MATTERS ARISING
https://doi.org/10.1038/s42256-020-00252-4

nature machine intelligence

Limited applicability of a COVID-19 specific mortality prediction rule to the intensive care setting

C. Dupuis[1]✉, E. De Montmollin[2,3], M. Neuville[4], B. Mourvillier[5], S. Ruckly[2] and J. F. Timsit[2,3]

ARISING FROM Yan et al. *Nature Machine Intelligence* https://doi.org/10.1038/s42256-020-0180-7 (2020)

We read the Article by Yan et al.[1] with great interest. The COVID-19 pandemic has established itself as a major burden on healthcare services worldwide. Scores or algorithms to optimize the use of healthcare resources are of paramount importance. Against this background, Yan et al. gathered samples from a cohort of 485 infected patients in the region of Wuhan, China with a high mortality rate of almost 40% and proposed a simple and operable decision rule based on lactic dehydrogenase (LDH), lymphocytes and high-sensitivity C-reactive protein (hs-CRP) to predict the occurrence of death in the following 10 days.

Since March 2020, France has also been confronted with the COVID-19 pandemic. The decision rule of Yan et al. could be used in our patients, but external repeatability would first be required. To validate the generalizability of the rule, we used data from Outcomerea, a French multicentre cohort of intensive care units (ICUs) involved in the management of patients critically ill with COVID-19. Methods for data collection and the quality of the database have been described in detail elsewhere[2]. Since the beginning of the COVID outbreak in France, a range of specific clinical and biological data for patients with COVID have also been recorded prospectively into this database.

We included 178 patients aged over 18 years who were admitted to the ICU from 1 March 2020 to 1 June 2020 with laboratory-confirmed COVID-19. Patients without a measurement of LDH, hs-CRP or lymphocytes during the first three days after ICU admission were excluded. The main characteristics of our cohort are reported in Table 1. Among the 178 patients, fever was the most common initial symptom (80.8%), followed by dyspnoea (74.2%), cough (63%) and fatigue (43.2%). The median time from symptoms onset to ICU admission was 10 days (range 7–12 days) and the median duration between hospital and ICU admission was 2 days (range 1–3 days). They had a median age of 61 years (range 52–69 years), a median Charlson comorbidity index of 1 (range 0–3) and a median sepsis-related organ failure assessment score (SOFA) score of 5 (range 4–8). The median LDH, hs-CRP levels and percentage of lymphocytes were 453 Ui l^{-1} (range 352–603 Ui l^{-1}), 166 mg l^{-1} (range 92.4–223 mg l^{-1}) and 9.6% (range 6.2–15%), respectively. The median ICU length of stay was 11 days (range 6–19 days). At days 14 and 28, the mortality rates were 18% and 34.2%. The results presented in Tables 2 and 3 show that the precision and accuracy of the decision rule were extremely low for the prediction of death.

Table 1 | Characteristics of the 178 patients of the Outcomerea database

Characteristics[a]	*N* = 178 patients
Age	61 [52; 69]
Sex (Male)	143 (80.4)
Body mass index (kg cm^{-2})	28.8 [25.6; 32.4]
Comorbidities	
At least one comorbidity	113 (63.4)
Charlson score	1 [0; 3]
Liver	6 (3.4)
Cardiovascular	50 (28)
Respiratory	23 (13)
Kidney	22 (12.4)
Immunosuppression	27 (15.2)
Symptoms on onset	
Fever	144 (80.8)
Cough	112 (63)
Fatigue	77 (43.2)
Dyspnoea	132 (74.2)
Diarrhoea	33 (18.6)
Chest distress	16 (9)
Anosmia	12 (6.8)
Arthralgia	17 (9.6)
Time from first symptoms to ICU admission (days)	10 [7; 12]
Time from hospital to ICU admission (days)	2 [1; 3]
Laboratory test on admission	
Neutrophils (×10^9 l^{-1})	5,950 [4,000; 9,200]
Lymphocytes (×10^9 l^{-1})	800 [580; 1,110]
Lymphocytes (%)	9.6 [6.2; 15]
High-sensitivity CRP (mg l^{-1})	165.6 [92.4; 223]

Continued

[1]Medical Intensive Care Unit, Gabriel Montpied University Hospital, Clermont-Ferrand, France. [2]UMR 1137 – IAME Team 5 – DeSCID: Decision Sciences in Infectious Diseases, Control and Care, INSERM/University Paris Diderot, Sorbonne Paris Cité, Paris, France. [3]Medical and Infectious Intensive Care Unit, Bichat Claude Bernard University Hospital, Paris, France. [4]Intensive Care Unit, Foch Hospital, Suresnes, France. [5]Medical Intensive Care Unit, Robert Debré University Hospital, Reims, France. ✉e-mail: cdupuis1@chu-clermontferrand.fr

