UK Health Security Agency

Is adenovirus infection associated with severe acute seronegative hepatitis in young children in the UK: protocol for a frequency matched case-control study

Short title: adenovirus and acute hepatitis in young children: UK case control study protocol

UKHSA R&D REGG Ref NR0320

Contents

Amendment history

Amendment No.	Protocol Version	Date issued	Author(s) of changes	Details of Changes made
1	2.0	6 May 2022	Sema Mandal Andre Charlett Hannah Emmett Ruth Simmons Monica Desai	Analysis plan updated with sensitivity analysis; inclusion of text to cover data from DA; appendices for DA protocol variations; updated data flow
2	2.2	13 May 2022	Sema Mandal Conall Watson Georgina Ireland Andre Charlett	Included HHV6; hypothesis updated; appendices for data dictionary and matching table
3	2.3	17 May 2022	Conall Watson Sema Mandal Georgina Ireland	DA sample testing and analysis. Matching within age groups. SARS-CoV-2 test window. Data dictionary. Reporting of previously-adenovirus tested controls.
4	2.4	18 May 2022	Claire Neill Conall Watson Sema Mandal Georgina Ireland	NI protocol variations. Minor edits for clarity.
5	2.5	18 May 2022	Sonia Ribeiro Sema Mandal Conall Watson	Edits for accessibility and clarity

Background and rationale

Following initial reports from Scotland (1), on 6th April 2022, UKHSA issued a briefing note (BN 2022/034) describing cases of acute hepatitis without detection of hepatitis viruses (A-E) in young children from England, Scotland, Northern Ireland and Wales since the beginning of 2022. An enhanced national incident was stood up due to the increasing incidence and severity of disease in some cases (2,3). For case ascertainment, clinicians were asked to report patients presenting since 1 January 2022 with an acute hepatitis (non hepatitis A-E) with serum transaminase >500 IU/L (AST or ALT), under 16 years old, and close contacts of any age of a confirmed case with an acute hepatitis (non hepatitis A-E viruses).

As part of the incident investigation, UKHSA has collected and reviewed the initial case reporting forms (linelist), pathogen testing data on cases from samples taken at or around the time of admission, laboratory surveillance data on new diagnoses of routinely reported pathogens (SGSS), statistical exceedances of pathogens reported through SGSS, NHS and clinical data on paediatric emergency admissions and registrations for transplants for compatible clinical syndromes, and undertaken open-ended toxicology screens on cases. In addition, UKHSA has interviewed parents of cases with a “trawling” questionnaire. Parallel investigations in Scotland, Wales and Northern Ireland have been undertaken with data sharing at a UK level.

Key findings from these investigations to date are: (i) most common clinical presentation is with gastro-intestinal symptoms plus jaundice; (ii) common viruses circulating in children are currently higher than in previous years and there is a marked increase of adenovirus, particularly in the 1 to 4 age group; (iii) adenovirus was the most common virus or bacteria detected in around 75% of cases tested; (iv) 11% of cases were positive for SARS-CoV-2 at admission between January and April but there was a high background rate of COVID-19 during the investigation period; v) adeno-associated virus 2 (AAV-2) and Human Herpes Virus type 6 (HHV6) were detected via metagenomics in explanted livers of cases; and (vi) the toxicology investigation has not identified a clear pattern in the cases that would point to this being caused by a toxin alone.

The investigations indicate that there may be an association between adenovirus and acute, severe seronegative (non hepatitis A-E) hepatitis in young children. The working hypothesis in order of best fit are:

1 A normal adenovirus infection, due to one of:

   a) Abnormal susceptibility or host response which allows adenovirus infection to progress more frequently to hepatitis (whether direct or immunopathological), for example from lack of exposure during the coronavirus (COVID-19) pandemic.

   b) An exceptionally large wave of normal adenovirus infections, causing a very rare or under-recognised complication to present more frequently.

   c) Abnormal susceptibility or host response to adenovirus due to priming by a prior infection with SARS-CoV-2 (including Omicron restricted) or another infection.

   d) Abnormal susceptibility or host response to adenovirus due to a coinfection with SARS-CoV-2 or another infection.

   e) Abnormal susceptibility or host response to adenovirus due to a toxin, drug or environmental exposure.

2 A novel variant adenovirus, with or without a contribution from a cofactor as listed above.

3 A post-infectious SARS-CoV-2 syndrome (including an Omicron restricted effect).

4 A drug, toxin or environmental exposure.

5 A novel pathogen either acting alone or as a coinfection.

6 A new variant of SARS-CoV-2

To test the primary hypothesis of adenovirus being associated with a severe, acute seronegative (non hepatitis A-E) hepatitis in children, we propose a matched case-control study.

