nature portfolio

Corresponding author(s): Yan Guo, Tie-lin Yang, Kun Zhang

Last updated by author(s): Aug 6, 2024

# Reporting Summary

Nature Portfolio 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 Portfolio 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<br>*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<br>*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 | Xcalibur (Thermo Scientific Corp.) was used for LC-MS data acquisition. |
| Data analysis | No custom code or mathematical algorithm was used in the methods. All statistical analyses were conducted in R (version 3.6.1) using published libraries and functions. Metabolomics data were processed using XCMS. Besides, we conducted OPLS-DA model, DSPC network analyses and pathway enrichment analyses using MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/). The effect size of statistical test was calculated using Computation of Effect Sizes (http://www.psychometrica.de/effect_size.html) website, with Cohen's D value was used for Student's T-test and Eta squared (η2) was used for Mann-Whitney U-test. |

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 Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All relevant data support the key findings of this study are available within the article and its Supplementary Information/Source Data file. The raw intensity values

and processed data matrix of metabolomics and proteomics are available in Supplementary Data. The mass spectrometry metabolomics data generated in this study have been deposited in the OMIX, China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences (https://ngdc.cncb.ac.cn/omix) under accession code OMIX005819. The mass spectrometry proteomics data generated in this study have been deposited in the ProteomeXchange (https://proteomecentral.proteomexchange.org) Consortium via iProX repository with the dataset identifier PXD054131. Raw data for clinical information are available from the corresponding author upon request. All data in this study is only allowed for academic use.

## Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

| | |
|---|---|
| Reporting on sex and gender | We didn't consider sex of participants before recruitment. As shown in Supplementary Data 1 in manuscript, the sex distribution of the sample was relatively balanced based on self-reported (female: 54.66% in the discovery cohort, 52% in the validation cohort). There were no sex-based analyses in our study and we corrected all metabolomics, proteomics and clinical data for sex as a covariate. |
| Reporting on race, ethnicity, or other socially relevant groupings | All participants included in this study were Chinese Han with age over 18 years. We adopted the following measures to control confounding variables in our analyses: 1) all patients were not on anticoagulants or antiplatelet drugs before sampling and fasting venous blood samples were obtained from each patient less than 12 hours after hospitalization; 2) patients with major diseases or known metabolic diseases were excluded, including type 2 diabetes, coronary artery disease, hypertension, obesity, kidney or liver diseases, hyperthyroidism, dyslipidemia, and cancers; 3) patients with chronic use of medications affecting metabolism (hormone replacement therapy and corticosteroid therapy) were excluded; 4) we mainly focused on traumatic fractures (falling during physical activity and biking: 73.8%; traffic accident: 12.9%; falling down from a high altitude: 9.4%; and crushing by the heavy object: 3.9%), and patients with osteoporotic fractures or anti-osteoporosis drug intake were excluded to minimize potential influence on metabolic backgrounds; 5) we focused on lower extremity fractures (femur: 50%; tibia or fibula: 20.29%; pelvis: 6.62%; knee: 9.85%; ankle: 6.03%; multiple bone fractures: 7.21%) and included only patients with mild trauma severity based on the Injury Severity Score (ISS) (definition of mild injury: ISS ≤16) to reduce the potential impact of trauma severity or mode; 6) all cases were patients with distal DVT of lower extremity (DVT involved the gastroc-soleal veins: 63.89%; DVT involved the axial calf veins: 10.32%; DVT involved both the axial calf veins and the gastroc-soleal veins: 25.79%) and the thrombi in all cases were relatively large in size (> 10 millimeters × 2 millimeters, length × width). |
| Population characteristics | Population characteristics/demographics information has been summarized within the Supplementary Data 1. The metabolites, proteomics and clinical parameters profile data were first corrected with “sex”, “age”, and “trauma time” as covariates with linear model. |
| Recruitment | The patients in our study were recruited as part of the study which was registered in the Chinese Clinical Trial Registry (ChiCTR) (Registration number: ChiCTR1800017754). We carried out a nested case-control study design using a prospective cohort, which enrolled ~4000 Chinese Han participants diagnosed with acute traumatic fractures at the Department of Trauma Surgery, Honghui Hospital (Xi’an, China) from October 2018 to October 2022. Based on this prospective cohort, we adopted strict inclusion/exclusion criteria to select samples for our study. The inclusion and exclusion criteria of patients have been described in manuscript Methods. |
| Ethics oversight | The study was approved by the Ethics Committee of Xi’an Jiaotong University Honghui Hospital. All patients were provided written, informed consent before participating in the study. |

