physiological considerations could provide further insight into the mechanisms of concussive injury. With improved insight, the selection of tests for use in low-resource environments could also be improved, and even in high-performance environments, where the under-reporting of concussions persists.[46]

The KD test was the most frequently used, potentially due to the ease of implementation in low-resource environments. While a wide range of sensitivity, specificity and AUC statistics were reported across the reports, measures of range, for example, 95% CIs, were not consistently reported. It is recommended that CIs should be consistently reported in future reports as they provide a more complete description of the performance of the test in a given sample, which is necessary if the test performance is exhibiting inter- and intra-study variations.

Limitations

Limitations of evidence

A slight complication of this review is that many of the interventions were not distinct but often encapsulated within each other. This is demonstrated in online supplemental material 12. Moreover, several methods were used as both index tests and gold standards within this review.

Limitations of review

This review only considered published journal articles; therefore, publication bias exists, and the analysis excludes preprints, conference proceedings and may exclude commercial projects. However, with the variation in quality evident even within this report, the conservative approach to selection taken here is justified. Language bias is also present within this work as only reports published in English were included. Additionally, only the first reviewer performed the risk of bias assessment, which reduces the reliability of the QUADAS-2 results.

The inclusion and exclusion criteria were altered for clarity and consistency through discussion between the first and second reviewers. This is a potential source of bias within the review as it is reviewer dependent. The finalised document describing these criteria can be found in online supplemental material 4.

The inclusion criteria which required that measurements were taken within a day of injury may have led to some distortion, in that reports which poorly reported the average time of injury measurements were included. This could lead to a bias in the included studies. This criterion was implemented to isolate methods which could be used in live sporting environments. This review is specific to the field of monitoring, assessing, diagnosing and predicting SR-mTBI. However, within this scope, the recommendations from this review are generalisable.

CONCLUSION

This review found that current sideline assessments to monitor and diagnose SR-mTBI are diverse. However, the average time between injury and measurement was not given in 90% of reports and female athletes are underrepresented within the studies. Performance varied considerably between different studies of the same test. The most frequently used performance summary metrics were all accuracy-related metrics (AUC, sensitivity and specificity). Overall, the validity of the results of these studies is hampered by poor and inconsistent reporting.

This analysis suggests that the field is moving towards using multimodal assessments which are combinations of previously existing assessments as these provide the optimal accuracy measures. Yet, these assessments can be resource intensive to implement and do not improve existing understanding of SR-mTBI, which is a key limitation in both diagnosis and the implementation of effective preventative measures. Providing additional insight into the mechanism of injury, both physiological and mechanical, could support the development of better assessment methods, and eventually a true gold standard for SR-mTBI.

Acknowledgements A subset analysis of this work was presented at the IRCOBI Cambridge conference in 2023. Nia Roberts provided advice and support as a librarian. Jane Clark proofread this work.

Contributors PH developed the idea for this work, with the guidance and under the supervision of AJ and JB. AJ and JB were involved in the planning of this work. AJ and JB supervised the entire work. PH performed the literature search. PH and LdAeB developed the screening criteria and screened the articles. PH wrote the article, with participation from JB and AJ, and is the guarantor of this work. All authors discussed the results and provided critical feedback.

Funding This work was funded by the Podium Institute for Sports Medicine and Technology, at the University of Oxford.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data are available upon reasonable request. Working documents from this review, including all the reports rejected at full text, can be found at: https://dx.doi.org/10.17605/OSF.IO/GAHTR. Additional data from this work is available upon request, to access please contact the corresponding author.