Objectives

Primary objective

To explore whether adenovirus infection is associated with acute, severe non A-E hepatitis in children 10 years and younger.

Secondary objectives

To explore whether SARS-CoV-2 concurrent or recent (in previous 2 weeks, or previous 90 days) PCR/LFD confirmed infection, as mono- or co-infection with adenovirus, is associated with acute, severe non A-E hepatitis in children 10 years and younger.

To explore whether HHV6 concurrent PCR confirmed infection as mono- or co-infection with adenovirus, is associated with acute, severe non A-E hepatitis in children 10 years and younger.

Methods

Sites

Cases are reported to UKHSA, Public Health Scotland (PHS), Public Health Wales (PHW), and the Public Health Agency (PHA) in Northern Ireland from NHS hospitals in the UK in response to the incident and case ascertainment alert. Hospital controls will be identified through residual EDTA blood samples from children ≤10 years of age who have been investigated for non-hepatitis acute illnesses in UK NHS hospitals. In England these will be patients who have been tested for meningococcal infection (and are meningococcal negative). In other UK nations, hospital controls may be identified by other routes; any protocol variations are summarised in Appendix 2.

Case definitions

The study case definitions are:

- case: a child 1 to 10 years old with an acute hepatitis (non A-E) and elevated transaminases (ALT/AST > 500 IU/L) reported to UK nations as part of the outbreak investigation
- control: a child 1 to 10 years of age who have had an EDTA blood sample (or plasma/serum sample, if there is no national sample repository for EDTA blood) taken on presentation to hospital for an acute illness, but not hepatitis

Controls will be frequency matched to cases on sex, age band (≤5 years, > 5 years), nation/supra-region, and month of sample.

Participants will be included from 1 January 2022.

Subject eligibility

Inclusion criteria

All below should be met for either controls or cases:

- age 1 to 10 years at sample date from 1 January 2022
- controls: have a residual EDTA blood sample (or plasma/serum sample, if there is no national sample repository for EDTA blood) drawn for pathogen testing for an acute non-hepatitis illness, which tested negative for that pathogen (for example tested negative for meningococcal infection for England controls)
- cases: have had adenovirus testing done on a blood/serum or other sample in a UK hospital at or around the time of admission
- England, Wales, Scotland or Northern Ireland resident

Exclusion criteria

If any of the below apply:

- age ≥11 years at sample date
- cases: not meeting the current outbreak case definitions and/or not tested for adenovirus
- controls: tested positive for the pathogen of interest (for example meningococcal infection) on EDTA/plasma/serum blood sample. Clinical testing for adenovirus was undertaken.
- controls: oncology/immunosuppressed patients, where known
- controls: acute hepatitis presentation
- not a resident of England, Scotland, Wales, or Northern Ireland

Procedures

Cases

Cases will be drawn from the UKHSA or devolved administration acute hepatitis non-A-E linelist and have documented evidence of adenovirus testing at or around the time of emergency department attendance or hospital admission. The minimal dataset for each case will include age, sex, postcode (to facilitate geospatial mapping and indices of deprivation), date of hospital sample collection (as a proxy for hospital admission and onset of symptoms), NHS number (for deduplication and linkage to other UKHSA held datasets for further demographics and pathogen tests).

Controls

Controls will be “recruited” from patients presenting with an acute illness for whom a blood EDTA sample was submitted to either the UKHSA Meningococcal Reference Unit (MRU) Manchester (England) for meningococcal PCR testing, or to a regional or national laboratory in one of the Devolved Administrations of Northern Ireland, Wales and Scotland.

Controls will be frequency matched to cases by age band, sex, supra-region (or devolved administration) and month of presentation. Cases will be placed in the categories formed by; aged ≤5 or aged >5 (with preferential matching within ages 1 to 2 and 3 to 5 years), males or females, North England or Midlands or southern England (including London), and calendar months January/February or March/April (see Appendix 3). Controls will be placed into the same categories, and within each category the controls will be allocated an unique number to form a category specific sampling frame. For each category, the number of cases will be calculated and 4 times this number of controls within the category will be selected using simple random sampling. If within a category there are fewer controls than 4 times the cases, then all controls will be selected.

The minimal dataset from the laboratory for each sample will include age, sex, postcode (to facilitate geospatial mapping and indices of deprivation), date of hospital sample collection (as a proxy for onset of symptoms), NHS number (for deduplication and linkage to other UKHSA held datasets).

