Patient-important outcomes other than mortality in contemporary ICU trials: a scoping review

Anders Granholm, Carl T. Anthon, Maj-Brit N. Kjær, Mathias Maagaard,

Benjamin S. Kaas-Hansen, Praleene Sivapalan, Olav L. Schjørring, Lars W. Andersen,

Ole Mathiesen, Thomas Strøm, Aksel K. G. Jensen, Anders Perner, Morten H. Møller

Supplemental Digital Content

Table of contents (continued next page)

Patient-important outcomes other than mortality in contemporary ICU trials: a scoping review

Supplemental Digital Content

Supplemental methods

Search strategy

We searched PubMed, as all 10 journals of interest are index in PubMed (and MEDLINE). We used the Cochrane Collaboration's highly sensitive filter for randomized controlled trials (1), and restricted the search to studies published from 2018 and onwards. No specific search terms for trials conducted in relevant populations were used, due to the limited number of journals and to avoid missing relevant results during the included time period. The search string was unaltered from the search string included in the published protocol (2).

Search string (search last conducted on 5 May 2022):

("The New England journal of medicine"[Journal] OR "JAMA"[Journal] OR "Lancet (London, England)"[Journal] OR "BMJ (Clinical research ed.)"[Journal] OR "The Lancet. Respiratory medicine"[Journal] OR "Intensive care medicine"[Journal] OR "American journal of respiratory and critical care medicine"[Journal] OR "Chest"[Journal] OR "Critical care medicine"[Journal] OR "Critical care (London, England)"[Journal])

AND

("2018/01/01"[Date - Publication]: "3000"[Date - Publication])

AND

(randomized controlled trial [pt] OR controlled clinical trial [pt] OR randomized [tiab] OR placebo [tiab] OR drug therapy [sh] OR randomly [tiab] OR trial [tiab] OR groups [tiab]) NOT (animals [mh] NOT humans [mh])

Supplemental Digital Content

Data extraction forms

Two data extraction forms were used, one for general trial data (1 row pr trial) and one for outcomes (1 row per outcome); the extraction forms were slightly adapted from the extraction form included in the published protocol (2) after pilot-testing, as described in the main text.

Extraction form for trial-level data:

Covidence ID	Initials of extractor	Notes	DOI	Journal	Author/trial acronym	Year	Countries	Number of centers

Number of patients	Intervention type(s) [drug/management/device]	Interventions	Full protocol available and referenced [y/n]	Restricted to patients with suspected or confirmed COVID-19 only [y/n]

Extraction form for outcome-level data:

Covidence ID	Initials of extractor	Outcome #	Outcome	Definition/ tool used	Type [days alive…/HRQoL/functional-cognitive-neurological/ordinal]	(Co-)primary outcome [y/n]	Prioritization of multiple outcome components [NR if not relevant]

Assessment time point incl. censoring/ truncation	Death - proportion [n/N - %]	Handling of death	Missing data - proportion [n/N - %]	Missing data handling strategy	Effect measure(s) used	Statistical analysis used

Data and code

As described in the main text, data were extracted independently and in duplicate and discrepancies double-checked and discussed where necessary. Following this, extractions were further categorized to facilitate analysis and presentation.

The final datasets used for analysis are available online at Zenodo (link: doi.org/10.5281/zenodo.6881171) and include the two files mentioned below. The briefly annotated analysis code (written in R) is appended at the end of this document under "Analysis code".

data general.csv:
Semi-colon separated file with general, trial-level data. Contains the following variables:

1. id (character): unique identifier for each trial, consisting of the trial's or the first authors name and the year of publication for the primary trial report.
2. original_found_outside_search (y/n): "y" if the primary trial report was included but not found in the search results (i.e., found via references from a secondary publication included in the search), otherwise "n".
3. number_of_studies_incl_original (numeric): number of publications included for each trial, counting both relevant studies (primary and secondary) identified in the search and eventual primary publications identified outside the search.
4. number_of_studies_contributing_incl_original (numeric): number of included publications contributing data; some secondary studies included in the search may be relevant but not provide additional relevant data than i.e. the primary publication (for example, secondary analyses or subgroup analyses without new relevant outcomes), these studies are not counted. This variable and the two above were used to create the study selection flowchart.
5. primary_doi (character): digital object identifier of the primary publication of each trial.
6. dois_contributing (character): digital object identifier of all included publications contributing data (separated by "and").
7. journal (character): name (abbreviated) of the journal where the primary report of each trial was published.
8. author_trial_name (character): name of the trial (if available) or alternatively of the first author of the primary trial report.
9. year (numeric): year of publication of the primary trial report.
10. countries_text (character): countries where patients were randomized in the trial, separated by ",".
11. countries_n (numeric): number of countries where patients were randomized in the trial.
12. centres_n (numeric): number of centers where patients were randomized in the trial.
13. patients_n (numeric): number of patients randomized in the trial.
14. intervention_type (Drug/Management/Device): type of intervention(s) and comparator(s) assessed in the trial.
15. intervention_text (character): intervention(s) and comparator(s) as text.
16. protocol_available (y/n): "y" if a full trial protocol was included with any of the included studies or clearly referenced in one of them and available (not considering trial registrations to be full protocols); "n" otherwise.
17. covid19 (y/n): "y" if the trial is restricted to patients with suspected or confirmed COVID-19 only.

data outcomes.csv:

Semi-colon separated file with outcome-level data (1 outcome per row). Contains the following variables:

1. id (character): unique identifier for each trial, as described above.
2. outcome_n (numeric): consecutive numbering for outcomes in each trial.
3. outcome (character): outcomes, categorized.
4. additional_comments (character): additional comments describing key additional information regarding the outcomes and their categorization (i.e., if two outcomes using the same title appear for the same trial, they describe their differences).
5. type (days alive..., functional/cognitive/neurological, HRQoL, ordinal/other): the type of outcome, categorised.
6. co_primary (y/n): "y" if the outcome is the primary outcome or a co-primary outcome of the trial; otherwise "n" (if an outcome is listed as the primary outcome in a secondary report but not in the primary trial report, this was "n").
7. prioritisation (character): prioritization of multiple outcome components; "NR" if not relevant.
8. time_point (character): the assessment time point of each outcome and any possible truncation mentioned.
9. death_pct (numeric): the percentage of patients that had died at the time of outcome assessment.
10. death_handling (character): how dead patients were handled in the analyses; categorized.
11. missing_pct (numeric): the percentage of patients with missing data for the outcome; dead patients and patients deliberately excluded were not considered to be missing, and dead patients were included in the calculations regardless of how death was handled in the analyses.
12. missing_handling (character): the missing data handling strategy, categorized. Important abbreviations: CCA: complete case analysis; LVCF: last value carried forward; MI: multiple imputations.
13. effect_measures (character): the effect measures used, abbreviated and separated by "and"; see **Table 4** in the main text for additional details.
14. stat (character) the statistical methods used, abbreviated and categorized and separate by "and"; see **Table 5** in the main text for additional details.

In both files, missing values are encoded as "NA". Some values were secondarily combined or separated into multiple values before analysis/presentation, as described in analysis code.

Supplemental Digital Content

Assumptions

The following assumptions/simplifications were made during data extraction:

- For the number of centers, we counted individual intensive care units as individual centers if reported as such, even if they were in the same hospital.
- For countries, we only considered countries where randomization took place, excluding countries where randomization was planned but did not happen.
- According to the protocol (2), we did not consider trial registrations as full study protocols.
- Generally, few details on censoring/truncation were reported; these were extracted where found.
- Missing data proportions were calculated not considering dead patients as missing (regardless of how death was handled) and not considering patients excluded on purpose (any post-randomization exclusion from the trial in general or exclusions from a particular analysis of subsets of patients, e.g., where patients on renal replacement therapy at baseline were excluded from analyses of the days alive without renal replacement therapy or where "return to work" outcomes were only calculated for those previously employed, and similar) or without consent as missing. Where no missing data was mentioned or could be identified, we assumed that data were complete.
- For missing data handling, we assumed that complete case analysis was used unless otherwise mentioned, and only considered the missing data handling strategy used in the primary analysis.
- For statistical methods used, we had to make assumptions on the methods used based on vague descriptions in many cases; a substantial number of studies only reported statistical methods overall (e.g., by outcome type or distribution) and not for each outcome. In these cases, we assumed the analysis method used out of those mentioned if the outcome distribution could be inferred from the text, the effect measures presented, or previous knowledge about typical distributions for similar outcomes.
- For outcomes such as "ventilator-free days" without further descriptions, we assumed that these only counted the days alive and not on ventilator, as is common practice (3-5).
- For outcome such as "days alive without mechanical ventilation", we assumed that death was not penalized if the authors provided a detailed description of the outcome without mentioning penalizing death; if no further description than the outcome name was provided, we considered the handling of dead patients as unclear.
- If an outcome was only reported in a subgroup of patients in a trial, we calculated the proportions of dead patients and missing values for the subgroup only (where possible).
- Components of composite outcomes or sub-components of complex outcomes (e.g., domains in tools used for assessing health-related quality of life) were only considered as separate outcomes if explicitly labelled as such by the authors or if a formal analysis comparing them was presented; if only descriptive data were presented secondarily to supplement the primary analysis, these components were not considered separate outcomes.
- Where individual components of health-related quality of life tools were reported with equal prioritization and considered a single outcome with multiple components, these were classified as a health-related quality of life outcome, although some components could also be considered functional.
- Days alive without respiratory support includes any type of respiratory support (different between trials), whereas days alive without (invasive) mechanical ventilation/ventilator-free days only includes invasive or non-invasive ventilation.
- Days alive and at home includes institution-free days.

7

- Outcomes with substantially overlapping definitions were grouped to facilitate analysis and presentation.

Supplemental Digital Content

Supplemental results

Supplemental Table 1: General outcome-level characteristics

Variable	All outcomes (n = 687)	Days alive without... (n = 337)	Functional/ cognitive/ neurological (n = 198)	HRQoL (n = 118)	Ordinal/ other (n = 34)	Missing values
Primary/ co-primary outcome	56 (8.2%)	36 (10.7%)	11 (5.6%)	5 (4.2%)	4 (11.8%)	0 (0.0%)
Prioritization of outcome components						0 (0.0%)
- Not relevant	625 (91.0%)	330 (97.9%)	178 (89.9%)	86 (72.9%)	31 (91.2%)	
- Equal prioritization	43 (6.3%)	0 (0.0%)	13 (6.6%)	27 (22.9%)	3 (8.8%)	
- Not relevant (individual categories/ domains/ components reported secondarily, but not considered or analyzed as primary outcomes)	18 (2.6%)	7 (2.1%)	6 (3.0%)	5 (4.2%)	0 (0.0%)	
- Three dichotomizations reported with equal prioritization	1 (0.1%)	0 (0.0%)	1 (0.5%)	0 (0.0%)	0 (0.0%)	
Assessment time point						2 (0.3%)
- Day 28[1]	224 (32.7%)	211 (63.0%)	4 (2.0%)	2 (1.7%)	7 (20.6%)	
- Day 180/6 months	133 (19.4%)	8 (2.4%)	73 (36.9%)	52 (44.1%)	0 (0.0%)	
- Day 90/3 months	87 (12.7%)	28 (8.4%)	38 (19.2%)	20 (16.9%)	1 (2.9%)	
- 1 year[2]	37 (5.4%)	1 (0.3%)	19 (9.6%)	17 (14.4%)	0 (0.0%)	
- Day 14[3]	27 (3.9%)	18 (5.4%)	0 (0.0%)	0 (0.0%)	9 (26.5%)	
- Day 60	27 (3.9%)	15 (4.5%)	8 (4.0%)	4 (3.4%)	0 (0.0%)	
- Day 30	26 (3.8%)	22 (6.6%)	3 (1.5%)	1 (0.8%)	0 (0.0%)	
- Hospital discharge/in-hospital	26 (3.8%)	0 (0.0%)	20 (10.1%)	4 (3.4%)	2 (5.9%)	
- ICU discharge[4]	23 (3.4%)	0 (0.0%)	19 (9.6%)	4 (3.4%)	0 (0.0%)	
- 90 days/3 months after ICU discharge	22 (3.2%)	0 (0.0%)	10 (5.1%)	10 (8.5%)	2 (5.9%)	
- Day 21[5]	19 (2.8%)	16 (4.8%)	0 (0.0%)	0 (0.0%)	3 (8.8%)	
- Day 7[6]	11 (1.6%)	6 (1.8%)	2 (1.0%)	0 (0.0%)	3 (8.8%)	
- 30 days after ICU discharge	4 (0.6%)	0 (0.0%)	0 (0.0%)	4 (3.4%)	0 (0.0%)	
- Day 25	3 (0.4%)	3 (0.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
- 6 months after hospital discharge	2 (0.3%)	0 (0.0%)	2 (1.0%)	0 (0.0%)	0 (0.0%)	
- 8 days (within ICU)	2 (0.3%)	2 (0.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
- Day 10	2 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (5.9%)	
- Day 15	2 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (5.9%)	
- Day 30 after end of study sedation (48 hours) or extubation	2 (0.3%)	2 (0.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
- Day 4	2 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (5.9%)	
- Day 7 after end of study sedation (48 hours) or extubation	2 (0.3%)	2 (0.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	

- 1 week after ICU discharge	1 (0.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (2.9%)	
- Day 29	1 (0.1%)	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Dead patients at the time of outcome assessment (%)	31.5 (22.0 - 41.4) [0.0 - 82.3]	30.8 (22.0 - 40.0) [7.5 - 68.9]	36.3 (20.5 - 42.8) [0.0 - 82.3]	28.9 (23.5 - 38.8) [2.5 - 49.2]	33.5 (18.9 - 39.8) [9.5 - 48.5]	165 (24.0%)
Handling of dead patients						66 (9.6%)
- Survivors only	227 (36.6%)	0 (0.0%)	121 (61.7%)	101 (88.6%)	5 (14.7%)	
- Dead = 0	163 (26.2%)	150 (54.2%)	5 (2.6%)	8 (7.0%)	0 (0.0%)	
- Included in scale/definition	94 (15.1%)	0 (0.0%)	66 (33.7%)	0 (0.0%)	28 (82.4%)	
- Actual number of days/value (no penalization of death)[7]	93 (15.0%)	93 (33.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
- Dead = worst value[8]	29 (4.7%)	23 (8.3%)	1 (0.5%)	5 (4.4%)	0 (0.0%)	
- Death penalized in some cases[9]	4 (0.6%)	4 (1.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
- Percentage of days out of the total number of days alive (not counting days where the patient was dead in either numerator or denominator)	3 (0.5%)	3 (1.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
- All included (no dead patients)	2 (0.3%)	0 (0.0%)	2 (1.0%)	0 (0.0%)	0 (0.0%)	
- Percentage of days out of the total number of days with possible condition (not counting days where the patient was dead or discharged from the ICU in either numerator or denominator)	2 (0.3%)	2 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
- Worst possible point score on days where the patient was dead	2 (0.3%)	2 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
- Dead patients included in analysis without having had the event	1 (0.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (2.9%)	
- Last value before death carried forward	1 (0.2%)	0 (0.0%)	1 (0.5%)	0 (0.0%)	0 (0.0%)	
Missingness of outcome (%)	0.4 (0.0 - 5.0) [0.0 - 53.6]	0.0 (0.0 - 0.5) [0.0 - 24.0]	3.0 (0.0 - 8.9) [0.0 - 53.6]	9.6 (7.3 - 18.9) [0.0 - 52.0]	0.2 (0.0 - 4.0) [0.0 - 12.6]	129 (18.8%)
Handling of missing data						110 (16.0%)
- Complete case analysis	258 (44.7%)	93 (29.0%)	88 (57.9%)	66 (80.5%)	11 (50.0%)	
- Complete case analysis (no missingness)[10]	247 (42.8%)	192 (59.8%)	40 (26.3%)	4 (4.9%)	11 (50.0%)	
- Multiple imputation	26 (4.5%)	7 (2.2%)	13 (8.6%)	6 (7.3%)	0 (0.0%)	

- Last value carried forward	16 (2.8%)	12 (3.7%)	4 (2.6%)	0 (0.0%)	0 (0.0%)
- Interpolation based on measured values and extrapolation using last value carried forward	5 (0.9%)	3 (0.9%)	0 (0.0%)	2 (2.4%)	0 (0.0%)
- Complete case analysis with logical imputation of intermittent missing repeated values	4 (0.7%)	4 (1.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
- Imputation, details unclear	4 (0.7%)	4 (1.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
- Model-based single imputation	4 (0.7%)	0 (0.0%)	4 (2.6%)	0 (0.0%)	0 (0.0%)
- Single imputation, worst-case imputation	4 (0.7%)	4 (1.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
- Complete case analysis and best-worst/worst-best case analysis	3 (0.5%)	1 (0.3%)	2 (1.3%)	0 (0.0%)	0 (0.0%)
- Multiple imputation and complete case analysis and best-worst/worst-best case analysis[11]	3 (0.5%)	0 (0.0%)	0 (0.0%)	3 (3.7%)	0 (0.0%)
- Categorized as "other or unknown" with other values	1 (0.2%)	0 (0.0%)	1 (0.7%)	0 (0.0%)	0 (0.0%)
- Partial single imputation, best case/normal imputation	1 (0.2%)	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
- Survivor average causal effects methods	1 (0.2%)	0 (0.0%)	0 (0.0%)	1 (1.2%)	0 (0.0%)

General outcome-level characteristics for al included trials and stratified by outcome category. Data are presented as *numbers (percentages)* or as *medians (interquartile range) [full range]*.

[1] includes 6 outcomes censored at ICU discharge/only assessed in ICU and 7 outcomes censored at hospital discharge.

[2] includes 8 outcomes planned to be assessed at 12 +/- 4 months after randomization, but actually assessed later than 16 months after randomization in some patients.

[3] includes 6 outcomes censored at ICU discharge.

[4] includes 3 outcomes assessed either at the day of ICU discharge or the next day.

[5] includes 16 outcomes censored at hospital discharge (patients assumed to be alive and without life support after hospital discharge).

[6] includes 1 outcome censored after patients had been off vasopressors for 4 hours.

[7] includes the actual point scores (no penalization of death) for the Delta Sepsis Support index outcome 2 times.

[8] includes 17 outcomes assessed on ordinal scales with death considered a value worse than all actual possible values, 5 time-to-liberation outcomes where dead patients were assigned the worst values, and 1 outcome assessed using a probabilistic index and tie method, with dead considered worse than all other values.

[9] includes 3 outcomes where dead patients were assigned 0 days/hours if they died while on life

support/vasopressors and 1 outcome where the actual number of days was used, except that days were not counted if the patient died within 48 hours of extubation or were intubated again.

[10] includes outcomes where no missing data were mentioned or identified and where complete case analysis was (assumed to be) used.

[11] Includes 2 outcomes where multiple imputation was the primary missing data handling strategy, and 1 outcome where complete case analysis was the primary missing data handling strategy; all with the other strategies as sensitivity analyses.

Abbreviations: HRQoL: health-related quality of life; ICU: intensive care unit.