ORCID iDs
Phoebe Haste http://orcid.org/0000-0001-5771-0765
Leonardo de Almeida e Bueno http://orcid.org/0000-0002-0347-4581
Antoine Jérusalem http://orcid.org/0000-0001-5026-8038
Jeroen Bergmann http://orcid.org/0000-0001-7306-2630

REFERENCES

1 Lefevre-Dognin C, Cogné M, Perdrieau V, *et al*. Definition and epidemiology of mild traumatic brain injury. *Neurochirurgie* 2021;67:218–21.
2 Bower A. 100 former rugby league players start legal fight with RFL over brain injuries. The Guardian 2023.
3 Bull A. US health body rules collision sports cause CTE in landmark change. The Guardian; 2022.
4 Goodger D. World rugby selects prevent biometrics as head impact monitoring system of choice. 2021.
5 Echemendia RJ, Meeuwisse W, McCrory P, *et al*. The Sport Concussion Assessment Tool 5th Edition (SCAT5): Background and rationale. *Br J Sports Med* 2017;51:848–50.
6 McCrory P, Feddermann-Demont N, Dvořák J, *et al*. What is the definition of sports-related concussion: a systematic review. *Br J Sports Med* 2017;51:877–87.
7 O'Connor KL, Peeters T, Szymanski S, *et al*. Individual Impact Magnitude vs. Cumulative Magnitude for Estimating Concussion Odds. *Ann Biomed Eng* 2017;45:1985–92.
8 Broglio SP, Schnebel B, Sosnoff JJ, *et al*. Biomechanical properties of concussions in high school football. *Med Sci Sports Exerc* 2010;42:2064–71.
9 Daly E, Pearce AJ, Finnegan E, *et al*. An assessment of current concussion identification and diagnosis methods in sports settings: a systematic review. *BMC Sports Sci Med Rehabil* 2022;14:125.
10 Moody JR, Feiss RS, Pangelinan MM. A systematic review of acute concussion assessment selection in research. *Brain Inj* 2019;33:967–73.
11 King D, Brughelli M, Hume P, *et al*. Assessment, management and knowledge of sport-related concussion: systematic review. *Sports Med* 2014;44:449–71.
12 Echemendia RJ, Burma JS, Bruce JM, *et al*. Acute evaluation of sport-related concussion and implications for the Sport Concussion Assessment Tool (SCAT6) for adults, adolescents and children: a systematic review. *Br J Sports Med* 2023;57:722–35.
13 Šimundić A-M. Measures of Diagnostic Accuracy: Basic Definitions. *EJIFCC* 2009;19:203–11.
14 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
15 Haste P, Jerusalem A, Bergmann J, *et al*. What is the accuracy of sports monitoring techniques for diagnosing tbi: a systematic review: protocol. 2022. Available: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022376560
16 Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
17 Broglio SP, Harezlak J, Katz B, *et al*. Acute Sport Concussion Assessment Optimization: A Prospective Assessment from the CARE Consortium. *Sports Med* 2019;49:1977–87.
18 Garcia G-GP, Lavieri MS, Jiang R, *et al*. A Data-Driven Approach to Unlikely, Possible, Probable, and Definite Acute Concussion Assessment. *J Neurotrauma* 2019;36:1571–83.
19 Le RK, Ortega J, Chrisman SP, *et al*. King-Devick Sensitivity and Specificity to Concussion in Collegiate Athletes. *J Athl Train* 2023;58:97–105.
20 Kahouadji S, Salamin P, Praz L, *et al*. S100B Blood Level Determination for Early Management of Ski-Related Mild Traumatic Brain Injury: A Pilot Study. *Front Neurol* 2020;11:856.
21 Wirsching A, Chen Z, Bevilacqua ZW, *et al*. Association of Acute Increase in Plasma Neurofilament Light with Repetitive Subconcussive Head Impacts: A Pilot Randomized Control Trial. *J Neurotrauma* 2019;36:548–53.
22 Rowson S, Duma SM. Brain injury prediction: assessing the combined probability of concussion using linear and rotational head acceleration. *Ann Biomed Eng* 2013;41:873–82.
23 Reyes J, Mitra B, McIntosh A, *et al*. An Investigation of Factors Associated With Head Impact Exposure in Professional Male and Female Australian Football Players. *Am J Sports Med* 2020;48:1485–95.
24 The under-representation of women in biomechanical sports-related concussion studies. Cambridge International Research Council on Biomechanics of Injury; 2023.
