

Supplementary table 1. Critical appraisal of statistical methodology of individual studies.

Reference	Number of samples	Post-injury sampling time	Type of analysis & biomarker classification*	Statistical method and critical appraisal
Hayes et al. (2002) ³⁵	1	>12 months after injury	cross-sectional diagnostic	student's t tests with unmatched subjects; tests for differences between groups; limited to one sample; only included
Davies et al. (2007) ³²	1	post-acute 2–52 wk post-injury chronic >52 wk	cross-sectional diagnostic	ANOVA for differences between groups; subjects; multiple linear regression; features in patients with SCI; limited to one sample; only included
Moghaddam et al. (2016) ⁴²	11	day 0 - 3 months	longitudinal predictive	prediction of AIS improvement; analysis for repeated measures
Tong et al. (2018) ⁴¹	4	day 0 – 4 weeks	longitudinal predictive	prediction of LEMS and “marked recovery = 2 or more AIS levels”; problematic in terms of scaling
Ahadi et al. (2015) ³⁴	3	24 - 72 hours	cross-sectional diagnostic	statistical analysis is not stated; illustrated only descriptively; p-values and control are reported in the parametric or non-parametric; blood sample collection only in
Kuhle et al. (2015) ²¹	15	within 12 h - 7 days	longitudinal diagnostic	mixed-effects linear regression; differences in AUC for the 7 days between the groups; Spearman's levels and motor and sensory s
Biglari et al. (2015) ³⁶	9	day 0 – 12 weeks	longitudinal prognostic	Mann–Whitney U-Test between improvement); Wilcoxon signed-rank test for cytokine levels within groups; analysis for repeated measures
Wolf et al. (2014) ³¹	1	within 24 h	cross-sectional diagnostic	Wilcoxon signed-rank test between neurologic lesions; small sample size; limited to one sample in the ac
De Mello Rieder et al. (2019) ³³	2	within 48 h and after 7 days	cross-sectional diagnostic	Biomarker concentrations were compared between Control using unpaired t-test with limited to two samples
Heller et al. (2017) ³⁷	10	4 hours - 12 weeks	longitudinal prognostic	Logistic regression modeling; p-values assessed by estimation of the AUC; testing; clinical endpoint = AIS conversion after injury – no data on neurologic endpoint were reported
Hassanshahi et al. (2013) ³⁹	4	3–6 hours - 3 months	longitudinal diagnostic	Repeated measures ANOVA / Friedman's test; measures; corrected for multiple comparisons; repeated measures ANOVA we

				and partial eta; bar charts don't show lower error bar
Kijima et al. (2019) ²⁶	1	within 3 days	cross-sectional prognostic	Statistical methods were described but difficult for the reader to distinguish between the animal trial data and the human data; limited to one sample; blood sample collection only in injury phase
Du et al. (2018) ²²	8	day 0 - 14 days	longitudinal prognostic	Statistics were described very briefly; measurements were included but not analyzed
Kwon et al. (2010) ²⁹	?	every 6 to 8 hours for the first 72 hours	cross-sectional prognostic	predictive modeling described but not for the reader, how many blood samples were used at 24 hours was used for statistical analysis; not report analysis for repeated measures; why the sample at 24 hours was used; blood sample collection only in injury phase
Kwon et al. (2017) ²⁸	1	24 hours	cross-sectional prognostic	Predictive modeling using a correlation matrix; sub-groups (cervical vs. thoracic) were used to limit the utility in acute patients; not be available; limited to one sample in the acute phase
Dalkilic et al. (2017) ³⁰	1	24 hours	cross-sectional prognostic	Linear discriminant analysis and logistic regression to predict AIS conversion; limited to one sample in the acute phase
Pouw et al. (2014) ²⁴	1	3 - 24 hours	cross-sectional diagnostic	Correlations and differences between groups; model was used; very small sample size; limited to one sample in the acute phase
Ungureanu et al. (2014) ³⁸	3-12	6 hours – 11 days	cross-sectional diagnostic	statistical analysis was not stated; which blood samples in the acute phase were used for the calculation; limited to one sample in the acute phase
Hayakawa et al. (2012) ²³	12	6 hours - 21 days	longitudinal diagnostic	Repeated measure ANOVA; results were not reported: no F value, no p value, sample size

* Biomarker classification: “diagnostic” if the studies reported on injury severity or the mere presence of SCI (when the control group consisted of patients without SCI) and “prognostic” if the included studies reported on neurological improvement by means of AIS grade conversion.