Supplemental Table 2: All reported outcomes

Outcome	n (total = 687)	% of outcomes in category	% of all outcomes
Days alive without…	337		49.1%
Days alive without (invasive) mechanical ventilation/ventilator-free days	126	37.4%	18.3%
Days alive without circulatory support	49	14.5%	7.1%
Days alive and out of ICU	43	12.8%	6.3%
Days alive without renal replacement therapy	40	11.9%	5.8%
Days alive and out of hospital	18	5.3%	2.6%
Days alive without delirium/coma	10	3.0%	1.5%
Days alive without respiratory support	9	2.7%	1.3%
Days alive without respiratory or circulatory support	8	2.4%	1.2%
Days alive without life support	7	2.1%	1.0%
Days alive without coma	5	1.5%	0.7%
Days alive without delirium	4	1.2%	0.6%
Delta Sepsis Support Index	4	1.2%	0.6%
Days alive without sedation	2	0.6%	0.3%
Dialysis catheter-free days	2	0.6%	0.3%
Hours alive without circulatory support	2	0.6%	0.3%
Time to ICU separation	2	0.6%	0.3%
Time to ventilator separation	2	0.6%	0.3%
Days alive without ECMO	1	0.3%	0.1%
Days alive without HFNO or NIV	1	0.3%	0.1%
Days alive and at home	1	0.3%	0.1%
Time to hospital discharge	1	0.3%	0.1%
Functional/cognitive/neurological	198		28.8%
CPC scale	7	3.5%	1.0%
Modified Rankin Scale score 4-6 (poor outcome)	7	3.5%	1.0%
Barthel Index	6	3.0%	0.9%
Katz Index	6	3.0%	0.9%
MoCA Blind	6	3.0%	0.9%
Alive and at home	5	2.5%	0.7%
CPC scale 1-2 (alive with good neurological outcome)	5	2.5%	0.7%
Medical Research Council muscle strength	5	2.5%	0.7%
WHODAS disability-free survival (<25%)	5	2.5%	0.7%
CPC scale 3-5 (poor outcome)	4	2.0%	0.6%
ICU mobility scale	4	2.0%	0.6%
Medical Research Council muscle strength <48 (ICU-acquired weakness)	4	2.0%	0.6%
SF-36 Physical Function Score	4	2.0%	0.6%
AD8 (cognition)	3	1.5%	0.4%
Difficulty in a daily activity	3	1.5%	0.4%
Disability score	3	1.5%	0.4%
Discharge destination (4 levels)	3	1.5%	0.4%
Galveston Orientation and Amnesia Test <75/100 (post-traumatic amnesia)	3	1.5%	0.4%
Hand grip strength	3	1.5%	0.4%
Impact on daily activities (any of 10 items)	3	1.5%	0.4%
IQCODE short version (cognitive decline)	3	1.5%	0.4%
Location of residence, % home	3	1.5%	0.4%
Modified Rankin Scale score	3	1.5%	0.4%
Return to work	3	1.5%	0.4%
Significant change in work duties	3	1.5%	0.4%

6-minute walking test	2	1.0%	0.3%
Adelaide Activities Profile	2	1.0%	0.3%
Barthel Index ≥85 (independent)	2	1.0%	0.3%
ECOG performance status	2	1.0%	0.3%
Functional independence (patient-reported)	2	1.0%	0.3%
Functional Status Score for the ICU	2	1.0%	0.3%
GOS	2	1.0%	0.3%
GOS-E	2	1.0%	0.3%
GOS-E 6-8 (good outcome)	2	1.0%	0.3%
GOS 4-5 (favorable neurologic outcome)	2	1.0%	0.3%
Lawton IADL	2	1.0%	0.3%
Medical Research Council muscle strength, lower limbs	2	1.0%	0.3%
Physical Function in the ICU Test	2	1.0%	0.3%
Timed up and go test	2	1.0%	0.3%
WHODAS no or mild disability	2	1.0%	0.3%
WHODAS score	2	1.0%	0.3%
Alive and at home off dialysis	1	0.5%	0.1%
Australian Labor Force Survey	1	0.5%	0.1%
Australian Labor Force Survey - returned to work	1	0.5%	0.1%
Clinical Frailty Scale >4	1	0.5%	0.1%
Coding test (translate numbers into figures using an answer key)	1	0.5%	0.1%
Cognitive function on individual level	1	0.5%	0.1%
Composite, days between randomization and death and age-adjusted TICS-T score	1	0.5%	0.1%
CPC scale (best during trial)	1	0.5%	0.1%
CPC scale 2-5	1	0.5%	0.1%
CPC scale 4-5	1	0.5%	0.1%
D-KEFS domains	1	0.5%	0.1%
D-KEFS total	1	0.5%	0.1%
Disability Rating Scale (categorized)	1	0.5%	0.1%
Disability Rating Scale >5	1	0.5%	0.1%
Discharge destination (5 levels)	1	0.5%	0.1%
Discharge destination (6 levels)	1	0.5%	0.1%
Discharge destination (7 levels)	1	0.5%	0.1%
GOS-E 1-4 (death or unfavorable neurological outcome)	1	0.5%	0.1%
GOS-E 5-8 (favorable outcome)	1	0.5%	0.1%
Hayling Sentence completion score	1	0.5%	0.1%
ICU mobility scale or 6-minute walk test	1	0.5%	0.1%
Institutional dependency	1	0.5%	0.1%
Karnofsky performance-status	1	0.5%	0.1%
Location/dead (5 levels)	1	0.5%	0.1%
Location/support (5 levels)	1	0.5%	0.1%
Location/support/dead (7 levels)	1	0.5%	0.1%
Logic thinking - Similarities score	1	0.5%	0.1%
Logical Memory I score	1	0.5%	0.1%
Logical Memory II score	1	0.5%	0.1%
Loss of employment (for patients who were fully or partially employed)	1	0.5%	0.1%
Memory score - Digit Span score	1	0.5%	0.1%
MoCA	1	0.5%	0.1%
MoCA change from baseline	1	0.5%	0.1%
Modified Rankin Scale score ≤2 (alive with good neurological outcome)	1	0.5%	0.1%

Modified Rankin Scale score ≤3 (alive with good neurological outcome)	1	0.5%	0.1%
Modified Rankin Scale score 5-6 (poor outcome)	1	0.5%	0.1%
Modified Rankin Scale score change	1	0.5%	0.1%
Mortality or poor neurological outcome	1	0.5%	0.1%
Muscular Disability Rating Scale (MDRS)	1	0.5%	0.1%
Not in hospital	1	0.5%	0.1%
Pfeffer FAQ (performance score)	1	0.5%	0.1%
Poor outcome	1	0.5%	0.1%
Presence of cognitive impairment based on multiple tests	1	0.5%	0.1%
Proportion of patients in paid employment at baseline who were unemployed at day 180	1	0.5%	0.1%
RAND-36 physical function	1	0.5%	0.1%
RBANS domains	1	0.5%	0.1%
RBANS total	1	0.5%	0.1%
Rey-Osterrieth complex figure test (copy)	1	0.5%	0.1%
Rey-Osterrieth complex figure test (immediate recall, delayed recall, recognition)	1	0.5%	0.1%
Rey auditory verbal learning test, immediate + recall	1	0.5%	0.1%
SF-36 Physical Function Score <20	1	0.5%	0.1%
Sit and stand (repetitions)	1	0.5%	0.1%
Skilled nursing home facility residence	1	0.5%	0.1%
Span of numbers	1	0.5%	0.1%
Stationary walk (repetitions)	1	0.5%	0.1%
TICS	1	0.5%	0.1%
TICS-T total score (age adjusted)	1	0.5%	0.1%
TICS categorized	1	0.5%	0.1%
Trailmaking A	1	0.5%	0.1%
Trailmaking B	1	0.5%	0.1%
Verbal fluency - Controlled Oral Word Association T-score	1	0.5%	0.1%
Verbal fluency test (S-words and animals)	1	0.5%	0.1%
HRQoL	118		17.2%
EQ-5D-5L index	19	16.1%	2.8%
EQ-5D-5L VAS	15	12.7%	2.2%
EQ-5D-3L index	12	10.2%	1.7%
EQ-5D-3L VAS	10	8.5%	1.5%
EQ-5D-5L components	9	7.6%	1.3%
HADS domains	7	5.9%	1.0%
SF-36	5	4.2%	0.7%
EQ-5D-3L components	3	2.5%	0.4%
EQ-5D index (unclear if 3L or 5L)	3	2.5%	0.4%
EQ-5D VAS (unclear if 3L or 5L)	3	2.5%	0.4%
Pain interference (5 levels)	3	2.5%	0.4%
PTSD - IES-R score	3	2.5%	0.4%
Self-rated health (5 levels)	3	2.5%	0.4%
SF-12	3	2.5%	0.4%
Assessment of Quality of Life 4D utility score	2	1.7%	0.3%
EQ-5D-3L index change from baseline	2	1.7%	0.3%
EQ-5D-5L components - "no problems"	2	1.7%	0.3%
PTSD - IES-R score, mean item score ≥1.6	2	1.7%	0.3%
PTSS >45 (post-traumatic stress-like symptoms)	2	1.7%	0.3%
EQ-5D-5L change from baseline	1	0.8%	0.1%
HADS anxiety/depression (≥11)	1	0.8%	0.1%
HADS total	1	0.8%	0.1%

PSS-SR (PTSD)	1	0.8%	0.1%
PSS-SR >18 (PTSD)	1	0.8%	0.1%
PTSD - IES-R score ≥33	1	0.8%	0.1%
PTSD - IES-R score >22	1	0.8%	0.1%
PTSD - IES domains	1	0.8%	0.1%
RAND-36 general health	1	0.8%	0.1%
VAS-A	1	0.8%	0.1%
Ordinal/other	34		4.9%
Clinical improvement (≥1 point on 8-level ordinal scale) or live discharge from hospital	6	17.6%	0.9%
Clinical improvement (8-level ordinal scale)	6	17.6%	0.9%
WHO ordinal scale for COVID-19	5	14.7%	0.7%
Clinical improvement (6-level ordinal scale)	3	8.8%	0.4%
Time to clinical improvement (6 level ordinal scale)	2	5.9%	0.3%
WHO ordinal scale for COVID-19 (only some levels)	2	5.9%	0.3%
Clinical status (5 levels)	1	2.9%	0.1%
Clinical status (6-level ordinal scale adapted from WHO ordinal scale)	1	2.9%	0.1%
Clinical status (6-level ordinal scale adapted from WHO ordinal scale) - cumulative	1	2.9%	0.1%
Clinical status (7-level ordinal scale)	1	2.9%	0.1%
ICU memory tool	1	2.9%	0.1%
ICU memory tool - components yes/no to each	1	2.9%	0.1%
ICU memory tool - number of memories	1	2.9%	0.1%
Modified Brice questionnaire	1	2.9%	0.1%
Progression on Government of India Ministry of Health and Welfare COVID-19 ordinal scale for moderate to severe or severe to death	1	2.9%	0.1%
Time to clinical recovery (7-level ordinal scale)	1	2.9%	0.1%

All outcomes in each category, sorted by frequency. All outcomes included in this table; **Table 2** in the main text is a simplified version of this table only including outcomes reported ≥5 times. **Supplemental Tables 3-5** includes all outcomes in each category in drug, management, and device trials, respectively. Of note, some of the different included outcomes are relatively similar based on the same tools (e.g., using different cut-offs), or sub-components of other included outcomes.

Abbreviations: 4D: 4 Dimensions; AD8: Ascertain Dementia 8-item informant questionnaire; COVID-19: corona-virus disease 2019; CPC: cerebral performance category; D-KEFS: Delis-Kaplan Executive Function System; ECMO: extracorporeal membrane oxygenation; ECOG: Eastern Cooperative Oncology Group; EQ-5D-3L: EuroQol 5-Dimension 3-Level instrument; EQ-5D-5L: EuroQol 5-Dimension 5-Level instrument; FAQ: Functional Activities Questionnaire; GOS: Glasgow Outcome Scale; GOS-E: Glasgow Outcome Scale Extended; HADS: Hospital Anxiety and Depression Scale; HFNO: high-flow nasal oxygen; HRQoL: health-related quality of life; IADL: Instrumental Activities of Daily Living scale; ICU: intensive care unit; IES-R: Impact of Event Scale – Revised; IQCODE: Informant Questionnaire on Cognitive Decline in the Elderly; MoCA: Montreal Cognitive Assessment; NIV: non-invasive ventilation; PSS-SR: PTSD Symptom Scale – Self Report; PTSD: post-traumatic stress disorder; PTSS: Post-Traumatic Symptom Scale; RAND-36: RAND corporation 36-Item Health Survey; RBANS: Repeatable Battery for the Assessment of Neuropsychological Status; SF-12: 12-Item Short Form Survey; SF-36: Short Form 36 Health Survey; TICS: Telephone Interview for Cognitive Status; VAS: visual analogue scale; VAS-A: Visual Analogue Scale for Anxiety; WHO: World Health Organization; WHODAS: World Health Organization Disability Assessment Schedule.

Supplemental Table 3: All reported outcomes in drug trials

Outcome	n (total = 454)	% of outcomes in category	% of all outcomes
Days alive without…	229		50.4%
Days alive without (invasive) mechanical ventilation/ventilator-free days	74	32.3%	16.3%
Days alive without circulatory support	37	16.2%	8.1%
Days alive and out of ICU	28	12.2%	6.2%
Days alive without renal replacement therapy	25	10.9%	5.5%
Days alive and out of hospital	14	6.1%	3.1%
Days alive without delirium/coma	9	3.9%	2.0%
Days alive without respiratory support	8	3.5%	1.8%
Days alive without respiratory or circulatory support	8	3.5%	1.8%
Days alive without life support	7	3.1%	1.5%
Days alive without coma	4	1.7%	0.9%
Days alive without delirium	4	1.7%	0.9%
Delta Sepsis Support Index	4	1.7%	0.9%
Hours without circulatory support	1	0.4%	0.2%
Time to ICU separation	2	0.9%	0.4%
Time to ventilator separation	2	0.9%	0.4%
Days alive without ECMO	1	0.4%	0.2%
Time to hospital discharge	1	0.4%	0.2%
Functional/cognitive/neurological	127		28.0%
CPC scale	5	3.9%	1.1%
Modified Rankin Scale score 4-6 (poor outcome)	3	2.4%	0.7%
Katz Index	4	3.1%	0.9%
MoCA Blind	3	2.4%	0.7%
Alive and at home	5	3.9%	1.1%
CPC scale 1-2 (alive with good neurological outcome)	1	0.8%	0.2%
Medical Research Council muscle strength	2	1.6%	0.4%
WHODAS disability-free survival (<25%)	5	3.9%	1.1%
CPC scale 3-5 (poor outcome)	3	2.4%	0.7%
Medical Research Council muscle strength <48 (ICU-acquired weakness)	2	1.6%	0.4%
SF-36 Physical Function Score	4	3.1%	0.9%
AD8 (cognition)	3	2.4%	0.7%
Difficulty in a daily activity	3	2.4%	0.7%
Disability score	3	2.4%	0.7%
Galveston Orientation and Amnesia Test <75/100 (post-traumatic amnesia)	3	2.4%	0.7%
Hand grip strength	2	1.6%	0.4%
Impact on daily activities (any of 10 items)	3	2.4%	0.7%
IQCODE short version (cognitive decline)	1	0.8%	0.2%
Location of residence, % home	3	2.4%	0.7%
Return to work	3	2.4%	0.7%
Significant change in work duties	3	2.4%	0.7%
6-minute walking test	2	1.6%	0.4%
Adelaide Activities Profile	2	1.6%	0.4%
ECOG performance status	2	1.6%	0.4%
GOS	2	1.6%	0.4%
GOS-E	1	0.8%	0.2%
GOS-E 6-8 (good outcome)	2	1.6%	0.4%
GOS 4-5 (favorable neurologic outcome)	2	1.6%	0.4%
Lawton IADL	1	0.8%	0.2%

Timed up and go test	1	0.8%	0.2%
WHODAS no or mild disability	2	1.6%	0.4%
WHODAS score	2	1.6%	0.4%
Australian Labor Force Survey	1	0.8%	0.2%
Australian Labor Force Survey - returned to work	1	0.8%	0.2%
Coding test (translate numbers into figures using an answer key)	1	0.8%	0.2%
Cognitive function on individual level	1	0.8%	0.2%
Composite, days between randomization and death and age-adjusted TICS-T score	1	0.8%	0.2%
CPC scale 2-5	1	0.8%	0.2%
CPC scale 4-5	1	0.8%	0.2%
D-KEFS domains	1	0.8%	0.2%
D-KEFS total	1	0.8%	0.2%
Discharge destination (5 levels)	1	0.8%	0.2%
GOS-E 1-4 (death or unfavorable neurological outcome)	1	0.8%	0.2%
Hayling Sentence completion score	1	0.8%	0.2%
ICU mobility scale or 6-minute walk test	1	0.8%	0.2%
Institutional dependency	1	0.8%	0.2%
Karnofsky performance-status	1	0.8%	0.2%
Location/dead (5 levels)	1	0.8%	0.2%
Location/support/dead (7 levels)	1	0.8%	0.2%
Logic thinking - Similarities score	1	0.8%	0.2%
Logical Memory I score	1	0.8%	0.2%
Logical Memory II score	1	0.8%	0.2%
Loss of employment (for patients who were fully or partially employed)	1	0.8%	0.2%
Memory score - Digit Span score	1	0.8%	0.2%
MoCA	1	0.8%	0.2%
MoCA change from baseline	1	0.8%	0.2%
Muscular Disability Rating Scale (MDRS)	1	0.8%	0.2%
Pfeffer FAQ (performance score)	1	0.8%	0.2%
Presence of cognitive impairment based on multiple tests	1	0.8%	0.2%
Proportion of patients in paid employment at baseline who were unemployed at day 180	1	0.8%	0.2%
RAND-36 physical function	1	0.8%	0.2%
RBANS domains	1	0.8%	0.2%
RBANS total	1	0.8%	0.2%
Rey-Osterrieth complex figure test (copy)	1	0.8%	0.2%
Rey-Osterrieth complex figure test (immediate recall, delayed recall, recognition)	1	0.8%	0.2%
Rey auditory verbal learning test, immediate + recall	1	0.8%	0.2%
SF-36 Physical Function Score <20	1	0.8%	0.2%
Skilled nursing home facility residence	1	0.8%	0.2%
Span of numbers	1	0.8%	0.2%
TICS	1	0.8%	0.2%
TICS-T total score (age adjusted)	1	0.8%	0.2%
TICS categorized	1	0.8%	0.2%
Trailmaking A	1	0.8%	0.2%
Trailmaking B	1	0.8%	0.2%
Verbal fluency - Controlled Oral Word Association T-score	1	0.8%	0.2%
Verbal fluency test (S-words and animals)	1	0.8%	0.2%
HRQoL	69		15.2%
EQ-5D-5L index	15	21.7%	3.3%
EQ-5D-5L VAS	10	14.5%	2.2%

EQ-5D-3L index	8	11.6%	1.8%
EQ-5D-3L VAS	4	5.8%	0.9%
EQ-5D-5L components	8	11.6%	1.8%
SF-36	2	2.9%	0.4%
EQ-5D-3L components	3	4.3%	0.7%
EQ-5D index (unclear if 3L or 5L)	3	4.3%	0.7%
EQ-5D VAS (unclear if 3L or 5L)	3	4.3%	0.7%
Pain interference (5 levels)	3	4.3%	0.7%
Self-rated health (5 levels)	3	4.3%	0.7%
SF-12	1	1.4%	0.2%
EQ-5D-5L components - "no problems"	2	2.9%	0.4%
PTSS >45 (post-traumatic stress-like symptoms)	2	2.9%	0.4%
EQ-5D-5L change from baseline	1	1.4%	0.2%
RAND-36 general health	1	1.4%	0.2%
Ordinal/other	29		6.4%
Clinical improvement (≥1 point on 8-level ordinal scale) or live discharge from hospital	6	20.7%	1.3%
Clinical improvement (8-level ordinal scale)	6	20.7%	1.3%
WHO ordinal scale for COVID-19	5	17.2%	1.1%
Clinical improvement (6-level ordinal scale)	3	10.3%	0.7%
Time to clinical improvement (6 level ordinal scale)	2	6.9%	0.4%
WHO ordinal scale for COVID-19 (only some levels)	2	6.9%	0.4%
Clinical status (5 levels)	1	3.4%	0.2%
Clinical status (6-level ordinal scale adapted from WHO ordinal scale)	1	3.4%	0.2%
Clinical status (6-level ordinal scale adapted from WHO ordinal scale) - cumulative	1	3.4%	0.2%
Modified Brice questionnaire	1	3.4%	0.2%
Progression on Government of India Ministry of Health and Welfare COVID-19 ordinal scale for moderate to severe or severe to death	1	3.4%	0.2%

All outcomes in each category in drug trials only, sorted in the same order as in **Supplemental Table 2** above (order of frequency across all trials). All outcomes in all trials are included in **Supplemental Table 2**. Of note, some of the different included outcomes are relatively similar based on the same tools (e.g., using different cut-offs), or sub-components of other included outcomes.

Abbreviations: see **Supplemental Table 2** above.

