Supplemental Content 5
Methods

Definitions

Potential Concussive Event (PCE). An event in which the head sustains a force or change in pressure that has the potential for injury.

Prevalence. For the purposes of this report, "prevalence" is the proportion of the sample with the particular sign, symptom, or deficit being discussed at the particular time point of measurement. When reporting data from studies used to support the definition, the "absolute prevalence" is used–the difference between the prevalence of the sign, symptom, or deficit in the "potentially concussed" group compared to controls (either self or control group).

Reliable Change Index/Estimate (RCI/RCE) (see Supplemental Content 14, Reliable Change Index). The Reliable Change Index (RCI) or Estimate (RCE) is a standardized difference score that was designed to assess the effects of a clinical intervention. Any change from one testing occasion to another is considered significant if the magnitude of the change is sufficiently large in proportion to the associated error variance of the test.[1]

Equivocal findings. Statistical significance of the difference between PCE and comparator scores was not always provided. When the data allowed, recovery curves were plotted to estimate differences, and clinical judgment was used to decide whether the observed differences were clinically significant. If a small difference was observed, but the clinical utility of the difference was unknown, the finding was considered "equivocal."

Abbreviations.
PCE – Potential Concussive Event
SSD – Sign, Symptom, Deficit

Inclusion/Exclusion Criteria

Inclusion Criteria

Population:	Having sustained a PCE Any age Awake at the time of neurologic or cognitive assessment (GCS 13 – 15) (In general, duration of loss of consciousness (LOC) was not specified in the included studies, but LOC > 30 min. was an exclusion criterion)
Comparators:	Comparison groups were required for Key Questions 1 and 2, and not for Key Questions 3 and 4. *Direct comparators only*

Individuals as their own controls (pre/post in same sample)
Uninjured volunteers
Patients who sustained traumas and not a PCE, or who have other pathologies, such as PTSD
Population norms

Not included as comparators

Patients with GCS 3 – 12
Patients with severe mental or neurological disorders such as bipolar disorder or Parkinson's disease
Patients with substance abuse

Measures: Sign. An objective, observed, or measurable parameter, like loss of consciousness or documented confusion or disorientation.
Symptom. A subjective complaint like a headache or dizziness.
Neurologic Deficit. A response to one or more components of an assessment of neuromotor function that is below the pre-injury or comparator response to the same component(s).
Cognitive Deficit. A response to one or more components of an assessment of cognitive function that is below the pre-injury or comparator response to the same component(s).

Timing: For Key Questions 1 and 2, assessments may be conducted at any time during the first 3 months post-injury. For Key Questions 3 and 4, measurements must be taken at a fixed time point uniformly for all participants.

Settings: All

Study Design: Prospective and retrospective cohort studies
Case-control studies
Before-after studies of a single cohort
Systematic reviews

Sample size: No minimum

Presentation of data: For signs, symptoms, and neurologic deficits, publications that combined a number of questions or tests that contribute to one single, discrete SSD were included. Those that combined questions or tests about more than one SSD and aggregate them as a single construct (composites) were excluded.

For cognitive deficits, measures reported in composites were included.

Exclusion Criteria

Penetrating head injury

Not English language
Ineligible study design (eg, abstract only, non-systematic review, case reports, editorial, letter)
Population ineligible (eg, severe TBI only, or mixed severities)
No SSD reported, or the only SSDs reported are part of the case definition
For Signs, Symptoms, and Neurologic Deficits, data are reported in aggregate or as an index
No time point reported at all.
For Cognitive Deficits data are self-reported
For Cognitive Deficits assessment instruments are not validated
Long-term follow-up, not relevant to concussion definition
Duplicate data (data reported in another included study)

Exclusion Criteria Specific to Key Questions 1 and 2

No or ineligible comparison group
Timeframe > 3 months

Exclusion Criteria Specific to Key Questions 3 and 4

No association reported
No fixed biologic time point for measurement

Protocol for identifying studies with most broad case definitions

The first task in defining concussion was to identify the occurrence of SSDs in samples of people who had sustained a PCE. However, if a sign was used as an inclusion criterion for a sample, the prevalence of that sign in the sample *as an outcome* does not accurately represent the estimated prevalence in the general PCE population. For example, sometimes LOC and PTA are used to define a case, and are a part of the inclusion criteria. If a paper uses LOC as an inclusion criterion, the count of people with LOC in the sample will overestimate the prevalence in the general PCE population.