Control data will be enriched by linkage to these NHS and UKHSA healthcare or surveillance datasets (including but not limited to Hospital Episode Statistics, NIMS (National Immunisation Management System, England), WIS (Wales Immunisation System), COVID-19 SGSS / USD, Wales National Laboratory System, and SGSS for other pathogens) to reduce missingness and misclassification of key covariates, before

being anonymised and unlinked for testing at a UKHSA or devolved administration laboratory. In England the linkage key will be held by a third party within UKHSA. REGG approval and Caldicott approval has been granted.

As samples are unlinked and anonymised, results cannot be given to participants and no opt-in is required. Residual samples will be stored securely by UKHSA or Devolved administration national laboratory and discarded 24 months after completion of testing and data analysis.

Any variation on procedures for devolved administrations is outlined in Appendix 2.

Laboratory tests

EDTA blood samples of at least 400µl from eligible controls will be unlinked and anonymised before being tested for adenovirus testing at a UKHSA laboratory or a national laboratory of a Devolved Administration. See Appendix 2 for variations for devolved administrations.

Adenovirus and HHV6 testing of samples from controls

Control samples will be tested for the presence of recent adenovirus, HHV6 (at the time of presentation / hospital admission) using validated PCR assays for each organism. Adenovirus panels from the National Institute for Biological Standards and Control (NIBSC) will be used to support this. Any adenovirus positive samples will be processed for typing at the UKHSA virus reference department or national laboratory of the Devolved Administration.

Adenovirus, HHV6 testing of samples from cases

Adenovirus, HHV6 testing of case samples is being done as part of clinical management in response to an outbreak at local hospitals and / or UKHSA or Devolved Administration national reference laboratories. This has been recommended by the UK national incident management teams. Any adenovirus positive samples will be processed for typing at the UKHSA virus reference department or national laboratory of the Devolved Administration.

Unlinking and anonymising controls

The following steps and attached flow chart (appendix 1) detail how unlinking and anonymising controls will be achieved in England:

- UKHSA Meningococcal Reference Unit (MRU) for meningococcal testing will share a list of EDTA blood samples (which are routinely collected for

meningococcal testing), and subsequently tested negative for meningococcal infection.

- a list of samples received for meningococcal testing and negative for meningococcal infection will be sent to UKHSA epi cell study team as potential controls. The list of controls will have full PII (Personal Identifiable Information) including name, DOB, NHS number, sex, age, address with postcode, to enable data enhancement (for ethnicity, date of hospital admission, COVID-19, test results and other test results) using established datasets (HES, NIMS, COVID-19 SGSS / USD and SGSS for other pathogens).
- list of controls will be frequency-matched to cases by age band, sex, and supra-region. Four controls to cases will be sought to increase power because of the small number of cases and the risk of insufficient residual sample in controls.
- list of selected controls will be given a unique study ID and name, date of birth, NHS number will be deleted from the dataset. Retained will be age, sex, ethnicity, postcode/region, and covid19 diagnosis information.
- the MRU laboratory number for each specimen and corresponding study ID for that individual will be stored separately.
- the list of MRU laboratory numbers for the matched controls will be sent back to MRU for specimen retrieval.
- the MRU will test the control samples for recent adenovirus, HHV-6 infection or send the control EDTA blood samples with the MRU laboratory number to a Colindale reference laboratory for adenovirus, HHV-6 testing.
- adenovirus results and the MRU laboratory number will be sent from the testing UKHSA laboratory to UKHSA's epi cell study team.
- the MRU laboratory number will be used to link to the relevant study number, once linked the laboratory number will be deleted leaving only the adenovirus result and study ID.
- the adenovirus results will be linked with the demographic information (age, sex, ethnicity, postcode/region) and COVID-19 diagnosis information using the study ID.
- all specimens held by the UKHSA MRU will be unlinked and anonymised prior to initiating laboratory processing for adenovirus testing. The resulting diagnosis from adenovirus testing will hold no clinical relevance to the controls.

Devolved administrations will develop protocols for unlinking and anonymising their controls (see Appendix 2).

Data management

Data on cases and controls will be collated by UKHSA epi cell study team and relevant national teams. The proposed data dictionary for England is in Appendix 4.

Data transfer

All England data will be transferred through secure email or using a secure web-based area managed by UKHSA. Anonymised data from devolved administrations will be shared with UKHSA following local linkage and cleaning procedures.

Storage and processing of data

Data on cases are stored on UKHSA servers, with access restricted to approved individual users in the study team.