Table 1 | Characteristics of the 178 patients of the Outcomerea database (continued)

Characteristics[a]	*N* = 178 patients
Ferritin (μg l^{-1})	1,213 [745; 2,008.8]
D-dimers (μg l^{-1})	1,300 [741; 3,706.8]
LDH (Ui l^{-1})	453 [352; 603]
Severity on admission	
$T > 39$ °C	68 (38.2)
Simplified Acute Physiology Score II	33.6 [25; 47]
Sepsis-related Organ Failure Assessment score	5 [4; 8]
Norepinephrine on admission	62 (34.8)
Glasgow Coma Scale < 15	37 (20.8)
Invasive mechanical ventilation on admission	83 (46.6)
PaO_2/FiO_2	179.2 [131.6; 243.6]
$PaO_2/FiO_2 < 200$ (ratio of arterial oxygen partial pressure (PaO_2 in mmHg) to fractional inspired oxygen (FiO_2))	157(88%)
Treatments on admission	
Lopinavir/ritonavir	72 (40.4)
Hydroxychloroquine	21 (11.8)
Corticosteroids	69 (38.8)
During ICU stay	
Invasive mechanical ventilation during ICU stay	115 (64.6)
Any nosocomial infection	60 (33.8)
Bacteremia	35 (19.6)
Hospital-acquired and ventilator-associated pneumonia	54 (30.4)
Outcomes	
ICU ventilatory free days	3 [1; 7]
ICU length of stay	11 [6; 19]
ICU death	58 (32.6)
Mortality at day 60	62 (34.8)

[a]Data are presented as *N* (%) or median [interquartile range, IQR].

Table 3 | Performance of the decision rule of Yan et al. on the French Outcomerea dataset

		Precision	Recall	F1 score	Support[a]
Day 7	**Survival**	1	0.14	0.24	23
	Death	0.06	1	0.11	155
	Accuracy			0.18	178
Day 14	**Survival**	0.87	0.14	0.24	23
	Death	0.21	0.92	0.35	155
	Accuracy			0.30	178
Day 28	**Survival**	0.83	0.16	0.27	23
	Death	0.37	0.93	0.53	155
	Accuracy			0.43	178

[a]Predicted number of patients.

The least bad results were obtained at day 28, with a precision of 37% (positive predictive value) and an accuracy of 43%, but a recall of 93% (negative predictive value). This decision rule lacked specificity in our preselected cohort of critically ill patients, which could compromise its routine use.

These results could be explained by the real specificity of our cohort. Indeed, only ~5% of patients with COVID-19 are admitted to ICU for acute hypoxemic respiratory failure (AHRF)[3]. Consequently, our ICU population did not include (1) the vast majority of pauci-symptomatic patients with very low LDH and hs-CRP serum levels and high lymphocyte counts (these patients have good outcomes) and (2) some of the most severely ill patients with high hs-CRP and LDH serum levels and low lymphocyte counts, who are not admitted to ICU because of therapeutic limitation (these patients have the worst outcomes). Thus, it is not surprising that the predictive rule of Yan et al. was not accurate in our cohort. However, their proposed biomarkers might be interesting for predicting ICU admission and also death for patients admitted to ICU, but with other thresholds. As a result, we believe that different rules should be adapted to different stages of the illness. For example, a decision tree could be rebuilt in the ICU to predict the occurrence of death. Furthermore, death might not be the most appropriate outcome—worsening of the disease could be better. Another decision rule could be built for patients admitted to the emergency room to predict worsening, that is, the occurrence of severe or critical types of COVID (COS-COVID)[4]. Finally, as already mentioned by Yan et al., we agree that, for the development of more rigorous prediction models, collaboration and sharing of well-documented individual data for COVID-19 are needed. The predictors already identified, such as LDH, hs-CRP and lymphocyte counts, should be considered as candidate predictors for new models[5].

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available in the Supplementary Information. Source data are provided with this paper.