Supplemental Table 4: All reported outcomes in management trials

Outcome	N (total = 189)	% of outcomes in category	% of all outcomes
Days alive without…	74		39.2%
Days alive without (invasive) mechanical ventilation/ventilator-free days	33	44.6%	17.5%
Days alive without circulatory support	10	13.5%	5.3%
Days alive and out of ICU	9	12.2%	4.8%
Days alive without renal replacement therapy	13	17.6%	6.9%
Days alive and out of hospital	4	5.4%	2.1%
Days alive without delirium/coma	1	1.4%	0.5%
Days alive without coma	1	1.4%	0.5%
Days alive without sedation	1	1.4%	0.5%
Dialysis catheter-free days	2	2.7%	1.1%
Functional/cognitive/neurological	64		33.9%
CPC scale	1	1.6%	0.5%
Modified Rankin Scale score 4-6 (poor outcome)	4	6.2%	2.1%
Barthel Index	6	9.4%	3.2%
Katz Index	2	3.1%	1.1%
MoCA Blind	3	4.7%	1.6%
CPC scale 1-2 (alive with good neurological outcome)	4	6.2%	2.1%
Medical Research Council muscle strength	3	4.7%	1.6%
CPC scale 3-5 (poor outcome)	1	1.6%	0.5%
ICU mobility scale	3	4.7%	1.6%
Medical Research Council muscle strength <48 (ICU-acquired weakness)	1	1.6%	0.5%
Discharge destination (4 levels)	3	4.7%	1.6%
Hand grip strength	1	1.6%	0.5%
IQCODE short version (cognitive decline)	2	3.1%	1.1%
Modified Rankin Scale score	3	4.7%	1.6%
Barthel Index ≥85 (independent)	2	3.1%	1.1%
Functional Status Score for the ICU	2	3.1%	1.1%
GOS-E	1	1.6%	0.5%
Lawton IADL	1	1.6%	0.5%
Medical Research Council muscle strength, lower limbs	2	3.1%	1.1%
Physical Function in the ICU Test	2	3.1%	1.1%
Timed up and go test	1	1.6%	0.5%
Alive and at home off dialysis	1	1.6%	0.5%
Clinical Frailty Scale >4	1	1.6%	0.5%
CPC scale (best during trial)	1	1.6%	0.5%
Disability Rating Scale (categorized)	1	1.6%	0.5%
Disability Rating Scale >5	1	1.6%	0.5%
Discharge destination (6 levels)	1	1.6%	0.5%
GOS-E 5-8 (favorable outcome)	1	1.6%	0.5%
Location/support (5 levels)	1	1.6%	0.5%
Modified Rankin Scale score ≤2 (alive with good neurological outcome)	1	1.6%	0.5%
Modified Rankin Scale score ≤3 (alive with good neurological outcome)	1	1.6%	0.5%
Modified Rankin Scale score 5-6 (poor outcome)	1	1.6%	0.5%
Mortality or poor neurological outcome	1	1.6%	0.5%
Not in hospital	1	1.6%	0.5%
Poor outcome	1	1.6%	0.5%
Sit and stand (repetitions)	1	1.6%	0.5%

Stationary walk (repetitions)	1	1.6%	0.5%
HRQoL	48		25.4%
EQ-5D-5L index	4	8.3%	2.1%
EQ-5D-5L VAS	5	10.4%	2.6%
EQ-5D-3L index	4	8.3%	2.1%
EQ-5D-3L VAS	6	12.5%	3.2%
EQ-5D-5L components	1	2.1%	0.5%
HADS domains	7	14.6%	3.7%
SF-36	3	6.2%	1.6%
PTSD - IES-R score	3	6.2%	1.6%
SF-12	2	4.2%	1.1%
Assessment of Quality of Life 4D utility score	2	4.2%	1.1%
EQ-5D-3L index change from baseline	2	4.2%	1.1%
PTSD - IES-R score, mean item score ≥1.6	2	4.2%	1.1%
HADS anxiety/depression (≥11)	1	2.1%	0.5%
HADS total	1	2.1%	0.5%
PSS-SR (PTSD)	1	2.1%	0.5%
PSS-SR >18 (PTSD)	1	2.1%	0.5%
PTSD - IES-R score ≥33	1	2.1%	0.5%
PTSD - IES-R score >22	1	2.1%	0.5%
PTSD - IES domains	1	2.1%	0.5%
Ordinal/other	3		1.6%
ICU memory tool	1	33.3%	0.5%
ICU memory tool - components yes/no to each	1	33.3%	0.5%
ICU memory tool - number of memories	1	33.3%	0.5%

All outcomes in each category in management trials only, sorted in the same order as in **Supplemental Table 2** above (order of frequency across all trials). All outcomes in all trials are included in **Supplemental Table 2**. Of note, some of the different included outcomes are relatively similar based on the same tools (e.g., using different cut-offs), or sub-components of other included outcomes.

Abbreviations: see **Supplemental Table 2** above.

Supplemental Table 5: All reported outcomes in device trials

Outcome	N (total = 44)	% of outcomes in category	% of all outcomes
Days alive without…	34		77.3%
Days alive without (invasive) mechanical ventilation/ventilator-free days	19	55.9%	43.2%
Days alive without circulatory support	2	5.9%	4.5%
Days alive and out of ICU	6	17.6%	13.6%
Days alive without renal replacement therapy	2	5.9%	4.5%
Days alive without respiratory support	1	2.9%	2.3%
Days alive without sedation	1	2.9%	2.3%
Hours without circulatory support	1	2.9%	2.3%
Days alive without HFNO or NIV	1	2.9%	2.3%
Days alive and at home	1	2.9%	2.3%
Functional/cognitive/neurological	7		15.9%
CPC scale	1	14.3%	2.3%
ICU mobility scale	1	14.3%	2.3%
Medical Research Council muscle strength <48 (ICU-acquired weakness)	1	14.3%	2.3%
Functional independence (patient-reported)	2	28.6%	4.5%
Discharge destination (7 levels)	1	14.3%	2.3%
Modified Rankin Scale score change	1	14.3%	2.3%
HRQoL	1		2.3%
VAS-A	1	100.0%	2.3%
Ordinal/other	2		4.5%
Clinical status (7-level ordinal scale)	1	50.0%	2.3%
Time to clinical recovery (7-level ordinal scale)	1	50.0%	2.3%

All outcomes in each category in device trials only, sorted in the same order as in **Supplemental Table 2** above (order of frequency across all trials). All outcomes in all trials are included in **Supplemental Table 2**. Of note, some of the different included outcomes are relatively similar based on the same tools (e.g., using different cut-offs), or sub-components of other included outcomes.

Abbreviations: see **Supplemental Table 2** above.

Supplemental Table 6: Effect measures used in drug trials

Effect measure	All outcomes (n = 454)	Days alive without... (n = 229)	Functional/ cognitive/ neurological (n = 127)	HRQoL (n = 69)	Ordinal/other (n = 29)
Mean difference	115 (25.3%)	68 (29.7%)	20 (15.7%)	27 (39.1%)	0 (0.0%)
Odds ratio	69 (15.2%)	24 (10.5%)	23 (18.1%)	0 (0.0%)	22 (75.9%)
Median difference	39 (8.6%)	33 (14.4%)	2 (1.6%)	4 (5.8%)	0 (0.0%)
Risk difference	36 (7.9%)	0 (0.0%)	31 (24.4%)	3 (4.3%)	2 (6.9%)
Hodges-Lehmann estimate	43 (9.5%)	30 (13.1%)	5 (3.9%)	6 (8.7%)	2 (6.9%)
Risk ratio	9 (2.0%)	0 (0.0%)	7 (5.5%)	2 (2.9%)	0 (0.0%)
Hazard ratio	6 (1.3%)	3 (1.3%)	1 (0.8%)	0 (0.0%)	2 (6.9%)
Incidence rate ratio/expected count ratio	8 (1.8%)	8 (3.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Probabilistic index	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Ratio of means	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive plus p-value only[1]	116 (25.6%)	66 (28.8%)	31 (24.4%)	18 (26.1%)	1 (3.4%)
Descriptive only[1]	39 (8.6%)	11 (4.8%)	17 (13.4%)	11 (15.9%)	0 (0.0%)
Descriptive for the entire trial only[1]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Unclear[2]	5 (1.1%)	5 (2.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Effect measured used for all outcomes and stratified by outcome type in drug trials only; effect measures are displayed in the same order as in **Table 3** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%.

[1] Outcomes only counted in one of these categories if no effect measure quantifying the actual differences between intervention arms was presented; thus, descriptive data (for each intervention)/p-values are not counted for outcomes using actual measures of differences. Descriptive data only includes graphical presentation of raw data.

[2] Outcomes where some effect measure quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this effect measure was.

Supplemental Table 7: Effect measures used in management trials

Effect measure	All outcomes (n = 189)	Days alive without... (n = 74)	Functional/ cognitive/ neurological (n = 64)	HRQoL (n = 48)	Ordinal/other (n = 3)
Mean difference	55 (29.1%)	27 (36.5%)	9 (14.1%)	18 (37.5%)	1 (33.3%)
Odds ratio	14 (7.4%)	8 (10.8%)	5 (7.8%)	1 (2.1%)	0 (0.0%)
Median difference	11 (5.8%)	6 (8.1%)	4 (6.2%)	1 (2.1%)	0 (0.0%)
Risk difference	13 (6.9%)	3 (4.1%)	7 (10.9%)	2 (4.2%)	1 (33.3%)
Hodges-Lehmann estimate	4 (2.1%)	4 (5.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Risk ratio	9 (4.8%)	1 (1.4%)	8 (12.5%)	0 (0.0%)	0 (0.0%)
Hazard ratio	2 (1.1%)	0 (0.0%)	2 (3.1%)	0 (0.0%)	0 (0.0%)
Incidence rate ratio/expected count ratio	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Probabilistic index	1 (0.5%)	1 (1.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Ratio of means	1 (0.5%)	1 (1.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive plus p-value only[1]	73 (38.6%)	27 (36.5%)	22 (34.4%)	24 (50.0%)	0 (0.0%)
Descriptive only[1]	14 (7.4%)	1 (1.4%)	11 (17.2%)	2 (4.2%)	0 (0.0%)
Descriptive for the entire trial only[1]	1 (0.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (33.3%)
Unclear[2]	4 (2.1%)	2 (2.7%)	2 (3.1%)	0 (0.0%)	0 (0.0%)
Unclear[2]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Effect measured used for all outcomes and stratified by outcome type in management trials only; effect measures are displayed in the same order as in **Table 3** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%.

[1] Outcomes only counted in one of these categories if no effect measure quantifying the actual differences between intervention arms was presented; thus, descriptive data (for each intervention)/p-values are not counted for outcomes using actual measures of differences. Descriptive data only includes graphical presentation of raw data.

[2] Outcomes where some effect measure quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this effect measure was.

Supplemental Table 8: Effect measures used in device trials

Effect measure	All outcomes (n = 44)	Days alive without... (n = 34)	Functional cognitive/ neurological (n = 7)	HRQoL (n = 1)	Ordinal/other (n = 2)
Mean difference	13 (29.5%)	12 (35.3%)	0 (0.0%)	1 (100.0%)	0 (0.0%)
Odds ratio	3 (6.8%)	2 (5.9%)	0 (0.0%)	0 (0.0%)	1 (50.0%)
Median difference	9 (20.5%)	9 (26.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Risk difference	1 (2.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (50.0%)
Hodges-Lehmann estimate	1 (2.3%)	1 (2.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Risk ratio	2 (4.5%)	0 (0.0%)	2 (28.6%)	0 (0.0%)	0 (0.0%)
Hazard ratio	1 (2.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (50.0%)
Incidence rate ratio/expected count ratio	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Probabilistic index	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Ratio of means	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive plus p-value only[1]	14 (31.8%)	9 (26.5%)	5 (71.4%)	0 (0.0%)	0 (0.0%)
Descriptive only[1]	3 (6.8%)	3 (8.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive for the entire trial only[1]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Unclear[2]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Effect measured used for all outcomes and stratified by outcome type in device trials only; effect measures are displayed in the same order as in **Table 3** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%.

[1] Outcomes only counted in one of these categories if no effect measure quantifying the actual differences between intervention arms was presented; thus, descriptive data (for each intervention)/p-values are not counted for outcomes using actual measures of differences. Descriptive data only includes graphical presentation of raw data.

[2] Outcomes where some effect measure quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this effect measure was.

Supplemental Table 9: Statistical methods used in drug trials

Statistical method	All outcomes (n = 454)	Days alive without... (n = 229)	Functional/ cognitive/ neurological (n = 127)	HRQoL (n = 69)	Ordinal/other (n = 29)
Non-parametric[1]	123 (27.1%)	101 (44.1%)	9 (7.1%)	11 (15.9%)	2 (6.9%)
Linear regression/ T-test[2]	129 (28.4%)	71 (31.0%)	27 (21.3%)	31 (44.9%)	0 (0.0%)
Chi-squared[3]	52 (11.5%)	1 (0.4%)	37 (29.1%)	8 (11.6%)	6 (20.7%)
Proportional odds/cumulative logistic regression	51 (11.2%)	24 (10.5%)	9 (7.1%)	0 (0.0%)	18 (62.1%)
Hodges-Lehmann estimator	43 (9.5%)	30 (13.1%)	5 (3.9%)	6 (8.7%)	2 (6.9%)
Bootstrapping	30 (6.6%)	25 (10.9%)	0 (0.0%)	5 (7.2%)	0 (0.0%)
Logistic regression	17 (3.7%)	0 (0.0%)	16 (12.6%)	0 (0.0%)	1 (3.4%)
Quantile regression	24 (5.3%)	20 (8.7%)	0 (0.0%)	4 (5.8%)	0 (0.0%)
Generalizing estimating equation with identity link	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Time-to-event analysis[4]	10 (2.2%)	7 (3.1%)	1 (0.8%)	0 (0.0%)	2 (6.9%)
Generalized linear (mixed) model, log-binomial	8 (1.8%)	0 (0.0%)	6 (4.7%)	2 (2.9%)	0 (0.0%)
Kryger-Jensen/Lange test	8 (1.8%)	4 (1.7%)	0 (0.0%)	4 (5.8%)	0 (0.0%)
Hurdle-negative binomial regression	6 (1.3%)	6 (2.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Poisson/negative binomial regression	2 (0.4%)	2 (0.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Binomial proportion test/exact binomial confidence interval	1 (0.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (3.4%)
Simple calculation	3 (0.7%)	0 (0.0%)	3 (2.4%)	0 (0.0%)	0 (0.0%)
Zero-one-inflated beta regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Beta-binomial regression	2 (0.4%)	2 (0.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized additive model with zero-inflated/zero-one-inflated beta distribution	1 (0.2%)	1 (0.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with zero-one-inflated beta distribution	2 (0.4%)	2 (0.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized linear model, identity-binomial	2 (0.4%)	0 (0.0%)	1 (0.8%)	0 (0.0%)	1 (3.4%)
Generalized linear (mixed) model, logit link and G-computation	1 (0.2%)	1 (0.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with logit link	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Probabilistic index	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive only[5]	39 (8.6%)	11 (4.8%)	17 (13.4%)	11 (15.9%)	0 (0.0%)
Descriptive for the entire trial only[5]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Unclear[6]	9 (2.0%)	7 (3.1%)	1 (0.8%)	1 (1.4%)	0 (0.0%)

Statistical methods used for all outcomes and stratified by outcome type in drug trials only; statistical methods are displayed in the same order as in **Table 4** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%. Frequentist/Bayesian variants of the same statistical methods summarised together.

[1] Includes all non-parametric tests of continuous data, e.g. Mann-Whitney U/Wilcoxon, van Elteren, Kruskal-Wallis, and the Wilcoxon rank sum test for clustered data using Rosner-Glynn-Lee method.

[2] Includes linear regression and all variants including T-tests, analysis of (co)-variance, linear mixed models, etc., also including these methods used on transformed data.

[3] Includes the Chi-squared test and related tests for contingency tables, e.g., Fisher's exact test, and the Cochran-Armitage test for trend.

[4] includes the Cox proportional hazards model, the Kaplan Meier estimator, the log-rank test, the Fine-Gray model, and unspecified survival analysis methods.

[5] Outcomes only counted in one of these categories if no inferential statistical method was used to compare intervention groups, i.e., if only descriptive data for each intervention group or for the full trial were presented.

[6] Outcomes where some statistical method quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this method was.

Supplemental Table 10: Statistical methods used in management trials

Statistical method	All outcomes (n = 189)	Days alive without... (n = 74)	Functional/ cognitive/ neurological (n = 64)	HRQoL (n = 48)	Ordinal/other (n = 3)
Non-parametric[1]	72 (38.1%)	36 (48.6%)	20 (31.2%)	15 (31.2%)	1 (33.3%)
Linear regression/ T-test[2]	49 (25.9%)	18 (24.3%)	10 (15.6%)	21 (43.8%)	0 (0.0%)
Chi-squared[3]	27 (14.3%)	0 (0.0%)	18 (28.1%)	8 (16.7%)	1 (33.3%)
Proportional odds/cumulative logistic regression	9 (4.8%)	8 (10.8%)	1 (1.6%)	0 (0.0%)	0 (0.0%)
Hodges-Lehmann estimator	4 (2.1%)	4 (5.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Bootstrapping	7 (3.7%)	3 (4.1%)	4 (6.2%)	0 (0.0%)	0 (0.0%)
Logistic regression	9 (4.8%)	0 (0.0%)	6 (9.4%)	2 (4.2%)	1 (33.3%)
Quantile regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with identity link	12 (6.3%)	9 (12.2%)	0 (0.0%)	3 (6.2%)	0 (0.0%)
Time-to-event analysis[4]	3 (1.6%)	0 (0.0%)	3 (4.7%)	0 (0.0%)	0 (0.0%)
Generalized linear (mixed) model, log-binomial	1 (0.5%)	0 (0.0%)	1 (1.6%)	0 (0.0%)	0 (0.0%)
Kryger-Jensen/Lange test	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Hurdle-negative binomial regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Poisson/negative binomial regression	1 (0.5%)	1 (1.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Binomial proportion test/exact binomial confidence interval	3 (1.6%)	0 (0.0%)	0 (0.0%)	2 (4.2%)	1 (33.3%)
Simple calculation	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Zero-one-inflated beta regression	3 (1.6%)	3 (4.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Beta-binomial regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized additive model with zero-inflated/zero-one-inflated beta distribution	1 (0.5%)	1 (1.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with zero-one-inflated beta distribution	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized linear model, identity-binomial	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized linear (mixed) model, logit link and G-computation	1 (0.5%)	0 (0.0%)	1 (1.6%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with logit link	1 (0.5%)	0 (0.0%)	0 (0.0%)	1 (2.1%)	0 (0.0%)
Probabilistic index	1 (0.5%)	1 (1.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive only[5]	14 (7.4%)	1 (1.4%)	11 (17.2%)	2 (4.2%)	0 (0.0%)
Descriptive for the entire trial only[5]	1 (0.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (33.3%)
Unclear[6]	6 (3.2%)	4 (5.4%)	2 (3.1%)	0 (0.0%)	0 (0.0%)

Statistical methods used for all outcomes and stratified by outcome type in management trials only; statistical methods are displayed in the same order as in **Table 4** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%. Frequentist/Bayesian variants of the same statistical methods summarised together.

[1] Includes all non-parametric tests of continuous data, e.g. Mann-Whitney U/Wilcoxon, van Elteren, Kruskal-Wallis, and the Wilcoxon rank sum test for clustered data using Rosner-Glynn-Lee method.

[2] Includes linear regression and all variants including T-tests, analysis of (co)-variance, linear mixed models, etc., also including these methods used on transformed data.

[3] Includes the Chi-squared test and related tests for contingency tables, e.g., Fisher's exact test, and the Cochran-Armitage test for trend.

[4] includes the Cox proportional hazards model, the Kaplan Meier estimator, the log-rank test, the Fine-Gray model, and unspecified survival analysis methods.

[5] Outcomes only counted in one of these categories if no inferential statistical method was used to compare intervention groups, i.e., if only descriptive data for each intervention group or for the full trial were presented.