Data linkage

Prior to anonymisation and unlinking, controls will be linked to other NHS and UKHSA healthcare or surveillance datasets held in UKHSA data stores to reduce missingness of demographic, laboratory and clinical data that are potential co-variates. These co-variates and linkage processes include: (i) ethnicity and date of hospital admission through linkage with Hospital Episode Statistics (HES) and Secondary Use Service (SUS); (ii) prior or concomitant COVID-19 infection confirmed by PCR through linkage to SGSS / unified COVID-19 dataset (UDS), and (iii) other pathogens via SGSS. Additional information may be available via HPZone records, but may not be available for all controls.

Devolved administrations will undertake data linkage using existing systems (see Appendix 2).

Analysis

Summary characteristics of the cases and controls will be compared.

Analyses will be conducted for:

- data limited to cases and controls with EDTA blood samples
- sensitivity analysis with cases and controls with non-whole blood samples (for example serum, faecal and respiratory)

We will also report on positivity of previously adenovirus-tested control group.

Measure of association

Odds ratio to compare the odds of testing adenovirus positive in the cases compared to the controls.

Multivariable analysis

Unconditional multivariable Firth penalised logistic regression analysis will be conducted with those demographic variables used in the frequency matching included to account for any residual confounding. The biases resulting from ignoring the frequency matched design within the analysis are relatively small and similar in magnitude to those seen on conditional analyses (4). It is reasonably well known that sparse data biases and possible separation affect odds ratio estimation in case-control studies with fewer than a few hundred cases, therefore Firth penalisation will be used as it adequately handles both of these problems (5). As the primary hypothesis being examined, the indicator predictor of adenovirus testing result will always be included in all regression models.

Effect modification with other co-infections, in particular SARS-CoV-2 and HHV6, will be explored by including interactions into models. Testing of model components will use the 5% significance level. There will be sufficient power for assessing the primary objective of an association with adenovirus, but it is noted for interactions there is potentially insufficient power to detect clinically important effect modification.

As a sensitivity analysis, a conditional logistic regression analysis will be performed using the frequency matched sets.

Limitations

There is a high background circulating prevalence of respiratory and faecal adenovirus in England and Wales during the period of study and we are unable to distinguish pulmonary from extrapulmonary disease in the controls, so we may not see a difference in prevalence between the cases and controls but the difference in response may be immune-mediated such that we cannot differentiate it with this study.

Similarly there is a high background circulating prevalence of SARS-CoV-2 during the period of study so we may not see a difference in prevalence between cases and controls but the mechanism of action may be immune-mediated such that we cannot differentiate it with this study. HHV6 may similarly be a common infection, with potential for an immune-mediated mechanism that cannot be observed within this study.

Data fields that we cannot collect may be involved in the explanatory pathway, such as nursery attendance or levels of mixing prior to illness.

MRU collects EDTA samples. It may be more difficult to collect controls with faecal or respiratory samples. This is not proposed in the current study but cases have tested positive on faecal/respiratory sampling.

The performance of adenovirus and other virus detection tests depends on the assay used (variable lower limits of detection), the type of sample (anatomical site) and timing of the sample with respect to the course of illness. These factors will vary between the cases, but less so between controls since testing of control samples will be done at designated labs using validated in-house or commercial assays on whole blood (EDTA) samples taken early in the control patient's admission episode. Adenovirus testing will be done separately by each UK nation but using as far as possible assays with similar performance.

The analysis is prone to sparse data bias. Sensitivity analysis using a penalised model may be used.

Withdrawal of participants

As residual unlinked anonymised samples will be used for controls, and cases have been ascertained as part of the response to a national public health incident, participants have not been formally recruited and consented to this study, so cannot withdraw. The justification for this approach is:

Cases: Response to a national outbreak of requires establishing enhanced surveillance with consistent reporting to enable a true reflection of disease trends, identification and distribution of risk characteristics, and estimation of burden of disease, under existing legal basis of Regulation 3 (see ethics section below). Understanding the true number of infected persons is important to identify and evaluate control and prevention measures for a national clinical and public health problem.

Controls: Patient identifiable information on controls will be removed so that their adenovirus and HHV6 test results cannot be linked back directly to them. There are no individual or public health benefits or interventions that would be prompted by knowledge of adenovirus and HHV6 testing results.

Number of subjects

Number of subjects is restricted by the number of cases reported in the incident and the number that have had adenovirus testing done on or around the time of admission. In England there are currently 65 cases who meet the inclusion criteria, of which 50 have had adenovirus testing on blood or serum sample. There is limited data on background population prevalence of adenovirus in this age group to estimate the probability of exposure in the control group, and there is seasonality. Data from respiratory samples reported in DataMart from a sentinel network of NHS laboratories suggests that during peak respiratory adenovirus periods positivity in under 5 year olds can be 10 -15%. In 2 subsets of paediatric samples, UKHSA MRU and Micropathology Ltd, a laboratory that undertakes adenovirus testing for NHS hospitals, the blood adenovirus positivity was found to be 10-11% in 2022 to-date.

