nature portfolio

Corresponding author(s): Charles Chiu

Last updated by author(s): 1/5/2023

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 *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

Viral PCR and hexon sequencing data were collected using manufacturer-provided software associated with the instrument. Metagenomic and whole-genome sequencing data were collected on Illumina sequencers (MiSeq or NextSeq Sequencing System) using manufacturer-provided software.

Data analysis

A cycle threshold cutoff of ≤40 was used to call a positive result by PCR. Sanger sequences were assembled and edited in Sequencher 5.2.4 (Gene Codes) and analyzed using the nucleotide BLAST aligner.

Binary base call (BCL) files generated by Illumina sequencers were simultaneously demultiplexed and converted to FASTQ files using bcl2fastq (version 2.20.0.422). Metagenomic sequencing data from all cases and controls were analyzed for viral nucleic acids using SURPI+ (v1.0.7-build.4), an automated bioinformatics pipeline for pathogen detection and discovery from metagenomic data that has been modified to incorporate enhanced filtering and classification algorithms. A threshold of ≥ 3 non-overlapping reads was used for calling a positive virus detection.

A custom script was used to assemble AdV and AAV2 genomes as follows. Briefly, raw FASTQ reads were filtered using BBDuk (version 38.87) for removal of adapters, primer sequences, and low-quality reads, and then HAdV-41 or AAV reads were identified by Bowtie2 alignment (parameters: -D 20 -R 3 -L11 -N 1) to a reference database consisting of 1,395 HAdV-41 or 3,600 AAV partial sequences / genomes, respectively. These aligned reads were then mapped to the HAdV-41 reference genome (accession DQ315364.2) or consecutively to AAV genomes 1-8. For all AAV genomes, the assembly with the highest breadth of coverage corresponded to the AAV2 reference genome (accession number NC_001401.2). Consensus assemblies were generated using iVar (parameters: -t 0.5 -m 1). AAV consensus genomes were further analyzed for shared mutations in the nucleotide and translated nucleotide (amino acid) sequences relative to the AAV2 reference by performing a multiple sequence alignment using the MAFFT algorithm (v7.388), followed by visualization of the alignment using Geneious

software (version 10.0.9).

Multiple sequence alignments were performed using MAFFT algorithm (v7.388) as implemented in Geneious (version 10.0.9). Nucleotide and amino acid phylogenetic trees were inferred using a maximum likelihood method with ultrafast bootstrap approximation as implemented in IQ-TREE (version 1.6.1). Trees were visualized using FigTree (version 1.4.4).

Statistical analyses were performed using the Python scipy package (version 1.5.2) and rstatix package (version 0.7.0) in R (version 4.0.3). Fisher's Exact Test was used to assess the association between variables. All statistical tests were conducted as two-sided at the 0.05 significance level.

Plots were generated using matplotlib (version 3.3.2), seaborn (version 0.11.0) and plotly (version 5.6.0) packages in Python software (version 3.7.12), Jupyter notebook (version 6.1.4), RStudio (version 1.4) and Adobe Illustrator (version 26.4.1) software.

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

Raw metagenomic sequencing data with human reads removed have been deposited in the NCBI Sequence Read Archive (Bioproject accession number PRJNA918667, under umbrella BioProject accession number PRJNA171119). Consensus genome files, alignment files, phylogenetic tree files and supplementary tables have been uploaded to a Zenodo data repository (doi: 10.5281/zenodo.7089581).