[6] Outcomes where some statistical method quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this method was.

Supplemental Digital Content

Supplemental Table 11: Statistical methods used in device trials

Statistical method	All outcomes (n = 44)	Days alive without... (n = 34)	Functional/ cognitive/ neurological (n = 7)	HRQoL (n = 1)	Ordinal/other (n = 2)
Non-parametric[1]	16 (36.4%)	15 (44.1%)	1 (14.3%)	0 (0.0%)	0 (0.0%)
Linear regression/ T-test[2]	7 (15.9%)	6 (17.6%)	0 (0.0%)	1 (100.0%)	0 (0.0%)
Chi-squared[3]	6 (13.6%)	0 (0.0%)	6 (85.7%)	0 (0.0%)	0 (0.0%)
Proportional odds/cumulative logistic regression	3 (6.8%)	2 (5.9%)	0 (0.0%)	0 (0.0%)	1 (50.0%)
Hodges-Lehmann estimator	1 (2.3%)	1 (2.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Bootstrapping	5 (11.4%)	4 (11.8%)	0 (0.0%)	0 (0.0%)	1 (50.0%)
Logistic regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Quantile regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with identity link	2 (4.5%)	2 (5.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Time-to-event analysis[4]	1 (2.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (50.0%)
Generalized linear (mixed) model, log-binomial	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Kryger-Jensen/Lange test	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Hurdle-negative binomial regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Poisson/negative binomial regression	2 (4.5%)	0 (0.0%)	2 (28.6%)	0 (0.0%)	0 (0.0%)
Binomial proportion test/exact binomial confidence interval	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Simple calculation	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Zero-one-inflated beta regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Beta-binomial regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized additive model with zero-inflated/zero-one-inflated beta distribution	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with zero-one-inflated beta distribution	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized linear model, identity-binomial	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized linear (mixed) model, logit link and G-computation	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with logit link	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Probabilistic index	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive only[5]	3 (6.8%)	3 (8.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive for the entire trial only[5]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Unclear[6]	6 (13.6%)	6 (17.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Statistical methods used for all outcomes and stratified by outcome type in device trials only; statistical methods are displayed in the same order as in **Table 4** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%. Frequentist/Bayesian variants of the same statistical methods summarised together.

[1] Includes all non-parametric tests of continuous data, e.g. Mann-Whitney U/Wilcoxon, van Elteren, Kruskal-Wallis, and the Wilcoxon rank sum test for clustered data using Rosner-Glynn-Lee method.

[2] Includes linear regression and all variants including T-tests, analysis of (co)-variance, linear mixed models, etc., also including these methods used on transformed data.

[3] Includes the Chi-squared test and related tests for contingency tables, e.g., Fisher's exact test, and the Cochran-Armitage test for trend.

[4] includes the Cox proportional hazards model, the Kaplan Meier estimator, the log-rank test, the Fine-Gray model, and unspecified survival analysis methods.

[5] Outcomes only counted in one of these categories if no inferential statistical method was used to compare intervention groups, i.e., if only descriptive data for each intervention group or for the full trial were presented.

[6] Outcomes where some statistical method quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this method was.

31

Supplemental Table 12: General trial-level characteristics by period

Variable	Published before 2020 (n = 76)	Published in 2020 or later (n = 91)
Intervention type		
- Drug	43 (56.6%)	54 (59.3%)
- Management	27 (35.5%)	24 (26.4%)
- Device	6 (7.9%)	13 (14.3%)
Journal		
- JAMA	20 (26.3%)	26 (28.6%)
- New England Journal of Medicine	25 (32.9%)	11 (12.1%)
- Intensive Care Medicine	8 (10.5%)	16 (17.6%)
- Critical Care Medicine	9 (11.8%)	7 (7.7%)
- The Lancet Respiratory Medicine	5 (6.6%)	11 (12.1%)
- Critical Care	4 (5.3%)	7 (7.7%)
- American Journal of Respiratory and Critical Care Medicine	1 (1.3%)	5 (5.5%)
- Chest	1 (1.3%)	5 (5.5%)
- The Lancet	2 (2.6%)	1 (1.1%)
- BMJ	0 (0.0%)	1 (1.1%)
- Circulation	1 (1.3%)	0 (0.0%)
- JAMA Neurology	0 (0.0%)	1 (1.1%)
Year	2018 (2018 - 2019) [2006 - 2019]	2021 (2020 - 2021) [2020 - 2022]
Countries (no.)	1 (1 - 2) [1 - 12]	1 (1 - 2) [1 - 15]
Centers (no.)	18 (5 - 34) [1 - 74]	12 (5 - 30) [1 - 393]
Patients (no.)	410 (158 - 953) [47 - 15802]	299 (140 - 666) [20 - 11052]
Outcomes included (no.)	3 (2 - 5) [1 - 43]	3 (1 - 5) [1 - 15]
Full protocol referenced and available	60 (78.9%)	74 (81.3%)
Any (co-)primary patient-important outcome other than mortality included	18 (23.7%)	36 (39.6%)

General trial-level characteristics for al included trials and stratified by period of publication of the primary trial report. Data are presented as *numbers (percentages)* or as *medians (interquartile range) [full range]*. Further details and abbreviations are presented in the **Table 1** legend in the primary text, which also includes a column with data for all trials.

Supplemental Table 13: General outcome-level characteristics by period

Variable	Published before 2020 (n = 325)	Published in 2020 or later (n = 362)	Missing values
Primary/ co-primary outcome	20 (6.2%)	36 (9.9%)	0 (0.0%)
Prioritization of outcome components			0 (0.0%)
- Not relevant	294 (90.5%)	331 (91.4%)	
- Equal prioritization	29 (8.9%)	14 (3.9%)	
- Not relevant (individual categories/ domains/ components reported secondarily, but not considered or analyzed as primary outcomes)	1 (0.3%)	17 (4.7%)	
- Three dichotomizations reported with equal prioritization	1 (0.3%)	0 (0.0%)	
Assessment time point			2 (0.3%)
- Day 28	103 (31.9%)	121 (33.4%)	
- Day 180/6 months	75 (23.2%)	58 (16.0%)	
- Day 90/3 months	38 (11.8%)	49 (13.5%)	
- 1 year	27 (8.4%)	10 (2.8%)	
- Day 14	4 (1.2%)	23 (6.4%)	
- Day 60	10 (3.1%)	17 (4.7%)	
- Day 30	16 (5.0%)	10 (2.8%)	
- Hospital discharge/in-hospital	12 (3.7%)	14 (3.9%)	
- ICU discharge	11 (3.4%)	12 (3.3%)	
- 90 days/3 months after ICU discharge	13 (4.0%)	9 (2.5%)	
- Day 21	0 (0.0%)	19 (5.2%)	
- Day 7	5 (1.5%)	6 (1.7%)	
- 30 days after ICU discharge	4 (1.2%)	0 (0.0%)	
- Day 25	0 (0.0%)	3 (0.8%)	
- 6 months after hospital discharge	2 (0.6%)	0 (0.0%)	
- 8 days (within ICU)	2 (0.6%)	0 (0.0%)	
- Day 10	0 (0.0%)	2 (0.6%)	
- Day 15	0 (0.0%)	2 (0.6%)	
- Day 30 after end of study sedation (48 hours) or extubation	0 (0.0%)	2 (0.6%)	
- Day 4	0 (0.0%)	2 (0.6%)	
- Day 7 after end of study sedation (48 hours) or extubation	0 (0.0%)	2 (0.6%)	
- 1 week after ICU discharge	1 (0.3%)	0 (0.0%)	
- Day 29	0 (0.0%)	1 (0.3%)	
Dead patients at the time of outcome assessment (%)	31.7 (24.0 - 40.6) [0.0 - 82.3]	30.4 (19.8 - 42.4) [2.5 - 80.8]	165 (24.0%)
Handling of dead patients			66 (9.6%)
- Survivors only	124 (41.3%)	103 (32.1%)	
- Dead = 0	76 (25.3%)	87 (27.1%)	
- Included in scale/definition	36 (12.0%)	58 (18.1%)	
- Actual number of days/value (no penalization of death)	50 (16.7%)	43 (13.4%)	
- Dead = worst value	9 (3.0%)	20 (6.2%)	
- Death penalized in some cases	0 (0.0%)	4 (1.2%)	
- Percentage of days out of the total number of days alive (not counting days where the patient was dead in either numerator or denominator)	1 (0.3%)	2 (0.6%)	
- All included (no dead patients)	2 (0.7%)	0 (0.0%)	
- Percentage of days out of the total number of days with possible condition (not counting days where the patient	0 (0.0%)	2 (0.6%)	

was dead or discharged from the ICU in either numerator or denominator)			
- Worst possible point score on days where the patient was dead	0 (0.0%)	2 (0.6%)	
- Dead patients included in analysis without having had the event	1 (0.3%)	0 (0.0%)	
- Last value before death carried forward	1 (0.3%)	0 (0.0%)	
Missingness of outcome (%)	0.2 (0.0 - 8.5) [0.0 - 53.6]	0.5 (0.0 - 4.8) [0.0 - 52.0]	129 (18.8%)
Handling of missing data			110 (16.0%)
- Complete case analysis	118 (45.0%)	140 (44.4%)	
- Complete case analysis (no missingness)	122 (46.6%)	125 (39.7%)	
- Multiple imputation	3 (1.1%)	23 (7.3%)	
- Last value carried forward	4 (1.5%)	12 (3.8%)	
- Interpolation based on measured values and extrapolation using last value carried forward	5 (1.9%)	0 (0.0%)	
- Complete case analysis with logical imputation of intermittent missing repeated values	0 (0.0%)	4 (1.3%)	
- Imputation, details unclear	1 (0.4%)	3 (1.0%)	
- Model-based single imputation	4 (1.5%)	0 (0.0%)	
- Single imputation, worst-case imputation	3 (1.1%)	1 (0.3%)	
- Complete case analysis and best-worst/worst-best case analysis	0 (0.0%)	3 (1.0%)	
- Multiple imputation and complete case analysis and best-worst/worst-best case analysis	0 (0.0%)	3 (1.0%)	
- Categorized as "other or unknown" with other values	1 (0.4%)	0 (0.0%)	
- Partial single imputation, best case/normal imputation	0 (0.0%)	1 (0.3%)	
- Survivor average causal effects methods	1 (0.4%)	0 (0.0%)	

General outcome-level characteristics for all included trials stratified by period of publication of the primary trial report. Data are presented as *numbers (percentages)* or as *medians (interquartile range) [full range]*. Further details and abbreviations are presented in the **Supplemental Table S1** legend, which also includes a column with data for all trials. Categories are presented in the same order as in **Supplemental Table S1** to ease comparison.

Supplemental Table 14: All reported outcomes in trials published before 2020

Outcome	n (total = 325)	% of outcomes in category	% of all outcomes
Days alive without…	147		45.2%
Days alive without (invasive) mechanical ventilation/ventilator-free days	55	37.4%	16.9%
Days alive without circulatory support	25	17.0%	7.7%
Days alive and out of ICU	21	14.3%	6.5%
Days alive without renal replacement therapy	17	11.6%	5.2%
Days alive and out of hospital	8	5.4%	2.5%
Days alive without delirium/coma	5	3.4%	1.5%
Days alive without respiratory or circulatory support	1	0.7%	0.3%
Days alive without life support	1	0.7%	0.3%
Days alive without coma	4	2.7%	1.2%
Days alive without delirium	2	1.4%	0.6%
Days alive without sedation	2	1.4%	0.6%
Dialysis catheter-free days	1	0.7%	0.3%
Hours without circulatory support	1	0.7%	0.3%
Time to ICU separation	1	0.7%	0.3%
Time to ventilator separation	1	0.7%	0.3%
Days alive and at home	1	0.7%	0.3%
Time to hospital discharge	1	0.7%	0.3%
Functional/cognitive/neurological	104		32.0%
CPC scale	5	4.8%	1.5%
Modified Rankin Scale score 4-6 (poor outcome)	4	3.8%	1.2%
Barthel Index	2	1.9%	0.6%
Katz Index	2	1.9%	0.6%
MoCA Blind	3	2.9%	0.9%
Alive and at home	1	1.0%	0.3%
CPC scale 1-2 (alive with good neurological outcome)	3	2.9%	0.9%
Medical Research Council muscle strength	3	2.9%	0.9%
WHODAS disability-free survival (<25%)	2	1.9%	0.6%
CPC scale 3-5 (poor outcome)	4	3.8%	1.2%
ICU mobility scale	2	1.9%	0.6%
Medical Research Council muscle strength <48 (ICU-acquired weakness)	3	2.9%	0.9%
AD8 (cognition)	3	2.9%	0.9%
Difficulty in a daily activity	3	2.9%	0.9%
Disability score	3	2.9%	0.9%
Discharge destination (4 levels)	1	1.0%	0.3%
Hand grip strength	3	2.9%	0.9%
Impact on daily activities (any of 10 items)	3	2.9%	0.9%
IQCODE short version (cognitive decline)	1	1.0%	0.3%
Location of residence, % home	3	2.9%	0.9%
Modified Rankin Scale score	1	1.0%	0.3%
Return to work	3	2.9%	0.9%
Significant change in work duties	3	2.9%	0.9%
Adelaide Activities Profile	2	1.9%	0.6%
Barthel Index ≥85 (independent)	2	1.9%	0.6%
ECOG performance status	2	1.9%	0.6%
Functional independence (patient-reported)	2	1.9%	0.6%
GOS	2	1.9%	0.6%
GOS-E	1	1.0%	0.3%
GOS 4-5 (favorable neurologic outcome)	2	1.9%	0.6%

Lawton IADL	1	1.0%	0.3%
Timed up and go test	2	1.9%	0.6%
WHODAS no or mild disability	2	1.9%	0.6%
WHODAS score	2	1.9%	0.6%
Alive and at home off dialysis	1	1.0%	0.3%
Australian Labor Force Survey	1	1.0%	0.3%
Australian Labor Force Survey - returned to work	1	1.0%	0.3%
CPC scale (best during trial)	1	1.0%	0.3%
D-KEFS domains	1	1.0%	0.3%
D-KEFS total	1	1.0%	0.3%
Discharge destination (6 levels)	1	1.0%	0.3%
Discharge destination (7 levels)	1	1.0%	0.3%
GOS-E 5-8 (favorable outcome)	1	1.0%	0.3%
ICU mobility scale or 6-minute walk test	1	1.0%	0.3%
Institutional dependency	1	1.0%	0.3%
Loss of employment (for patients who were fully or partially employed)	1	1.0%	0.3%
MoCA	1	1.0%	0.3%
MoCA change from baseline	1	1.0%	0.3%
Modified Rankin Scale score ≤2 (alive with good neurological outcome)	1	1.0%	0.3%
Modified Rankin Scale score ≤3 (alive with good neurological outcome)	1	1.0%	0.3%
Muscular Disability Rating Scale (MDRS)	1	1.0%	0.3%
RAND-36 physical function	1	1.0%	0.3%
RBANS domains	1	1.0%	0.3%
RBANS total	1	1.0%	0.3%
Sit and stand (repetitions)	1	1.0%	0.3%
Skilled nursing home facility residence	1	1.0%	0.3%
Stationary walk (repetitions)	1	1.0%	0.3%
HRQoL	70		21.5%
EQ-5D-5L index	10	14.3%	3.1%
EQ-5D-5L VAS	4	5.7%	1.2%
EQ-5D-3L index	7	10.0%	2.2%
EQ-5D-3L VAS	5	7.1%	1.5%
EQ-5D-5L components	1	1.4%	0.3%
HADS domains	7	10.0%	2.2%
SF-36	3	4.3%	0.9%
EQ-5D-3L components	1	1.4%	0.3%
EQ-5D index (unclear if 3L or 5L)	2	2.9%	0.6%
Pain interference (5 levels)	3	4.3%	0.9%
PTSD - IES-R score	3	4.3%	0.9%
Self-rated health (5 levels)	3	4.3%	0.9%
SF-12	3	4.3%	0.9%
Assessment of Quality of Life 4D utility score	2	2.9%	0.6%
EQ-5D-3L index change from baseline	2	2.9%	0.6%
EQ-5D-5L components - "no problems"	2	2.9%	0.6%
PTSD - IES-R score, mean item score ≥1.6	2	2.9%	0.6%
PTSS >45 (post-traumatic stress-like symptoms)	2	2.9%	0.6%
EQ-5D-5L change from baseline	1	1.4%	0.3%
HADS total	1	1.4%	0.3%
PSS-SR (PTSD)	1	1.4%	0.3%
PSS-SR >18 (PTSD)	1	1.4%	0.3%
PTSD - IES-R score >22	1	1.4%	0.3%

PTSD - IES domains	1	1.4%	0.3%
RAND-36 general health	1	1.4%	0.3%
VAS-A	1	1.4%	0.3%
Ordinal/other	4		1.2%
ICU memory tool	1	25.0%	0.3%
ICU memory tool - components yes/no to each	1	25.0%	0.3%
ICU memory tool - number of memories	1	25.0%	0.3%
Modified Brice questionnaire	1	25.0%	0.3%

All outcomes in each category in trials with the primary trial report published before 2020, sorted by frequency. All outcomes in all trials are included in **Supplemental Table 2**. Of note, some of the different included outcomes are relatively similar based on the same tools (e.g., using different cut-offs), or sub-components of other included outcomes.

Abbreviations: see **Supplemental Table 2** above.