Considering the limited case numbers, we are planning a study with 3 to 4 frequency-matched controls per case to allow for any attrition due to sample. Prior data indicate that the probability of exposure among controls is 10-15%. We assume a correlation coefficient for exposure between matched cases and controls is 0.2.

Table 1 below shows the sample sizes required by varying the probability of exposure among controls and odds ratio (OR) with 4 controls per case, alpha 5%, power 80% and correlation coefficient 0.2 (6).

Table 1: Sample size calculation with different probabilities of exposure among controls (alpha 5%, power 80% and correlation coefficient 0.2)

Exposure amongst controls	Minimum OR to detect	Case patients	Matched controls (4:1)	Total subjects
5%	2.0	380	1,140	1,520
10%	2.0	210	630	840
15%	2.0	155	465	620
5%	3.0	125	375	500
10%	3.0	73	219	292
15%	3.0	56	168	224
5%	4.0	69	207	276
10%	4.0	42	126	168
15%	4.0	33	99	132
5%	5.0	46	138	184
10%	5.0	29	87	116
15%	5.0	24	72	96

With the projected available sample size of 80 cases and 320 controls (in England), alpha 5%, power 80%, the minimum detectable odds ratio will be 2.7.

Compliance with guidelines

As a public health body, UKHSA data collection role is strictly governed. All data will be collected and handled in accordance with UKHSA guidelines and policy:

- recommendations of the UKHSA Caldicott committee.
- General Data Protection Act (GDPR) and Data Protection Act 2018.
- Human Rights Act.
- Section 3 of the Health Service Regulations 2002.

Devolved administrations will handle data in line with UK guidelines and policy.

Ethical approval

This study has been given UKHSA Research and Public Health Ethics Governance Group (REGG) approval and Caldicott approval.

This study is part of enhanced surveillance which is being performed as part of UKHSA's responsibility to investigate the risk factors associated with severe acute non A-E hepatitis in children of unclear aetiology, as part of the national enhanced incident response. This information will inform national recommendations on testing and management of cases of acute seronegative hepatitis in children. This work has been identified as a public health priority and is being undertaken as part of the UK response to the national increase in cases with this clinical syndrome. The results will be used to provide an evidence base to inform national guidance and public health policy. As such, this work falls outside of the Health Research Authority remit for ethical review. This is in accordance with the revised guidance in the Governance Arrangements for Research Ethics Committees (GAfREC) that was released in September 2011.

UKHSA has legal permission, provided by Regulation 3 of The Health Service (Control of Patient Information) Regulations 2002, to collect and process information without explicit consent.

UKHSA Research and Development and the Sponsors were consulted and confirmed that the work would be covered by Regulation 3 and hence does not require external research ethics approval.

Devolved administrations will ensure confidential data are collected and processed according to their legal permissions.

Participant confidentiality

Personal data collected for the purposes of this study may include name, date of birth, age, sex, ethnicity, NHS number, address, contact information. The only people with access to this information will be the surveillance staff, or regulatory authorities who may wish to check the surveillance is being carried out according to appropriate guidelines. Data will only be used for the purposes of this surveillance, stored in secure UKHSA facilities with restricted access.