25 Kilianski J, Peeters S, Debad J, *et al*. Plasma creatine kinase B correlates with injury severity and symptoms in professional boxers. *J Clin Neurosci* 2017;45:100–4.
26 Hubbard R, Stringer G, Peterson K, *et al*. The King-Devick test in mixed martial arts: the immediate consequences of knock-outs, technical knock-outs, and chokes on brain functions. *Brain Inj* 2019;33:349–54.
27 Hecimovich M, King D, Dempsey A, *et al*. In situ use of the King-Devick eye tracking test and changes seen with sport-related concussion: saccadic and blinks counts. *Phys Sportsmed* 2019;47:78–84.
28 Thomas R, Lynch CE, Debad J, *et al*. Plasma von Willebrand Factor Is Elevated Hyperacutely After Mild Traumatic Brain Injury. *Neurotrauma Rep* 2023;4:655–62.
29 Galetta KM, Barrett J, Allen M, *et al*. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology (ECronicon)* 2011;76:1456–62.
30 Monroe DC, Thomas EA, Cecchi NJ, *et al*. Salivary S100 calcium-binding protein beta (S100B) and neurofilament light (NfL) after acute exposure to repeated head impacts in collegiate water polo players. *Sci Rep* 2022;12:3439.
31 Huibregtse ME, Sweeney SH, Stephens MR, *et al*. Association Between Serum Neurofilament Light and Glial Fibrillary Acidic Protein Levels and Head Impact Burden in Women's Collegiate Water Polo. *J Neurotrauma* 2023;40:1130–43.
32 Fogarty AE, Guay CS, Simoneau G, *et al*. Head Motion Predicts Transient Loss of Consciousness in Human Head Trauma: A Case-Control Study of Mixed Martial Artists. *Am J Phys Med Rehabil* 2019;98:859–65.
33 King D, Hume P, Gissane C, *et al*. Use of the King-Devick test for sideline concussion screening in junior rugby league. *J Neurol Sci* 2015;357:75–9.
34 King D, Hume PA, Clark TN, *et al*. Use of the King-Devick test for the identification of concussion in an amateur domestic women's rugby union team over two competition seasons in New Zealand. *J Neurol Sci* 2020;418:117162.
35 Falvey É, Tucker R, Fuller G, *et al*. Head injury assessment in rugby union: clinical judgement guidelines. *BMJ Open Sport Exerc Med* 2021;7:e000986.
36 Reyes J, Mitra B, Makdissi M, *et al*. Visible Signs of Concussion and Cognitive Screening in Community Sports. *J Neurotrauma* 2022;39:122–30.
37 Hecimovich M, Murphy M, Chivers P, *et al*. Evaluation and Utility of the King-Devick With Integrated Eye Tracking as a Diagnostic Tool for Sport-Related Concussion. *Orthop J Sports Med* 2022;10:23259671221142255.
38 King D, Gissane C, Hume PA, *et al*. The King-Devick test was useful in management of concussion in amateur rugby union and rugby league in New Zealand. *J Neurol Sci* 2015;351:58–64.
39 Giza CC, Hovda DA. The new neurometabolic cascade of concussion. *Neurosurgery* 2014;75 Suppl 4:S24–33.
40 Koerte IK, Schultz V, Sydnor VJ, *et al*. Sex-Related Differences in the Effects of Sports-Related Concussion: A Review. *J Neuroimaging* 2020;30:387–409.
41 Torday JS, Baluška F. Why control an experiment? *EMBO Reports* 2019;20.
42 Harmon KG, Drezner JA, Gammons M, *et al*. American Medical Society for Sports Medicine position statement: concussion in sport. *Br J Sports Med* 2013;47:15–26.
43 Brennan JH, Mitra B, Synnot A, *et al*. Accelerometers for the Assessment of Concussion in Male Athletes: A Systematic Review and Meta-Analysis. *Sports Med* 2017;47:469–78.
44 Miller LE, Urban JE, Stitzel JD. Validation performance comparison for finite element models of the human brain. *Comput Methods Biomech Biomed Engin* 2017;20:1273–88.
45 Fuller GW, Cross MJ, Stokes KA, *et al*. King-Devick concussion test performs poorly as a screening tool in elite rugby union players: a prospective cohort study of two screening tests versus a clinical reference standard. *Br J Sports Med* 2019;53:1526–32.
46 Longworth T, McDonald A, Cunningham C, *et al*. Do rugby league players under-report concussion symptoms? A cross-sectional