Supplemental Digital Content

Supplemental Table 15: All reported outcomes in trials published in 2020 or later

Outcome	n (total = 362)	% of outcomes in category	% of all outcomes
Days alive without…	190		52.5%
Days alive without (invasive) mechanical ventilation/ventilator-free days	71	37.4%	19.6%
Days alive without circulatory support	24	12.6%	6.6%
Days alive and out of ICU	22	11.6%	6.1%
Days alive without renal replacement therapy	23	12.1%	6.4%
Days alive and out of hospital	10	5.3%	2.8%
Days alive without delirium/coma	5	2.6%	1.4%
Days alive without respiratory support	9	4.7%	2.5%
Days alive without respiratory or circulatory support	7	3.7%	1.9%
Days alive without life support	6	3.2%	1.7%
Days alive without coma	1	0.5%	0.3%
Days alive without delirium	2	1.1%	0.6%
Delta Sepsis Support Index	4	2.1%	1.1%
Dialysis catheter-free days	1	0.5%	0.3%
Hours without circulatory support	1	0.5%	0.3%
Time to ICU separation	1	0.5%	0.3%
Time to ventilator separation	1	0.5%	0.3%
Days alive without ECMO	1	0.5%	0.3%
Days alive without HFNO or NIV	1	0.5%	0.3%
Functional/cognitive/neurological	94		26.0%
CPC scale	2	2.1%	0.6%
Modified Rankin Scale score 4-6 (poor outcome)	3	3.2%	0.8%
Barthel Index	4	4.3%	1.1%
Katz Index	4	4.3%	1.1%
MoCA Blind	3	3.2%	0.8%
Alive and at home	4	4.3%	1.1%
CPC scale 1-2 (alive with good neurological outcome)	2	2.1%	0.6%
Medical Research Council muscle strength	2	2.1%	0.6%
WHODAS disability-free survival (<25%)	3	3.2%	0.8%
ICU mobility scale	2	2.1%	0.6%
Medical Research Council muscle strength <48 (ICU-acquired weakness)	1	1.1%	0.3%
SF-36 Physical Function Score	4	4.3%	1.1%
Discharge destination (4 levels)	2	2.1%	0.6%
Galveston Orientation and Amnesia Test <75/100 (post-traumatic amnesia)	3	3.2%	0.8%
IQCODE short version (cognitive decline)	2	2.1%	0.6%
Modified Rankin Scale score	2	2.1%	0.6%
6-minute walking test	2	2.1%	0.6%
Functional Status Score for the ICU	2	2.1%	0.6%
GOS-E	1	1.1%	0.3%
GOS-E 6-8 (good outcome)	2	2.1%	0.6%
Lawton IADL	1	1.1%	0.3%
Medical Research Council muscle strength, lower limbs	2	2.1%	0.6%
Physical Function in the ICU Test	2	2.1%	0.6%
Clinical Frailty Scale >4	1	1.1%	0.3%
Coding test (translate numbers into figures using an answer key)	1	1.1%	0.3%
Cognitive function on individual level	1	1.1%	0.3%
Composite, days between randomization and death and age-adjusted TICS-T score	1	1.1%	0.3%

CPC scale 2-5	1	1.1%	0.3%
CPC scale 4-5	1	1.1%	0.3%
Disability Rating Scale (categorized)	1	1.1%	0.3%
Disability Rating Scale >5	1	1.1%	0.3%
Discharge destination (5 levels)	1	1.1%	0.3%
GOS-E 1-4 (death or unfavorable neurological outcome)	1	1.1%	0.3%
Hayling Sentence completion score	1	1.1%	0.3%
Karnofsky performance-status	1	1.1%	0.3%
Location/dead (5 levels)	1	1.1%	0.3%
Location/support (5 levels)	1	1.1%	0.3%
Location/support/dead (7 levels)	1	1.1%	0.3%
Logic thinking - Similarities score	1	1.1%	0.3%
Logical Memory I score	1	1.1%	0.3%
Logical Memory II score	1	1.1%	0.3%
Memory score - Digit Span score	1	1.1%	0.3%
Modified Rankin Scale score 5-6 (poor outcome)	1	1.1%	0.3%
Modified Rankin Scale score change	1	1.1%	0.3%
Mortality or poor neurological outcome	1	1.1%	0.3%
Not in hospital	1	1.1%	0.3%
Pfeffer FAQ (performance score)	1	1.1%	0.3%
Poor outcome	1	1.1%	0.3%
Presence of cognitive impairment based on multiple tests	1	1.1%	0.3%
Proportion of patients in paid employment at baseline who were unemployed at day 180	1	1.1%	0.3%
Rey-Osterrieth complex figure test (copy)	1	1.1%	0.3%
Rey-Osterrieth complex figure test (immediate recall, delayed recall, recognition)	1	1.1%	0.3%
Rey auditory verbal learning test, immediate + recall	1	1.1%	0.3%
SF-36 Physical Function Score <20	1	1.1%	0.3%
Span of numbers	1	1.1%	0.3%
TICS	1	1.1%	0.3%
TICS-T total score (age adjusted)	1	1.1%	0.3%
TICS categorized	1	1.1%	0.3%
Trailmaking A	1	1.1%	0.3%
Trailmaking B	1	1.1%	0.3%
Verbal fluency - Controlled Oral Word Association T-score	1	1.1%	0.3%
Verbal fluency test (S-words and animals)	1	1.1%	0.3%
HRQoL	48		13.3%
EQ-5D-5L index	9	18.8%	2.5%
EQ-5D-5L VAS	11	22.9%	3.0%
EQ-5D-3L index	5	10.4%	1.4%
EQ-5D-3L VAS	5	10.4%	1.4%
EQ-5D-5L components	8	16.7%	2.2%
SF-36	2	4.2%	0.6%
EQ-5D-3L components	2	4.2%	0.6%
EQ-5D index (unclear if 3L or 5L)	1	2.1%	0.3%
EQ-5D VAS (unclear if 3L or 5L)	3	6.2%	0.8%
HADS anxiety/depression (≥11)	1	2.1%	0.3%
PTSD – IES-R score ≥33	1	2.1%	0.3%
Ordinal/other	30		8.3%
Clinical improvement (≥1 point on 8-level ordinal scale) or live discharge from hospital	6	20.0%	1.7%
Clinical improvement (8-level ordinal scale)	6	20.0%	1.7%
WHO ordinal scale for COVID-19	5	16.7%	1.4%

Supplemental Digital Content

Clinical improvement (6-level ordinal scale)	3	10.0%	0.8%
Time to clinical improvement (6 level ordinal scale)	2	6.7%	0.6%
WHO ordinal scale for COVID-19 (only some levels)	2	6.7%	0.6%
Clinical status (5 levels)	1	3.3%	0.3%
Clinical status (6-level ordinal scale adapted from WHO ordinal scale)	1	3.3%	0.3%
Clinical status (6-level ordinal scale adapted from WHO ordinal scale) - cumulative	1	3.3%	0.3%
Clinical status (7-level ordinal scale)	1	3.3%	0.3%
Progression on Government of India Ministry of Health and Welfare COVID-19 ordinal scale for moderate to severe or severe to death	1	3.3%	0.3%
Time to clinical recovery (7-level ordinal scale)	1	3.3%	0.3%

All outcomes in each category in trials with the primary trial report published in 2020 or later, sorted by frequency. All outcomes in all trials are included in **Supplemental Table 2**. Of note, some of the different included outcomes are relatively similar based on the same tools (e.g., using different cut-offs), or sub-components of other included outcomes.

Abbreviations: see **Supplemental Table 2** above.

Supplemental Digital Content

Supplemental Table 16: Effect measures used in trials published before 2020

Effect measure	All outcomes (n = 325)	Days alive without... (n = 147)	Functional/ cognitive/ neurological (n = 104)	HRQoL (n = 70)	Ordinal/other (n = 4)
Mean difference	94 (28.9%)	43 (29.3%)	20 (19.2%)	30 (42.9%)	1 (25.0%)
Odds ratio	15 (4.6%)	5 (3.4%)	9 (8.7%)	1 (1.4%)	0 (0.0%)
Median difference	15 (4.6%)	8 (5.4%)	6 (5.8%)	1 (1.4%)	0 (0.0%)
Risk difference	32 (9.8%)	3 (2.0%)	22 (21.2%)	5 (7.1%)	2 (50.0%)
Hodges-Lehmann estimate	27 (8.3%)	16 (10.9%)	5 (4.8%)	6 (8.6%)	0 (0.0%)
Risk ratio	12 (3.7%)	0 (0.0%)	10 (9.6%)	2 (2.9%)	0 (0.0%)
Hazard ratio	2 (0.6%)	0 (0.0%)	2 (1.9%)	0 (0.0%)	0 (0.0%)
Incidence rate ratio/expected count ratio	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Probabilistic index	1 (0.3%)	1 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Ratio of means	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive plus p-value only[1]	113 (34.8%)	63 (42.9%)	27 (26.0%)	23 (32.9%)	0 (0.0%)
Descriptive only[1]	12 (3.7%)	4 (2.7%)	6 (5.8%)	2 (2.9%)	0 (0.0%)
Descriptive for the entire trial only[1]	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (25.0%)
Unclear[2]	4 (1.2%)	4 (2.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Effect measured used for all outcomes in trials with the primary trial report published before 2020, stratified by outcome type in drug trials only; effect measures are displayed in the same order as in **Table 3** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%.

[1] Outcomes only counted in one of these categories if no effect measure quantifying the actual differences between intervention arms was presented; thus, descriptive data (for each intervention)/p-values are not counted for outcomes using actual measures of differences. Descriptive data only includes graphical presentation of raw data.

[2] Outcomes where some effect measure quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this effect measure was.

Supplemental Table 17: Effect measures used in trials published in 2020 or later

Effect measure	All outcomes (n = 362)	Days alive without... (n = 190)	Functional/ cognitive/ neurological (n = 94)	HRQoL (n = 48)	Ordinal/other (n = 30)
Mean difference	89 (24.6%)	64 (33.7%)	9 (9.6%)	16 (33.3%)	0 (0.0%)
Odds ratio	71 (19.6%)	29 (15.3%)	19 (20.2%)	0 (0.0%)	23 (76.7%)
Median difference	44 (12.2%)	40 (21.1%)	0 (0.0%)	4 (8.3%)	0 (0.0%)
Risk difference	18 (5.0%)	0 (0.0%)	16 (17.0%)	0 (0.0%)	2 (6.7%)
Hodges-Lehmann estimate	21 (5.8%)	19 (10.0%)	0 (0.0%)	0 (0.0%)	2 (6.7%)
Risk ratio	8 (2.2%)	1 (0.5%)	7 (7.4%)	0 (0.0%)	0 (0.0%)
Hazard ratio	7 (1.9%)	3 (1.6%)	1 (1.1%)	0 (0.0%)	3 (10.0%)
Incidence rate ratio/expected count ratio	8 (2.2%)	8 (4.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Probabilistic index	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Ratio of means	1 (0.3%)	1 (0.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive plus p-value only[1]	90 (24.9%)	39 (20.5%)	31 (33.0%)	19 (39.6%)	1 (3.3%)
Descriptive only[1]	44 (12.2%)	11 (5.8%)	22 (23.4%)	11 (22.9%)	0 (0.0%)
Descriptive for the entire trial only[1]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Unclear[2]	5 (1.4%)	3 (1.6%)	2 (2.1%)	0 (0.0%)	0 (0.0%)

Effect measured used for all outcomes in trials with the primary trial report published in 2020 or later, stratified by outcome type in drug trials only; effect measures are displayed in the same order as in **Table 3** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%.

[1] Outcomes only counted in one of these categories if no effect measure quantifying the actual differences between intervention arms was presented; thus, descriptive data (for each intervention)/p-values are not counted for outcomes using actual measures of differences. Descriptive data only includes graphical presentation of raw data.

[2] Outcomes where some effect measure quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this effect measure was.

Supplemental Table 18: Statistical methods used in trials published before 2020

Statistical method	All outcomes (n = 325)	Days alive without... (n = 147)	Functional/ cognitive/ neurological (n = 104)	HRQoL (n = 70)	Ordinal/other (n = 4)
Non-parametric[1]	102 (31.4%)	78 (53.1%)	11 (10.6%)	12 (17.1%)	1 (25.0%)
Linear regression/ T-test[2]	97 (29.8%)	41 (27.9%)	23 (22.1%)	33 (47.1%)	0 (0.0%)
Chi-squared[3]	56 (17.2%)	0 (0.0%)	44 (42.3%)	10 (14.3%)	2 (50.0%)
Proportional odds/cumulative logistic regression	12 (3.7%)	5 (3.4%)	7 (6.7%)	0 (0.0%)	0 (0.0%)
Hodges-Lehmann estimator	27 (8.3%)	16 (10.9%)	5 (4.8%)	6 (8.6%)	0 (0.0%)
Bootstrapping	6 (1.8%)	3 (2.0%)	3 (2.9%)	0 (0.0%)	0 (0.0%)
Logistic regression	8 (2.5%)	0 (0.0%)	5 (4.8%)	2 (2.9%)	1 (25.0%)
Quantile regression	2 (0.6%)	2 (1.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with identity link	8 (2.5%)	5 (3.4%)	0 (0.0%)	3 (4.3%)	0 (0.0%)
Time-to-event analysis[4]	6 (1.8%)	3 (2.0%)	3 (2.9%)	0 (0.0%)	0 (0.0%)
Generalized linear (mixed) model, log-binomial	9 (2.8%)	0 (0.0%)	7 (6.7%)	2 (2.9%)	0 (0.0%)
Kryger-Jensen/Lange test	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Hurdle-negative binomial regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Poisson/negative binomial regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Binomial proportion test/exact binomial confidence interval	3 (0.9%)	0 (0.0%)	0 (0.0%)	2 (2.9%)	1 (25.0%)
Simple calculation	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Zero-one-inflated beta regression	3 (0.9%)	3 (2.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Beta-binomial regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized additive model with zero-inflated/zero-one-inflated beta distribution	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with zero-one-inflated beta distribution	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized linear model, identity-binomial	1 (0.3%)	0 (0.0%)	1 (1.0%)	0 (0.0%)	0 (0.0%)
Generalized linear (mixed) model, logit link and G-computation	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with logit link	1 (0.3%)	0 (0.0%)	0 (0.0%)	1 (1.4%)	0 (0.0%)
Probabilistic index	1 (0.3%)	1 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive only[5]	12 (3.7%)	4 (2.7%)	6 (5.8%)	2 (2.9%)	0 (0.0%)
Descriptive for the entire trial only[5]	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (25.0%)
Unclear[6]	12 (3.7%)	10 (6.8%)	1 (1.0%)	1 (1.4%)	0 (0.0%)

Statistical methods used for all outcomes in trials with the primary trial report published before 2020, stratified by outcome type in drug trials only; statistical methods are displayed in the same order as in **Table 4** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%. Frequentist/Bayesian variants of the same statistical methods summarised together.
[1] Includes all non-parametric tests of continuous data, e.g. Mann-Whitney U/Wilcoxon, van Elteren, Kruskal-Wallis, and the Wilcoxon rank sum test for clustered data using Rosner-Glynn-Lee method.

[2] Includes linear regression and all variants including T-tests, analysis of (co)-variance, linear mixed models, etc., also including these methods used on transformed data.

[3] Includes the Chi-squared test and related tests for contingency tables, e.g., Fisher's exact test, and the Cochran-Armitage test for trend.

[4] includes the Cox proportional hazards model, the Kaplan Meier estimator, the log-rank test, the Fine-Gray model, and unspecified survival analysis methods.

[5] Outcomes only counted in one of these categories if no inferential statistical method was used to compare intervention groups, i.e., if only descriptive data for each intervention group or for the full trial were presented.

[6] Outcomes where some statistical method quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this method was.

Supplemental Digital Content

Supplemental Table 19: Statistical methods used in trials published in 2020 or later

Statistical method	All outcomes (n = 362)	Days alive without... (n = 190)	Functional/ cognitive/ neurological (n = 94)	HRQoL (n = 48)	Ordinal/other (n = 30)
Non-parametric[1]	109 (30.1%)	74 (38.9%)	19 (20.2%)	14 (29.2%)	2 (6.7%)
Linear regression/ T-test[2]	88 (24.3%)	54 (28.4%)	14 (14.9%)	20 (41.7%)	0 (0.0%)
Chi-squared[3]	29 (8.0%)	1 (0.5%)	17 (18.1%)	6 (12.5%)	5 (16.7%)
Proportional odds/cumulative logistic regression	51 (14.1%)	29 (15.3%)	3 (3.2%)	0 (0.0%)	19 (63.3%)
Hodges-Lehmann estimator	21 (5.8%)	19 (10.0%)	0 (0.0%)	0 (0.0%)	2 (6.7%)
Bootstrapping	36 (9.9%)	29 (15.3%)	1 (1.1%)	5 (10.4%)	1 (3.3%)
Logistic regression	18 (5.0%)	0 (0.0%)	17 (18.1%)	0 (0.0%)	1 (3.3%)
Quantile regression	22 (6.1%)	18 (9.5%)	0 (0.0%)	4 (8.3%)	0 (0.0%)
Generalizing estimating equation with identity link	6 (1.7%)	6 (3.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Time-to-event analysis[4]	8 (2.2%)	4 (2.1%)	1 (1.1%)	0 (0.0%)	3 (10.0%)
Generalized linear (mixed) model, log-binomial	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Kryger-Jensen/Lange test	8 (2.2%)	4 (2.1%)	0 (0.0%)	4 (8.3%)	0 (0.0%)
Hurdle-negative binomial regression	6 (1.7%)	6 (3.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Poisson/negative binomial regression	5 (1.4%)	3 (1.6%)	2 (2.1%)	0 (0.0%)	0 (0.0%)
Binomial proportion test/exact binomial confidence interval	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (3.3%)
Simple calculation	3 (0.8%)	0 (0.0%)	3 (3.2%)	0 (0.0%)	0 (0.0%)
Zero-one-inflated beta regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Beta-binomial regression	2 (0.6%)	2 (1.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized additive model with zero-inflated/zero-one-inflated beta distribution	2 (0.6%)	2 (1.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with zero-one-inflated beta distribution	2 (0.6%)	2 (1.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized linear model, identity-binomial	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (3.3%)
Generalized linear (mixed) model, logit link and G-computation	2 (0.6%)	1 (0.5%)	1 (1.1%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with logit link	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Probabilistic index	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive only[5]	44 (12.2%)	11 (5.8%)	22 (23.4%)	11 (22.9%)	0 (0.0%)
Descriptive for the entire trial only[5]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Unclear[6]	9 (2.5%)	7 (3.7%)	2 (2.1%)	0 (0.0%)	0 (0.0%)

Statistical methods used for all outcomes in trials with the primary trial report published in 2020 or later, stratified by outcome type in drug trials only; statistical methods are displayed in the same order as in **Table 4** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%. Frequentist/Bayesian variants of the same statistical methods summarised together.

[1] Includes all non-parametric tests of continuous data, e.g. Mann-Whitney U/Wilcoxon, van Elteren, Kruskal-Wallis, and the Wilcoxon rank sum test for clustered data using Rosner-Glynn-Lee method.

[2] Includes linear regression and all variants including T-tests, analysis of (co)-variance, linear mixed models, etc., also including these methods used on transformed data.

[3] Includes the Chi-squared test and related tests for contingency tables, e.g., Fisher's exact test, and the Cochran-Armitage test for trend.

[4] includes the Cox proportional hazards model, the Kaplan Meier estimator, the log-rank test, the Fine-Gray model, and unspecified survival analysis methods.

[5] Outcomes only counted in one of these categories if no inferential statistical method was used to compare intervention groups, i.e., if only descriptive data for each intervention group or for the full trial were presented.

[6] Outcomes where some statistical method quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this method was.

Supplemental Digital Content

Supplemental Table 20: General trial-level characteristics by COVID-19 status

Variable	Not restricted to patients with suspected or confirmed COVID-19 (n = 144)	Restricted to patients with suspected or confirmed COVID-19 (n = 23)
Intervention type		
- Drug	77 (53.5%)	20 (87.0%)
- Management	51 (35.4%)	0 (0.0%)
- Device	16 (11.1%)	3 (13.0%)
Journal		
- JAMA	36 (25.0%)	10 (43.5%)
- New England Journal of Medicine	34 (23.6%)	2 (8.7%)
- Intensive Care Medicine	23 (16.0%)	1 (4.3%)
- Critical Care Medicine	15 (10.4%)	1 (4.3%)
- The Lancet Respiratory Medicine	12 (8.3%)	4 (17.4%)
- Critical Care	10 (6.9%)	1 (4.3%)
- American Journal of Respiratory and Critical Care Medicine	4 (2.8%)	2 (8.7%)
- Chest	5 (3.5%)	1 (4.3%)
- The Lancet	3 (2.1%)	0 (0.0%)
- BMJ	0 (0.0%)	1 (4.3%)
- Circulation	1 (0.7%)	0 (0.0%)
- JAMA Neurology	1 (0.7%)	0 (0.0%)
Year	2019 (2018 - 2021) [2006 - 2022]	2021 (2021 - 2022) [2020 - 2022]
Countries (no.)	1 (1 - 2) [1 - 15]	1 (1 - 4) [1 - 10]
Centers (no.)	14 (5 - 31) [1 - 168]	16 (8 - 71) [1 - 393]
Patients (no.)	352 (150 - 740) [20 - 15802]	250 (106 - 650) [47 - 2097]
Outcomes included (no.)	3 (1 - 5) [1 - 43]	4 (1 - 4) [1 - 13]
Full protocol referenced and available	113 (78.5%)	21 (91.3%)
Any (co-)primary patient-important outcome other than mortality included	38 (26.4%)	16 (69.6%)

General trial-level characteristics for al included trials and stratified by whether the trial was restricted to patients with suspected or confirmed COVID-19. Data are presented as *numbers (percentages)* or as *medians (interquartile range) [full range]*.

Further details and abbreviations are presented in the **Table 1** legend in the primary text, which also includes a column with data for all trials.