Target dates

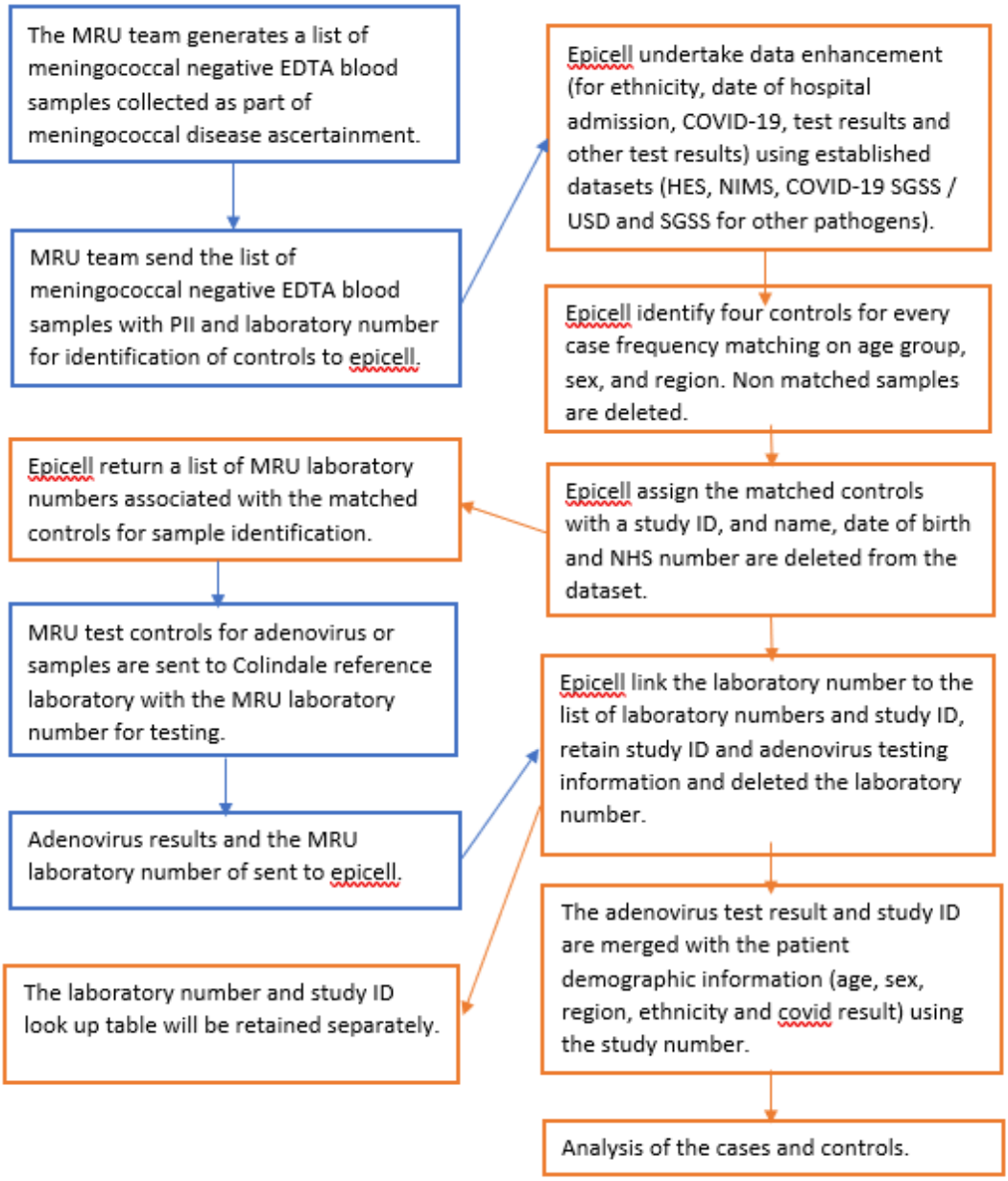
- cases and controls to be identified and matched: 13 May 2022
- control samples to be tested: 20 May 2022
- completion of preliminary analysis: 14 June 2022

References

1. Marsh K, Tayler R, Pollock L, Roy K, Lakha F, Ho A, Henderson D, Divala T, Currie S, Yirrell D, Lockhart M, Rossi, M K. and Phin, N. 'Investigation into cases of hepatitis of unknown aetiology among young children, Scotland, 1 January 2022 to 12 April 2022'. Eurosurveillance 2022: volume 27, page 2200318
2. UK Health Security Agency 'Investigation into acute hepatitis of unknown aetiology in children in England'. Technical briefing 25 April 2022; updated 6 May 2022
3. European Centre for Disease Prevention and Control. ECDC Update: Hepatitis of unknown origin in children April 2022
4. Kuo C-L, Duan Y, Grady J Unconditional or Conditional Logistic Regression Model for Age-Matched Case–Control Data? Frontiers in Public Health 2018 vol 6
5. Greenland S, Mansournia M A, Altman D G. Sparse data bias: a problem hiding in plain sight BMJ 2016; 352 :i1981 doi:10.1136/bmj.i1981
6. Stevenson M, Nunes T, Heuer C, Marshall J, Sanchez J, Thornton R, Reiczigel J, Robison-Cox J, Sebastiani P, Solymos P, Yoshida K, Firestone S. epiR: An R package for the analysis of epidemiological data. R package version 0.9-69.

Appendix 1. Data flow (UKHSA)

.



Note: same flow and processes applies to HHV6 results

Colour coding:

UKHSA MRU

UKHSA Epi cell

Appendix 2. Protocol variations for devolved administrations

A2.1 Scotland

Control specimen samples will be sourced from archives of nuclei acid extracted from whole blood.