The following rubric was developed to select which publications contained sign and symptom outcome data that could be used to formulate the definition:

- If the sign or symptom was required as an inclusion criterion for a case, data about the prevalence of that sign or symptom could not be used to derive the definition.
- If the sign or symptom was used to distinguish a "concussed" sample from a "non-concussed" sample (eg, in a hospital sample, if LOC was used to constitute the sample of "concussed" patients and No LOC was used to constitute the sample of "other injury" patients), data about the prevalence of that sign or symptom could not be used to derive the definition.
- If a case definition was ambiguous (eg, it is not clear if the sign or symptom must be present or may be present in order to identify a case), data about the prevalence of that sign or symptom could not be used to derive the definition.

- If a sign or symptom was one of a list that may or may not be present to identify a case (eg, a case was identified by LOC and/or PTA and/or one or more of headache, dizziness, disorientation, nausea, vomiting, fatigue, etc.), data about the prevalence of that sign or symptom could be used to derive the definition.

Utility of Data

Composite vs Individual Measures. To define concussion, it was necessary to identify the set of essential attributes of concussion. Some of the publications included in this review present data for the presence of individual SSDs, and other publications report data as composites. Composite "scores" are values from more than one test or question aggregated into a single score. Some composite measures combine a number of questions or tests that contribute to one single, discrete SSD; others combine questions or tests about more than one SSD, and aggregate them as a single construct. The former would be useful information for the definition of concussion, whereas the latter would not; aggregating more than one SSD into a composite eliminates the ability to know the proportion of patients with each, distinct SSD, and obscures the information necessary for a definition. For example, a study might report that some proportion of a sample had 3 or more symptoms, from a list of possible symptoms, within 1 day post-injury. Reporting the information in this way does not provide information about what proportion of the sample had which specific symptoms (eg, the essential attributes).

Proportions vs Average Scores. Similarly, the more prevalent a particular SSD in a sample of persons who sustained a PCE, the more "essential" the attribute (the SSD). Averages for an entire group disguise the count of how many persons had the SSD. Therefore, for the task of defining concussion, results reported in some form of proportion were considered more useful than those reported as mean differences.

Fixed vs Range Time Points. Some publications took measures at explicitly fixed time points. Others targeted a fixed time point, and reported some measure of variance for the actual data collection time points (eg, targeted for 1 week post-injury, with a report of the average and standard deviation for actual days from post-injury to measurement). Some publications collected data over a designated range (eg, within 3 weeks post-injury). For the purposes of defining concussion, explicitly fixed time points provided the most useful information about the occurrence of SSDs over time. For the purposes of this report, time-pints were:

- pre-event (eg, pre-season measures),
- immediately following the event,
- some specified short-term length of time post-event (eg, 5 min., 15 min., 3 hrs.),
- within 48 hrs. post-event, and
- at 1-day intervals from 48 hrs. forward.

Reporting Group Differences. Only studies with comparison groups were included in the analysis of the occurrences of SSDs over time (Key Questions 1 and 2). The comparator may have been measures from the same people taken pre-injury, or measures from a

control group. Three methods of indicating a difference between groups for any given test were used in the included studies:
(1) The proportion of PCE participants with a particular SSD compared to the proportion of Controls with the SSD (sometimes analyzed for statistical significance, and sometimes simply a report of proportions);
(2) The proportion of participants with a PCE who had a clinically significant difference from Control measures based on a reliable change index (RCI); and
(3) A statistically significant difference in mean scores between PCE and Control groups.

REFERENCE

1. Hinton-Bayre AD, Geffen GM, Geffen LB, McFarland KA, Friis P. Concussion in contact sports: reliable change indices of impairment and recovery. *J Clin Exp Neuropsychol.* 1999;21(1):70-86.