Supplemental Table 21: General outcome-level characteristics by COVID-19 status

Variable	Not restricted to patients with suspected or confirmed COVID-19 (n = 603)	Restricted to patients with suspected or confirmed COVID-19 (n = 84)	Missing values
Primary/ co-primary outcome	40 (6.6%)	16 (19.0%)	0 (0.0%)
Prioritization of outcome components			0 (0.0%)
- Not relevant	547 (90.7%)	78 (92.9%)	
- Equal prioritization	43 (7.1%)	0 (0.0%)	
- Not relevant (individual categories/ domains/ components reported secondarily, but not considered or analyzed as primary outcomes)	12 (2.0%)	6 (7.1%)	
- Three dichotomizations reported with equal prioritization	1 (0.2%)	0 (0.0%)	
Assessment time point			2 (0.3%)
- Day 28	195 (32.4%)	29 (34.5%)	
- Day 180/6 months	128 (21.3%)	5 (6.0%)	
- Day 90/3 months	79 (13.1%)	8 (9.5%)	
- 1 year	37 (6.2%)	0 (0.0%)	
- Day 14	18 (3.0%)	9 (10.7%)	
- Day 60	26 (4.3%)	1 (1.2%)	
- Day 30	24 (4.0%)	2 (2.4%)	
- Hospital discharge/in-hospital	25 (4.2%)	1 (1.2%)	
- ICU discharge	23 (3.8%)	0 (0.0%)	
- 90 days/3 months after ICU discharge	22 (3.7%)	0 (0.0%)	
- Day 21	0 (0.0%)	19 (22.6%)	
- Day 7	8 (1.3%)	3 (3.6%)	
- 30 days after ICU discharge	4 (0.7%)	0 (0.0%)	
- Day 25	3 (0.5%)	0 (0.0%)	
- 6 months after hospital discharge	2 (0.3%)	0 (0.0%)	
- 8 days (within ICU)	2 (0.3%)	0 (0.0%)	
- Day 10	0 (0.0%)	2 (2.4%)	
- Day 15	0 (0.0%)	2 (2.4%)	
- Day 30 after end of study sedation (48 hours) or extubation	2 (0.3%)	0 (0.0%)	
- Day 4	0 (0.0%)	2 (2.4%)	
- Day 7 after end of study sedation (48 hours) or extubation	2 (0.3%)	0 (0.0%)	
- 1 week after ICU discharge	1 (0.2%)	0 (0.0%)	
- Day 29	0 (0.0%)	1 (1.2%)	
Dead patients at the time of outcome assessment (%)	31.5 (22.1 - 42.4) [0.0 - 82.3]	32.5 (20.3 - 38.0) [9.5 - 68.9]	165 (24.0%)
Handling of dead patients			66 (9.6%)
- Survivors only	223 (40.4%)	4 (5.8%)	
- Dead = 0	147 (26.6%)	16 (23.2%)	
- Included in scale/definition	66 (12.0%)	28 (40.6%)	
- Actual number of days/value (no penalization of death)	88 (15.9%)	5 (7.2%)	
- Dead = worst value	13 (2.4%)	16 (23.2%)	
- Death penalized in some cases	4 (0.7%)	0 (0.0%)	
- Percentage of days out of the total number of days alive (not counting days where the patient was dead in either numerator or denominator)	3 (0.5%)	0 (0.0%)	

- All included (no dead patients)	2 (0.4%)	0 (0.0%)	
- Percentage of days out of the total number of days with possible condition (not counting days where the patient was dead or discharged from the ICU in either numerator or denominator)	2 (0.4%)	0 (0.0%)	
- Worst possible point score on days where the patient was dead	2 (0.4%)	0 (0.0%)	
- Dead patients included in analysis without having had the event	1 (0.2%)	0 (0.0%)	
- Last value before death carried forward	1 (0.2%)	0 (0.0%)	
Missingness of outcome (%)	0.4 (0.0 - 5.8) [0.0 - 53.6]	0.4 (0.0 - 2.4) [0.0 - 24.0]	129 (18.8%)
Handling of missing data			110 (16.0%)
- Complete case analysis	223 (44.2%)	35 (47.9%)	
- Complete case analysis (no missingness)	217 (43.1%)	30 (41.1%)	
- Multiple imputation	22 (4.4%)	4 (5.5%)	
- Last value carried forward	16 (3.2%)	0 (0.0%)	
- Interpolation based on measured values and extrapolation using last value carried forward	5 (1.0%)	0 (0.0%)	
- Complete case analysis with logical imputation of intermittent missing repeated values	4 (0.8%)	0 (0.0%)	
- Imputation, details unclear	4 (0.8%)	0 (0.0%)	
- Model-based single imputation	4 (0.8%)	0 (0.0%)	
- Single imputation, worst-case imputation	3 (0.6%)	1 (1.4%)	
- Complete case analysis and best-worst/worst-best case analysis	2 (0.4%)	1 (1.4%)	
- Multiple imputation and complete case analysis and best-worst/worst-best case analysis	1 (0.2%)	2 (2.7%)	
- Categorized as "other or unknown" with other values	1 (0.2%)	0 (0.0%)	
- Partial single imputation, best case/normal imputation	1 (0.2%)	0 (0.0%)	
- Survivor average causal effects methods	1 (0.2%)	0 (0.0%)	

General outcome-level characteristics for all included trials stratified by whether the trial was restricted to patients with suspected or confirmed COVID-19. Data are presented as *numbers (percentages)* or as *medians (interquartile range) [full range]*.

Further details and abbreviations are presented in the **Supplemental Table S1** legend, which also includes a column with data for all trials. Categories are presented in the same order as in **Supplemental Table S1** to ease comparison.

Supplemental Table 22: All reported outcomes in trials not restricted to COVID-19

Outcome	n (total = 603)	% of outcomes in category	% of all outcomes
Days alive without…	288		47.8%
Days alive without (invasive) mechanical ventilation/ventilator-free days	110	38.2%	18.2%
Days alive without circulatory support	42	14.6%	7.0%
Days alive and out of ICU	40	13.9%	6.6%
Days alive without renal replacement therapy	37	12.8%	6.1%
Days alive and out of hospital	16	5.6%	2.7%
Days alive without delirium/coma	10	3.5%	1.7%
Days alive without respiratory support	1	0.3%	0.2%
Days alive without respiratory or circulatory support	2	0.7%	0.3%
Days alive without life support	3	1.0%	0.5%
Days alive without coma	5	1.7%	0.8%
Days alive without delirium	4	1.4%	0.7%
Delta Sepsis Support Index	4	1.4%	0.7%
Days alive without sedation	2	0.7%	0.3%
Dialysis catheter-free days	2	0.7%	0.3%
Hours without circulatory support	2	0.7%	0.3%
Time to ICU separation	2	0.7%	0.3%
Time to ventilator separation	2	0.7%	0.3%
Days alive without ECMO	1	0.3%	0.2%
Days alive without HFNO or NIV	1	0.3%	0.2%
Days alive and at home	1	0.3%	0.2%
Time to hospital discharge	1	0.3%	0.2%
Functional/cognitive/neurological	198		32.8%
CPC scale	7	3.5%	1.2%
Modified Rankin Scale score 4-6 (poor outcome)	7	3.5%	1.2%
Barthel Index	6	3.0%	1.0%
Katz Index	6	3.0%	1.0%
MoCA Blind	6	3.0%	1.0%
Alive and at home	5	2.5%	0.8%
CPC scale 1-2 (alive with good neurological outcome)	5	2.5%	0.8%
Medical Research Council muscle strength	5	2.5%	0.8%
WHODAS disability-free survival (<25%)	5	2.5%	0.8%
CPC scale 3-5 (poor outcome)	4	2.0%	0.7%
ICU mobility scale	4	2.0%	0.7%
Medical Research Council muscle strength <48 (ICU-acquired weakness)	4	2.0%	0.7%
SF-36 Physical Function Score	4	2.0%	0.7%
AD8 (cognition)	3	1.5%	0.5%
Difficulty in a daily activity	3	1.5%	0.5%
Disability score	3	1.5%	0.5%
Discharge destination (4 levels)	3	1.5%	0.5%
Galveston Orientation and Amnesia Test <75/100 (post-traumatic amnesia)	3	1.5%	0.5%
Hand grip strength	3	1.5%	0.5%
Impact on daily activities (any of 10 items)	3	1.5%	0.5%
IQCODE short version (cognitive decline)	3	1.5%	0.5%
Location of residence, % home	3	1.5%	0.5%
modified Rankin Scale score	3	1.5%	0.5%
Return to work	3	1.5%	0.5%
Significant change in work duties	3	1.5%	0.5%

6-minute walking test	2	1.0%	0.3%
Adelaide Activities Profile	2	1.0%	0.3%
Barthel Index >=85 (independent)	2	1.0%	0.3%
ECOG performance status	2	1.0%	0.3%
Functional independence (patient-reported)	2	1.0%	0.3%
Functional Status Score for the ICU	2	1.0%	0.3%
GOS	2	1.0%	0.3%
GOS-E	2	1.0%	0.3%
GOS-E 6-8 (good outcome)	2	1.0%	0.3%
GOS 4-5 (favorable neurologic outcome)	2	1.0%	0.3%
Lawton IADL	2	1.0%	0.3%
Medical Research Council muscle strength, lower limbs	2	1.0%	0.3%
Physical Function in the ICU Test	2	1.0%	0.3%
Timed up and go test	2	1.0%	0.3%
WHODAS no or mild disability	2	1.0%	0.3%
WHODAS score	2	1.0%	0.3%
Alive and at home off dialysis	1	0.5%	0.2%
Australian Labor Force Survey	1	0.5%	0.2%
Australian Labor Force Survey - returned to work	1	0.5%	0.2%
Clinical Frailty Scale >4	1	0.5%	0.2%
Coding test (translate numbers into figures using an answer key)	1	0.5%	0.2%
Cognitive function on individual level	1	0.5%	0.2%
Composite, days between randomization and death and age-adjusted TICS-T score	1	0.5%	0.2%
CPC scale (best during trial)	1	0.5%	0.2%
CPC scale 2-5	1	0.5%	0.2%
CPC scale 4-5	1	0.5%	0.2%
D-KEFS domains	1	0.5%	0.2%
D-KEFS total	1	0.5%	0.2%
Disability Rating Scale (categorized)	1	0.5%	0.2%
Disability Rating Scale >5	1	0.5%	0.2%
Discharge destination (5 levels)	1	0.5%	0.2%
Discharge destination (6 levels)	1	0.5%	0.2%
Discharge destination (7 levels)	1	0.5%	0.2%
GOS-E 1-4 (death or unfavorable neurological outcome)	1	0.5%	0.2%
GOS-E 5-8 (favorable outcome)	1	0.5%	0.2%
Hayling Sentence completion score	1	0.5%	0.2%
ICU mobility scale or 6-minute walk test	1	0.5%	0.2%
Institutional dependency	1	0.5%	0.2%
Karnofsky performance-status	1	0.5%	0.2%
Location/dead (5 levels)	1	0.5%	0.2%
Location/support (5 levels)	1	0.5%	0.2%
Location/support/dead (7 levels)	1	0.5%	0.2%
Logic thinking - Similarities score	1	0.5%	0.2%
Logical Memory I score	1	0.5%	0.2%
Logical Memory II score	1	0.5%	0.2%
Loss of employment (for patients who were fully or partially employed)	1	0.5%	0.2%
Memory score - Digit Span score	1	0.5%	0.2%
MoCA	1	0.5%	0.2%
MoCA change from baseline	1	0.5%	0.2%
Modified Rankin Scale score ≤2 (alive with good neurological outcome)	1	0.5%	0.2%

Modified Rankin Scale score ≤3 (alive with good neurological outcome)	1	0.5%	0.2%
modified Rankin Scale score 5-6 (poor outcome)	1	0.5%	0.2%
modified Rankin Scale score change	1	0.5%	0.2%
Mortality or poor neurological outcome	1	0.5%	0.2%
Muscular Disability Rating Scale (MDRS)	1	0.5%	0.2%
Not in hospital	1	0.5%	0.2%
Pfeffer FAQ (performance score)	1	0.5%	0.2%
Poor outcome	1	0.5%	0.2%
Presence of cognitive impairment based on multiple tests	1	0.5%	0.2%
Proportion of patients in paid employment at baseline who were unemployed at day 180	1	0.5%	0.2%
RAND-36 physical function	1	0.5%	0.2%
RBANS domains	1	0.5%	0.2%
RBANS total	1	0.5%	0.2%
Rey-Osterrieth complex figure test (copy)	1	0.5%	0.2%
Rey-Osterrieth complex figure test (immediate recall, delayed recall, recognition)	1	0.5%	0.2%
Rey auditory verbal learning test, immediate + recall	1	0.5%	0.2%
SF-36 Physical Function Score <20	1	0.5%	0.2%
Sit and stand (repetitions)	1	0.5%	0.2%
Skilled nursing home facility residence	1	0.5%	0.2%
Span of numbers	1	0.5%	0.2%
Stationary walk (repetitions)	1	0.5%	0.2%
TICS	1	0.5%	0.2%
TICS-T total score (age adjusted)	1	0.5%	0.2%
TICS categorized	1	0.5%	0.2%
Trailmaking A	1	0.5%	0.2%
Trailmaking B	1	0.5%	0.2%
Verbal fluency - Controlled Oral Word Association T-score	1	0.5%	0.2%
Verbal fluency test (S-words and animals)	1	0.5%	0.2%
HRQoL	113		18.7%
EQ-5D-5L index	16	14.2%	2.7%
EQ-5D-5L VAS	13	11.5%	2.2%
EQ-5D-3L index	12	10.6%	2.0%
EQ-5D-3L VAS	10	8.8%	1.7%
EQ-5D-5L components	9	8.0%	1.5%
HADS domains	7	6.2%	1.2%
SF-36	5	4.4%	0.8%
EQ-5D-3L components	3	2.7%	0.5%
EQ-5D index (unclear if 3L or 5L)	3	2.7%	0.5%
EQ-5D VAS (unclear if 3L or 5L)	3	2.7%	0.5%
Pain interference (5 levels)	3	2.7%	0.5%
PTSD - IES-R score	3	2.7%	0.5%
Self-rated health (5 levels)	3	2.7%	0.5%
SF-12	3	2.7%	0.5%
Assessment of Quality of Life 4D utility score	2	1.8%	0.3%
EQ-5D-3L index change from baseline	2	1.8%	0.3%
EQ-5D-5L components - "no problems"	2	1.8%	0.3%
PTSD - IES-R score, mean item score ≥1.6	2	1.8%	0.3%
PTSS >45 (post-traumatic stress-like symptoms)	2	1.8%	0.3%
EQ-5D-5L change from baseline	1	0.9%	0.2%
HADS anxiety/depression (≥11)	1	0.9%	0.2%
HADS total	1	0.9%	0.2%

PSS-SR (PTSD)	1	0.9%	0.2%
PSS-SR >18 (PTSD)	1	0.9%	0.2%
PTSD - IES-R score ≥33	1	0.9%	0.2%
PTSD - IES-R score >22	1	0.9%	0.2%
PTSD - IES domains	1	0.9%	0.2%
RAND-36 general health	1	0.9%	0.2%
VAS-A	1	0.9%	0.2%
Ordinal/other	4		0.7%
ICU memory tool	1	25.0%	0.2%
ICU memory tool - components yes/no to each	1	25.0%	0.2%
ICU memory tool - number of memories	1	25.0%	0.2%
modified Brice questionnaire	1	25.0%	0.2%

All outcomes in each category in trials not restricted to patients with suspected or confirmed COVID-19, sorted by frequency. All outcomes in all trials are included in **Supplemental Table 2**. Of note, some of the different included outcomes are relatively similar based on the same tools (e.g., using different cut-offs), or sub-components of other included outcomes.

Abbreviations: see **Supplemental Table 2** above.

Supplemental Digital Content

Supplemental Table 23: All reported outcomes in trials restricted to COVID-19

Outcome	n (total = 84)	% of outcomes in category	% of all outcomes
Days alive without…	49		58.3%
Days alive without (invasive) mechanical ventilation/ventilator-free days	16	32.7%	19.0%
Days alive without circulatory support	7	14.3%	8.3%
Days alive and out of ICU	3	6.1%	3.6%
Days alive without renal replacement therapy	3	6.1%	3.6%
Days alive and out of hospital	2	4.1%	2.4%
Days alive without respiratory support	8	16.3%	9.5%
Days alive without respiratory or circulatory support	6	12.2%	7.1%
Days alive without life support	4	8.2%	4.8%
Functional/cognitive/neurological	0		0.0%
HRQoL	5		6.0%
EQ-5D-5L index	3	60.0%	3.6%
EQ-5D-5L VAS	2	40.0%	2.4%
Ordinal/other	30		35.7%
Clinical improvement (≥1 point on 8-level ordinal scale) or live discharge from hospital	6	20.0%	7.1%
Clinical improvement (8-level ordinal scale)	6	20.0%	7.1%
WHO ordinal scale for COVID-19	5	16.7%	6.0%
Clinical improvement (6-level ordinal scale)	3	10.0%	3.6%
Time to clinical improvement (6 level ordinal scale)	2	6.7%	2.4%
WHO ordinal scale for COVID-19 (only some levels)	2	6.7%	2.4%
Clinical status (5 levels)	1	3.3%	1.2%
Clinical status (6-level ordinal scale adapted from WHO ordinal scale)	1	3.3%	1.2%
Clinical status (6-level ordinal scale adapted from WHO ordinal scale) - cumulative	1	3.3%	1.2%
Clinical status (7-level ordinal scale)	1	3.3%	1.2%
Progression on Government of India Ministry of Health and Welfare COVID-19 ordinal scale for moderate to severe or severe to death	1	3.3%	1.2%
Time to clinical recovery (7-level ordinal scale)	1	3.3%	1.2%

All outcomes in each category in trials restricted to patients with suspected or confirmed COVID-19, sorted by frequency. All outcomes in all trials are included in **Supplemental Table 2**. Of note, some of the different included outcomes are relatively similar based on the same tools (e.g., using different cut-offs), or sub-components of other included outcomes.

Abbreviations: see **Supplemental Table 2** above.

Supplemental Table 24: Effect measures used in trials not restricted to COVID-19

Effect measure	All outcomes (n = 603)	Days alive without... (n = 288)	Functional/ cognitive/ neurological (n = 198)	HRQoL (n = 113)	Ordinal/other (n = 4)
Mean difference	166 (27.5%)	95 (33.0%)	29 (14.6%)	41 (36.3%)	1 (25.0%)
Odds ratio	45 (7.5%)	16 (5.6%)	28 (14.1%)	1 (0.9%)	0 (0.0%)
Median difference	44 (7.3%)	35 (12.2%)	6 (3.0%)	3 (2.7%)	0 (0.0%)
Risk difference	48 (8.0%)	3 (1.0%)	38 (19.2%)	5 (4.4%)	2 (50.0%)
Hodges-Lehmann estimate	45 (7.5%)	34 (11.8%)	5 (2.5%)	6 (5.3%)	0 (0.0%)
Risk ratio	20 (3.3%)	1 (0.3%)	17 (8.6%)	2 (1.8%)	0 (0.0%)
Hazard ratio	6 (1.0%)	3 (1.0%)	3 (1.5%)	0 (0.0%)	0 (0.0%)
Incidence rate ratio/expected count ratio	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Probabilistic index	1 (0.2%)	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Ratio of means	1 (0.2%)	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive plus p-value only[1]	194 (32.2%)	94 (32.6%)	58 (29.3%)	42 (37.2%)	0 (0.0%)
Descriptive only[1]	56 (9.3%)	15 (5.2%)	28 (14.1%)	13 (11.5%)	0 (0.0%)
Descriptive for the entire trial only[1]	1 (0.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (25.0%)
Unclear[2]	8 (1.3%)	6 (2.1%)	2 (1.0%)	0 (0.0%)	0 (0.0%)

Effect measured used for all outcomes in trials not restricted to patients with suspected or confirmed COVID-19, stratified by outcome type in drug trials only; effect measures are displayed in the same order as in **Table 3** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%.

[1] Outcomes only counted in one of these categories if no effect measure quantifying the actual differences between intervention arms was presented; thus, descriptive data (for each intervention)/p-values are not counted for outcomes using actual measures of differences. Descriptive data only includes graphical presentation of raw data.

[2] Outcomes where some effect measure quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this effect measure was.