Recruitment in May 2022 will be included within the March and April band.

Data sources include:

The Turas Vaccination Management Tool (TVMT).

The Electronic Communication of Surveillance in Scotland (ECOSS) database.

Scottish Index of Multiple Deprivation

A2.2 Wales

Recruitment of controls in Wales

Recruitment criteria are identical to UKHSA with the following changes outlined:

Controls will be recruited from residual EDTA blood samples collected from those aged below 6 years at specimen date for an acute non-hepatitis illness which tested negative for that pathogen (for example tested negative for meningococcal infection).

Inclusion criteria:

Resident of Wales

AND

aged above 12 months and below 6 years old

Exclusion criteria:

insufficient volume (<400 microlitres) in EDTA sample to test for adenovirus

Matching of controls to cases

Those that meet the control criteria will form a sampling frame, and controls will be frequency-matched to cases on sex and month of sample collection (January - February, March - April*).

Welsh cases have a range of 1 to 5 years (inclusive) and so controls will be selected from a pool of the same age group (see inclusion criteria). Finer age-category matching will be explored.

Up to 6 controls per case will be evaluated with the aim of achieving a final ratio of 1:3.

*One case tested on 2 May 2022 has been included in the March to April band.

Testing of controls in Wales

Welsh control samples will be tested for adenovirus and HHV6/7 locally. This will be performed using a combination of a qualitative commercial multiplex assay, an in-house pan adenovirus assay and a quantitative commercial adenovirus assay.

A2.3 Northern Ireland

Procedures:

Cases

Cases will be drawn from the Northern Ireland acute hepatitis non-A-E linelist and will include those who have documented evidence of adenovirus testing at or around the time of emergency department attendance or hospital admission.

Cases will be placed in the categories formed by; aged from 1 to 5 years inclusive or aged 6 to 10 years inclusive, males or females, according to region (as being from Northern Ireland), and sample date according to calendar months of January, February, March, April or May.

The minimal dataset for each case will include age, sex, deprivation quintile (according to postcode), date of hospital sample collection (as a proxy for hospital admission and onset of symptoms), and study ID number (used for anonymised reporting of findings to the UKHSA study team).

Controls

Controls will be “recruited” from patients presenting with an acute illness for whom a blood EDTA sample was submitted to the Regional Virology Laboratory (RVL) in Northern Ireland. Controls will be frequency matched to cases by age band, sex, region (as Northern Ireland), and month of sample collection.

Controls will be placed into the same categories as cases, and within each category the controls will be allocated a unique number to form a category specific sampling frame. For each category, the number of cases will be calculated and 4 times this number of controls within the category will be selected using simple random sampling. If within a category there are fewer controls than 4 times the cases, then all controls will be selected.

The minimal dataset for each control will include (as per cases) age, sex, deprivation quintile (allocated according to postcode), date of hospital sample collection (as a proxy for hospital admission and onset of symptoms), and unique study ID number (used for anonymised reporting by the RVL to the Public Health Agency and UKHSA study team).

As samples are unlinked and anonymised, results cannot be given to participants and no opt-in is required. Residual samples will be stored securely by the RVL and will be discarded 12-months after completion of testing and data analysis.

Laboratory tests

EDTA blood plasma samples of at least 400µl from eligible controls will be unlinked and anonymised before being tested for adenovirus testing at the RVL.

Adenovirus testing of samples from controls

Plasma samples from control samples will be tested for the presence of recent adenovirus infection (at the time of presentation / hospital admission) using validated the CE-marked Altona adenovirus PCR assay usually used to provide the normal clinical service. Any adenovirus positive samples will be referred for typing at UKHSA Colindale.

Adenovirus testing of samples from cases

Adenovirus testing of case samples is being done as part of clinical management in response to an outbreak. This has been recommended by the UK national incident management teams. Plasma samples will have been tested using the CE-marked Altona adenovirus PCR assay in Regional Virology Lab (RVL).

