

## 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 [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

### 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   |
|-------------------------------------|---|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> 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) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> 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

*Symlur Transcript and Analytics tool for extraction of tweets.*

Data analysis

*Microsoft Excel.*

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/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

*The data used in this study is publicly available on Twitter, Facebook and Macmillan.org.uk.*

### 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

## Life sciences study design

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

Sample size	<i>Not applicable</i>
Data exclusions	<i>Not applicable</i>
Replication	<i>Not applicable</i>
Randomization	<i>Not applicable</i>
Blinding	<i>Not applicable</i>

## Behavioural & social sciences study design

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

Study description	<i>Collection, classification and frequency analysis of qualitative data posted about lung cancer on three different social media platforms.</i>
Research sample	<i>Posts made by registered users of three social media platforms i.e. Twitter, Lung Cancer Support Group on Facebook and lung cancer discussion forum on Macmillan.org.uk.</i>
Sampling strategy	<i>Sampling of posts on the social media platforms: For the sample of posts from Twitter, the sampling was done through the use of hashtags and use of Symplur's Transcript and Analytics tool, for which the sampling procedure is unknown. For the Lung Cancer Support Group on Facebook and lung cancer discussion forum on Macmillan.org.uk, all posts were included in the sample from the one month period.</i>
Data collection	<i>Data was collected manually by one researcher using a computer. The researcher was not blind to the experimental condition that was being investigated.</i>
Timing	<i>Data posted during October 2017, was collected from all three social media platforms in December 2017.</i>
Data exclusions	<i>No data was excluded during the collection stage. Non English language posts were excluded during the categorisation stage.</i>
Non-participation	<i>No participants dropped out or declined to participate.</i>
Randomization	<i>Participants were not allocated to experimental groups other than social media platform from which the data was sourced.</i>

## Ecological, evolutionary & environmental sciences study design

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

Study description	<i>Not applicable</i>
Research sample	<i>Not applicable</i>
Sampling strategy	<i>Not applicable</i>

Data collection	<i>Not applicable</i>
Timing and spatial scale	<i>Not applicable</i>
Data exclusions	<i>Not applicable</i>
Reproducibility	<i>Not applicable</i>
Randomization	<i>Not applicable</i>
Blinding	<i>Not applicable</i>
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

## Field work, collection and transport

Field conditions	<i>Not applicable</i>
Location	<i>Not applicable</i>
Access and import/export	<i>Not applicable</i>
Disturbance	<i>Not applicable</i>

## 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	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used	<i>Not applicable</i>
Validation	<i>Not applicable</i>

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	<i>Not applicable</i>
Authentication	<i>Not applicable.</i>
Mycoplasma contamination	<i>Not applicable</i>

Commonly misidentified lines  
(See [ICLAC](#) register)

*Not applicable*

## Palaeontology

Specimen provenance

*Not applicable*

Specimen deposition

*Not applicable*

Dating methods

*Not applicable*

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

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

*Study did not involve laboratory animals.*

Wild animals

*Not applicable*

Field-collected samples

*Not applicable*

Ethics oversight

*Not applicable*

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

*See above*

Recruitment

*Participants have a user account for the social media platforms included in the study.*

Ethics oversight

*Relevant ethical guidelines were applied and approval was provided by the University of Edinburgh IRB.*

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

*Not applicable*

Study protocol

*Not applicable*

Data collection

*Not applicable*

Outcomes

*Not applicable*

## 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 <i>May remain private before publication.</i>	<i>Not applicable</i>
Files in database submission	<i>Not applicable</i>
Genome browser session (e.g. <a href="#">UCSC</a> )	<i>Not applicable</i>

## Methodology

Replicates	<i>Not applicable</i>
Sequencing depth	<i>Not applicable</i>
Antibodies	<i>Not applicable</i>
Peak calling parameters	<i>Not applicable</i>
Data quality	<i>Not applicable</i>
Software	<i>Not applicable</i>

## 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	<i>Not applicable</i>
Instrument	<i>Not applicable</i>
Software	<i>Not applicable</i>
Cell population abundance	<i>Not applicable</i>
Gating strategy	<i>Not applicable</i>

- 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	<i>Not applicable</i>
Design specifications	<i>Not applicable</i>
Behavioral performance measures	<i>Not applicable</i>

## Acquisition

Imaging type(s)	<input type="text" value="Not applicable"/>
Field strength	<input type="text" value="Not applicable"/>
Sequence & imaging parameters	<input type="text" value="Not applicable"/>
Area of acquisition	<input type="text" value="Not applicable"/>
Diffusion MRI	<input type="checkbox"/> Used <input checked="" type="checkbox"/> Not used
Parameters	<input type="text" value="Not applicable."/>

## Preprocessing

Preprocessing software	<input type="text" value="Not applicable"/>
Normalization	<input type="text" value="Not applicable"/>
Normalization template	<input type="text" value="Not applicable"/>
Noise and artifact removal	<input type="text" value="Not applicable"/>
Volume censoring	<input type="text" value="Not applicable"/>

## Statistical modeling & inference

Model type and settings	<input type="text" value="Not applicable"/>
Effect(s) tested	<input type="text" value="Not applicable"/>
Specify type of analysis:	<input type="checkbox"/> Whole brain <input type="checkbox"/> ROI-based <input type="checkbox"/> Both
Anatomical location(s)	<input type="text" value="Not applicable"/>
Statistic type for inference (See <a href="#">Eklund et al. 2016</a> )	<input type="text" value="Not applicable."/>
Correction	<input type="text" value="Not applicable"/>

## Models & analysis

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Functional and/or effective connectivity
<input checked="" type="checkbox"/>	<input type="checkbox"/> Graph analysis
<input checked="" type="checkbox"/>	<input type="checkbox"/> Multivariate modeling or predictive analysis
Functional and/or effective connectivity	<input type="text" value="Not applicable"/>
Graph analysis	<input type="text" value="Not applicable"/>
Multivariate modeling and predictive analysis	<input type="text" value="Not applicable"/>