Supplemental Table 25: Effect measures used in trials restricted to COVID-19

Effect measures	All outcomes (n = 84)	Days alive without... (n = 49)	Functional/ cognitive/ neurological (n = 0)	HRQoL (n = 5)	Ordinal/other (n = 30)
Mean difference	17 (20.2%)	12 (24.5%)	-	5 (100.0%)	0 (0.0%)
Odds ratio	41 (48.8%)	18 (36.7%)	-	0 (0.0%)	23 (76.7%)
Median difference	15 (17.9%)	13 (26.5%)	-	2 (40.0%)	0 (0.0%)
Risk difference	2 (2.4%)	0 (0.0%)	-	0 (0.0%)	2 (6.7%)
Hodges-Lehmann estimate	3 (3.6%)	1 (2.0%)	-	0 (0.0%)	2 (6.7%)
Risk ratio	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Hazard ratio	3 (3.6%)	0 (0.0%)	-	0 (0.0%)	3 (10.0%)
Incidence rate ratio/expected count ratio	8 (9.5%)	8 (16.3%)	-	0 (0.0%)	0 (0.0%)
Probabilistic index	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Ratio of means	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Descriptive plus p-value only[1]	9 (10.7%)	8 (16.3%)	-	0 (0.0%)	1 (3.3%)
Descriptive only[1]	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Descriptive for the entire trial only[1]	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Unclear[2]	1 (1.2%)	1 (2.0%)	-	0 (0.0%)	0 (0.0%)

Effect measured used for all outcomes in trials restricted to patients with suspected or confirmed COVID-19, stratified by outcome type in drug trials only; effect measures are displayed in the same order as in **Table 3** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%.

[1] Outcomes only counted in one of these categories if no effect measure quantifying the actual differences between intervention arms was presented; thus, descriptive data (for each intervention)/p-values are not counted for outcomes using actual measures of differences. Descriptive data only includes graphical presentation of raw data.

[2] Outcomes where some effect measure quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this effect measure was.

Supplemental Table 26: Statistical methods used in trials not restricted to COVID-19

Statistical method	All outcomes (n = 603)	Days alive without… (n = 288)	Functional/ cognitive/ neurological (n = 198)	HRQoL (n = 113)	Ordinal/other (n = 4)
Non-parametric[1]	195 (32.3%)	138 (47.9%)	30 (15.2%)	26 (23.0%)	1 (25.0%)
Linear regression/ T-test[2]	168 (27.9%)	83 (28.8%)	37 (18.7%)	48 (42.5%)	0 (0.0%)
Chi-squared[3]	80 (13.3%)	1 (0.3%)	61 (30.8%)	16 (14.2%)	2 (50.0%)
Proportional odds/cumulative logistic regression	26 (4.3%)	16 (5.6%)	10 (5.1%)	0 (0.0%)	0 (0.0%)
Hodges-Lehmann estimator	45 (7.5%)	34 (11.8%)	5 (2.5%)	6 (5.3%)	0 (0.0%)
Bootstrapping	21 (3.5%)	17 (5.9%)	4 (2.0%)	0 (0.0%)	0 (0.0%)
Logistic regression	25 (4.1%)	0 (0.0%)	22 (11.1%)	2 (1.8%)	1 (25.0%)
Quantile regression	18 (3.0%)	16 (5.6%)	0 (0.0%)	2 (1.8%)	0 (0.0%)
Generalizing estimating equation with identity link	14 (2.3%)	11 (3.8%)	0 (0.0%)	3 (2.7%)	0 (0.0%)
Time-to-event analysis[4]	11 (1.8%)	7 (2.4%)	4 (2.0%)	0 (0.0%)	0 (0.0%)
Generalized linear (mixed) model, log-binomial	9 (1.5%)	0 (0.0%)	7 (3.5%)	2 (1.8%)	0 (0.0%)
Kryger-Jensen/Lange test	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Hurdle-negative binomial regression	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Poisson/negative binomial regression	3 (0.5%)	1 (0.3%)	2 (1.0%)	0 (0.0%)	0 (0.0%)
Binomial proportion test/exact binomial confidence interval	3 (0.5%)	0 (0.0%)	0 (0.0%)	2 (1.8%)	1 (25.0%)
Simple calculation	3 (0.5%)	0 (0.0%)	3 (1.5%)	0 (0.0%)	0 (0.0%)
Zero-one-inflated beta regression	3 (0.5%)	3 (1.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Beta-binomial regression	2 (0.3%)	2 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized additive model with zero-inflated/zero-one-inflated beta distribution	2 (0.3%)	2 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with zero-one-inflated beta distribution	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Generalized linear model, identity-binomial	1 (0.2%)	0 (0.0%)	1 (0.5%)	0 (0.0%)	0 (0.0%)
Generalized linear (mixed) model, logit link and G-computation	1 (0.2%)	0 (0.0%)	1 (0.5%)	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with logit link	1 (0.2%)	0 (0.0%)	0 (0.0%)	1 (0.9%)	0 (0.0%)
Probabilistic index	1 (0.2%)	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Descriptive only[5]	56 (9.3%)	15 (5.2%)	28 (14.1%)	13 (11.5%)	0 (0.0%)
Descriptive for the entire trial only[5]	1 (0.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (25.0%)
Unclear[6]	21 (3.5%)	17 (5.9%)	3 (1.5%)	1 (0.9%)	0 (0.0%)

Statistical methods used for all outcomes in trials not restricted to patients with suspected or confirmed COVID-19, stratified by outcome type in drug trials only; statistical methods are displayed in the same order as in **Table 4** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%. Frequentist/Bayesian variants of the same statistical methods summarised together.

[1] Includes all non-parametric tests of continuous data, e.g. Mann-Whitney U/Wilcoxon, van Elteren,

Kruskal-Wallis, and the Wilcoxon rank sum test for clustered data using Rosner-Glynn-Lee method.

[2] Includes linear regression and all variants including T-tests, analysis of (co)-variance, linear mixed models, etc., also including these methods used on transformed data.

[3] Includes the Chi-squared test and related tests for contingency tables, e.g., Fisher's exact test, and the Cochran-Armitage test for trend.

[4] includes the Cox proportional hazards model, the Kaplan Meier estimator, the log-rank test, the Fine-Gray model, and unspecified survival analysis methods.

[5] Outcomes only counted in one of these categories if no inferential statistical method was used to compare intervention groups, i.e., if only descriptive data for each intervention group or for the full trial were presented.

[6] Outcomes where some statistical method quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this method was.

Supplemental Table 27: Statistical methods used in trials restricted to COVID-19

Statistical method	All outcomes (n = 84)	Days alive without… (n = 49)	Functional/ cognitive/ neurological (n = 0)	HRQoL (n = 5)	Ordinal/other (n = 30)
Non-parametric[1]	16 (19.0%)	14 (28.6%)	-	0 (0.0%)	2 (6.7%)
Linear regression/ T-test[2]	17 (20.2%)	12 (24.5%)	-	5 (100.0%)	0 (0.0%)
Chi-squared[3]	5 (6.0%)	0 (0.0%)	-	0 (0.0%)	5 (16.7%)
Proportional odds/cumulative logistic regression	37 (44.0%)	18 (36.7%)	-	0 (0.0%)	19 (63.3%)
Hodges-Lehmann estimator	3 (3.6%)	1 (2.0%)	-	0 (0.0%)	2 (6.7%)
Bootstrapping	21 (25.0%)	15 (30.6%)	-	5 (100.0%)	1 (3.3%)
Logistic regression	1 (1.2%)	0 (0.0%)	-	0 (0.0%)	1 (3.3%)
Quantile regression	6 (7.1%)	4 (8.2%)	-	2 (40.0%)	0 (0.0%)
Generalizing estimating equation with identity link	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Time-to-event analysis[4]	3 (3.6%)	0 (0.0%)	-	0 (0.0%)	3 (10.0%)
Generalized linear (mixed) model, log-binomial	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Kryger-Jensen/Lange test	8 (9.5%)	4 (8.2%)	-	4 (80.0%)	0 (0.0%)
Hurdle-negative binomial regression	6 (7.1%)	6 (12.2%)	-	0 (0.0%)	0 (0.0%)
Poisson/negative binomial regression	2 (2.4%)	2 (4.1%)	-	0 (0.0%)	0 (0.0%)
Binomial proportion test/exact binomial confidence interval	1 (1.2%)	0 (0.0%)	-	0 (0.0%)	1 (3.3%)
Simple calculation	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Zero-one-inflated beta regression	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Beta-binomial regression	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Generalized additive model with zero-inflated/zero-one-inflated beta distribution	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with zero-one-inflated beta distribution	2 (2.4%)	2 (4.1%)	-	0 (0.0%)	0 (0.0%)
Generalized linear model, identity-binomial	1 (1.2%)	0 (0.0%)	-	0 (0.0%)	1 (3.3%)
Generalized linear (mixed) model, logit link and G-computation	1 (1.2%)	1 (2.0%)	-	0 (0.0%)	0 (0.0%)
Generalizing estimating equation with logit link	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Probabilistic index	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Descriptive only[5]	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Descriptive for the entire trial only[5]	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)
Unclear[6]	0 (0.0%)	0 (0.0%)	-	0 (0.0%)	0 (0.0%)

Statistical methods used for all outcomes in trials restricted to patients with suspected or confirmed COVID-19, stratified by outcome type in drug trials only; statistical methods are displayed in the same order as in **Table 4** in the main text that includes all trials (i.e., decreasing order of use in all trials). As many outcomes were assessed multiple times, the numbers/percentages do not sum to the total number of outcomes/100%. Frequentist/Bayesian variants of the same statistical methods summarised together.

[1] Includes all non-parametric tests of continuous data, e.g. Mann-Whitney U/Wilcoxon, van Elteren, Kruskal-Wallis, and the Wilcoxon rank sum test for clustered data using Rosner-Glynn-Lee method.

[2] Includes linear regression and all variants including T-tests, analysis of (co)-variance, linear mixed models, etc., also including these methods used on transformed data.

[3] Includes the Chi-squared test and related tests for contingency tables, e.g., Fisher's exact test, and the Cochran-Armitage test for trend.

[4] includes the Cox proportional hazards model, the Kaplan Meier estimator, the log-rank test, the Fine-Gray model, and unspecified survival analysis methods.

[5] Outcomes only counted in one of these categories if no inferential statistical method was used to compare intervention groups, i.e., if only descriptive data for each intervention group or for the full trial were presented.

[6] Outcomes where some statistical method quantifying differences between intervention groups was used, but where it was not stated or possible to reasonably infer what this method was.

Supplemental Digital Content

PRISMA-ScR checklist

Completed PRISMA-ScR (6) checklist.

Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	1
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	3-4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	5-6
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	7
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	7
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	7-10
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	8-11, Supplemental Digital Content
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	8-10, Supplemental Digital Content
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	8-10
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	10-11, Supplemental Digital Content
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	10-11, Supplemental Digital Content

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
Critical appraisal of individual sources of evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	11/not applicable
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	11
RESULTS			
Selection of sources of evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	12, Fig. 1, Supplemental Digital Content
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	12, Tables, Supplemental Digital Content
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	Not applicable
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	12-15, Tables, Supplemental Digital Content
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	12-15, Tables, Supplemental Digital Content
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	16-18
Limitations	20	Discuss the limitations of the scoping review process.	18-19
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	20
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	2

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.

* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.

† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).

‡ The frameworks by Arksey and O'Malley and Levac and colleagues and the JBI guidance refer to the process of data extraction in a scoping review as data charting [see the PRISMA-ScR manuscript for references[3]].

§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).

References

1. Lefebvre C, Glanville J, Briscoe S, et al: Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, et al (Eds). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane 2021. Available at: https://www.training.cochrane.org/handbook. Accessed: May 10, 2021.

2. Granholm A, Kaas-Hansen BS, Kjær MN, et al: Patient-important outcomes other than mortality in recent ICU trials: protocol for a scoping review. *Acta Anaesthesiol Scand* 2021; 65:1002-1007.

3. Schoenfeld DA, Bernard GR: Statistical evaluation of ventilator-free days as an efficacy measure in clinical trials of treatments for acute respiratory distress syndrome. *Crit Care Med* 2002; 30:1772–1777.

4. Yehya N, Harhay MO, Curley MAQ, et al: Reappraisal of ventilator-free days in critical care research. Am J Respir Crit Care Med 2019; 200:828–836.

5. Contentin L, Ehrmann S, Giraudeau B: Heterogeneity in the Definition of Mechanical Ventilation Duration and Ventilator-Free Days. *Am J Respir Crit Care Med* 2014; 189:998–1002.

6. Tricco AC, Lillie E, Zarin W, et al: PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med* 201; 169: 467-473.

Analysis code

The remainder of this document contains the code used to conduct the analyses and create all tables included in the manuscript and this supplement. The paths in the first code chunk will have to be updated to match local paths if re-run by the reader. See "Data and code" above for additional details. The R code was written by Anders Granholm (andersgran@gmail.com) using the R Markdown format (rmarkdown.rstudio.com) and has been rendered with syntax highlighting, but without results included, to be included in this supplement.

Load libraries, define functions, load data

```r
library(tidyverse)
library(rlang)
library(skimr)

dta_gen_path <- "" # Path to "data general.csv"
dta_out_path <- "" # Path to "data outcomes.csv"
output_path <- "" # Path to folder where csv-files will be saved
```

Function to create stratified tables:

```r
strat_table <- function(.dta, strat_var, ..., round_n = 1, pct_suffix = "%",
                        fct_prefix = "- ", to = " - ", num_no_dec = FALSE) {
  vars <- quos(...) # Capture variables
  n_vars <- length(vars) # Number of variables
  # Get stratification variable options
  strat_var <- enquo(strat_var)
  strat_lvls <- .dta %>% pull(!!strat_var)
  if (is.factor(strat_lvls)) {
    strat_lvls <- levels(strat_lvls)
  } else {
    strat_lvls <- sort(unique(strat_lvls))
  }
  n_strat <- length(strat_lvls)

  n <- nrow(.dta)

  # Create base for results
  res_base <- as_tibble(matrix(rep(NA_character_, 3 + n_strat), nrow = 1),
                    .name_repair = "minimal")
  names(res_base)[c(1, 2, 3 + n_strat)] <- c("Variable",
                                        paste0("All (n = ", n, ")"),
                                        "Missing values")
  # Set column names
  strat_names <- as.character(strat_lvls)
  strat_n <- vector(mode = "numeric", n_strat)
  for (i in 1:n_strat){ # Set column names
    strat_n <- .dta %>% filter(!!strat_var == strat_lvls[i]) %>% nrow()
    names(res_base)[2 + i] <- paste0( strat_names[i], " (n = ", round(strat_n,
0), ")")
  }
```

Supplemental Digital Content

```r
  # Go through all variables and fill in data
  res <- vector(mode = "list", length = n_vars)
  for (i in 1:n_vars){
    cur_res <- res_base
    cur_var <-.dta %>% pull(!!vars[[i]])

    # General stuff
    m <- sum(is.na(cur_var)) # Number of missing values for current variable
    strat_var_m <- .dta %>% filter(!is.na(!!vars[[i]])) %>% pull(!!strat_var) #
Select outcome categories matching non-missing values for cur_var
    cur_var <- cur_var[!is.na(cur_var)] # Select only non-missing values for cu
r_var
    cur_res[1, 1] <- rlang::as_label(vars[[i]])

    if (is.numeric(cur_var)){ # Do this for numeric variables
      # All patients
      dec <- if (num_no_dec) 0 else round_n
      q <- quantile(cur_var, probs = 0.25 * 0:4) # Median and IQR and range
      cur_res[1, 2] <- paste0(format(round(q[3], dec), nsmall = dec), " (",
                              format(round(q[2], dec), nsmall = dec), to,
                              format(round(q[4], dec), nsmall = dec), ") [",
                              format(round(q[1], dec), nsmall = dec), to,
                              format(round(q[5], dec), nsmall = dec), "]")

      # For each strata
      for (j in 1:n_strat){
        q <- quantile(cur_var[strat_var_m == strat_lvls[j]], probs = 0.25 * 0:4
, na.rm = TRUE)
        cur_res[1, 2 + j] <- paste0(format(round(q[3], dec), nsmall = dec), " (
",
                                    format(round(q[2], dec), nsmall = dec), to,
                                    format(round(q[4], dec), nsmall = dec), ")
[",
                                    format(round(q[1], dec), nsmall = dec), to,
                                    format(round(q[5], dec), nsmall = dec), "]"
)
      }

      # Add missingness
      cur_res[1, 3 + n_strat] <- paste0(m, " (", format(round(m / n * 100, roun
d_n) , nsmall = round_n), pct_suffix, ")")

    } else if (is.logical(cur_var)) {  # End numeric vars / start logical vars
      # All patients
      d <- cur_var
      cur_res[1, 2] <- paste0(sum(d), " (", format(round(sum(d) / length(d) * 1
00, round_n), nsmall = round_n), pct_suffix, ")")

      # For each strata
      for (j in 1:n_strat){
        d <- cur_var[strat_var_m == strat_lvls[j]]
```

```r
      d <- d[!is.na(d)] # Remove NAs introduced by missing levels of the stra
tification variable
      cur_res[1, 2 + j] <- paste0(sum(d), " (", format(round(sum(d) / length(
d) * 100, round_n), nsmall = round_n), pct_suffix, ")")
    }

    # Add missingness
    cur_res[1, 3 + n_strat] <- paste0(m, " (", format(round(m / n * 100, roun
d_n) , nsmall = round_n), pct_suffix, ")")

  } else if(is.factor(cur_var)) { # End logical vars / start factor vars
    lvls <- levels(cur_var) # Get factor levels

    # All patients
    cur_res[1, 2:(2 + n_strat)] <- "" # Emtpy fields

    # Add missingness
    cur_res[1, 3 + n_strat] <- paste0(m, " (", format(round(m / n * 100, roun
d_n) , nsmall = round_n), pct_suffix, ")")

    # For each level of the factor
    for (j in 1:length(lvls)) {
      cur_res <- bind_rows(cur_res, res_base) # Add a new row for each factor
level
      cur_res[1 + j, 1] <- paste0(fct_prefix, lvls[j]) # Add name of this lev
el
      cur_res[1 + j, 3 + n_strat] <- "" # No text here

      # For all patients
      d <- cur_var == lvls[j] # Find data matching the level of the factor
      d <- d[!is.na(d)] # Remove NAs introduced by missing levels of the stra
tification variable
      cur_res[1 + j, 2] <- paste0(sum(d), " (", format(round(sum(d) / length(
d) * 100, round_n), nsmall = round_n), pct_suffix, ")")

      # For each strata
      for (k in 1:n_strat) {
        dk <- cur_var[strat_var_m == strat_lvls[k]]
        d <- dk == lvls[j] # Find data matching levels of the factor
        d <- d[!is.na(d)] # Remove NAs introduced by missing levels of the st
ratification variable
        cur_res[1 + j, 2 + k] <- paste0(sum(d), " (", format(round(sum(d) / l
ength(d) * 100, round_n), nsmall = round_n), pct_suffix, ")")
      }

    } # End factor levels
  } # End factor vars

  # Save current result in the list
  res[[i]] <- cur_res
} # End looping through all vars
```

Supplemental Digital Content

```
  # Return all results merged
  bind_rows(res)
}
```

Load and skim general data:

```
dta_gen <- read_csv2(dta_gen_path) %>%
  mutate(across(c(journal, intervention_type), factor),
         across(c(original_found_outside_search, protocol_available, covid19),
~ case_when(.x == "y" ~ TRUE,

.x == "n" ~ FALSE,

TRUE ~ NA)))
 skim(dta_gen)
```

Load and skim outcome data:

```
dta_out <- read_csv2(dta_out_path) %>%
  mutate(co_primary = case_when(co_primary == "y" ~ TRUE,
                                co_primary == "n" ~ FALSE,
                                TRUE ~ NA),
         across(c(outcome, type, prioritisation, time_point, death_handling, mi
ssing_handling), factor)) %>%
  mutate(across(c(where(is.factor), -time_point), fct_infreq))
skim(dta_out)
```

Trial-level characteristics

```
nrow(dta_gen)
dta_gen <- dta_gen %>%
  mutate(across(c(journal, intervention_type), fct_infreq)) %>%
  full_join(count(dta_out, id, name = "outcomes_n"), by = "id") %>%
  full_join( # Add new variable - any (co-)primary patient-important outcome ot
her than mortality included
    dta_out %>% group_by(id) %>% summarise(any_co_primary = sum(co_primary) > 0
),
    by = "id")
nrow(dta_gen)

dta_gen %>%
  strat_table(strat_var = intervention_type,
              journal, year, countries_n, centres_n, patients_n, outcomes_n, pr
otocol_available,
              any_co_primary, covid19,
              num_no_dec = TRUE) %>%
  select(-`Missing values`) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Trial level results.csv"))
```