Unlinking and anonymising controls

The following steps detail how unlinking and anonymising controls will be achieved in N. Ireland:

- the Regional Virology Lab (RVL) will identify meningococcal test requests with negative results between 1 January 2022 to 1 May 2022
- relevant sample details will populate a spreadsheet including specimen number of the sample, sample date, the patients' date of birth, sex, deprivation quintile (according to postcode) and COVID-19 status
- this will be converted the following categories: unique study ID, month of sample test, age, sex, which will be encrypted with password protection and shared with the Public Health Agency (PHA), including no PII, for purposes of matching
- PHA will assign frequency matched controls to cases using this list according to age band, sex and date of sample. Four controls to cases will be used to increase power and efficiency because of the small number of cases and the risk of insufficient residual sample in controls
- the list will be returned to RVL, for sample retrieval and testing
- the RVL laboratory number for each specimen and corresponding study ID for that individual will be stored separately
- the master spreadsheet will have the specimen number removed after the stored samples have been assigned the study ID, thereby anonymization is complete, prior to samples being tested for adenovirus
- RVL will test the control samples for recent adenovirus infection
- the adenovirus results will be linked by RVL with the demographic information (age, sex, deprivation quintile) and COVID-19 diagnosis information using the study ID
- the spreadsheet will be encrypted with password protection and emailed to PHA.

Data Management

Data on cases and controls will be collated by the PHA case-control study team.

1. Data transfer

Anonymised data will be shared with the UKHSA EpiCell following local processing as detailed above, and appropriate cleaning procedures.

2. Storage and processing of data

Data on cases are stored on PHA servers, with access restricted to approved individual users in the study team.

3. Data Linkage

Data will have been linked by the RVL according to the processes above, based on laboratory specimen number, which will have been removed prior to testing.

The PHA will receive data for controls based on unique study ID and linkage will be with cases based on the minimum dataset provided by RVL.

Appendix 3. England frequency matching table

Region	Age group	Month (grouped)	Sex	Number of cases	Number of controls required*
Midlands	1-5	1-2	M	2	8
Midlands	1-5	3-4	F	5	20
Midlands	1-5	3-4	M	5	20
North	1-5	1-2	F	3	12
North	1-5	1-2	M	2	8
North	1-5	3-4	F	5	20
North	1-5	3-4	M	6	24
North	6-10	3-4	F	1	4
North	6-10	3-4	M	2	8
South	1-5	1-2	F	5	20
South	1-5	1-2	M	2	8
South	1-5	3-4	F	10	40
South	1-5	3-4	M	6	24
South	6-10	1-2	F	1	4
South	6-10	3-4	M	4	16

*number of controls required based on 1:4 matching to cases where possible

Appendix 4. Proposed data dictionary

Column	Description
CC_ID	ID
Type	Case/Control
Region_grouped	DA Name: Scotland/Wales/N Ireland
ethnicity_grouped	Ethnicity: White/not White
Ethnicity_detailed	White/Black/Asian/Mixed/Other/Not known
Age_grouped	1-5/6-10
Sex	M/F
Age	years (integer)
Postcode	
Deprivation_decile	1 = most deprived; 10 = least deprived
Adenovirus	Any sample: Detected/Not Detected
Adenovirus - level unknown sample	Detected/Not Detected
If Adenovirus vireamia, peak DNA levels (copies/ml)	number
Adenovirus - Stool Sample	Detected/Not Detected
Adenovirus - Respiratory Sample	Detected/Not Detected
Adenovirus - Blood Sample	Detected/Not Detected
Adenovirus - Other Sample	Detected/Not Detected
SampleDate_case_adeno	Date of case adenovirus blood sample
SampleDate_case_HHV6	Date of case HHV6 sample
EarliestPresentationDate	Date of earliest presentation
SampleDate_control	Date of control sample
HHV6	Detected/Not Detected/NK
Week	ISO week (integer)
Month	Month of presentation/sample (integer: 1,2,3,4)
COVIDTEST_14day_result	result (Detected/Not Detected) in 2 weeks prior to illness or at time of admission (up to 2 days after). Detected reported over not detected results if both within that period.
COVIDTEST_90day_result	Result (Detected/Not Detected) in 90days prior to illness or at time of admission (up to 2 days after). Detected reported over not detected results if both within that period.
COVIDTEST_firstpos_date	Date of first ever SARS-CoV-2 positive test
COVIDVAX1_date	dose 1: date

Column	Description
COVIDVAX1_type	dose 1 brand name: Pfizer/Moderna/AstraZeneca
COVIDVAX2_date	dose 2: date
COVIDVAX2_type	dose 2 brand name: Pfizer/Moderna/AstraZeneca

About the UK Health Security Agency

UKHSA is responsible for protecting every member of every community from the impact of infectious diseases, chemical, biological, radiological and nuclear incidents and other health threats. We provide intellectual, scientific and operational leadership at national and local level, as well as on the global stage, to make the nation health secure.

UKHSA is an executive agency, sponsored by the Department of Health and Social Care.

Version 2.0

Published: May 2022
Publishing reference: GOV-12291

UKHSA supports the UN Sustainable Development Goals