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Sex and gender were not considered in study design. Analysis was performed on aggregated data..
Population characteristics	In this study, 27 samples (21 whole blood, 2 plasma, 1 liver tissue, 1 nasopharyngeal swab, and 2 stool sample(s)) from 16 children with acute severe hepatitis of unknown etiology were analyzed. All children met the clinical case definition established by the CDC, including lack of a confirmed etiology, liver enzyme levels (aspartate aminotransferase (AST) or alanine aminotransferase (ALT)) >500 UL, age<10 years, and onset on or after October 1, 2021. Cases were enrolled from 6 states (Alabama, California, Florida, Illinois, North Carolina, and South Dakota) from October 1, 2021 to May 22, 2022. All 16 cases had positive testing for HAdV from blood, and thus HAdV infection was over-represented compared to the overall affected population, of which HAdV is detected in 45-90%. The median age of affected children was 3 years; 56% were female and 44% male. Controls (n=113) consisted of 78 whole blood, 1 serum, and 34 plasma samples. Many controls were enrolled from California (n=54) and Georgia (n=24) to be geographically similar to the cases, with the remaining controls enrolled from Ohio (n=12), Texas (n=14), and Washington (n=9). Sixty-nine of 113 controls (61.0%) were collected over the same time frame as the cases (i.e., collected between October 1, 2021 – May 22, 2022). Differences in age and gender between cases and controls were non-significant. The 113 controls consisted of 42 patients (37%) without hepatitis, 30 (26.5%) patients with acute hepatitis (ALT > 100 U/L) of defined etiology, 23 (21%) patients with acute gastroenteritis (12 with positive HAdV stool testing), and 18 (16%) blood donors.
Recruitment	This was a retrospective observational case-control study using all available samples from cases and controls. A severe acute hepatitis case enrolled in this study was a person under investigation (PUI) by local, state, or federal public health agencies, defined as a person <10 years of age with elevated (>500 U/L) aspartate aminotransferase (AST) or alanine aminotransferase (ALT), an unknown etiology for the hepatitis, and onset on or after October 1, 20218. All cases (n=16) were hospitalized with acute elevation in liver enzymes, asparate aminotransferase (AST) or alanine aminotransferase (ALT), and one or more of the following symptoms on presentation: nausea, vomiting, jaundice, generalized weakness, and abdominal pain. Cass were selected for inclusion into the current study if blood samples had previously tested positive for adenovirus by clinical testing, resulting in a selection bias that made case-control comparisons related to adenovirus detection not meaningful. Blood samples from cases meeting PUI criteria but testing negative for adenovirus were not available. Controls from University of California, San Francisco (UCSF) in California (n=54) included children hospitalized with hepatitis of defined etiology or another inflammatory or non-inflammatory condition. These controls were selected to be geographically similar (located within the same state) to cases from California. Remnant whole blood samples from controls from UCSF (n=54) were retrospectively biobanked and aliquoted with addition of 2X DNA/RNA Shield (Zymo Research) in a 1:1 ratio by volume and stored at -80°C until use. Controls from Children's Healthcare of Atlanta (n=24) in Georgia included children hospitalized with hepatitis of defined etiology or another inflammatory or non-inflammatory condition (n=18) or blood donors (n=6). These controls were selected to be geographically similar (located within a neighboring state with similar demographic characteristics) to the cases from

Alabama and Florida. Collected samples were stored at -80°C until use.

Serum or plasma samples from children enrolled in the New Vaccine Surveillance Network (NVSN) were also included as controls. Three groups of children were selected: children admitted for acute gastroenteritis who tested positive for adenovirus in the stool (n=12), children admitted for acute gastroenteritis who tested negative for adenovirus in the stool (n=11), and blood donors (n=12). These children were enrolled from three sites (Seattle, Houston, and Cincinnati) from June 2021 to May 2022. Collected samples were stored at -80°C until use.

Ethics oversight

Remnant clinical samples from cases with acute severe hepatitis were collected and analyzed under "no subject contact" protocols with waiver of informed consent approved by the institutional review boards (IRBs) of University of Alabama, Birmingham, California Department of Public Health, New York State Department of Health, and CDC. Whole blood samples from pediatric controls (age < 18) from Children's Healthcare of Atlanta were prospectively collected and analyzed under a protocol approved by the Emory IRB (STUDY00000723); parents or guardians of these children provided oral consent for study enrollment and collection and analysis of their samples. Remnant whole blood samples from pediatric controls (age < 18) at University of California, San Francisco (UCSF) were collected, biobanked, and analyzed under a "no subject contact" protocol with waiver of informed consent approved by the UCSF IRB (protocol no. 11-05519). A subset of the control samples was provided by the CDC from children enrolled in the National Vaccine Surveillance Network (NVSN) study. Approval for the NVSN study was obtained from the institutional review board at each participating site and from the Centers for Disease Control and Prevention (protocol no. 6164). Parents or guardians of eligible children provided written informed consent for participant enrollment. Blood specimens were also collected as leftover samples from clinical procedures.

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

No sample size was calculated. All available samples were processed. 27 samples (21 whole blood, 2 plasma, 1 liver tissue, 1 nasopharyngeal swab, and 2 stool sample(s)) from 16 pediatric patients with acute severe hepatitis of unknown etiology were analyzed. Affected children included in this analysis were admitted to tertiary care hospitals in 6 states (Alabama, California, Florida, Illinois, North Carolina, and South Dakota) from October 1, 2021 to May 22, 2022.
Controls (n=113) consisted of 78 whole blood, 1 serum, and 34 plasma samples. The majority of controls were enrolled from California (n=54) and Georgia (n=24) to be geographically similar to the cases, with the remaining controls enrolled from Ohio (n=12), Texas (n=14), and Washington (n=9). The controls consisted of a mix of remnant convenience samples (California) and samples from patients prospectively enrolled in a clinical (Georgia) or population-level surveillance (Ohio, Texas, and Washington) study.

Data exclusions

No data were excluded from the analysis.

Replication

Samples were processed only once but analyses were done on independent sample cohorts from 6 states (Alabama, California, Florida, Illinois, North Carolina, and South Dakota) and independent control cohorts from California, Georgia, Ohio, Texas and Washington state.

Randomization

Randomization is not applicable as this is a retrospective observational case-control study that used all available samples from cases (pediatric acute severe hepatitis) and controls.

Blinding

Researchers were blinded during processing as all samples were assigned a random unique ID. Cases and controls were analyzed in parallel. Samples were unblinded during analyses to interpret and visualize results and to determine associations between detection of virus and severe acute hepatitis.

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

Methods

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