Supplemental Digital Content

Outcomes overall and by outcome type

Simplify/recode outcome data

Simplify prioritization:

```
# First, print and save all original prioritisations, counted:
dta_out %>%
  count(prioritisation) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Original prioritisations counted.csv")
)

# First, print and save all original assessment time-points by outcome types, c
ounted:
dta_out %>%
  count(prioritisation, type) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Original prioritisations by type count
ed.csv"))

# Then recode to simplify:
dta_out <- dta_out %>%
  mutate(prioritisation = case_when(prioritisation %in% c("NR (individual categ
ories/domains secondarily reported but not analysed)", "NR (individual componen
ts reported secondarily)") ~ "NR (individual categories/domains/components repo
rted secondarily, but not considered or analysed as primary outcomes)",
                                    TRUE ~ as.character(prioritisation))) %>%
  mutate(prioritisation = fct_infreq(factor(prioritisation)))

# Count new values overall:
dta_out %>%
  count(prioritisation)
```

Simplify assessment time points:

```
# First, print and save all original assessment time-points, counted:
dta_out %>%
  count(time_point) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Original assessment time-points counte
d.csv"))

# Then, print and save all original assessment time-points by outcome types, co
unted:
dta_out %>%
  count(time_point, type) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Original assessment time-points by typ
e counted.csv"))

# Then recode to simplify:
```

Supplemental Digital Content

```r
dta_out <- dta_out %>%
  mutate(time_point = case_when(time_point %in% c("day 28", "day 28 censored at
hospital discharge", "day 28 (but only in-ICU)", "day 28 censored at ICU discha
rge") ~ "day 28",
                                time_point %in% c("1 year", "12 months", "12 +/
- 4 months (>16 months in some patients)", "day 365") ~ "1 year",
                                time_point %in% c("day 90", "3 months") ~ "day
90/3 months",
                                time_point %in% c("day 180", "6 months") ~ "day
180/6 months",
                                time_point %in% c("day 14", "day 14 censored at
ICU discharge") ~ "day 14",
                                time_point %in% c("day 21", "day 21 (assumed al
ive and without life support after hospital discharge)") ~ "day 21",
                                time_point %in% c("hospital discharge", "in-hos
pital") ~ "hospital discharge/in-hospital",
                                time_point %in% c("day 7", "day 7 censored afte
r patient has been off pressors for 4 hours") ~ "day 7",
                                time_point %in% c("ICU discharge", "day of ICU
discharge or the following day") ~ "ICU discharge",
                                time_point %in% c("3 months after ICU discharge
", "90 days post ICU discharge") ~ "90 days/3 months after ICU discharge",
                                TRUE ~ as.character(time_point))) %>%
  mutate(time_point = fct_infreq(factor(time_point)))

# Count new values overall:
dta_out %>%
  count(time_point)
```

Simplify/recode death handling:

```r
# First, print and save all original death handling values, counted:
dta_out %>%
  count(death_handling) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Original death handling values counted
.csv"))

# Then, print and save all original death handling values by outcome types, cou
nted:
dta_out %>%
  count(death_handling, type) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Original death handling values by type
counted.csv"))

# Then recode to simplify:
dta_out <- dta_out %>%
  mutate(death_handling = case_when(death_handling %in% c("Dead = 0 days", "Dea
d = 0") ~ "Dead = 0",
                                    death_handling %in% c("If dead while on lif
e support = 0 days, otherwise actual number of days", "Dead = 0 hours, but only
```

```
if on pressors at the time of dead", "Actual number of days, except days not co
unted if dead within 48 h after extubation OR intubated again") ~ "Death penali
sed in some cases",
                                  death_handling %in% c("Included in scale",
"Included in definition") ~ "Included in scale/definition",
                                  death_handling %in% c("Dead = worst value",
"Dead = worst value (time to liberation outcome)", "Dead = worst value (tie met
hod)", "Dead = worst value, ordinal") ~ "Dead = worst value",
                                  death_handling %in% c("Actual number of day
s", "Actual point score (no penalisation of death)") ~ "Actual number of days/v
alue (no penalisation of death)",
                                  TRUE ~ as.character(death_handling))) %>%
   mutate(death_handling = fct_infreq(factor(death_handling)))

# Count new values overall:
dta_out %>%
   count(death_handling)
```

Recode missing data handling to separate CCA into CCA with missing data and without missing data:

```
# First, print and save all original missing handling values, counted:
dta_out %>%
   count(missing_handling) %>%
   print() %>%
   write_csv2(file = paste0(output_path, "Original missing handling values count
ed.csv"))

# Then, print and save all original missing handling values by outcome types, c
ounted:
dta_out %>%
   count(missing_handling, type) %>%
   print() %>%
   write_csv2(file = paste0(output_path, "Original missing handling values by ty
pe counted.csv"))

# Then recode to simplify/add new category:
dta_out <- dta_out %>%
   mutate(missing_handling = if_else(missing_pct == 0 & missing_handling == "CCA
",
                                  "CCA (no missingness)",
                                  as.character(missing_handling))) %>%
   mutate(missing_handling = if_else(missing_handling %in% c("MI and Best-worst/
worst-best and CCA", "CCA and MI and Best-worst/worst-best"), "MI and CCA and B
est-worst/worst-best", missing_handling)) %>%
   mutate(missing_handling = fct_infreq(factor(missing_handling)))

# Count new values overall:
dta_out %>%
   count(missing_handling)
```

Supplemental Digital Content

Outcomes overall - general results

Full table (only):

```
dta_out %>%
  strat_table(strat_var = type,
              co_primary, prioritisation, time_point,
              death_pct, death_handling, missing_pct, missing_handling) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "General outcome-level results - ALL.cs
v"))
```

Most used outcomes

Full table, stratified by outcome type:

```
l <- list()

for (i in seq_along(levels(dta_out$type))) {
  cur_level <- levels(dta_out$type)[i]
  cur_dta <- dta_out %>%
    filter(type == cur_level)
  l[[i]] <- bind_rows(tibble(outcome = paste("#####", cur_level, "#####"),
                             n = nrow(cur_dta),
                             pct_type = "100.0%",
                             pct_all = paste0(format(round(nrow(cur_dta) / nrow
(dta_out) * 100, digits = 1), nsmall = 1), "%")),
                      cur_dta %>%
                        count(outcome) %>%
                        mutate(pct_type = paste0(format(round(n / sum(n) * 100,
digits = 1), nsmall = 1), "%"),
                               pct_all = paste0(format(round(n / nrow(dta_out)
* 100, digits = 1), nsmall = 1), "%")))
}

bind_rows(l) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Most frequent outcomes by type - ALL.c
sv"))
```

Infrequent categories (<5) lumped together, stratified by outcome type:

```
l <- list()

for (i in seq_along(levels(dta_out$type))) {
  cur_level <- levels(dta_out$type)[i]
  cur_dta <- dta_out %>%
    filter(type == cur_level)
  l[[i]] <- bind_rows(tibble(outcome = paste("#####", cur_level, "#####"),
                             n = nrow(cur_dta),
                             pct = "100.0%",
                             pct_all = paste0(format(round(nrow(cur_dta) / nrow
(dta_out) * 100, digits = 1), nsmall = 1), "%")),
```

Supplemental Digital Content

```
                    cur_dta %>%
                      mutate(outcome = fct_lump_min(outcome, 5)) %>%
                      count(outcome) %>%
                      mutate(pct = paste0(format(round(n / sum(n) * 100, digi
ts = 1), nsmall = 1), "%"),
                             pct_all = paste0(format(round(n / nrow(dta_out)
* 100, digits = 1), nsmall = 1), "%")))
}

bind_rows(l) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Most frequent outcomes by type - min 5
.csv"))
```

Note: some outcomes are not exactly similar but relatively similar; this is not directly considered in these tables.

Effect measures and statistical methods

Function to capitalize first letter only (for formatting):

```
cap_first <- function(x) {
  ifelse(is.na(x), NA,
         paste0(str_to_upper(str_sub(x, 1, 1)), str_sub(x, 2, -1)))
}
```

Extract and print all different effect measures:

```
all_effect_measures <- dta_out %>%
  mutate(eff_list = str_split(effect_measures, " and ")) %>%
  pull(eff_list) %>%
  unlist() %>%
  cap_first() %>%
  table(useNA = "always") %>%
  sort(decreasing = TRUE) %>%
  names()

# Move "Descriptive" and NA to the end
all_effect_measures <- c(na.omit(all_effect_measures[!str_detect(all_effect_mea
sures, "Descriptive") & !is.na(all_effect_measures)]), na.omit(all_effect_measu
res[str_detect(all_effect_measures, "Descriptive")]), NA) %>%
  unique() %>%
  print()
```

Extract and print all different statistical methods used:

```
# First, print all possible values before merging:
dta_out %>%
  mutate(eff_list = str_split(stat, " and ")) %>%
  pull(eff_list) %>%
  unlist() %>%
  cap_first() %>%
  table(useNA = "always") %>%
```

```r
  sort(decreasing = TRUE) %>%
  names()

# Then merge selected values
dta_out <- dta_out %>%
  mutate(stat = str_replace_all(stat, fixed("GAM with ZIB distribution", ignore
_case = TRUE), "GAM with ZIB/ZOIB distribution"),
         stat = str_replace_all(stat, fixed("GAM with ZOIB distribution", ignor
e_case = TRUE), "GAM with ZIB/ZOIB distribution"),
         stat = str_replace_all(stat, fixed("POLR", ignore_case = TRUE), "POLR
(including Bayesian cumulative logistic regression)"),
         stat = str_replace_all(stat, fixed("Bayesian cumulative logistic regre
ssion model", ignore_case = TRUE), "POLR (including Bayesian cumulative logisti
c regression)"),
         stat = str_replace_all(stat, fixed("Bootstrapping", ignore_case = TRUE
), "Bootstrapping (all variants, including Bayesian bootstrapping)"))

# Filter and print all values after merging:
all_stat <- dta_out %>%
  mutate(eff_list = str_split(stat, " and ")) %>%
  pull(eff_list) %>%
  unlist() %>%
  cap_first() %>%
  table(useNA = "always") %>%
  sort(decreasing = TRUE) %>%
  names()

cat("\n\nAfter merging:\n\n")

# Move "Descriptive" and NA to the end
all_stat <- c(na.omit(all_stat[!str_detect(all_stat, "Descriptive") & !is.na(al
l_stat)]), na.omit(all_stat[str_detect(all_stat, "Descriptive")]), NA) %>%
  unique() %>%
  print()
```

Function to tabulate effect measures and statistical methods:

```r
stat_eff_single <- function(l, vals, lab) {
  n <- length(l)
  res <- character(length(vals))
  for (i in seq_along(vals)) {
    cur_val <- vals[i]
    if (!is.na(cur_val)) {
      contains <- sapply(l, function(x) cur_val %in% cap_first(x))
    } else {
      contains <- sapply(l, function(x) all(is.na(x)))
      vals[i] <- "Unclear"
    }

    res[i] <- if (length(contains) > 0) {
      paste0(sum(contains), " (", format(round(mean(contains) * 100, digits = 1
), nsmall = 1), "%)")
```

```
    } else {
      "-"
    }

  }

  setNames(data.frame(vals, res), c("Value", paste0(lab, " (n = ", n, ")"))))
}

stat_eff_table <- function(.dta, .var, vals) {
  .dta$x <- str_split(.dta[[as_label(ensym(.var))]], " and ")

  res <- list()
  res[[1]] <- stat_eff_single(.dta$x, vals, "All")

  for (i in seq_along(levels(.dta$type))) {
    cur_lvl <- levels(.dta$type)[i]
    cur_dta <- .dta %>%
      filter(type == cur_lvl)
    res[[1 + i]] <- stat_eff_single(cur_dta$x, vals, cur_lvl)
  }

  res %>%
    reduce(full_join, by = "Value")
}
```

Table with effect measures:

```
dta_out %>%
  stat_eff_table(effect_measures, all_effect_measures) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Effect measures.csv"))
```

Table with statistical methods:

```
dta_out %>%
  stat_eff_table(stat, all_stat) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Statistical methods.csv"))
```

Separate tables according to each type of intervention

Most frequent outcomes (stratified by type), effect measures, statistical methods:

```
for (it in levels(dta_gen$intervention_type)) {

  # Subset datasets with outcomes
  cur_ids <- dta_gen %>%
    filter(intervention_type == it) %>%
    pull(id)
  cur_dta_out <- dta_out %>%
    filter(id %in% cur_ids)
```

Supplemental Digital Content

```r
# Most frequent outcomes by outcome type
l <- list()

for (i in seq_along(levels(dta_out$type))) {
  cur_level <- levels(dta_out$type)[i]
  cur_dta <- cur_dta_out %>%
    filter(type == cur_level)
  l[[i]] <- bind_rows(tibble(outcome = paste("#####", it, "-", cur_level, "##
###"),
                             n = nrow(cur_dta),
                             pct_type = "100.0%",
                             pct_all = paste0(format(round(nrow(cur_dta) / nr
ow(cur_dta_out) * 100, digits = 1), nsmall = 1), "%")),
                      cur_dta %>%
                        count(outcome) %>%
                        mutate(pct_type = paste0(format(round(n / sum(n) * 10
0, digits = 1), nsmall = 1), "%"),
                               pct_all = paste0(format(round(n /nrow(cur_dta_
out) * 100, digits = 1), nsmall = 1), "%")))
}

bind_rows(l) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Most frequent outcomes by type - ",
it, " - ALL.csv"))


# Table with effect measures
cur_dta_out %>%
  stat_eff_table(effect_measures, all_effect_measures) %>%
  setNames(c(paste("Value - ", it), names(.)[-1])) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Effect measures - ", it, ".csv"))


# Table with statistical methods
cur_dta_out %>%
  stat_eff_table(stat, all_stat) %>%
  setNames(c(paste("Value - ", it), names(.)[-1])) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Statistical methods - ", it, ".csv")
)
}
```

Separate tables according to COVID-19 or not

Trial-/outcome-level characteristics, most frequent outcomes (stratified by type), effect measures, statistical methods:

```r
# General trial-level results
dta_gen %>%
```

```r
  strat_table(strat_var = covid19,
              intervention_type, journal, year, countries_n, centres_n, patient
s_n, outcomes_n,
              protocol_available, any_co_primary,
              num_no_dec = TRUE) %>%
  select(-`Missing values`) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Trial level results - COVID-19.csv"))

# General outcome-level results
dta_out %>%
  mutate(covid19 = id %in% (dta_gen %>% filter(covid19) %>% pull(id))) %>%
  strat_table(strat_var = covid19,
              co_primary, prioritisation, time_point,
              death_pct, death_handling, missing_pct, missing_handling) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "General outcome-level results - COVID-
19.csv"))

# Outcomes, effect measures, statistical methods:
for (it in unique(dta_gen$covid19)) {

  # Subset datasets with outcomes
  cur_ids <- dta_gen %>%
    filter(covid19 == it) %>%
    pull(id)
  cur_dta_out <- dta_out %>%
    filter(id %in% cur_ids)


  # Most frequent outcomes by outcome type
  l <- list()

  for (i in seq_along(levels(dta_out$type))) {
    cur_level <- levels(dta_out$type)[i]
    cur_dta <- cur_dta_out %>%
      filter(type == cur_level)
    l[[i]] <- bind_rows(tibble(outcome = paste("##### COVID-19 ", it, "-", cur_
level, "#####"),
                               n = nrow(cur_dta),
                               pct_type = "100.0%",
                               pct_all = paste0(format(round(nrow(cur_dta) / nr
ow(cur_dta_out) * 100, digits = 1), nsmall = 1), "%")),
                        cur_dta %>%
                          count(outcome) %>%
                          mutate(pct_type = paste0(format(round(n / sum(n) * 10
0, digits = 1), nsmall = 1), "%"),
                                 pct_all = paste0(format(round(n /nrow(cur_dta_
out) * 100, digits = 1), nsmall = 1), "%")))
  }

  bind_rows(l) %>%
```

```
    print() %>%
    write_csv2(file = paste0(output_path, "Most frequent outcomes by type - COV
ID-19 ", it, " - ALL.csv"))


  # Table with effect measures
  cur_dta_out %>%
    stat_eff_table(effect_measures, all_effect_measures) %>%
    setNames(c(paste("Value - ", it), names(.)[-1])) %>%
    print() %>%
    write_csv2(file = paste0(output_path, "Effect measures - COVID-19 ", it, ".
csv"))


  # Table with statistical methods
  cur_dta_out %>%
    stat_eff_table(stat, all_stat) %>%
    setNames(c(paste("Value - ", it), names(.)[-1])) %>%
    print() %>%
    write_csv2(file = paste0(output_path, "Statistical methods - COVID-19 ", it
, ".csv"))
}
```

Separate tables according to year (<2020 vs. >=2020)

Trial-/outcome-level characteristics, most frequent outcomes (stratified by type), effect measures, statistical methods:

```
# Add dichotomized year variable
dta_gen$year_2020later <- dta_gen$year >= 2020
table(dta_gen$year_2020later)

# General trial-level results
dta_gen %>%
  strat_table(strat_var = year_2020later,
              intervention_type, journal, year, countries_n, centres_n, patient
s_n, outcomes_n,
              protocol_available, any_co_primary,
              num_no_dec = TRUE) %>%
  select(-`Missing values`) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "Trial level results - Year.csv"))

# General outcome-level results
dta_out %>%
  mutate(year_2020later = id %in% (dta_gen %>% filter(year_2020later) %>% pull(
id))) %>%
  strat_table(strat_var = year_2020later,
              co_primary, prioritisation, time_point,
              death_pct, death_handling, missing_pct, missing_handling) %>%
  print() %>%
  write_csv2(file = paste0(output_path, "General outcome-level results - Year.c
```

```r
sv"))

# Outcomes, effect measures, statistical methods:

for (it in unique(dta_gen$year_2020later)) {
  lab <- ifelse(it, "Year 2020-2022", paste0("Year ", min(dta_gen$year), "-2019
"))

  # Subset datasets with outcomes
  cur_ids <- dta_gen %>%
    filter(year_2020later == it) %>%
    pull(id)
  cur_dta_out <- dta_out %>%
    filter(id %in% cur_ids)


  # Most frequent outcomes by outcome type
  l <- list()

  for (i in seq_along(levels(dta_out$type))) {
    cur_level <- levels(dta_out$type)[i]
    cur_dta <- cur_dta_out %>%
      filter(type == cur_level)
    l[[i]] <- bind_rows(tibble(outcome = paste("##### ", lab, "-", cur_level, "
#####"),
                              n = nrow(cur_dta),
                              pct_type = "100.0%",
                              pct_all = paste0(format(round(nrow(cur_dta) / nr
ow(cur_dta_out) * 100, digits = 1), nsmall = 1), "%")),
                        cur_dta %>%
                          count(outcome) %>%
                          mutate(pct_type = paste0(format(round(n / sum(n) * 10
0, digits = 1), nsmall = 1), "%"),
                                 pct_all = paste0(format(round(n /nrow(cur_dta_
out) * 100, digits = 1), nsmall = 1), "%")))
  }

  bind_rows(l) %>%
    print() %>%
    write_csv2(file = paste0(output_path, "Most frequent outcomes by type - ",
lab, " - ALL.csv"))


  # Table with effect measures
  cur_dta_out %>%
    stat_eff_table(effect_measures, all_effect_measures) %>%
    setNames(c(paste("Value - ", it), names(.)[-1])) %>%
    print() %>%
    write_csv2(file = paste0(output_path, "Effect measures - ", lab, ".csv"))


  # Table with statistical methods
```

Supplemental Digital Content

```
  cur_dta_out %>%
    stat_eff_table(stat, all_stat) %>%
    setNames(c(paste("Value - ", it), names(.)[-1])) %>%
    print() %>%
    write_csv2(file = paste0(output_path, "Statistical methods - ", lab, ".csv"
))
}
```

Log date and session info

```
date()
sessionInfo()
```