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

# 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 | With the strict inclusion/exclusion criteria, a total of 680 participants were enrolled in our study. For discovery cohort, 252 patients diagnosed with incident pt-DVT were selected as cases, and 328 patients without DVT were selected as controls, which were enrolled from October 2018 to December 2020. Subsequently, from March 2021 to July 2021, 50 pt-DVT patients and 50 controls were selected with the same recruitment criteria as a separate validation cohort. No statistical power calculations were used to determine the same size, and we enrolled as much patients based on our inclusion/exclusion criteria. It is difficult to perform sample size calculations for untargeted metabolomics and proteomics data due to the lack of pilot data and the highly multivariate nature of the data. It is generally accepted that over 100 samples within a study is a minimum sample size to derive sufficient statistical power. (Trivedi, D. K., Hollywood, K. A. & Goodacre, R. New Horizons in Translational Medicine Metabolomics for the masses: The future of metabolomics in a personalized world. New Horizons Transl. Med. 3, 294–305 (2017).) |
| Data exclusions | No data was excluded. |

| | |
|---|---|
| Replication | For metabolomics analysis, the metabolites were retained for analyses only if they were measured in all three batches of LC-MS detection. The results of prediction model with metabolites and clinical parameters were replicated in an independent validation population. In addition, the results of proteomics analyses revealed changes in the corresponding proteins of the metabolic pathways identified by metabolomics analyses. The above results confirmed that our findings were replicated. |
| Randomization | In this study, we conducted a nested case-control design. For discovery cohort, 252 patients were selected as cases and 328 patients were selected as controls. Subsequently, 50 pt-DVT patients and 50 controls were selected as validation cohort. With the same recruitment criteria, the controls were randomly selected from all participants matched by age and sex according to a 1:1 propensity matching at group level. The order of proteomics and metabolomics data acquisition by LC-MS is randomly disrupted. |
| Blinding | The samples in this study were labeled with an numeric ID whose annotation was kept blinded during data collection and analyses. We declare that the investigators conducting untargeted metabolomics and proteomics analyses were blinded to group allocation. |

# 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 |
| ☐ | ☒ Clinical data |
| ☒ | ☐ Dual use research of concern |
| ☒ | ☐ Plants |

## Methods

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

## Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJEguidelines for publication of clinical research and a completedCONSORT checklist must be included with all submissions.

| | |
|---|---|
| Clinical trial registration | The patients in our study were recruited as part of the study which was registered in the Chinese Clinical Trial Registry (ChiCTR) (Registration number: ChiCTR1800017754). It can be available in the WHO International Clinical Trials Registry Platform (ICTRP) (https://trialsearch.who.int/). |
| Study protocol | The study protocol can be available from the corresponding author (Prof. Yang, yangtielin@xjtu.edu.cn) upon reasonable request. Requests including a formal research proposal indicating the use of data and planned analyses, and it will be processed within two weeks. |
| Data collection | All participants enrolled in this study were diagnosed with acute traumatic fractures at the Department of Trauma Surgery, Honghui Hospital (Xi'an, China) from October 2018 to October 2022. For discovery cohort, 252 patients diagnosed with incident pt-DVT were selected as cases, and 328 patients without DVT were selected as controls, which were enrolled from October 2018 to December 2020. Subsequently, from March 2021 to July 2021, 50 pt-DVT patients and 50 controls were selected with the same recruitment criteria as a separate validation cohort. For metabolomics analysis, 580 plasma samples from the discovery cohort were completely randomized and analyzed in 2 batches from March 2020 to March 2021. The 100 plasma samples in the validation cohort were randomized and analyzed in a separate batch in August 2021. The proteomics analysis of 183 subjects was conducted in June 2022. |
| Outcomes | In this study, we identified the deep vein thrombosis events after trauma as the outcome. |

## Plants

| | |
|---|---|
| Seed stocks | NA |
| Novel plant genotypes | NA |
| Authentication | NA |