Table 2 | Confusion matrix for the French Outcomerea dataset

	Day 7			Day 14			Day 28		
	True label								
	Survival	**Death**	**All**	**Survival**	**Death**	**All**	**Survival**	**Death**	**All**
Predicted label									
Survival	23	0	23	20	3	23	19	4	23
Death	146	9	155	122	33	155	98	57	155
All	169	9	178	142	36	178	117	61	178

Received: 26 May 2020; Accepted: 1 October 2020;
Published online: 12 November 2020

References

1. Yan, L. et al. An interpretable mortality prediction model for COVID-19 patients. *Nat. Mach. Intell.* **2**, 283–288 (2020).
2. Dupuis, C. et al. Effect of transfusion on mortality and other adverse events among critically ill septic patients: an observational study using a marginal structural Cox model. *Crit. Care Med.* **45**, 1972–1980 (2017).
3. Wu, Z. & McGoogan, J. M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72,314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* **323**, 1239–1242 (2020).
4. Jin, X. et al. Core outcome set for clinical trials on coronavirus disease 2019 (COS-COVID). *Engineering* https://doi.org/10.1016/j.eng.2020.03.002 (2020).
5. Wynants, L. et al. Prediction models for diagnosis and prognosis of Covid-19: systematic review and critical appraisal. *BMJ* **369**, m1328 (2020).

Author contributions

C.D and J.T.conceived and drafted the letter. C.D. and S.R. analysed the data. C.D., E.M., M.N and B.M. collected data.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s42256-020-00252-4.

Correspondence and requests for materials should be addressed to C.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Dupuis Claire

Last updated by author(s): Jun 2, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data were collected using MySQL
Data analysis	Data were analyzed using SAS® (Version 9.4; SAS Institute, Cary, NC, USA) and R (Version 3.4.0; R Core Team, Wien, Austria)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data are available in the Supplementary Information

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	102 patients
Data exclusions	No
Replication	No
Randomization	No
Blinding	No

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*

Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access & import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	*Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Validation	*Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.*

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	*State the source of each cell line used.*
Authentication	*Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.*

Mycoplasma contamination	*Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.*
Commonly misidentified lines (See ICLAC register)	*Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Palaeontology and Archaeology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	*For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	We analysed 102 cases of confirmed COVID-19 admitted between the 1st march 2020 and 16 April 2020 in one of 3 ICUs participating in the outcomerea database (Bichat university hospital (APHP, France), Foch hospital (Suresnes, France), Clermont Ferrand University hospital (France). Those patients were included only if the had hsCRP, LDH, Lymphocytes measurments on admission and outcome at Day 28. Only complete case analyses were performed. They had a median age of 59 years [range 52-67], a median Charlson comorbidity index of 1 [range 0-2] and a median SOFA score of 1 [range 0-3]. The median LDH and CRP levels and median percentage of lymphocytes were 503 Ui/L [range 367.7-651], 174 mg/L range [86.5-245] and 9% [range 5.3-13.8], respectively. At days 14 and 28, mortality rates were 15% and 32%, respectively.
Recruitment	It is a retrospective analysis of the prospective outcomerea cohort, including most of the patients with COVID-19 admitted in one of its participating ICU. We only included patients from 3 hospitals (Bichat university hospital, APHP - Clermont Ferrand university hospital, Foch hospital (France)), having LDH, hsCRP and lymphocytes recorded on admission and outcome at day 28 available at the time of the analysis.
Ethics oversight	In accordance with French law, the OutcomeReaTM database was declared to the "Commission Nationale de l'Informatique et des Libertés" (number 999262). The objectives of this data collection were approved by the institutional review board (number 5891) of the Clermont-Ferrand University Hospital (Clermont- Ferrand, France).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	In accordance with French law, the OutcomeReaTM database was declared to the "Commission Nationale de l'Informatique et des Libertés" (number 999262). The objectives of this data collec-tion were approved by the institutional review board (number 5891) of the Clermont-Ferrand University Hospital (Clermont- Ferrand, France).

Study protocol	Na
Data collection	Data came from the Outcomerea database, a French prospective cohort
Outcomes	Outcome at Day 7, 14 and 28 were recorded (death or alive)

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☐	Public health
☐	☐	National security
☐	☐	Crops and/or livestock
☐	☐	Ecosystems
☐	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☐	☐	Demonstrate how to render a vaccine ineffective
☐	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐	Increase transmissibility of a pathogen
☐	☐	Alter the host range of a pathogen
☐	☐	Enable evasion of diagnostic/detection modalities
☐	☐	Enable the weaponization of a biological agent or toxin
☐	☐	Any other potentially harmful combination of experiments and agents

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*

Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g.*

Normalization template

original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.

Noise and artifact removal

Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference
(See Eklund et al. 2016)

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.