study of elite teams based in Australia. *BMJ Open Sport Exerc Med* 2021;7:e000860.
47 Campolettano ET, Gellner RA, Smith EP, *et al*. Development of a Concussion Risk Function for a Youth Population Using Head Linear and Rotational Acceleration. *Ann Biomed Eng* 2020;48:92–103.
48 Espinoza TR, Hendershot KA, Liu B, *et al*. A Novel Neuropsychological Tool for Immersive Assessment of Concussion and Correlation with Subclinical Head Impacts. *Neurotrauma Rep* 2021;2:232–44.
49 Fuller GW, Tucker R, Starling L, *et al*. The performance of the World Rugby Head Injury Assessment Screening Tool: a diagnostic accuracy study. *Sports Med Open* 2020;6:2.
50 Fuller GW, Miles J, Tucker R, *et al*. Diagnostic Utility of New SCAT5 Neurological Screen Sub-tests. *Sports Med Open* 2021;7:14.
51 Fuller GW, Gardner A, Tucker R, *et al*. Expansion of cognitive testing for off-field concussion screening in elite rugby players: A cohort study. *J Sci Med Sport* 2021;24:1204–10.
52 Dhawan PS, Leong D, Tapsell L, *et al*. King-Devick Test identifies real-time concussion and asymptomatic concussion in youth athletes. *Neurol Clin Pract* 2017;7:464–73.
53 Echemendia RJ, Thelen J, Meeuwisse W, *et al*. The Utility of the King-Devick Test in Evaluating Professional Ice Hockey Players With Suspected Concussion. *Clin J Sport Med* 2022;32:265–71.
54 Hecimovich M, King D, Dempsey AR, *et al*. The King-Devick test is a valid and reliable tool for assessing sport-related concussion in Australian football: A prospective cohort study. *J Sci Med Sport* 2018;21:1004–7.
55 Krause DA, Hollman JH, Breuer LT, *et al*. Validity Indices of the King-Devick Concussion Test in Hockey Players. *Clin J Sport Med* 2022;32:e313–5.
56 Molloy JH, Murphy I, Gissane C. The King-Devick (K-D) test and concussion diagnosis in semi-professional rugby union players. *J Sci Med Sport* 2017;20:708–11.
57 Naidu D, Borza C, Kobitowich T, *et al*. Sideline Concussion Assessment: The King-Devick Test in Canadian Professional Football. *J Neurotrauma* 2018;35:2283–6.
58 Galetta KM, Morganroth J, Moehringer N, *et al*. Adding Vision to Concussion Testing: A Prospective Study of Sideline Testing in Youth and Collegiate Athletes. *J Neuroophthalmol* 2015;35:235–41.
59 McCrea M. Standardized Mental Status Testing on the Sideline After Sport-Related Concussion. *J Athl Train* 2001;36:274–9.
60 Putukian M, Echemendia R, Dettwiler-Danspeckgruber A, *et al*. Prospective clinical assessment using Sideline Concussion Assessment Tool-2 testing in the evaluation of sport-related concussion in college athletes. *Clin J Sport Med* 2015;25:36–42.
61 Randolph C, Millis S, Barr WB, *et al*. Concussion symptom inventory: an empirically derived scale for monitoring resolution of symptoms following sport-related concussion. *Arch Clin Neuropsychol* 2009;24:219–29.
62 Shahim P, Tegner Y, Marklund N, *et al*. Neurofilament light and tau as blood biomarkers for sports-related concussion. *Neurology (ECronicon)* 2018;90:e1780–8.